

視覚の持続性を利用したマテリアルディスプレイとその生成モデルを活用した設計

湯蓋 康平^{*1}

渡辺 義浩^{*1}

Abstract – マテリアルディスプレイは実物体を用いることで、解像度やダイナミックレンジなどの制約を緩和し、現実に近い質感提示を可能にする。特に質感提示の重要な要素として、物体表面の立体構造であるテクスチャが挙げられる。従来手法では、回転スクリーンに実素材を配置し、そこへプロジェクタから光を投影することで、現実にも忠実なテクスチャを提示した。しかし、同手法で提示されるテクスチャは、スクリーン上に配置した実素材そのものであるため、提示可能なテクスチャの種類は回転スクリーンに配置した実素材に限定される。そこで本稿では、回転スクリーン上に基底となる素材を配置し、視覚の持続性を用いて重ね合わせることで、少数の素材から複数のテクスチャを提示するマテリアルディスプレイの実現を試みる。このとき、プロジェクタからバイナリ形式のマスク画像を投影することで少数の素材を重ね合わせる。本手法の実現に向けて、テクスチャ特徴量に基づき、提示目標となる複数のテクスチャ画像から少数の基底画像とマスク画像を取得する基底分解手法を提案する。さらに、より多くのテクスチャ提示のための追加最適化によるデノイズ手法とテクスチャ生成モデルを用いた提示目標の最適化手法を提案する。最後に、テクスチャ生成モデルを用いて連続的に変化するテクスチャ提示手法を提案する。

Keywords : 質感再現, 次元削減, ディスプレイ

1 はじめに

現実の物体は、光沢や色、表面構造など、様々な視覚的質感を持つ。そのため、視覚的な質感を正確に再現し、提示する技術は、プロダクトデザインにおける完成イメージの共有などの応用が期待できる。また、幅広い応用を可能にするためには、長時間の観察でも疲労が生じにくく、特別な装置の装着を必要としないことが重要な要件である。本稿では、特に物体表面の立体構造であるテクスチャに注目する。

高精度な質感提示を行う研究はこれまでに幅広く行われており、その中の一つとしてマテリアルディスプレイがある。マテリアルディスプレイは、実物体を用いることで、解像度やダイナミックレンジなどの制約を緩和し、現実に近い質感提示を行う。その中でも、回転スクリーンに実物体を配置し、回転運動に同期してプロジェクタからパターン光を照射することでテクスチャ提示を行うディスプレイが提案された [1]。同ディスプレイでは、プロジェクタから高速にパターン光を投影することで視覚刺激が消失した後も一定時間の間、視覚情報が残る視覚の持続性を利用し、提示像におけるテクスチャの空間的な分布を動的に変更することができる。しかし、同システムで提示できるテクスチャは、スクリーン上に配置した実素材そのものである。そのため、回転スクリーンに配置していないテクスチャを提示することは困難である。

そこで、本稿では回転スクリーン上に基底となる素

材を配置し、視覚の持続性を用いてそれらを重ね合わせることで少数の素材から複数のテクスチャを提示するマテリアルディスプレイの実現を試みる。具体的には、回転スクリーン上の基底となる素材に対してプロジェクタからバイナリ形式のマスク画像を投影することで、回転スクリーン上の素材を空間的に重ね合わせる。

本稿では、回転スクリーンに配置する基底素材およびマスク画像を求めるためにテクスチャ特徴量を用いる。テクスチャ特徴量とは、テクスチャ画像における統計的特徴を数値化する指標である。同一のテクスチャ特徴量を持つ画像間では、画素値としては異なる画像でも同じカテゴリのテクスチャと知覚される。そのため、実物体を用いて特定のテクスチャそのものを提示するのではなく、そのテクスチャ特徴量を満たす外観を再現することで、スクリーン上に配置した素材数に制限されず多くのテクスチャを提示できる可能性がある。

本ディスプレイの実現に向けて、四つの手法を提案する。まず、テクスチャ特徴量を用いて提示目標のテクスチャ画像から回転スクリーンに配置する基底素材とマスク画像を取得する基底分解手法を提案する。一方、提案する基底分解手法では、テクスチャ同士の組み合わせやテクスチャ数によってノイズやテクスチャの崩れが生じる課題がある。そこで、ノイズ低減のための追加最適化手法を提案する。

さらに、基底分解を行う前段階で提示目標とするテ

^{*1}東京科学大学

クスチャ画像の特徴を近づけることで、テクスチャの崩れを低減する手法を提案する。そのために、本稿ではテクスチャ画像が持つ多様な外観を統一的に扱う手法であるテクスチャ生成モデルに着目する。テクスチャ生成モデルは、大量のテクスチャ画像を学習することで、共通の潜在空間を獲得し、一つのモデルから多様なテクスチャを生成する手法である。その中でも、色や光沢とともに表面構造を含めて表現する Spatially Varying Bidirectional Reflectance Distribution Function (SVBRDF) を対象とした生成モデルがある [2]。同テクスチャ生成モデルの潜在空間において、潜在ベクトルをテクスチャ特徴量に基づいて操作することで異なるテクスチャ画像間の大域的な特徴を近づけることができる。本稿では、大域的な特徴を近づけたテクスチャ画像を事前に作成し、基底分解を行うことでテクスチャの崩れを低減する。

また、デザイン検討や質感提示では、異なるテクスチャを離散的に切り替えて比較するだけでなく、それらの間に位置する外観を連続的に確認したい場合がある。特に、表面の粗さや模様スケールといったテクスチャ属性が連続的に変化する過程を観察できれば、素材外観の変化をより直感的に把握できる。テクスチャ生成モデルでは、二つのテクスチャ画像に対応する潜在ベクトルから、その中間的な外観を持つテクスチャ画像を生成することができる。そこで、本稿ではテクスチャ生成モデルを用いて、連続的に変化するテクスチャ提示手法を提案する。

実験では、シミュレーションによるテクスチャ再構成を行い、少数の基底素材から複数のテクスチャ提示が可能であることを確認した。また、テクスチャ生成モデルを用いることで、連続的に変化するテクスチャ提示が可能であることを確認した。

2 関連研究

2.1 微細なテクスチャ提示を行うディスプレイ

これまで、微細なテクスチャを提示する様々な手法が提案されてきた。Nishi らは、投影対象に微小振動を与え、プロジェクタから高速にパターン光を投影することで、微細なテクスチャを提示する手法を提案した [3]。具体的には、振動する投影対象の奥行き位置と同期してプロジェクタから分割したデプス画像を高速に投影することでテクスチャを提示する。しかし、同手法で提示されるテクスチャの水平解像度はプロジェクタの解像度に制限される。

また、回転スクリーンに実物体を配置し、プロジェクタから光を投影することで現実に忠実なテクスチャを提示するマテリアルディスプレイが提案された [1]。同方式では、提示目標のテクスチャを持つ実素材を直接スクリーンに配置するため、プロジェクタに起因す

る解像度やダイナミックレンジの制約を受けない。さらに、同システムを螺旋状のスクリーンを持つ Swept Volume Display と組み合わせることで立体的な形状を提示する手法も提案されている [4]。しかし、同システムで提示されるテクスチャは、スクリーン上に配置した実素材そのものである。そのため、回転スクリーンに配置していないテクスチャを提示することは困難である。

2.2 基底分解

複数の画像を、少数の画像の組み合わせによって表現できる場合がある。基底分解とは、複数の画像を効率よく表現するために必要となる少数の基底画像を取得する手法である。基底分解の例として、主成分分析 (PCA) がある [5]。PCA は複数の入力画像をベクトルとして扱い、分散が最大となる直交軸を求めることで、入力画像を少数の基底画像と係数の線形結合で表現する手法である。一方、PCA により得られる基底画像や係数には負の値が含まれるため、実物体を用いて質感提示を行うマテリアルディスプレイでの利用は難しい。

上記の制約に対して、非負値行列因子分解 (NMF) は、入力画像から非負値の基底画像と非負値の係数を取得する手法である [6]。同手法は、非負値制約により、各基底画像が加算的に入力画像を再構成するため、視覚の持続性に基づく基底画像の重ね合わせに適している。しかし、NMF は顔画像のような空間的に局所的なパーツを基底として分解する傾向があり、画像全体に均一な特徴を持つテクスチャ画像に対しては、適切な基底を取得することが困難である [7]。

2.3 テクスチャ特徴量

テクスチャ特徴量はテクスチャ画像が持つ統計的性質を数値化したものである。テクスチャ特徴量を用いた研究分野の一つに、テクスチャ合成がある。テクスチャ合成とは、あるテクスチャ画像の統計的性質を保持したまま、視覚的には同一の新たなテクスチャ画像を生成する手法である。その代表的なアプローチとして、目標画像と出力画像のテクスチャ特徴量の差が最小となるように出力画像を更新することで目標画像と同じカテゴリに属する別のテクスチャ画像を生成する手法が知られている。

Zhou らは、テクスチャ画像をパッチに分割し、各パッチを VGG ネットワークに入力して得られた特徴ベクトルをテクスチャ特徴量として扱う [8]。同手法では、テクスチャ画像とランダムノイズ画像のテクスチャ特徴量の差分を Guided Correspondence Loss (GCLoss) とし、GCLoss を最小化するようにランダムノイズ画像を更新することでテクスチャ合成を行う。

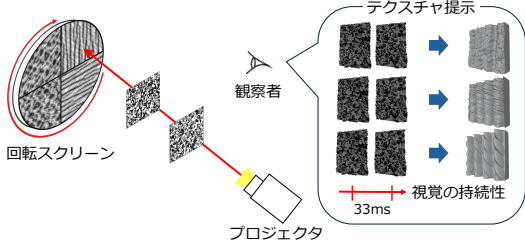


図1 視覚の持続性を利用した複数のテクスチャ提示.

Fig.1 A method for displaying multiple textures using the persistence of vision.

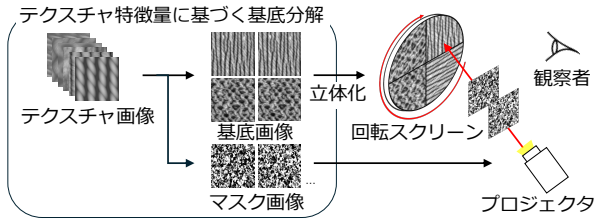


図2 テクスチャ特徴量を用いた基底分解手法.

Fig.2 Basis decomposition method using texture features.

2.4 テクスチャ生成モデル

テクスチャ生成モデルは、大量のテクスチャ画像を学習することで、テクスチャの外観を統一的に生成可能な潜在空間を獲得し、一つのモデルから多様なテクスチャを生成する手法である。特に、物体表面の反射特性を記述するSVBRDFを扱うテクスチャ生成モデルの一つとして、MaterialGANがある[2]。MaterialGANは、入力として与えられた撮影画像に対して、それを再現するSVBRDFを推定することが可能である。具体的には、潜在ベクトルから生成されたSVBRDFに基づくレンダリング画像と、入力との差が最小となるように潜在ベクトルを最適化することで、入力画像に対応するSVBRDFを推定する。また、MaterialGANの潜在空間では、二つの潜在ベクトルを線形補間することで、その中間的な外観を持つSVBRDFを生成できる。

3 提案手法

3.1 視覚の持続性を用いたテクスチャ提示

本節では、回転スクリーンに配置した基底素材とマスク画像によりテクスチャ提示を行う手法を提案する。図1に示すように、回転スクリーンに配置した基底素材に対して、プロジェクタからバイナリ形式のマスク画像を投影する。このとき、スクリーン上に配置された少数の基底素材は視覚の持続性により空間的に重なって知覚される。具体的には、視覚の時間分解能を示す臨界融合周波数(CFF)よりも高速にプロジェクタからマスク画像を投影することで実現する。

3.2 テクスチャ特徴量を用いた基底分解手法

本稿では、テクスチャの凹凸をデプス画像、法線マップ、陰影画像で表現し、図2に示すように提示目標と

するテクスチャ画像からテクスチャ特徴量を用いてデプス形式の基底画像とマスク画像を取得する。その後、基底画像を立体化し、基底素材を作成する。

以下では、デプス画像を例に、基底画像およびマスク画像による再構成の定式化について説明する。デプス画像は、各画素が三次元空間内の奥行きを表す画像である。まず、提示目標とするデプス形式のテクスチャ画像の集合を $V^{\text{depth}} = [v_1^{\text{depth}}; v_2^{\text{depth}}; \dots; v_N^{\text{depth}}]$ 、基底画像の集合を $H^{\text{depth}} = [h_1^{\text{depth}}; h_2^{\text{depth}}; \dots; h_n^{\text{depth}}]$ 、マスク画像の集合を $W = [w_1^{\text{depth}}; w_2^{\text{depth}}; \dots; w_n^{\text{depth}}]$ と定義する。このとき、 w_k^i は i 番目のテクスチャ画像を表現するための基底画像 h_k^{depth} に対応するマスク画像、 n, N は基底画像および提示目標のテクスチャ画像の枚数である。マスク画像はバイナリ形式であり、基底画像に重ねられる。各画素において、マスク画像の画素値が1に対応する基底画像の画素が提示され、複数の基底画像が空間的に組み合わせることでテクスチャ画像を再構成する。そのため、基底画像とマスク画像からなる再構成画像 $\tilde{V}^{\text{depth}} = [\tilde{v}_1^{\text{depth}}; \tilde{v}_2^{\text{depth}}; \dots; \tilde{v}_N^{\text{depth}}]$ について、 $\tilde{v}_i^{\text{depth}}$ は次式で定義される。

$$\tilde{v}_i^{\text{depth}} = \sum_{k=1}^n w_k^i h_k^{\text{depth}} \quad (1)$$

基底分解では、 V^{depth} と \tilde{V}^{depth} のテクスチャ特徴量の差分を計算し、その差分が最小となるように勾配降下法により H^{depth} と W を更新する。本稿では、Zhouらが提案したGCLossをテクスチャ特徴量の評価指標として用いる[8]。

基底画像を立体化する観点では、デプス画像のみを用いることで基底素材を作成可能である。しかし、人間が知覚するテクスチャの印象は奥行き情報だけでなく、表面の傾きによる陰影の影響も大きい。そこで、本稿ではデプス画像に加えて法線マップと陰影画像を用い、GCLossで評価する。法線マップは、物体表面の法線ベクトルをRGB成分で表現した画像であり、微細な傾きを表現する。また、陰影画像はテクスチャの凹凸によって生じる明暗の分布を表現した画像である。

まず、 \tilde{V}^{depth} の x, y 方向の勾配を計算し、法線方向を推定することで、再構成画像の法線マップ $\tilde{V}^{\text{normal}}$ を生成する。次に、陰影画像の計算のために平行光の照射方向を定義し、 $\tilde{V}^{\text{normal}}$ との内積を計算することで、陰影画像 \tilde{V}^{shade} を作成する。同様の手順により、提示目標に対しても V^{depth} 、 V^{normal} および V^{shade} を用意する。

最後に、 V^{depth} と \tilde{V}^{depth} 、 V^{normal} と $\tilde{V}^{\text{normal}}$ および V^{shade} と \tilde{V}^{shade} の各画像に対して、GCLossを計算し、その和を損失関数とする。したがって、GCLossを L_{GC} とすると基底分解は以下の最適化問題として

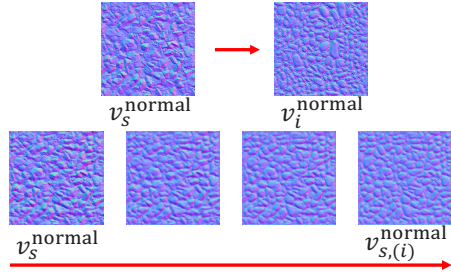


図3 大域的な特徴は保持されたままテクスチャが変化する。

Fig. 3 Texture changes while global features are preserved.

定式化される。

$$\begin{aligned} \min_{W; H^{\text{depth}}} \quad & \sum_{i=1}^N f_{LGC} v_i^{\text{depth}}; \tilde{v}_i^{\text{depth}} \\ & + L_{GC} v_i^{\text{normal}}; \tilde{v}_i^{\text{normal}} \\ & + L_{GC} v_i^{\text{shade}}; \tilde{v}_i^{\text{shade}} g \quad (2) \\ \text{s.t.} \quad & 0 \leq H^{\text{depth}} \leq 1; W \geq f_0; 1g \quad (3) \end{aligned}$$

得られたデプス形式の基底画像は、画素値を高さとして、立体形状へ変換することで基底素材を作成する。本稿では、公開されているオンラインツール [9] を用いて、基底画像を STL 形式の立体モデルへ変換し、回転スクリーンに配置可能な基底素材を作成する。

3.3 追加最適化によるデノイズ手法

基底分解によって得られる再構成画像では、マスク画像にバイナリ制約を課していることに起因し、粒状のノイズが生じる場合がある。そこで本節では、基底分解における再構成画像に生じるノイズを低減するための追加最適化手法を導入する。

まず、基底分解によって得られた再構成画像 \tilde{V}^{depth} に対してデノイズフィルタを適用し、高周波ノイズを抑制した参照画像 $\tilde{V}^{\text{dn}} = [\tilde{v}_1^{\text{dn}}; \tilde{v}_2^{\text{dn}}; \dots; \tilde{v}_N^{\text{dn}}]$ を用意する。次に、 \tilde{V}^{dn} と \tilde{V}^{depth} で、L2 距離および SSIM を計算し、その差が最小となるように H^{depth} のみを最適化する。そのため、最適化問題は以下の式で定式化される。

$$\begin{aligned} \min_{H^{\text{depth}}} \quad & \sum_{i=1}^N f_1 \tilde{v}_i^{\text{depth}} \tilde{v}_i^{\text{dn}} \\ & + \sum_{i=1}^N \text{SSIM}(\tilde{v}_i^{\text{depth}}; \tilde{v}_i^{\text{dn}}) g \quad (4) \end{aligned}$$

本稿では、デノイズフィルタとしてメディアンフィルタを用い、 $f_1 = 0.2$ 、 $f_2 = 0.8$ とした。

3.4 テクスチャ生成モデルを用いた提示目標の最適化

上述の基底分解手法では、複数のテクスチャ画像を少数の基底画像の組み合わせで再構成するため、提示目標のテクスチャ画像間で特徴が大きく異なる場合、再構成結果にテクスチャの崩れが生じることがある。一方で、提示目標間の大域的な特徴が近い場合、各テ

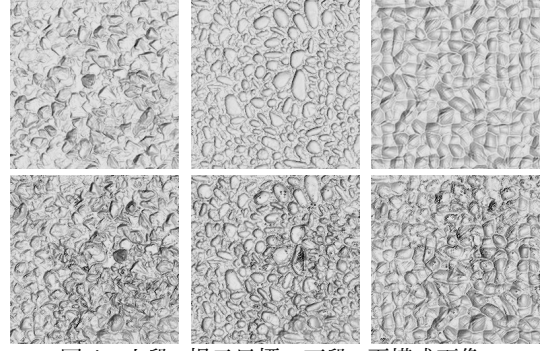


図4 上段: 提示目標. 下段: 再構成画像.

Fig. 4 Top: Target images. Bottom: Reconstructed images.

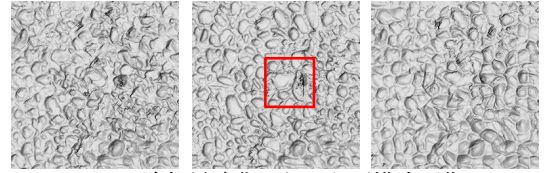


図5 追加最適化を行った再構成画像.

Fig. 5 Reconstructed images after additional optimization.

クスチャに共有する構造を同一の基底で再構成しやすくなる。

そこで、提示目標とするテクスチャ画像の集合から基準となるテクスチャ画像を選択し、MaterialGAN [2] を用いて提示目標間の大域的な特徴を合わせる手法を提案する。具体的には、提示目標とするテクスチャ画像の集合 V^{normal} から基準となるテクスチャ画像 v_s^{normal} を選択し、 v_s^{normal} に対応する潜在ベクトル z_s をそれ以外の各テクスチャ画像 v_i^{normal} ($i \notin s$) に向けて最適化を行う。ここで、MaterialGAN の生成器を $G(\cdot)$ とすると、ある潜在ベクトル z から生成される法線マップ v^{normal} は以下の式で表される。

$$v^{\text{normal}} = G(z) \quad (5)$$

本稿では、GCLoss を損失関数として用い、 v_i^{normal} ($i \notin s$) に対して、次式で定義される最適化問題を解くことにより z_s を更新する。

$$\min_{z_s} L_{GC} G(z_s); v_i^{\text{normal}}; \delta i \geq f_1; \dots; N g n f s g \quad (6)$$

この最適化により、図3下段の右に示すように v_s^{normal} の大域的な特徴は初期状態からほぼ変化せず、局所的なテクスチャの模様のみが変化した生成結果 $v_{s,(t)}^{\text{normal}}$ が得られる。これは、GCLoss が局所的な統計量に基づく指標であり、画像全体の大域的な特徴を強く拘束しない性質を持つためである。最後に、 $v_{s,(t)}^{\text{normal}}$ における勾配を積分することでデプス画像を生成し、3.2 節と同様に陰影画像を作成することで提示目標を得る。

3.5 テクスチャ生成モデルを用いた連続的に変化するテクスチャ提示手法

本節では、テクスチャ生成モデルを用いて、連続的に変化するテクスチャ提示手法について説明する。具体

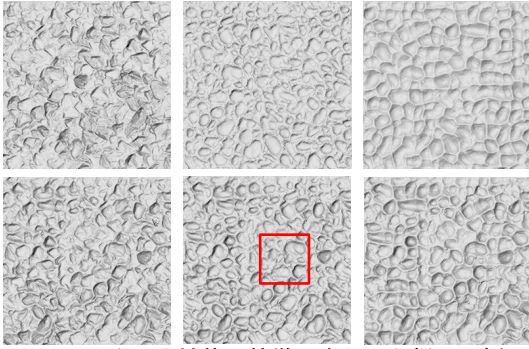


図 6 上段: 大域的な特徴を合わせた提示目標.
下段: 再構成画像.

Fig. 6 Top: Target textures with aligned global features. Bottom: Reconstructed images.

的には、まず連続的に変化させたい二つのテクスチャ画像 v_0^{normal} および v_1^{normal} を用意する。次に、3.4 節で述べた手法によりテクスチャ特徴量を用いて v_0^{normal} を v_1^{normal} に向けて最適化する。この最適化の過程では、図 3 の下段に示すように v_0^{normal} から $v_{0,(1)}^{\text{normal}}$ の外観へと徐々に変化する中間的な画像が得られる。本稿では、あらかじめ用意した二つのテクスチャ画像と、これらの最適化過程で得られる中間的な生成画像を提示目標として基底分解を行うことで、連続的に変化するテクスチャを提示する。

4 実験

4.1 概要

本節では、提案手法の有効性を確認するために二つの実験を行う。一つ目の実験では、基底分解によって求めた基底画像とマスク画像を用い、複数のテクスチャの再構成結果を確認し、追加最適化およびテクスチャ生成モデルを用いて大域的な特徴を合わせた結果との比較を行う。加えて、基底画像を立体化した基底素材を用い、Blender による 3DCG 上での再構成結果を確認する。

二つ目の実験では、テクスチャ生成モデルを用いることで、連続的に変化するテクスチャ提示が可能であるかを検証する。この実験においても、画像上での再構成結果と、Blender による 3DCG 上での再構成結果を確認する。

画像上での再構成では、基底画像とマスク画像を用い、画素値の線形和によってテクスチャ画像を再構成する。提示目標のテクスチャ画像、基底画像、およびマスク画像の解像度は 256×256 pixel とした。なお、提示目標および再構成結果の可視化には陰影画像を用いた。実験で用いるデプス画像および法線マップは、MatSynth データセットから選択した [10]。一つ目の実験では、提示目標として図 4 に示す画像を用い、基底画像は 2 枚とした。二つ目の実験では、提示目標として図 9 に示す画像を用い、基底画像は 4 枚とした。このとき、MaterialGAN を用いて図 9 の上段におけ

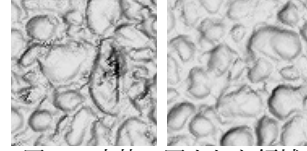


図 7 左: 図 5 の赤枠で囲まれた領域. 右: 図 6 の赤枠で囲まれた領域.

Fig. 7 Left: Region enclosed by the red box in Fig.5. Right: Region enclosed by the red box in Fig.6.

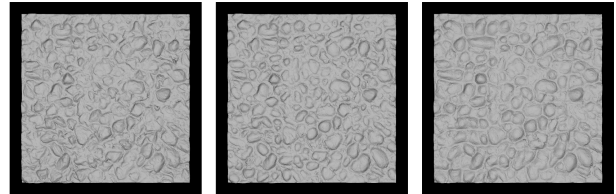


図 8 Blender での再構成結果.

Fig. 8 Reconstruction results in Blender.

る両端の画像から中間の 6 枚の画像を作成した。

Blender によるテクスチャの再構成では、3.2 節で説明した方法により作成した基底素材を用い、光源からマスク画像を投影する。その後、マスク画像が投影された各基底素材をレンダリングし、得られたレンダリング画像を画像上で重ね合わせることでテクスチャを再構成する。

4.2 テクスチャ特徴量を用いた複数のテクスチャ再構成

基底分解によって得られた基底画像とマスク画像から再構成した結果を図 4 に示す。図 4 より、再構成画像において、提示目標のテクスチャ画像に含まれる模様を再現していることを確認できる。

次に、追加最適化によるデノイズを行った再構成結果を図 5 に示す。図 4 と図 5 を比較すると、再構成画像中に見られた粒状のノイズが減少していることがわかる。

さらに、大域的特徴を合わせた提示目標を用いたときの再構成結果を図 6 に示す。また、図 5 および図 6 における左から二番目の再構成画像について、赤枠で囲った領域を図 7 に示す。図 7 より、大域的な特徴を合わせていない再構成結果では、テクスチャが潰れている箇所がある。一方、大域的な特徴を合わせることで、再構成画像に生じるテクスチャの崩れが低減していることを確認できる。

次に、Blender による再構成結果を図 8 に示す。図 8 より、提示目標のテクスチャの特徴を確認できる。

4.3 テクスチャ生成モデルを用いた連続的に変化するテクスチャ提示

基底分解によって得られた基底画像とマスク画像から再構成した結果を図 9 に示す。図 9 より、提示目標のテクスチャを再現できており、斜め方向のテクスチャが徐々に変化している様子が確認できる。

