Vol.2025-CG-200 No.43
Vol.2025-DCC-41 No.43
Vol.2025-CVIM-243 No.43
2025/11/7

微分可能レンダリングと拡散モデルによる正面復元画像を参照とした泉崎横穴3Dモデルの仮想修復

鶴貝 拓海[1,a)]　藤村 友貴[1]　舩冨 卓哉[1,2]　森本 哲郎[3]　高松 淳[4]　大石 岳史[3]　朽津 信明[5]
池内 克史[3]　向川 康博[1]

概要：本研究では微分可能レンダリングと拡散モデルを利用し，正面復元画像を参照して泉崎横穴の 3D モデルを仮想修復する手法を提案する．正面復元画像は撮影位置や照明環境などの条件が不明で，かつ横穴の一部しか捉えていないため 3D モデル全体のテクスチャ生成には直接利用できなかった．本研究ではこれに対し，正面復元画像を手がかりに，拡散モデルを用いてテクスチャを生成・補完する．具体的には微分可能レンダリングで画像に写る部分を変換し，その変換前後のペア画像を教師として ControlNet を再学習させ，正面復元画像のスタイルを再現する生成モデルを構築する．本手法により生成した画像を用いることで，一枚の正面復元画像から泉崎横穴の 3D モデル全体のテクスチャを仮想修復する．

キーワード：Stable Diffusion，ControlNet，微分可能レンダリング，泉崎横穴

1. 序論

文化財のデジタルデータを用いた仮想修復は，その恒久的な保存と活用を実現するうえで重要な取り組みである．3D スキャナを用いて取得した 3D モデルは，文化財の形状を正確に記録することが可能である．一方，経年によって失われた築造当時の色彩や模様といったテクスチャの仮想修復も重要な課題の一つである．本研究で対象とする泉崎横穴もまた，壁画が人為的に損傷され，本来の姿は失われている．この失われた壁画の姿を取り戻す試みとして，朽津ら [7] による先行研究が存在する．朽津らは，現存する壁画の顔料（ベンガラ）の科学的な色計測と，損傷前に撮影されていた白黒写真の見え方を組み合わせることで，考古学的・科学的考証に基づいて正面の壁画を復元した．本研究では，これを「**正面復元画像**」と呼ぶものとする．

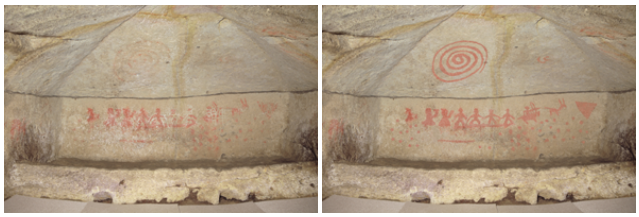
しかし，この正面復元画像は，3D モデル全体の仮想修復を目指すうえで制約がある．第一に，この正面復元画像は単一視点から見た画像であるが，3D モデルに対する正確なカメラパラメータや光源といった撮影条件が不明である．そのため，3D モデルにそのままテクスチャとして貼り付けることができない．第二に，正面復元画像は正面のテクスチャしか含んでおらず，側面や背面のテクスチャは含まれていない．これは，一般的なテクスチャの欠損の問題とは異なり，大部分の情報がそもそも存在しないという問題である．

そこで本稿では，現存する 3D モデルに対して，一枚の正面復元画像だけを手掛かりに泉崎横穴全体のテクスチャを仮想修復することを目的として，以下の二つの技術を組み合わせた手法を提案する．まず，撮影条件が不明であるという課題に対し，微分可能レンダリング [3] を用いて 3D モデルのレンダリング結果が正面復元画像と一致するように，カメラパラメータや光源を自動で最適化する．次に，情報の欠損という課題に対して，生成モデルの一種である拡散モデルを適用する．正面復元画像から壁画の画風やスタイルを学習させ，それをガイドとして側面や背面など，情報が存在しない部分のテクスチャのスタイルを維持しつつ，整合性の取れたテクスチャを生成・補完する．

本稿では，この一連の手法によって，泉崎横穴 3D モデルの仮想修復を実現するプロセスとその結果を論じる．第 2 章では，本研究の対象である泉崎横穴と，利用可能な資料について詳述する．続く第 3 章で提案手法の詳細を述べ，第 4 章で実験結果を示し，第 5 章で結論を述べる．

[1] 奈良先端科学技術大学院大学
Nara Institute of Science and Technology
[2] 京都大学
[3] 東京大学
[4] マイクロソフト
[5] 東京文化財研究所
[a)] tsurugai.takumi.tw6@naist.ac.jp

Vol.2025-CG-200 No.43
Vol.2025-DCC-41 No.43
Vol.2025-CVIM-243 No.43
2025/11/7



(a) 2007 年の泉崎横穴の壁画　　(b) 復元された壁画（正面復元画像）

図 1　朽津らの研究 [7] で用いられた画像

(a) 3D モデル全体の外観　　(b) 3D モデルの壁画の一部分

図 2　2025 年に計測された 3D モデル

2. 泉崎横穴

2.1 歴史的背景と現状

本研究で対象とする泉崎横穴は，福島県泉崎村に現存する，今から約 1400 年前の古墳時代後期に造られた国指定史跡の横穴墓である．1933 年の道路工事中に偶然発見された 7 基の横穴墓のうちの 1 基であり，その内部の壁面や天井には，赤い顔料（ベンガラ）で描かれた壁画が存在する．この壁画は，人物や馬などの動物，渦巻きなどの文様が描かれた装飾古墳の一例であり，発見当時は東北地方初の事例として大きな注目を集めた．その後，他の横穴墓は現存していないが，この壁画を持つ 1 基のみがその学術的価値の高さから国史跡として保存されている．しかし，1960 年代に人為的に傷つけられ，壁画は損傷された状態にある．本研究では，この壁画が描かれた玄室部分を仮想修復の対象とする．

2.2 仮想修復のために利用する情報

本横穴の仮想修復にあたり，間接的・直接的に利用する情報をまとめる．

2.2.1 1933 年に撮影された白黒写真

泉崎横穴が発見された直後の 1933 年に撮影された白黒写真は，壁画の当初の姿を伝えるための極めて重要な歴史的資料 [8] である．この写真の最大の価値は，1969 年に人為的に損傷を受ける以前の壁画の状態を記録している点にある．これにより，現在では失われてしまった壁画の形状や文様の全体像を把握することができる．しかし，これは白黒写真であるため，顔料の色彩に関する情報は含まれていない．

2.2.2 2007 年に作成された正面復元画像

2007 年に朽津ら [7] によって，失われた壁画の姿を復元する重要な研究が行われた．図 1(a) は泉崎横穴の壁画で，図 1(b) は復元後の壁画である．彼らは，図 1(a) の現状の壁画の全体像を撮影し，その後，代表的な色領域に対して分光放射輝度計を用いて分光反射率情報を求めた．次に，画像データに対して各色領域を構成する画素の平均表色値を計算し，これが分光反射率情報に基づいて算出された対応する領域の表色値に一致するように，それぞれの画素値を変換し，任意照明下での正確な色情報を求めた．

最後に，壁画の損傷部分の復元を，1933 年に撮影された白黒写真と当時撮影した画像と比較して行った．欠損部分の色は，白黒写真でもともと顔料の存在が確認される部分は顔料の色で，それ以外の部分は周囲の岩の色で埋めた．また，白黒写真など発掘当初の情報には確認されるが，現在は認識しづらくなっている文様には，現在の壁画上で顔料の残存を一部にでも確認できた箇所についてのみ顔料の色で再現した．これにより，図 1(b) の失われた壁画の姿を復元した．本研究ではこれを**正面復元画像**と呼ぶ．この復元は，科学的根拠に基づき失われた壁画の姿を画像として復元した重要な成果であるが，復元されたのは玄室の正面から見た画像のみである．

2.2.3 2025 年に計測された 3D モデル

本研究で利用する 3D モデルの外観を図 2(a) に示す．このモデルは，2025 年 3 月に泉崎横穴の玄室にて 3D スキャナ（FARO FocusS 150）を用いて計測したデータを基に構築したものである．図 2(b) に示す通り，壁面の凹凸といった正確な三次元形状を捉えている一方で，スキャン時に取得した頂点カラーは，損傷があり壁画本来の色彩が一部失われた現状を記録したものとなっている．

3. 提案手法

3.1 手法全体の流れ

本研究では，一枚の正面復元画像を用いて，泉崎横穴の 3D モデル全体のテクスチャを仮想修復するパイプラインを提案する．図 3 にその全体像を示す．図 3(a) では，微分可能レンダリングを用いて生成モデルの学習の基盤となる教師データを生成する．図 3(b) では，その教師データで ControlNet を学習させ，壁画のスタイルを任意の視点に適用できるスタイル変換モデルを構築する．図 3(c) では，学習済みモデルが生成した多視点テクスチャを，再度微分可能レンダリングを用いて 3D モデル全体にシームレスに統合し，仮想修復を完成させる．

3.2 微分可能レンダリングによる撮影条件の推定

本段階の目的は，後続の生成モデルの学習に不可欠な教師データとなる，入力画像と目標画像のペアを生成することである．

Vol.2025-CG-200 No.43
Vol.2025-DCC-41 No.43
Vol.2025-CVIM-243 No.43
2025/11/7

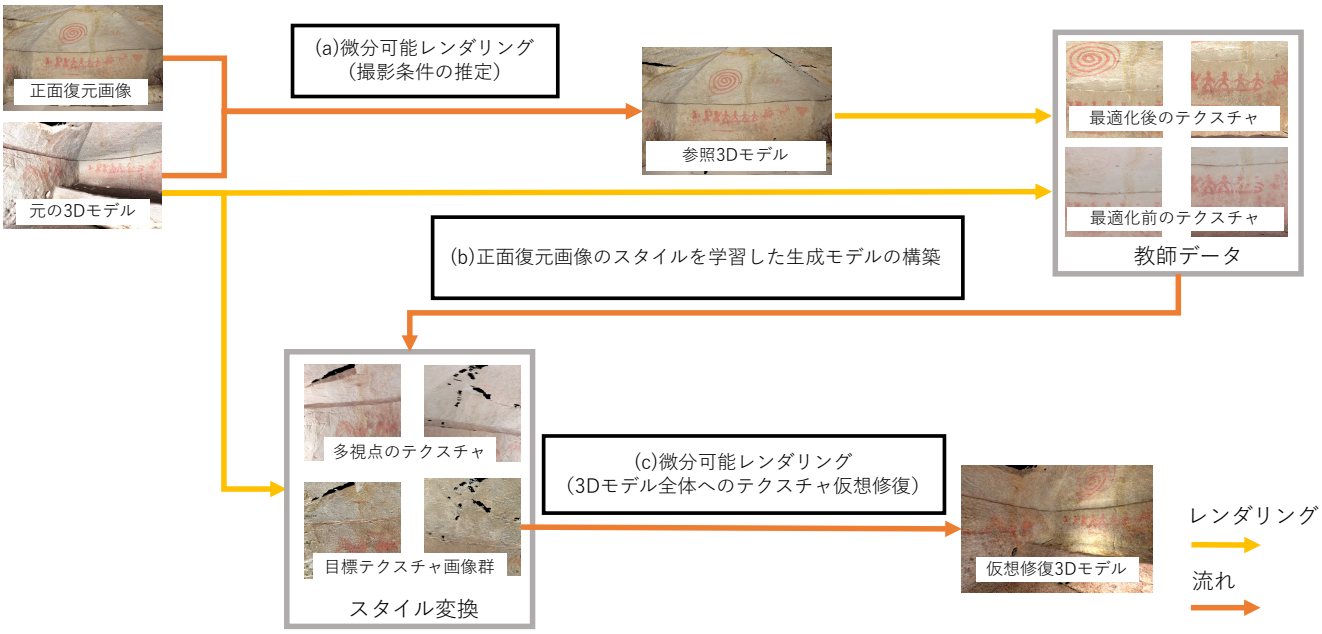


図 3 提案手法の全体像．本手法は，(a) 撮影条件の推定，(b) 正面復元画像のスタイルを学習した生成モデルの構築，(c) 3D モデル全体へのテクスチャ仮想修復，という 3 つの主要なステップで構成される．各ステップは本文中のセクションと対応している．

3.2.1 カメラパラメータの最適化

まず，3D モデルをレンダリングする際の視点と，参照画像である正面復元画像の視点を一致させるカメラパラメータの最適化である．この最適化には，レンダリングプロセス全体を微分可能として扱う微分可能レンダリングの技術を用いる．これにより，画像間の誤差を基にカメラパラメータを勾配降下法で自動的に更新することが可能となる．なお，最適化の効率と安定性を高めるため，初期カメラパラメータは PnP アルゴリズムを用いて事前に推定しておく．具体的には，3D モデルと正面復元画像から手動で取得した 6 点の対応点に基づき大まかなカメラパラメータを算出し，これを最適化の初期値とする．

この初期値から，SSIM[5] を指標とした微分可能レンダリングによる最適化を行う．3D モデルのカメラパラメータ，注視点，画角，そしてカメラ自身の回転などをまとめたものを変数 P とし，現在のパラメータ P でレンダリングをしグレースケールに変換したものと，目標となる正面復元画像をグレースケールに変換したものとの間の構造的類似性を SSIM を用いて評価し，その損失 L_{SSIM} を計算する．SSIM は，単純なピクセル差ではなく，画像の局所的な輝度，コントラスト，構造を比較するため，より人間の知覚に近い形で画像の類似性を評価できる指標である．SSIM は値が 1.0 に近いほど類似度が高いことを示すため，これを最大化すべき損失関数として扱う．そのため SSIM の損失関数 $L_{SSIM}(1)$ を以下のように定義する．

$$L_{SSIM}(I_1, I_2) = 1 - SSIM(I_1, I_2) \tag{1}$$

この損失を最小化する最適なカメラパラメータ P^* を，勾配降下法に基づくオプティマイザである Adam[2] を用いて探索する．この最適化問題は，次式のように表される．

$$P^* = \underset{P}{\operatorname{argmin}}\, L_{SSIM}(I_{gray}(M, P), I_{t\_gray}) \tag{2}$$

ここで，$I_{gray}(M, P)$ は 3D モデル M をパラメータ P でレンダリングしたグレースケール画像，I_{t_gray} は目標となるグレースケール画像である．このプロセスにより，3D モデルが正面復元画像と幾何学的に整合する，最適なカメラパラメータが自動的に推定される．

3.2.2 頂点カラーおよび光源環境の最適化

次に，推定したカメラパラメータ P^* を固定し，正面復元画像の見た目（スタイル情報）を 3D モデルの頂点カラー C に適用する．ここでも微分可能レンダリングを用い，レンダリング結果が正面復元画像 I_t とピクセル単位で可能な限り一致するように，頂点カラー C と仮想的な光源パラメータ L を同時に最適化する．損失関数には式 (3) に示す平均二乗誤差（MSE）を用いる．

$$L_{MSE}(1, I_2) = \frac{1}{N}\sum_{p=1}^{N}(I_1(p) - I_2(p))^2 \tag{3}$$

このピクセル単位の損失 L_{MSE} を最小化する最適な頂点カラー C^* と光源パラメータ L^* を，式 (4) によって求める．

$$\{C^*, L^*\} = \underset{C,L}{\operatorname{argmin}}\, L_{MSE}(R_{full}(M, P^*, C, L), I_t) \tag{4}$$

ここで，R_{full} は全てのパラメータを用いてフルカラーでレンダリングした画像である．この最適化により，正面復

元画像のスタイル情報が適用された新たな頂点カラー C^* が得られる．このステップの最終的な成果物は，「元の 3D モデル」と，正面復元画像のスタイルが適用された「**参照 3D モデル**」である．この 2 つの 3D モデルを同じカメラパラメータ P^* でレンダリングすることで，生成モデルの学習に用いる入力画像と目標画像のペアが生成される．両者は空間的に完全に対応が取れているため，生成モデルは純粋なスタイル変換のみを効率的に学習できる．

3.3 正面復元画像のスタイルを学習したテクスチャ生成モデルの構築

本段階では前段で得られた教師データを用いて，正面復元画像が持つ独特のスタイルを再現する生成モデルを構築する．ここでの目的は，元の 3D モデルが持つ基本的なテクスチャに対し，正面復元画像特有のスタイルを付与できる，スタイル変換モデルを完成させることである．

モデルの構築には, 画像生成モデルである Stable Diffusion を基盤とした．このモデルは Rombach らが提案した Latent Diffusion Model[4] の考えに基づいている．さらに，その生成プロセスを精密に制御するために，Zhang らが提案した ControlNet[6] を採用する．ControlNet は，事前学習済みの Stable Diffusion モデルに対し，入力画像の構造情報を保持したまま，目標とするスタイルを適用するタスクに非常に優れているため，今回の目的に適している．

具体的な学習では，元の 3D モデルからレンダリングした入力画像（最適化前のテクスチャ）と，参照 3D モデルからレンダリングした目標画像（スタイルが適用されたテクスチャ）のペアを教師データとして用いる．ControlNet は，入力画像から目標画像を生成できるようにファインチューニングされ，その過程で両画像の差であるスタイルのみを抽出して学習する．

この学習プロセスを通して，泉崎横穴の復元スタイルを任意の視点の画像に適用できる，汎用的なスタイル変換モデルを作成する．このモデルが，最終段階である 3D モデル全体の仮想修復において中心的な役割を果たす．

3.4 微分可能レンダリングによる 3D モデル全体へのテクスチャ仮想修復

最終段階では，学習させたスタイル変換モデルと微分可能レンダリングを組み合わせ，3D モデル全体のテクスチャを生成し，一つの連続したテクスチャへと統合する．このプロセスは，生成モデルによる多視点テクスチャの生成と微分可能レンダリングによる最終統合という 2 つのステップで構成される．

まず，生成モデルによる多視点テクスチャ生成を行う．3D モデルの側面や背面など，正面復元画像に含まれていない全方位の領域を対象として，元の 3D モデルのテクスチャをレンダリングする．次に，これらの画像を図 3(b) のプロセスで学習した ControlNet モデルに入力し，スタイル変換を適用する．これにより，3D モデルの全方位の領域に対応した，スタイル付きの目標テクスチャ画像群が生成される．

しかし，これらの生成されたスタイル付きの目標テクスチャを単純に 3D モデルに適用しただけでは，視点間の境界に継ぎ目が生じ，自然な見た目にはならない．この継ぎ目の問題を解決するため，本手法ではミニバッチを用いた微分可能レンダリングによる最適化を行う．これは，最適化の各ステップでランダムに選ばれた複数の視点（ミニバッチ）からの誤差を同時に計算し，その平均損失を最小化するように頂点カラー C_f を更新する．これにより，複数の視点から見える頂点は，それぞれの目標テクスチャからの制約を同時に受けて，矛盾の少ない色へと収束していく．このプロセスを繰り返すことで，テクスチャ間の継ぎ目が自然になり，全体として一貫性のあるテクスチャが構築される．このミニバッチ方式は，継ぎ目を解消する上で本質的な役割を果たすと同時に，全視点を一度に考慮するよりも計算コストを削減できるという利点も持つ．この最適化問題は，次式のように定式化される．

$$C_f^* = \underset{C_f}{\operatorname{argmin}} \mathbb{E}_{P' \sim P_m} \left[L_{MSE}(R(M, P', C_f), I_{AI}(P')) \right] \tag{5}$$

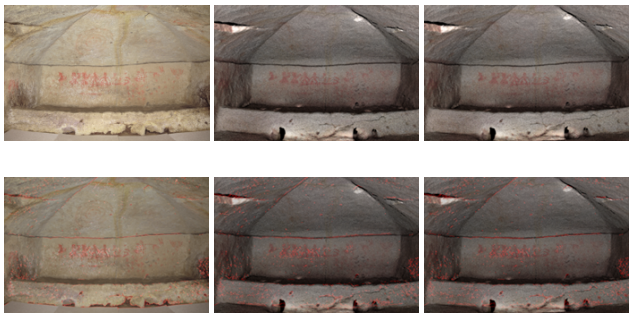
ここで，C_f は最適化対象となる 3D モデル全体の頂点カラーを示し，C_f^* は 3D モデル全体の最終的な頂点カラーである．P_m は全方位の視点パラメータの集合，P' はそこからサンプリングされたミニバッチ内の視点パラメータを示す．$\mathbb{E}$ はミニバッチに対する期待値を表し，R はレンダリング画像，I_{AI} は学習済み ControlNet が生成した目標画像である．損失関数 L_{MSE} には平均二乗誤差を用いる．このミニバッチを用いた最適化を繰り返すことで，各頂点カラーが複数の視点からの制約を同時に受けて更新されるため，生成モデルが生成したテクスチャ間の継ぎ目における不整合が自然に解消される．これにより，計算リソースを効率的に使用しつつ，多視点からの損失を全体として最小化し，全体として一貫性の取れた「**仮想修復 3D モデル**」を構築する．

4. 実験と結果

4.1 実験設定

本実験で用いる 3D モデルは，3D 点群スキャナ（FARO FocusS 150）を用いて取得した泉崎横穴のデータである．同モデルは 2,877,462 頂点，5,693,220 面の三角形メッシュで構成され，各頂点に現状の色彩情報が記録されている．参照画像には，朽津らの研究 [3] による 1,000 × 666 ピクセルの正面復元画像（PNG 形式）を用いた．スタイル変換モデルの基盤として，Stable Diffusion (runwayml/stable-diffusion-

Vol.2025-CG-200 No.43
Vol.2025-DCC-41 No.43
Vol.2025-CVIM-243 No.43
2025/11/7



(a) 正面復元画像　(b) 最適化前の画像　(c) 最適化後の結果

図 4　撮影条件推定の定性結果
正面復元画像のエッジ（赤色）と各画像の比較

v1-5) および ControlNet (lllyasviel/control_v11f1e_sd15_tile) の事前学習済みモデルを利用した．

実装環境として，ハードウェアには NVIDIA A100 80GB PCIe（VRAM 81920MB）を搭載したワークステーションを用いた．主要なソフトウェアライブラリとバージョンは，Mitsuba 3（ver. 3.5.2），Dr.Jit（ver. 0.4.6）[1]，PyTorch（ver. 2.4.1+cu121）である．

各ステップにおける主要なハイパーパラメータは以下の通り設定した．撮影条件推定では，オプティマイザに Adam を用い，学習率 0.005 で 500 イテレーションの最適化を行った．続く頂点カラー最適化でも同様に Adam を用い，学習率 0.002 で 1000 イテレーション実行した．ControlNet の学習では，133 枚の画像ペアを用い，バッチサイズを 1，学習率を 2.0×10^{-6} とし，512 × 512 ピクセルの画像で 500 エポックの学習を行った．最終的なテクスチャ統合の最適化では，540 枚の多視点画像から学習した ControlNet を用いてスタイル付きの目標テクスチャ画像群を生成し，これを目標とし，学習率 0.01，バッチサイズ 4 で 100 エポックで行った．

4.2　撮影条件推定の結果

4.2.1　カメラパラメータの最適化の結果

本節では，3.2 節で述べた微分可能レンダリングによる撮影条件最適化の結果を示す．本実験の目的は，3D モデルのレンダリング結果が，参照画像である正面復元画像と幾何学的に一致することを示すことである．その検証のため，3.2.1 節の手法に基づき，SSIM 損失を指標とした最適化を 500 回実行した．図 4 に，本手法による撮影条件推定の定性的な結果を示す．画像比較から，最適化前のレンダリング画像 (図 4(a)) は，目標となる正面復元画像 (図 4(b)) に対し，天井の高さや壁の輪郭にわずかなズレがあることがわかる．一方，最適化を適用した後のレンダリング画像 (図 4(c)) は，視覚的にほぼ完全に一致している．

この幾何学的な一致の精度は，図 4 のエッジ重ね合わせ画像でより明確に確認できる．正面復元画像のエッジ（赤色）は，最適化前は壁の境界や段差から外れている箇所が

(a) 正面復元画像　(b) 最適化後のレンダリング画像　(c) (a) と (b) の差分画像

図 5　頂点カラー及び光源環境の最適化結果

(a) ControlNet への入力となる最適化前のテクスチャ

(b) 正面復元画像のスタイルが適用されたテクスチャ

図 6　本最適化プロセスによって生成された
生成モデル学習用の教師データペアの例

散見されるが，最適化後には重なっており，高精度な位置合わせが達成されたことを示している．

以上の定性評価から，提案手法によって 3D モデルと正面復元画像の撮影条件を高い精度で一致させられることが示された．この正確に推定されたカメラパラメータは，次のステップである頂点カラーおよび光源環境の最適化における重要な基盤となる．

4.2.2　頂点カラー及び光源環境の最適化の結果

本節では，前節で推定したカメラパラメータを固定し，3D モデルの頂点カラーと光源環境を最適化した結果を示す．この最適化の目的は，正面復元画像の色彩や質感を 3D モデルの頂点カラーへ適用し，後続のスタイル変換モデルの学習に用いる教師データのペアを生成することである．

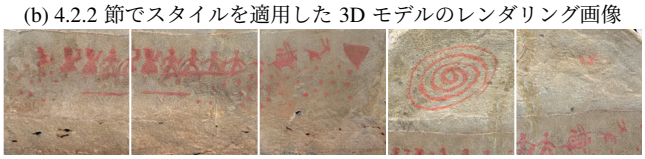
図 5 に，目標となる正面画像と，本手法による最適化後のレンダリング画像の比較を示す．最適化後の画像 (5(b)) は，目標画像の複雑な陰影や色彩を高い忠実度で再現していることがわかる．これを裏付けるように，両画像のピクセル単位の差分を可視化した差分画像 (5(c)) では，全体がほぼ黒一色となっており，誤差が極めて小さいことが視覚的に確認できる．

この最適化プロセスの最終的な成果物は，図 6 に示す生成モデル学習用の教師データである．図 6(a) が ControlNet への入力となる最適化前のテクスチャ，図 6(b) が学習目標となる正面復元画像のスタイルが適用された画像である．両者は同一のカメラパラメータからレンダリングされているため空間的に完全に一致しており，これにより，生成モデルは純粋なスタイル変換のみを効率的に学習することが可能となる．以上の結果から，提案手法によって正面復元

Vol.2025-CG-200 No.43
Vol.2025-DCC-41 No.43
Vol.2025-CVIM-243 No.43
2025/11/7

(a) 入力となる壁画のテクスチャ

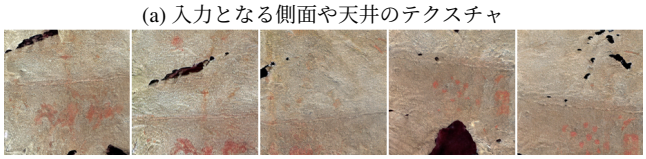
(b) 4.2.2 節でスタイルを適用した 3D モデルのレンダリング画像



(c) ControlNet によるスタイル変換後の出力画像

図 7　壁画部分におけるスタイル変換結果

(a) 入力となる側面や天井のテクスチャ



(b) ControlNet によるスタイル変換後の出力画像

図 8　側面や天井におけるスタイル変換結果

画像の見た目を 3D モデルに忠実に適用し，スタイル変換モデルの学習に不可欠な，高品質かつ完全に位置整合した教師データを生成できることが示された．

4.3　拡散モデルによるテクスチャ生成結果

本節では，4.2 節で生成した教師データペアを用いて学習させた，スタイル変換 ControlNet モデルの性能を評価する．評価項目は，モデルが正面復元画像の持つ特徴的なスタイルを学習できたか，そして，そのスタイルを側面や天井などの学習データには含まれていない未観測の領域に対しても適切に適用できるか，の 2 点である．

まず，学習データと類似する壁画部分へのスタイル変換性能を検証する．図 7 に，側面や天井に描かれた壁画部分でのスタイル変換結果を示す．上段がモデルへの入力となるテクスチャ，中段が比較の基準となる，4.2.2 節でスタイルを適用した 3D モデルのレンダリング画像，下段がスタイル変換後のテクスチャである．

この評価項目は，スタイル変換の出力画像 (図 7(c)) が，比較基準となる 3D モデルのレンダリング結果 (図 7(b)) とどれだけ視覚的に一致するかという点にある．生成モデルの出力 (図 7(c)) は，比較基準となる 3D モデルのレンダリング結果 (図 7(b)) の持つ複雑な陰影や質感を非常によく再現している．

この結果は，モデルが単なる色の変換を学習したのではなく，陰影や質感の変化といった，3 次元的なサーフェスを想起させる壁画のスタイルそのものを学習したことを示唆している．これにより，本モデルが泉崎横穴の壁画スタイルを任意の視点・形状に適用できる，表現力を獲得したことが示された．

次に，モデルの汎化性能をより評価するため，学習データに含まれない側面や天井の壁画を入力した場合を検証する．図 8 に示すように，復元壁画が存在しない入力 (図 8(a)) に対しても，モデルは破綻することなく，周囲と調和する自然な岩肌のテクスチャ (図 8(b)) を生成した．重要なのは，不自然な文様を生成することなく，学習した環境テクスチャを適切に適用している点であり，これはモデルが汎化性能を持つことを示している．

以上の 2 つの評価から，本研究で学習させた ControlNet モデルは，泉崎横穴の壁画スタイルを正しく学習し，それを類似箇所や全く異なる未知の領域に対しても安定して適用できる，高い性能と汎用性を持つことが実証された．

4.4　仮想修復の最終結果

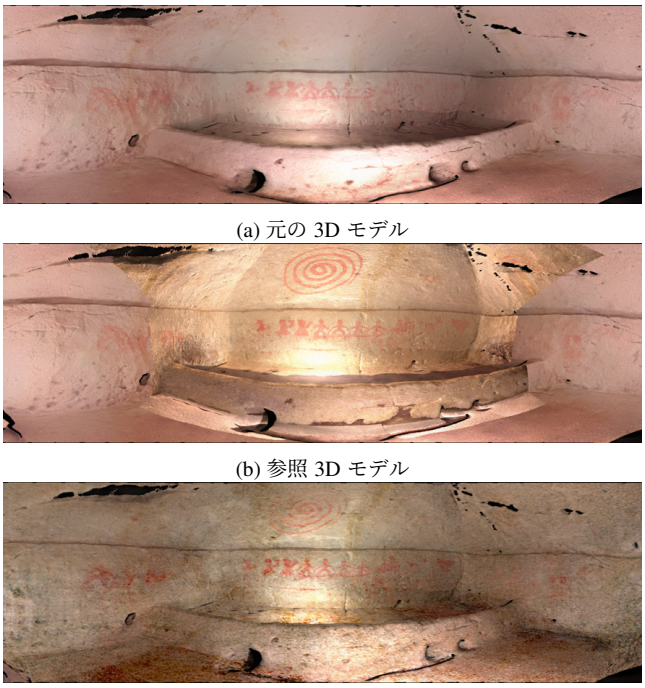
まず，図 9 において，本研究の元の 3D モデル，参照 3D モデル，仮想修復 3D モデルのパノラマによる比較を示す．図 9(a) は「元の 3D モデル」であり，壁画はかすれて不鮮明である．図 9(b) は，参照 3D モデルである．このモデルは正面の見た目こそ忠実に再現しているが，利用できる情報が正面のみであるため，側面や天井部分ではテクスチャが破綻，あるいは欠損している．これに対し，図 9(c) に示す本手法の最終結果である仮想修復 3D モデルでは，生成モデルによるテクスチャ生成を経て，側面や天井に至るまで全体として一貫性のある自然なテクスチャが生成されていることがわかる．この比較により，単純なテクスチャの適用だけでは不完全であり，本研究の生成モデルを用いた全面的なテクスチャ生成が不可欠であることが示される．

次に，図 10 では，最終結果として得られる仮想修復 3D モデルの品質と，テクスチャの連続性を検証する．これは，仮想修復された横穴内部を異なる 4 つの視点からレンダリングしたものである．壁の角や天井との境界部分にもテクスチャの破綻や継ぎ目は見られず，単一の連続した空間として，自然な見た目が実現できている．これは，手法の最終段階で行った微分可能レンダリングによるミニバッチ学習が効果的に機能したことを示している．

以上の結果から，本研究が単一の正面復元画像を手掛かりに，3D モデル全体の仮想修復という目的を達成したことを確認した．

5.　結論

本研究では，一枚の正面復元画像のみを手掛かりとし



(a) 元の 3D モデル

(b) 参照 3D モデル

(c) 仮想修復 3D モデル

図 9 仮想修復プロセスの各段階における 3D モデルのパノラマ展開図の比較

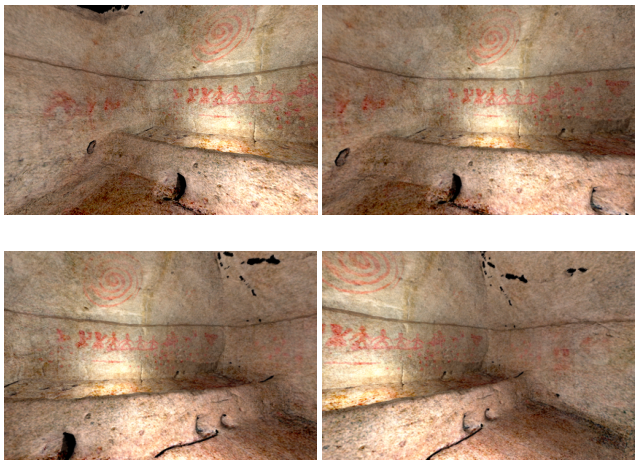


図 10 仮想修復 3D モデルの多視点レンダリング結果

て，3D モデル全体を仮想修復するという課題に取り組んだ．この目的のため，微分可能レンダリングと拡散モデルを組み合わせた手法を提案した．本手法では，まず微分可能レンダリングを用いて 3D モデルと参照画像を精密に位置合わせし，スタイル変換モデル学習のための高品質な教師データを生成する．次に，この教師データで ControlNet をファインチューニングし，対象の文化財のスタイルを適用できる生成モデルを構築した．最終的に，このモデルが生成した多視点テクスチャを，再度微分可能レンダリングを用いて 3D モデル全体にシームレスに統合した．

実験の結果，本手法は泉崎横穴の 3D モデルに対し，正面だけでなく側面や天井に至るまで，一貫したスタイルを持つテクスチャを生成し，その全体像を仮想的に復元することに成功した．単一の画像から 3D モデル全体のテクスチャを生成・統合するこの一連のパイプラインは，本研究の主要な貢献である．完成した 3D モデルとパノラマ展開図は，本手法の技術的な有効性を示すと共に，文化財の新たな鑑賞・分析の可能性を提示するものである．

一方で，本研究には今後の課題も残されている．生成されるテクスチャの解像度や計算コストの問題は，より高精細な仮想的な復元や幅広い応用を目指す上で改善が必要である．また，生成モデルが生成する内容は考古学的妥当性を必ずしも保証するものではなく，専門家の知見を反映させるインタラクティブな仕組みも望まれる．将来的には，これらの課題を克服し，本手法を他の文化財へ応用するだけでなく，形状とテクスチャを同時に復元する，より統合的な仮想修復技術へと発展させることが期待される．

謝辞

計測実験に協力していただき，復元画像を提供していただいた福島県泉崎村教育委員会に感謝します．なお，本研究は JSPS 科研費 JP23H00499 の助成を受けたものです．

参考文献

[1] Jakob, W., Speierer, S., Roussel, N. and Vicini, D.: Dr.Jit: A Just-In-Time Compiler for Differentiable Rendering, *Transactions on Graphics (Proceedings of SIGGRAPH)*, Vol. 41, No. 4 (online), DOI: 10.1145/3528223.3530099 (2022).

[2] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, *arXiv preprint arXiv:1412.6980*, (online), available from ⟨https://arxiv.org/abs/1412.6980⟩ (2014).

[3] Li, T.-M., Aittala, M., Durand, F. and Lehtinen, J.: Differentiable Monte Carlo ray tracing through edge sampling, *ACM Trans. Graph.*, Vol. 37, No. 6 (online), DOI: 10.1145/3272127.3275109 (2018).

[4] Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B.: High-Resolution Image Synthesis with Latent Diffusion Models, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695 (2022).

[5] Wang, Z., Bovik, A., Sheikh, H. and Simoncelli, E.: Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600–612 (online), DOI: 10.1109/TIP.2003.819861 (2004).

[6] Zhang, L., Rao, A. and Agrawala, M.: Adding Conditional Control to Text-to-Image Diffusion Models, *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3836–3847 (2023).

[7] 朽津信明，三橋　徹，池内克史：泉崎横穴における損傷壁画の画像復元，日本文化財科学会第 24 回大会 (2007).

[8] 　文化庁：史跡泉崎横穴の取組，古墳壁画の保存活用に関する検討会 装飾古墳ワーキンググループ（第 9 回）(2013).