

AI 生成画像と夢の視覚表現の類似性研究 ——アニメーションにおける生成美学について

東京藝術大学大学院映像研究科アニメーション専攻

学籍番号 3225357

スウ ナン／Zou Nan

AI 生成画像と夢の視覚表現の類似性研究——アニメーションにおける生成美学について

A Study on the Similarities Between AI-Generated Images and Dream Visual Expression: Generative Aesthetics in Animation

要旨：

本研究は、夢と人工知能生成画像の比較分析を通じて、両者の生成メカニズムおよび視覚表現における類似性を検討するものである。あわせて、筆者の作品『新ルネサンス』および『天眼』の制作実践を基に、AI 生成過程において生じる不安定性、認識の偏差および制御不能性が、いかに生成的美学の手法へと転化されうるのかを分析し、アニメーション制作における新たな視覚表現の可能性を提示する。

1. 序論

これまでの自身のアニメーション制作において、私は無意識のうちに「夢」と人工知能によって生成された画像とを結びつけてきた。当初は両者の関係性について明確に意識していなかったが、制作を重ねる中で、両者のあいだにはある種の共通性が存在するのではないかと感じるようになった。

AI 生成画像はしばしば、曖昧で、漂流し、不安定で、時に未完成のように見える視覚状態を呈する。一方、夢の中における視覚体験もまた、変形や断裂、不連続性に満ちている。これらはいずれも、現実と虚構のあいだに位置する「非現実感（Unreality）」を帯びた視覚表象であり、現実世界を忠実に再現する写実とも、完全に現実から切り離された純粋な幻想とも異なる性質を持っている。

夢における視覚経験は、現実世界の単純な複製ではなく、記憶、感情、外部からの刺激が再編成され、誇張された結果として立ち現れるものである。そのため、夢の映像には明確な時間的論理や因果関係が欠如し、登場する人物や空間は、どこか見覚えがありながらも判別しがたい姿で現れる。また、現実世界における物理法則が容易に破綻する点も、夢の特徴の一つである。

同様の現象は、AI 生成画像においても観察される。画面内の主体は一見すると合理的に見えるものの、その形態を安定して維持することができず、部分的なディテールは極めて写実的であるにもかかわらず、全体的な関係性は揺らぎ続ける。生成された画像は、「完成しつつある状態」と「制御不能な変形」とのあいだを往復するように存在している。

こうした観察を通して、私は夢と AI 生成画像との関係性を、表層的な印象ではなく、その生成メカニズムの次元から捉える必要があると考えるようになった。夢は理性的な意識による能動的な創作ではなく、無意識の状態において既存の視覚経験が断片化され、再構成される過程によって生じるものである。一方、生成式 AI による画像生成もまた、現実世界の物理構造や因果関係を「理解」した上で行われるものではなく、大量の画像データの学習を通じて、統計的に既存の視覚パターンを再配置することで成立している。

両者はいずれも、世界に対する認知的推論を基盤とするのではなく、確率、連想、記憶の痕跡によって像を生成する点において共通している。

以上の背景を踏まえ、本論文では「夢」と「AI 生成画像」を並置し、視覚生成のメカニズムおよび美学的特性の観点から、生成的映像表現がアニメーション制作にどのように応

用されうるのかを考察する。具体的には、夢の生成原理およびアニメーションにおける夢的視覚表現の分析、さらに AI 画像生成の原理的構造の整理を行い、それらを私自身の制作作品である『新ルネサンス』および『天眼』の実践と結びつける。これらを通じて、AI 生成過程において顕在化する不安定性、誤認識、制御不能性が、夢、記憶、幻覚といった主題を表現するための有効な美的手法となりうることを明らかにすることを目的とする。

2. AI 生成画像と夢の生成メカニズム

2-1. AI による画像生成のメカニズム：復元・推論・補完

AI を用いて画像を生成する際、多くの場合、生成中の画像がモザイク状から徐々に鮮明になっていく過程が観察される。AI 画像生成（例えば Midjourney で使用されるモデル）の核心原理は拡散モデルである。人間が物体構造や因果関係を理解して描画するのとは異なり、拡散モデルは直接「描く」のではなく、段階的にノイズが加えられた実際の画像を復元する方法を学習する。

学習段階では、モデルは大量の画像に反復的に接触し、異なるノイズ強度の下で元の画像を最も可能性の高い形態に予測することを学ぶ。生成段階では、この過程が逆に用いられる：モデルは完全にランダムなノイズ画像から開始し、テキストプロンプトなどの条件下で、複数回の反復を通じてノイズを除去し、モザイク状の状態から全体の構図、主体の輪郭、最終的なディテールへと徐々に顕現させる。

情報が不確定な場合、モデルは推論を行い、統計的に補完を行う。例えば img2img では、元画像がぼやけていたり抽象的であったり、プロンプトと矛盾していたり、文脈情報が不足している場合、AI は推論を行い、モデル内で頻度の高い物体を用いて補完する。補完は現実の復元を目的とするのではなく、統計的な意味で合理的な結果を目指す。これが AI 生成時に見られる「AI 幻覚（hallucination）」の正体である。

また、画像から画像、あるいは連続するフレームを用いた変換では、隣接するフレーム間の差が極めて小さくても、生成結果に顕著な変化が生じる場合がある。各フレームは潜在空間内で独立に意味的リサンプリングが行われ、潜在空間内部の意味的境界は視覚的連続性を保証しない。この「個別フレームは合理的だが、全体は制御不能」という特性が、AI 生成画像における動的映像でのチラツキ、不安定性、漂流の視覚状態を生む。

2-2. 夢の生成メカニズム：記憶の断片・連想・非線形再構成

フロイトは夢を「潜在意識への広い橋」と表現し、夢の内容は記憶や欲望の断片が加工・再構成されることで生じるとした。彼は夢の構成メカニズムとして、凝縮（複数のイメージの統合）や転置（忘れられた細部を異なる形で再現する）を挙げ、夢における時間・空間の混沌、イメージの曖昧さの特徴を説明した。

これらのメカニズムにより、夢の映像は現実世界の論理構造から逸脱し、感情や連想の運行に従う。心理学や美術史の研究において、夢は非理性的かつ非線形の映像生成過程として理解される。覚醒時の論理的・因果的認知とは異なり、夢は記憶の断片、情動体験、外部刺激が相互作用し、無意識下で再構成される。

サルバドール・ダリの『目覚める直前、ザクロの木を飛ぶミツバチが引き起こす夢』（図1）は、このメカニズムを視覚的に表現した典型例である。作品内の飛ぶミツバチ、ザクロ、突如現れる虎や銃剣は、現実における連続的な物語からではなく、微小な現実刺激（ミツバチの羽音）によって誘発され、夢の中で増幅・変形・再構築され、高度に緊張感のある、論理的に断絶した画面が生まれる。夢の中のイメージは現実世界の因果関係に従わないが、知覚的には「夢のリアリティ」として自己完結する。



図1

夢における人物や場面は、既視感があるが具体的には判別できない状態で現れる。空間は一見連続しているように見えるが、瞬時に断裂や飛躍が生じる。この不安定さは夢の欠陥ではなく、夢という生成システムの本質であり、形式の安定性を放棄しながら意味を生成する性質に由来する。

2-3. AI 生成と夢の共通特性

2-3-1. 生成メカニズム

AI 生成画像の拡散過程と夢の生成メカニズムには多くの類似性がある。いずれも既存の視覚経験を断片化・再組み合わせし、既視感がありつつも新しいイメージを構築する点で共通している。

フロイトの転置および凝縮を AI 画像生成に対応させて整理すると以下ようになる。

転置：意味が不確定な場合、生成の重心を高頻度・安定した視覚要素に移すことで、現実と虚構の間に位置する幻覚的なイメージを生む。

凝縮：AI は生成時、複数の学習サンプルから視覚的特徴を圧縮して 1 つの結果に反映させるため、そのイメージは既視感を持つが特定の何かではない印象を与える。

生成の観点から見ると、AI 映像と夢は偶然に似ているのではなく、「現実に基づく組み合わせ」と「非確定的生成」という論理基盤を共有しており、確率的・連想的な映像生成により、夢のような非現実的・幻覚的な視覚体験が生まれる。

2-3-2. 現実と虚構の間に位置する視覚表現

因果律の不在

夢と AI 生成映像の顕著な共通点は、現実世界の物理的因果関係に従わない点である。夢においては光源方向、物体のスケール、空間の深度、人物と環境の関係が現実と一致しないことが多いが、知覚上は成立する。同様に AI 生成画像も、一見整合しているように見えて、細部では因果の断絶（光源矛盾、遠近関係の混乱、物体の支えの欠如など）が生じる。これは生成過程が物理的世界の理解に基づかず、視覚パターンの統計的関連に依存しているためである。

例えば筆者の修士一年次作品『天眼』では、Midjourney で生成した広告中の女性看護師像を用いた。プロンプトは「1980 年代、眼科診療所で働く眼鏡をかけた中国人若年女性看護師の胸像」である（図 2）。この条件下で生成された画像は一見合理的に見えるが、眼鏡の位置が浮いていたり、目の透視が不自然であったり、読めない漢字の視力表など論理的に破綻した要素が存在する。



図 2

夢における視覚表現でも、今敏のアニメ『パプリカ』において、人物や場面が途切れなく変化し、空間構造が崩壊・再構築され、現実の因果関係が完全に停止する様子が描かれる（図 3）。映画のクライマックスでは、街中にテレビ、自由の女神像、人形、鳥居など、本来同一画面に存在し得ない物体や人物が雑多に混在している。これにより、夢も AI 生成映像も、現実の因果連鎖ではなく、知覚・連想に基づく一時的に成立する視覚環境を作り出す。



図 3

過学習（Overfitting）

AI の過学習は、モデルが学習データに過度に依存し、未学習のテストデータに対して性能が低下する現象を指す。生成モデルが特定サンプルに過度依存すると、入力条件が変化しても出力は類似した視覚特徴を示し、固定化されたテンプレートのような画像が繰り返し

生成される。この「似て非なる反復性」は、夢における人物の顔の曖昧さやアイデンティティの混同体験と一致し、現実と虚構の境界的視覚特性を強化する。

AI 生成ツールが大規模に普及し始めた初期段階では、指のねじれ、余分な肢体、テクスチャの混乱、顔の特徴の異常などの明確な欠陥がしばしば見られる。これらは過学習によるテンプレート化現象と同様であり、一定の構造や特徴への固定化依存に起因する（加えて学習データの不足やモデルの能力不足も影響する）。しかし、これらの初期的欠陥は独自の視覚言語へと進化し、超現実的な奇異性と不安を伴う夢幻的質感を保持したまま、新たな芸術表現の可能性を生む。

3. 夢における生成的視覚——『新ルネサンス』『天眼』における AI 生成美術の作品内応用

3-1. 『新ルネサンス』——人工知能による回想録の試み

私は、AIGC（AI Generated Content）を創作手段として用いる場合、それ自体は表現したい内容に奉仕すべきであると考えている。

私が AI 生成画像に初めて触れたのは、2023 年初頭に『新ルネサンス』の構想を練っていた時である。本作は人工知能による回想録を描いた作品であり、映像表現においても AI 生成による構築が可能かどうかを試みた。

使用した AI ソフトウェアは Midjourney と Stable Diffusion である。Midjourney は背景素材の生成に使用し、Stable Diffusion は主に人物像の生成に用いた。Stable Diffusion ではダ・ヴィンチの LoRA をトレーニングし、自身を俳優としてモデルに投影した（図 4）。しかし、AI モデルの不安定性により、「ダ・ヴィンチ」の姿は常に激しく点滅し、一瞬たりとも固定された形態を保持することはなかった。

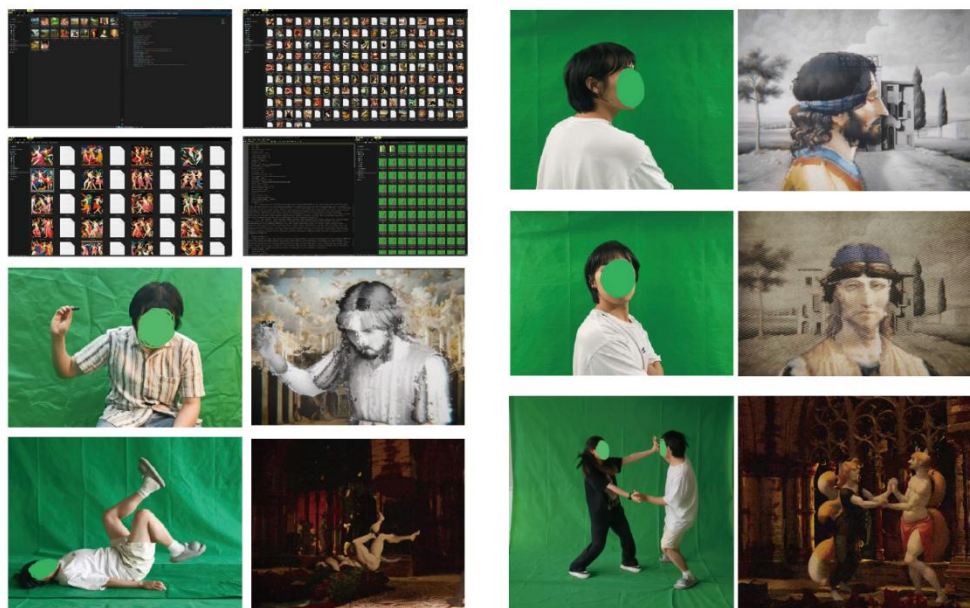


図 4

本文 2.1 で述べたように、各フレームのトレースが毎回異なるのは、画像から画像への変換が原画の連続的視覚復元を行っているのではなく、潜在空間において各フレームが独立して意味的に再サンプリングされるためである。潜在空間は物理的連続性ではなく意味的クラスタリングに基づき組織されているため、連続するフレーム間に微細な差しかなくとも、生成過程において意味的判断のジャンプが生じ、視覚結果の不連続が発生する。

不連続で点滅するトレースは、AI アニメーション制作において克服すべき大きな課題である。アニメーションは「連続性」を要求する一方で、モデルが得意とするのは「単フレームの合理性」である。Sora や Runway などの映像生成ソフトウェアは、アップデートや改良の過程で「より安定した画面」を重要な売りとしている。

しかし、このような不連続・不安定な画面表現こそが、『新ルネサンス』において必要とされるものである。本作は、形態が固定され、アイデンティティが明確なダ・ヴィンチを再現することを目指していない。むしろ、AI が歴史的人物を回想・再構築する際に直面する認知的困難を表現することを試みている。各フレームが絶えず再判断・再生成を行うため、ダ・ヴィンチの姿は常に点滅し変化し続け、一貫性を保つことはない。AI にとって、ダ・ヴィンチに確定的な姿は存在せず、生成を重ねるごとに接近することはできても、決して「記憶」することはできない。この不安定性は、AI の夢のような断片化された記憶方式を反映しており、作品の核心的な視覚的特徴を構成している。

3-2. 『天眼』における幻境／偽りの神

私の修士一年次作品『天眼』は、幼少期に中医の民間診療所へ近視治療に連れて行かれた経験を描いたものである。その診療所では、科学的根拠のない偽中医に対し、信者が神に対して抱くような信仰が存在しており、私はまるで集団主義的な幻覚の中にいるかのように感じた。人々は、包装され偽造された神を盲目的に崇拝していたのである。

『新ルネサンス』と『天眼』の制作期間は三年にわたるが、この間に AI ソフトウェアは飛躍的に発展し、映像制作の時点ではより安定したビデオ→ビデオの AI トレースツールも存在した。しかし私は意図的に初期の Stable Diffusion を使用し、点滅するような質感を作り出すことを選んだ。

プロンプトのみで画像・映像を生成する形式には慣れておらず、全工程を AI に任せることに対して不安を感じた。自分が画面を完全にコントロールできないように思えたからである。しかし、3D→トレースの手法を用いることで、この問題はある程度解消された。一方で、建模・カメラワーク・アニメーションを自分で調整でき、他方で AI が予期せぬ表現を補完してくれる。この手法は、私が現実と虚構の間を揺れ動く創作理念を反映したものである。

グリッチ

アニメーション内では、幻境に登場する場面の描写に比較的多くの尺を割いた。制作過程は、まず Cinema4D で 3D シーンと人物・カメラアニメーションを構築し、次に Stable Diffusion で画像→画像変換によるトレースを行うという方法である。この制作方法は『新ルネサンス』と類似しているが、実写→トレースではなく、3D モデリングレンダリング→トレースへと置き換えた点が異なる。

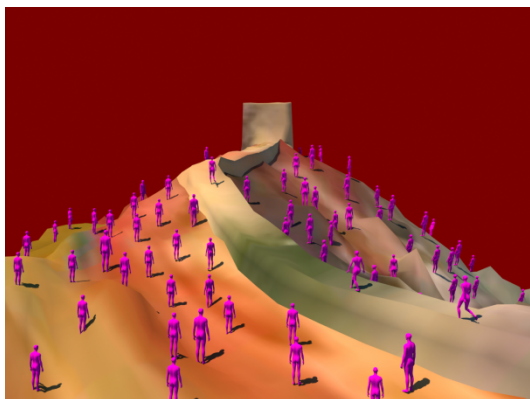


図 5、図 6

本作における各フレームの点滅は、『新ルネサンス』で付与された意味とは若干異なる。フレーム間の不連続性や認識の誤りは、私が AI の制御不能性と捉える部分である。私はこの創作技法上の「制御不能性」を用いて、幼少期の自分が集団の幻境に入り込むことを制御できなかった状況を表現した。

AI 幻覚 (hallucination)

AI の認識誤り（人工知能幻覚）は、本作の多くの場面にも現れている。私はこの認識誤りを積極的に利用し、興味深い場面を生み出すことを選んだ。例えば、巨人のモデルをレンダリングしたが、AI に「山」という指示を与えると、AI は先入観としてこれを山と認識する。その結果、画面は倒れた人間のようにも見え、山の向光面は赤、被光面は青という表現になり、人体構造は弱化され、岩層のテクスチャに置き換えられる（図 7、8）。

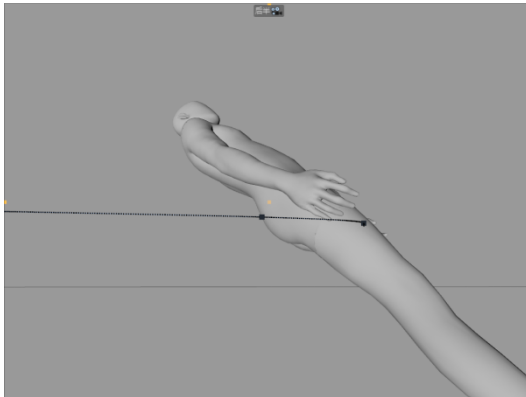


図 7、図 8

運動の補完

本作には「太極拳を打つ」アニメーションが登場し、最高権力者の象徴として機能する。この制作過程は、現実の太極動作を忠実に再現するものではなく、生成的 AI を意図的に介入させ、現実文脈から逸脱した身体運動を構築するものである。

まず Cinema4D で人物素体モデルを作成し、太極動作に基づく 4 つの主要ポーズを設定し、静的キーフレームとしてレンダリングした。次に、この 4 枚のキーフレームを AI 映像生成ソフト「可灵」に入力し、モデルが潜在空間上でキーフレーム間の動作を推論・補完することで、約 15 秒の連続動作映像を生成した。

生成された「太極拳」は現実世界のいかなる太極拳体系にも属さず、動作のリズムや重心移動、技と技の連結はいずれも検証可能な技法論理を欠く。言い換えれば、視覚上は「合理的に見える」が、身体知識の観点では成立しない運動映像であり、生成モデルによって構築

された「偽太極」とみなすことができる。この現実と虚構の間にある身体運動表現により、映像内で儀式的な雰囲気を帯びつつ、技芸の内在的根拠を欠いた空虚な形態が現れるのである。

私はこの「偽太極」の生成手法を通じて、作品内における偽造された「神」の像を構築した。生成的 AI は単なる技術的ツールとしてではなく、その確率的補完および意味的推論メカニズムを通じて、この偽神像の生成に能動的に関与しており、映像自体が信仰の生成・複製・誤認の隠喩となっている。

4. 結論

AI 生成画像のメカニズムと夢の生成メカニズムを並置して分析することで、両者は動作の論理および視覚結果のレベルにおいて顕著な類似性を有することが確認できる。拡散モデルによる潜在空間での視覚情報の確率的推論・補完であれ、夢における記憶の断片や感情連想による非線形な再構成であれ、いずれも現実世界の物理構造や因果関係の再現を目的とせず、現実と虚構の間に位置する映像状態を生成することを目指している。まさにこの「不安定」「不完全」な生成過程において、生成的映像は独自の美学的可能性を獲得するのである。

アニメーション制作において、生成的 AI はしばしば修正・克服すべき技術的対象として捉えられる。点滅・漂移・認識誤り・不連続性などは一般に欠陥と見なされる。しかし、本稿で『新ルネサンス』および『天眼』の制作実践を分析した結果、これらのいわゆる「問題」は必ずしも排除されるべきものではないことが示された。むしろ、創作テーマが夢・記憶・幻覚・集団意識などの不安定な状態を指向する場合、AI 生成過程において露呈する不確定性や制御不能性は、むしろ有効な視覚表現手段となり得るのである。

『新ルネサンス』においては、画像→画像トレースによって生じる形象の点滅やアイデンティティの漂移が、AI が歴史的人物を安定的に再現できないことの視覚的メタファーとして転化された。一方、『天眼』においては、AI の認識誤り、運動補完、生成バイアスが、歪みのある空虚な集団幻境を構築するために利用された。生成プロセスへの主体的介入・制御を通じて、AI はもはや創作者を置き換える自動化ツールではなく、不可予測性を有する生成システムとして、創作者と共に映像の構築過程に参加する存在となる。

したがって、アニメーション制作における AI 生成映像の意義は、効率向上や現実再現にとどまらない。創作者がもはや画面の安定性や完璧性を単純に追求せず、生成過程の不確定

性を活用可能な美学的資源として捉える時、AI 生成画像は「生成的美学」としてアニメーション制作に介入し得る。この美学は現実の再現を目的とせず、生成メカニズムの理解と翻訳を通じて、映像が夢・記憶・認知の境界において表現可能な可能性を探るものである。

文献

- ・ フロイト（2016）『夢的解析』方厚昇訳、浙江文芸出版社

作品

- ・ 今 敏（2006）『パプリカ』
- ・ スウナン（2023）『新ルネサンス』
- ・ スウナン（2025）『天眼』