

I-023

大規模言語モデルを用いた制作者の意図に基づくオブジェクトの 自動選定・配置システムの提案

西居有花[†] 石戸莞楽[†] 中村亮太[†]

1. はじめに

3DCG ソフトウェアによるバーチャル空間は、ゲーム開発から映画製作、建築設計まで、さまざまな分野で利用されている。その一方で、バーチャル空間の生成には相当のスキルと時間が必要とされる。具体的には、3DCG ソフトウェアの操作に関する前提知識を理解していることが求められ、数多く存在するオブジェクトの中から、自分が意図している特定のオブジェクトを見つけ出すだけでも膨大な時間がかかることがある。さらに、このオブジェクトを探し、配置する作業は 3DCG ソフトウェアを扱い慣れていない人にとっては複雑である。

本研究では、3DCG ソフトウェアで構築した空間内で、大規模言語モデルである LLM (Large Language Model) を活用して、ユーザから入力されたテキストから制作意図を推定し、推定結果に基づいて、オブジェクトのリストの中から最も適したものを自動的に選択し、配置するシステムを提案する。本システムにより、3DCG ソフトウェアに詳しくない制作者でも、対話しながら容易にバーチャル空間を制作することが可能になる。これによって、バーチャル空間の制作に対する敷居が大幅に下がり、より多くの人々がバーチャル空間を活用できるようになることが期待される。

2. 関連研究

本研究に関連するシステムとして「Versy.ai」という生成 AI が存在する[1]。Versy.ai はユーザがシンプルなテキストプロンプトを入力すると、プロンプトの情報を基にバーチャル空間を生成するという「Text-to-Space」技術を搭載している。Text-to-Space 技術を活用すれば、手間と時間がかかる仮想空間の手動作成を大幅に削減することが可能である。さらに、Versy.ai は、CRM、AI、IoT などのさまざまなデータソースへ簡単に接続できる柔軟性を持っており、ダイナミックなバーチャル体験の構築に活用できるという特長がある。また、Unreal Engine 公式の講演では、和室を自動配置するプロシージャル技法が公開されている[3]。プロシージャル技法を使用するために、この講演ではオブジェクトの制作から行なっている。そのため点群を利用することにより、オブジェクト同士がぶつからないように、また適した位置関係で配置することができる。

しかし、Text-to-Space 技術やプロシージャル技法においては、使用する際にオブジェクトを事前に学習させる必要があり、新たな外部のオブジェクトを追加する場合には、その都度、再学習が必要となる。そのため、容易に外部のオブジェクトを使用することができないという課題がある。そこで本研究では、外部のオブジェクトに対応でき、再学習の必要がないという条件を満たす Auto 3D Selector システムの開発を行った。石戸ら[2]による研究では、マルチモーダルモデルの CLIP を用いた 3D モデルの自動選択手法が提案されている。しかし CLIP を用いると説明文を生成する際にオブジェクトの情報だけではなく”on_top_of_a_tiled_floor”という背景の情報を含んでしまい、精度の荒さが目立ってしまう(図 1)。この問題を防ぐため、本研究では CLIP を用いずに LLM で代用した。自身で送るプロンプトを設定することにより、オブジェクトのみに絞って説明文を取得できる(図 2)。



図 1 CLIP を用いた場合

図 2 LLM を用いた場合

3. Auto 3D Selector システムの提案

ユーザが意図するオブジェクトに関連する単語を入力することで、オブジェクトを自動で配置する Auto 3D Selector システム(図 3)を提案する。本システムは、3DCG ソフトウェアに詳しくない制作者でも容易にバーチャル空間を制作することを目的としている。使い方は、表示されている質問の通りテキストを入力した後に送信ボタンを押す、というとてもシンプルな操作だ(図 4)。送信ボタンを押すと自動でオブジェクトが目の前に配置されるシステムになっている。

[†] 武蔵野大学データサイエンス学部
Faculty of Data Science, Musashino University

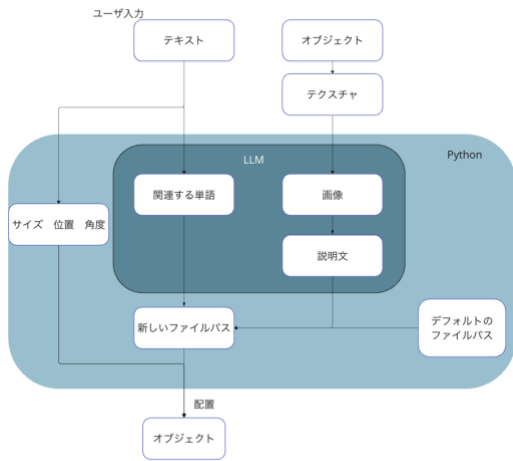


図 3 システム構成図

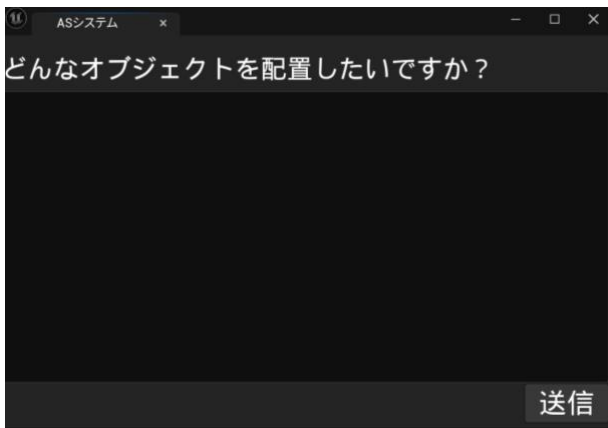


図 4 テキスト入力画面



図 5 無意味なファイル名の例

4. オブジェクト自動選定能の実装

オブジェクト自動選定機能を実装するための手順を示す。

4.1 オブジェクトの画像変換

まず、3DCG ソフトウェアにダウンロードされているオブジェクトを画像キャプションモデル (ThumbnailToTextureTool) に送信し、そのテキストチャを取得する。取得したテキストチャを PNG 形式でエクスポートする。この工程では、オブジェクトを画像に変換することで、後述のマルチモーダル AI に読み込ませるための準備を行う。

4.2 画像の説明文生成

次に、OpenAI の API を使用して、4.1 で得られた画像と「Please answer based on the characteristics of the objects in the photo, such as color and material. No background explanation needed. Please answer using English words only. Please replace blanks and spaces with "_". Please don't answer otherwise.」というテキストメッセージを LLM に送信し、それぞれの画像の説

明文をリスト取得する。LLM からは、画像に描画されたオブジェクトの色や素材など、詳細な説明が含まれた結果が得られる。例えば、図 2 のような白いイスのオブジェクトの画像を読み込ませた場合、"white_plastic_chair_wooden_legs" のように、「白い」、「プラスチック」、「イス」、「木の足」という情報が得られ、画像からテキストへの変換が行われる。この工程により、外部からダウンロードしたオブジェクトにも対応可能となる。

4.3 ファイル名の変換

上記 4.2 で得た画像の説明文を元に、元々のオブジェクトのファイル名を変更する。これにより、オブジェクトとファイル名の対応付けを行う。後述のオブジェクトの自動配置を行う際、このファイル名を使用するのでこの時点でファイル名を変更する。特に外部からダウンロードしたオブジェクトは、"staticmesh_1" や "chai_1" などの無意味なファイル名になっていることが多いため、この処理が必要となる (図 5)。

4.4 関連単語の取得

ユーザが使用したいオブジェクトの特徴を LLM に送信し、それに関連する単語の一覧を取得する。例えば、「イス」と入力した場合、"chair" だけでなく、"sofa" や "stool", "ottoman" などの同じ役割を果たす単語も取得される。また、「白いイス」と入力した場合は、"white_chair" だけでなく、"ivory_chair" や "snow_stool" といった、色や素材を限定した単語も返されるようになっている。関連する単語を取得することで、ファイル名が完全に一致していなくてもおおよそ意図しているオブジェクトが該当するようになる。

4.5 ファイルパスの取得

4.4 で取得した単語が含まれているファイル名を、PC のファイルの中から全て取り出し、そのファイルパスを取得する。3DCG ソフトウェアの検索バーで「white_chair」と入力して検索しても、4.2 で「ivory_chair」とファイル名が変更された場合には検索に引っかからない。その現象を解消するためにこの工程を行なっている。

5. オブジェクト自動配置機能の実装

上記4.5で取得したファイルパスを使用して、該当するオブジェクトを3DCGソフトウェア上に自動配置する。今回は、サイズ、位置座標、回転の角度は事前に指定しておいた。

6. 評価実験

本システムの有効性を検証するため、19代から30代の男女14名を対象に評価実験を行った。実験参加者には、本システムを使ってオブジェクトを配置した場合と、システムを使わずに手で配置した場合の2つの方法で、指定されたオブジェクトを配置してもらった。実際行ってもらった流れを図6～図9に示す。図6は、システム実行前の3DCG空間である。ここに、ユーザが意図するオブジェクトを配置していく。システムを使わなかった場合、図7のようにダウンロードされている膨大な複数のオブジェクトの中から見つけなければならなかった。次に、Auto 3D Selectorシステムを使用し、「白」という単語を入力した結果が図8である。白い紙袋、白い椅子、白いテーブルなど、「白」に関連するオブジェクトが自動で選定され、空間内に配置されている。次に、「白いイス」とより限定した単語を入力した結果が図9である。図8とは違い、白いイスのみに絞り込んで配置することに成功している。この流れを実験参加者に実際行ってもらった。



図6 実行する前の画面

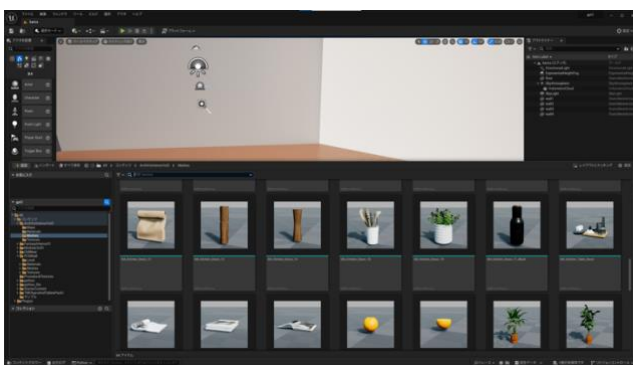


図7 システムを使わなかった場合



図8 「白」の実行画面



図9 「白いイス」の実行画面

その後、以下の4つの質問について5段階評価でアンケートを実施した。

- Q1. 3DCGソフトウェアの使い方が分からなくても操作できたか
- Q2. システムを使用した際に、自身が意図していたオブジェクトが配置されたか
- Q3. システムを使わずにオブジェクトを配置するには、手間暇がかかると感じたか
- Q4. 実際にあれば便利だと思うか

Q1の結果にはばらつきが見られた。(図10)これはソフトウェアを扱ったことがある人や全く触ったことのない人がいるという結果だ。Q2については、85%の実験参加者が「とてもそう思う」または「ややそう思う」と肯定的に回答した。

(図11)一方で、約15%の実験参加者は「ややそう思わない」と回答しており、その理由を尋ねたところ、「白いイス」と入力した際に、単純なイスではなくソファに近いオブジェクトが配置されたことや、黄色いクッションが含まれていたことが挙げられた。この問題が起こる原因は、「Armchair_White_Yellow_Pillow_Wooden_Flame」といったファイル名にある。例えば「白いイス」と入力した際だけでなく、「黄色い椅子」と入力してもこのオブジェクトが配置されてしまう可能性がある。より意図するオブジェクトを選

択するためには、複数のオブジェクトを提示し、その中から選べるようなシステムの導入や、4.2 の LLM から得られる結果の改善により、限定したオブジェクトが配置できるようになると考えられる。Q3 と Q4 については、それぞれ約 85%、約 90%の実験参加者が「とてもそう思う」、「ややそう思う」と回答しており (図 12, 図 13)、本システムによってオブジェクト配置の手間が大幅に削減できること、また 3DCG ソフトウェアに詳しくない人でも扱いやすいことが示された。このシステムは入力したテキストに関連したオ

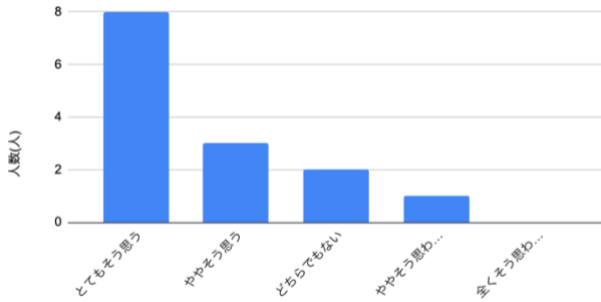


図 10 Q1 のアンケート結果

(Q1. 3DCG ソフトウェアの使い方が分からなくても操作できたか)

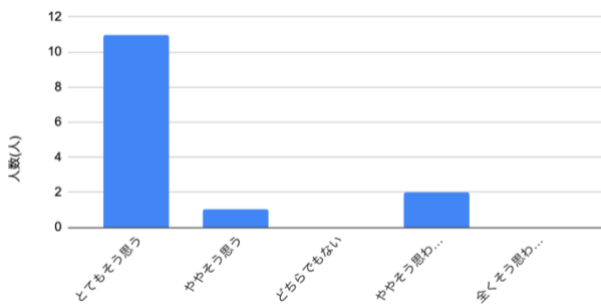


図 11 Q2 のアンケート結果

(Q2. システムを使用した際に、自身が意図していたオブジェクトが配置されたか)

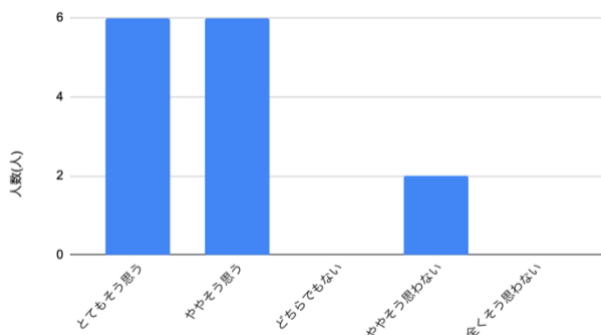


図 12 Q3 のアンケート結果

(Q3. システムを使わずにオブジェクトを配置するには、手間暇がかかると感じたか)

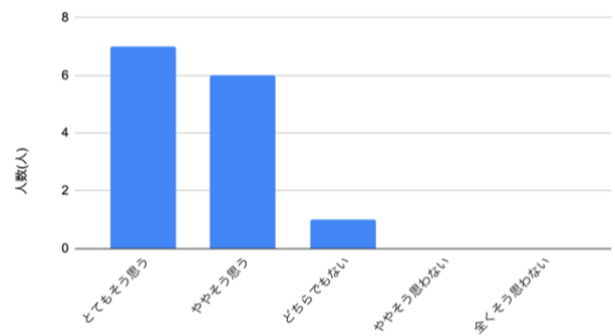


図 13 Q4 のアンケート結果

(Q4. 実際にあれば便利だと思うか)

ブジェクトは配置できるが、そのオブジェクトの移動、回転の角度は事前に指定していたため変更ができない。

7. おわりに

本稿では、3DCG ソフトウェアで構築したバーチャル空間内で、LLM を活用してユーザの制作意図を推定し、それに基づいて適切なオブジェクトを自動選定・配置する Auto 3D Selector システムを提案した。本システムにより、3DCG ソフトウェアを使いこなせない人でも、簡単にオブジェクトを配置できるようになった。さらに、外部オブジェクトにも対応させることで、システムの汎用性を大幅に向上させた。評価実験の結果、本システムの有効性が確認された一方で、オブジェクトの選定精度にはまだ改善の余地があることがわかった。また、システムの実行に複数の工程を要するため、処理に時間がかかるという課題も明らかになった。今後は、これらの問題を解決するために、LLM から得られる結果の改善や、ユーザインタフェースの改良などを進めていく予定である。将来的には、本システムをベースとして、誰もが直感的に扱えるバーチャル空間制作ツールの開発につなげていきたい。そしてこのシステムの限界であった複数のオブジェクト同士の位置関係も考慮した自動配置ができるシステムも開発していきたいと考えている。

謝辞

本研究は JSPS 科研費 JP21K01749 の助成を受けたものである。

参考文献

- [1] "Text-to-Space Create Virtual Experiences with Versy.ai". future en. <https://futureen.com/tool/versy-ai/> (参照 2024-6-14) .
- [2] 石戸莞楽, 中村亮太. 「3次元 Web 空間制作支援のための大規模言語モデルを用いた 3D モデル自動選択手法の提案」2024IPS J 全国大会.
- [3] "プロシージャル技法による背景の自動生成 『1,000 の和室』". エピック・ゲームズ・ジャパン Epic Games Japan. <https://www.slideshare.net/slideshow/ue4fest-sqexjapanese-styleroom/142312903> (参照 2024-6-14) .