

XR（クロスリアリティ）時代における、

映像表現および映像制作の可能性について考察する

大阪芸術大学 映像学科 准教授 佐藤貴雄

本研究では、生成AI技術を活用し、コンセプトアート制作から3D空間の作成、撮影における多様なアングルハンティングまでを対象として、VRやARも視野に入れた映像制作の可能性を検証することを目的とした。そのため、複数の生成AIアプリケーションを用い、とりわけ制作プロセス上の有効性を探る検証を行った。

まず、テスト用空間のビジュアルイメージを生成するため、実在しても不自然ではない現実的な空間から、ファンタジー色の強い幻想的な空間まで、コンセプトアートとして多様なパターンの画像を作成した。現在、画像生成AIは日々更新され、種類も把握しきれないほど増えているため、すべてを試すことは難しいが、いくつかの候補を比較検討した。その結果、「Gemini (nano banana)」と「Midjourney」という2種類の画像（動画）生成AIを中心に使用した。現実的な表現を得意とする「Gemini」と、アーティスト的な表現を得意とする「Midjourney」は、表現の方向性が比較的明確であり、この対比が生成サンプルの幅を広げると考えたためである。生成された画像は多岐にわたり、さまざまなイメージをプロンプトとして与えることで、手間をかけずに多様なシーンをテスト作成できた。言葉だけでなく、実際に視覚化して検討できる点は非常に有用かつ効果的であった。

次に、画像としてコンセプトアートを生成・作成した後、それらを3Dデータへ変換する作業が必要となる。画像からの3Dデータ変換についても、画像生成と同様に無数の3Dデータ生成AIアプリケーションが存在するため、いくつかをテストして検討した。大別すると、静止画像から3Dデータを生成するものと、動画から生成するものに分けられる。まず、静止画から生成可能な「Tripo3D」をテストした。これは1枚の静止画だけでなく、前面、背面、右面など複数方向からの画像入力が可能で、高精度な3Dデータを作成でき、4Kサイズのテクスチャーも生成可能なアプリケーションである。しかし、本検証では期待した結果を得られなかった。理由として、このアプリケーションが主に人物やキャラクターなどのオブジェクト生成を想定しており、外向きに情報が連続するような空間をうまく認識できないためと考えられる。建築物であれば十分に生成できたことから、サイズの問題ではないと思われる。

続いて、動画から3Dデータを生成する「Vid2scene」をテストした。空間内でのカメラ移動（前方へのドリー移動や横方向へのパン）を設定した動画を生成し、それを用いて3Dデータ生成を行ったが、こちらも不十分な結果に終わった。元の動画が一見問題ないように見えても、生成AIによって作成された動画では、時間の経過と

ともに空間の形状や位置関係、さらにはレンズ特性までもが少しずつ変化してしまい、「立体としての矛盾」が生じている可能性がある。

その後、静止画から3Dデータを生成する「Marble」をテストした。これは前述のアプリケーションとは異なり、入力された静止画像をいったん画像を構成するプロンプトに分解し、それを用いて空間を再生成・再構成する手法を採用している点が特徴である。本検証では申し分のない3Dデータが生成できたため、「Marble」をメインにデータ作成を行うことにした。さらに、ARや撮影時のプレビズ（事前シミュレーション）用途として、実際の教室のアセットを準備した。空間のキャプチャーには、将来的に授業へ組み込むことも考慮し、学生でも気軽に使用できるスマートフォン用アプリケーションを検討した。その結果、3Dスキャンアプリケーションとして「Scaniverse」と「Luma3D」を候補に挙げ、双方をテストした。

「Scaniverse」は教室のような広い空間では、スキャンを続ける間にズレが生じやすく、比較的小規模な空間に向いているようであった。一方で「Luma3D」はズレが比較的目立ちにくいと感じられた。もちろん、テストに使用したスマートフォンの性能やスキャン方法の影響も考えられるが、今回は「Luma3D」をメインとした。このようにして作成した3Dデータ（Gaussian Splatting）を、UnrealEngineにプラグインを介して読み込み、ステージを作成したうえで、「Unreal Vcam」というスマートフォン用のバーチャルカメラアプリケーションを使用した。これにより、スマートフォン自体を3D空間を撮影できるカメラとして利用できる。通常の撮影と同様にスマートフォンを動かすと、3D空間内のカメラが連動し、記録や録画も可能である。さまざまなステージでテストした結果、完成形の映像にそのまま用いるには、今回作成した形式の3Dデータ（Gaussian Splatting）の特性上、アーティファクト（ノイズ）が多く、現時点では実撮影には適さないと考えられた。一方で、直感的な操作でその空間に立って眺めるように検討できる点は、カメラワークのアングルハンティングや舞台セットの検討など、映像プランニングにおいて非常に有用であると考えられる。以上の検討結果から、映像制作において生成AIは、従来、手間と時間を要していたコンセプトアート作成から、撮影のための舞台セット準備、映像プランニングにおけるアイデア出しや検討までを、AIをアシスタントとして活用することで効率化し得ることが示唆された。言葉や絵だけでは見えにくかった映像の可能性を、実際に目に見える形でブレインストーミング的に広げられる点で、非常に有用な技術である。今後の映像制作および教育現場においても、新しい映像表現の発見や映像体験につながる可能性があるため、引き続き応用を考察したい。