

多数のテクスチャ提示が可能なマテリアルディスプレイのための 基底抽出手法の検証

湯蓋 康平^{*1}渡辺 義浩^{*1}

Abstract – 現実の物体を実際に近い形状や質感で提示するディスプレイは、デザイン産業などでの応用が期待できる。これまでに、実物体をスクリーンに直接配置することで、物体表面の微細な幾何構造であるテクスチャを提示するディスプレイが提案された。一方、同手法では配置された実物体によって提示されるテクスチャの種類が限定される。そこで、時間統合的にスクリーンに配置された素材を重ね合わせることで空間的に配置可能な素材数に制限がある中でも多数のテクスチャを提示する手法を提案する。そのために、画像群から主要な特徴を持つ少数の基底画像と重みを抽出する基底抽出においてテクスチャ画像に特化した画像類似度指標を用いる手法を提案する。検証では既存手法と比べ、多数のテクスチャ画像を再構成できることを確認した。

Keywords : 質感操作, 質感再現, 次元削減

1 はじめに

実世界の物体を実際に近い形状や質感で提示できるディスプレイはプロダクトデザインや、ネットショッピングなどにおいて体験価値の向上が期待できる。特に、現実的な質感を提示するためには、物体の表面構造による反射特性の再現が重要な要素である [1]。物体の表面構造を現実空間で再現することにより、自然な視差効果が得られ、物体の表面特性や奥行きを正確に知覚させることができる。本稿では、このように物体表面の構造を再現し、視覚的に現実に近い質感を提示するディスプレイをマテリアルディスプレイと定義する。

マテリアルディスプレイの一例として、Asahina らは、螺旋状のスクリーンに実素材を直接貼り付けることで、プロジェクターから階調投影をせずとも、物体表面の微細な幾何構造 (テクスチャ) を提示可能な 3D ディスプレイを提案した [2]。また、同ディスプレイでは、知覚的残像効果を利用し、複数の実素材を時間統合することでスクリーンに配置していない色を提示することも可能である [3]。しかし、同手法では提示されるテクスチャがスクリーン上に配置した実素材に依存するため、提示可能なテクスチャの種類が限定される。

また、できるだけ少ない要素の組み合わせで多くのデータを表現するための手法として基底抽出と呼ばれる手法がある。同手法では、画像群から主要な特徴を持つ少数の基底画像と重みを抽出し、不要な次元を削減することが可能である。

そのため、上記で述べた 3D ディスプレイにおける

空間的に配置可能な素材数の制約を踏まえ、本稿では時間統合的にスクリーンに配置された素材を重ね合わせることで空間的に配置可能な素材数に制限がある中でも多数のテクスチャを提示する手法を提案する。この目的を達成するために、テクスチャに特化した画像類似度指標を用いた基底抽出手法を提案する。ここで、本手法における時間統合は基底抽出によって得られた基底素材に対して、重みをプロジェクターの照射強度として瞬間的に照射することで観察者には複数の基底素材が重なって知覚されることを意味する。そのため、基底抽出によって得られる値は非負の値である必要がある。

また、本提案手法では提示されるテクスチャが人間の知覚上、参照とするテクスチャと同一に知覚される場合、テクスチャ構造上の崩れを許容し、同一カテゴリ、または類似したテクスチャの提示に限定することで多数のテクスチャ提示が可能であると考えられる。

検証では、基底画像と重みを用いて画素値の線形和によってテクスチャ画像を再構成し、既存手法である Non-negative Matrix Factorization (NMF)[4] と提案手法で再構成可能なテクスチャ画像枚数について比較した。検証の結果、画像類似度指標を用いた本提案手法が、既存手法に比べて多くのテクスチャ画像を再構成できることを確認した。

2 関連研究

2.1 マテリアルディスプレイ

コンピュータグラフィックスの進歩によって従来よりも現実的な質感提示が可能になっている。しかし、提示方法の制約により観察者に負担をかけずに運動視

^{*1}東京科学大学

差や両眼視差によるリアルな質感提示を行うことは難しい。そのため、物体の表面構造を現実空間で再現することで実物体の質感の全てもしくは特定の要素を提示するための様々なディスプレイが提案されている。

例えば、コロイド溶液膜を超音波で振動させることで透明度を操作するディスプレイが提案されている [5]。しかし、この手法では提示面が平面に限定されており、空間的に変化する幾何構造を再現することは困難である。

また、Miyashita らは空間的に配置した複数の実素材とストロボ光の組み合わせで鏡面反射特性や色を操作する時間統合型のマテリアルディスプレイを提案した [6]。同手法ではゾートロープとソーマトロープの時間統合原理を応用している。具体的には、人間の時間分解能を示す臨界融合周波数 (CFF) は低輝度光で約 15 Hz、高輝度光で約 60 Hz とされており、CFF よりも高速に提示像を切り替えることで、人間の目には複数の像が重なって知覚される。そのため、実物体を時間統合的に重ね合わせ、スクリーンに配置していない素材を合成することができる。しかし、同手法では鏡面反射特性と色の合成を目的として素材が選定されているため、空間的に変化するテクスチャを時間統合によって操作することはできない。

次に、ポリュメトリックディスプレイの一例として、螺旋型のスクリーンに実素材を貼り付けることで、プロジェクタからは階調投影をせずとも、テクスチャのリアルな提示が可能な 3D ディスプレイが提案されている [2]。この手法では、Miyashita らの手法と同様に時間統合によって複数の実素材を合成し、スクリーン上にない色を持つテクスチャの提示も可能である [3]。しかし、提示可能なテクスチャはスクリーンに配置された実素材に依存する。また、配置可能な実素材の数と立体形状を提示するためのボクセル解像度にはトレードオフの関係があるため、提示可能なテクスチャの種類が限られる。

さらに、投影対象に奥行き方向の微小な高速振動を与え、奥行き位置と同期して高速に画像を投影することでテクスチャ提示を行う手法が提案されている [7]。同手法では毛皮のような細かい 3 次元の表面テクスチャを提示できるが、振動による奥行き範囲にテクスチャが制限される。

2.2 基底抽出

基底抽出は与えられた高次元データから主要な特徴を持つ低次元の基底を取得するための手法である。ここでの基底とは与えられたデータを表現するための基本単位の集合であり、基底の組み合わせでデータを再構成または近似することが可能である。

例えば、主成分分析 [8] はデータの分散が最大となる

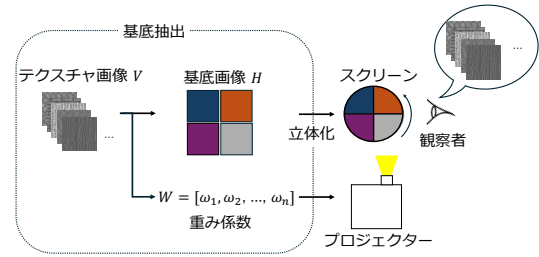


図1 提案手法のパイプライン
Fig. 1 Pipeline of the proposed method

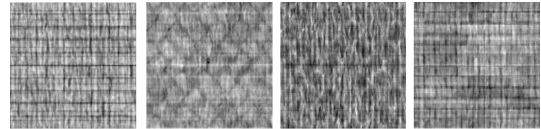


図2 DISTs を用いた基底抽出によって得られた基底画像
Fig. 2 Basis image obtained by basis extraction using DISTs

軸を基底とすることで次元削減を行う。一方、スパースコーディング [9] は基底と重み係数の線形結合でデータを表現することを仮定し、次元削減を行う。さらに、ニューラルネットワークを用いた基底抽出手法としてオートエンコーダが提案されている [10]。同手法では、エンコーダ部分で次元削減を行い、基底を取得することが可能である。しかし、上記の手法で抽出される値には負の値が含まれる場合があるため、実物体と光照射による重みとして再現することができない。

上記の手法に対して、NMF は非負の値のみを要素とする行列から非負の要素を持つ基底行列と重み行列を抽出する手法である [4]。具体的には行列 V をデータ行列、行列 W を重み行列、行列 H を基底行列とし、次式で表される。

$$V \simeq WH \quad (1)$$

さらに、損失関数を式 (2) とし、式 (3)、(4) の乗法更新則によって単一画素ごとに重み行列と基底行列が求められる。

$$\min_{W, H} \|V - WH\|^2 \quad (2)$$

$$W_{ij} \leftarrow W_{ij} \frac{(VH^T)_{ij}}{(WHH^T)_{ij}} \quad (3)$$

$$H_{ij} \leftarrow H_{ij} \frac{(W^TV)_{ij}}{(W^TWH)_{ij}} \quad (4)$$

2.3 画像類似度指標

画像類似度指標は 2 枚の画像間の類似度を測るための指標であり、画像の品質評価や、画像生成の分野で広く利用されている。また、使用する画像の特徴によって様々な手法が提案されている [11]。

まず Structural SIMilarity (SSIM) は人間の知覚的類似性に基づく画像類似度指標である [12]。 C_0, C_1, C_2, a, b, c を定数とし、画像 x, y から得られる平

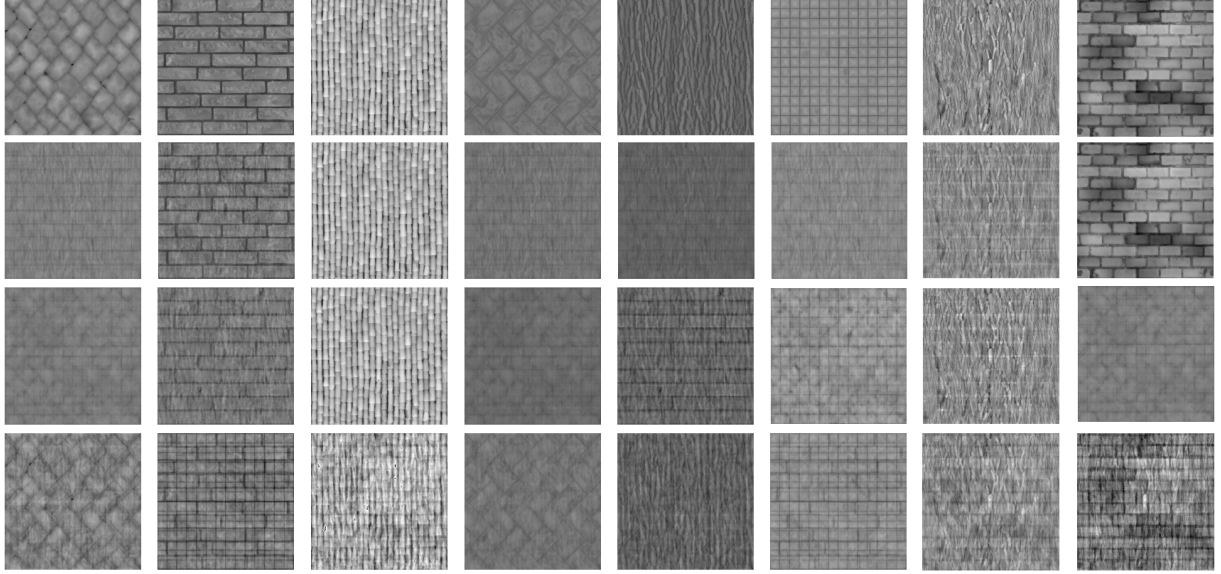


図3 再構成画像比較, (1行目) 参照テクスチャ画像, (2行目)NMF, (3行目)SSIM, (4行目)DISTS

Fig. 3 Reconstructed image comparison, (line 1) reference texture image, (line 2) NMF, (line 3) SSIM, (line 4) DISTS

均値 μ_x , μ_y , 分散 σ_x , σ_y , および共分散 σ_{xy} から輝度項, コントラスト項, 構造項をそれぞれ式 (5), (6), (7) として計算し, 統合した式 (8) によって全体の類似度が求められる.

$$l(x, y) = \frac{2\mu_x\mu_y + C_0}{(\mu_x)^2 + (\mu_y)^2 + C_0} \quad (5)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_1}{(\sigma_x)^2 + (\sigma_y)^2 + C_1} \quad (6)$$

$$s(x, y) = \frac{\sigma_{xy} + C_2}{\sigma_x\sigma_y + C_2} \quad (7)$$

$$q(x, y) = (l(x, y))^a (c(x, y))^b (s(x, y))^c \quad (8)$$

ただし, 式 (7) は単一画素ごとに計算されるため, テクスチャ画像のような画像全体の構造的相関関係を計算する必要がある場合には適していない [13].

また, Deep Image Structure and Texture Similarity (DISTS) は VGG 畳み込みニューラルネットワーク (CNN) を用いて特徴マップから類似度を計算する手法として提案されている [13]. 同手法では, 構造的な相関に基づいて類似度を計算するため, 特にテクスチャ画像において優れた性能を示している. 具体的には, CNN を用いて取得された特徴マップを基に, SSIM における式 (5) と式 (6) を計算し, 統合した式 (9) により類似度を算出する.

$$D(x, y; \alpha, \beta) = \sum_{k=0}^m \sum_{l=1}^{n_k} \left(\alpha_{kl} l(\tilde{x}_l^{(k)}, \tilde{y}_l^{(k)}) + \beta_{kl} c(\tilde{x}_l^{(k)}, \tilde{y}_l^{(k)}) \right) \quad (9)$$

ここで, m は畳み込み層の数, n_k は k 番目の畳み込み層から得られた特徴マップの数を表し, $\tilde{x}_l^{(k)}$, $\tilde{y}_l^{(k)}$ は

画像 x , y を CNN に入力した際の k 番目の畳み込み層における l 番目の特徴マップを意味する. また, α , β は被験者実験の結果をもとに, 人間の知覚評価と一致するように調整された定数 α_{kl} , β_{kl} の集合である.

2.4 テクスチャ特徴量

グラム行列は画像一枚からテクスチャ特徴量として計算されるテクスチャ評価指標である [14]. 同手法では, CNN における特徴マップ F の内積からグラム行列 G が計算される, 例えば, k 番目の畳み込み層における s , t 番目の特徴マップの内積によって得られるグラム行列 G_{st}^k は式 (10) で求めることができる. ここで, M_k は特徴マップの要素数を表す.

$$G_{st}^k = \sum_{u=1}^{M_k} F_{su}^k F_{tu}^k \quad (10)$$

また, グラム行列を用いてテクスチャ画像を生成する手法が提案されている. 具体的には, テクスチャ画像 x とランダムに初期化した画像 \hat{x} をそれぞれ CNN に入力し, 畳み込み層ごとに得られる各グラム行列, G_{st}^k , \hat{G}_{st}^k の差を損失関数とする式 (12) が最小化するように勾配降下法によって画像 \hat{x} を更新することで画像生成を行う. 以下に計算式を示す.

$$E_k = \frac{1}{4N_k^2 M_k^2} \sum_{s,t} \left(G_{st}^k - \hat{G}_{st}^k \right)^2 \quad (11)$$

$$L(x, \hat{x}) = \sum_{k=0}^L w_k E_k \quad (12)$$

ここで, N_k は k 番目の畳み込み層における特徴マップの数, L は畳み込み層の総数, w_k は重み係数を示す.

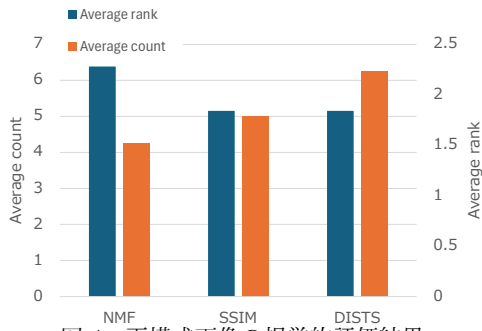


図4 再構成画像の視覚的評価結果
Fig.4 Results of visual evaluation of reconstructed images

3 提案手法

3.1 概要

本稿では図1に示すような手法を提案する. 具体的には回転スクリーンに配置されたテクスチャ特徴を持つ基底素材とプロジェクターの照射強度による重みを時間統合的に組み合わせることで, スクリーンに配置した基底素材数よりも多くのテクスチャ提示を行う. このとき, 時間統合はプロジェクターからの光をCFFを超える速度で照射することにより, 人間の目には複数の基底素材が重なり, 再構成されたテクスチャとして知覚されることを意味する.

さらに, より少ない基底素材数で多くのテクスチャを提示するために, 提示目標のテクスチャ画像群から画像類似度指標に基づいて基底抽出を行う. この基底抽出によって得られた基底画像を立体物へ変換し, 基底素材を作成する. この際, スクリーンに配置する基底素材と光照射による重みは現実空間で再現される必要があるため, 基底抽出では式(1)および式(14)が成り立つ必要がある.

3.2 画像類似度指標を用いた基底抽出手法

本稿では図1に示す基底抽出手法において基底画像と重みの線形和によってテクスチャ画像が再構成され, 全ての値が非負の制約のもと, 画像類似度指標を用いることを提案する. 具体的には画像類似度指標を損失関数 f とし, ランダムに初期化された基底画像 H と重み行列 W を勾配降下法によって損失関数が最大化するように値を更新することで基底抽出を行う. また, 本手法で用いる画像類似度指標は0~1の値を取り, 類似度が高いほど大きい値が出力される. 以下に, 本手法における計算式を示す.

$$\max_{W, H} f(V, WH) \quad (13)$$

$$W \geq 0, H \geq 0 \quad (14)$$

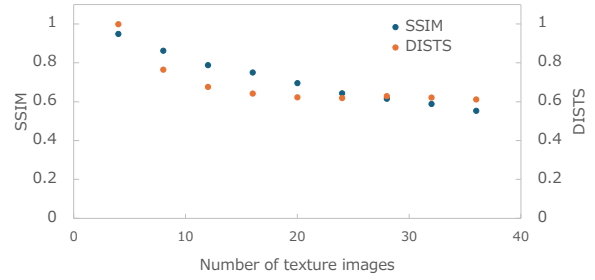


図5 テクスチャ画像枚数変更に伴う類似度推移
Fig.5 Similarity transition with change in number of texture images

4 検証

4.1 基底抽出による基底画像と重みの取得

NMF, SSIM, DISTSを用いた基底抽出を行い, テクスチャ画像を再構成可能か検証した. このとき, 使用したテクスチャ画像はPolyHeaven [15]とtexture.com [16]から選定し, 図3に示す8枚を使用した. また, 図1にあるように, 基底画像を立体化することを想定してデプス画像を選択し, 基底画像の枚数を4と設定した.

次に, DISTSを用いた基底抽出手法で取得した基底画像を図2に示す. 図2より, 基底画像は複数のテクスチャ画像が重なったような画像となり, 重みはおよそ0.1~0.5の範囲の値を示した. また, 他手法においても同様の基底画像と重みが得られた.

さらに, 各基底抽出手法によって得られた基底画像と重みを画素値の線形和で再構成した画像を図3に示す. NMFを用いた基底抽出手法では, 再構成画像のうち複数枚がテクスチャ画像と大きく異なる結果となった. 一方, SSIMおよびDISTSを用いた基底抽出手法では再構成された画像のうちのほとんどで参照したテクスチャ画像の模様が確認された.

4.2 再構成画像の視覚評価

前小節で抽出された基底画像と重みによる再構成画像について, 4人の被験者に協力を依頼し, 2種類の視覚的評価実験を行った. まず1つ目は, 各テクスチャ画像と手法ごとの再構成画像を比較し, 再構成画像にテクスチャ画像の特徴を確認できるかを判定するよう被験者に指示した. 次に, 再構成画像が参照したテクスチャ画像を再構成できていると判断された枚数の合計から, 再構成画像の平均枚数を算出した.

2つ目は, テクスチャ画像と各基底抽出手法で得られた再構成画像3枚を比較し, テクスチャ画像に近い特徴を持つと思われる再構成画像の順位を付けるよう指示した. その結果から, 各手法の平均順位を求めた. このとき, どちらの実験でも観察者の位置や環境光の条件などで差が極力発生しないように調整した.

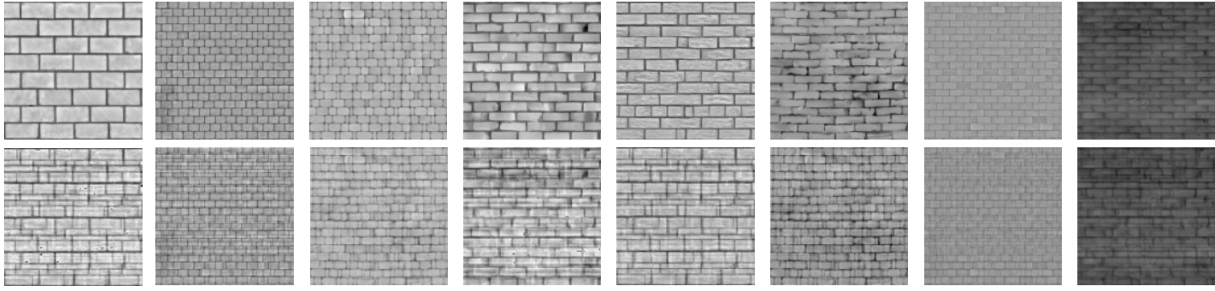


図6 再構成失敗例. 上段にテクスチャ画像, 下段に再構成画像を示す. 明瞭な線が重なり合うことでテクスチャ画像の特徴が失われている.

Fig. 6 Example of reconstruction failure. The texture image is shown in the upper panel and the reconstructed image in the lower panel. The overlapping of distinct lines has lost some of the features of the texture image.

次に, 視覚的評価実験の結果を図4に示す. まず, 1つ目の実験ではDISTSを用いた基底抽出手法による再構成画像が最も多くのテクスチャ画像の特徴を再構成できていることがわかった. 一方, 2つ目の実験ではSSIMおよびDISTSを用いた基底抽出手法による再構成画像が同順位を示し, NMFによる再構成画像は他手法よりも低い順位に評価された.

4.3 画像枚数と種類の変更による検証

テクスチャ画像枚数を増やした場合のテクスチャ画像と再構成画像の類似度推移について検証した. まず, 基底画像を4枚とし, テクスチャ画像を4枚ずつ変化させ, 基底抽出を行った. このとき, テクスチャ画像の選択方法は最も多い枚数の36枚パターンから基底抽出を行った際に各テクスチャ画像と再構成画像で類似度を計算し, 最も類似度が低い4枚のテクスチャ画像を除外することで選択される. 全てのパターンで基底抽出を行った後, 各テクスチャ画像と再構成画像の類似度を計算し, その平均値をプロットした結果を図5に示す. また, NMFについては前小節において明らかに他2手法に比べて再構成可能なテクスチャ画像枚数が少ないため検証から除いた.

結果より, SSIMを用いた基底抽出手法ではテクスチャ画像枚数の増加に伴って線形的に類似度が低下する傾向があるのに対し, DISTSを用いた基底抽出手法では非線形的に類似度が変化していることがわかる.

次に, さまざまなテクスチャ画像を用いて基底抽出を行った結果, 再構成失敗例について図6に示す. このとき, DISTSを用いて基底画像枚数を4として基底抽出を行った. 図6より, レンガ模様の画像では, 明瞭な線が重なり合うことで, テクスチャ画像の特徴が失われていることを確認した.

5 考察

5.1 画像類似度指標を用いた基底抽出

検証の結果, テクスチャ類似度指標であるDISTSを用いた基底抽出手法が最も多くのテクスチャ画像を再構成できることを確認した. 一方, 使用するテクスチャ画像によって再構成結果が変化するため, 画像選定について十分に検討する必要がある. 解決策として, 類似度の高い画像同士をテクスチャ画像として使用することが考えられるが, 類似度が高すぎる場合, 異なる画像でも人間には区別がつかない可能性がある. そのため, 異なる画像として許容できるセットを選定する必要がある.

5.2 グラム行列を用いた基底抽出

追加検証として, グラム行列を用いた基底抽出を行った. 具体的には, 式(12)において, テクスチャ類似度指標による基底抽出手法と同様に, テクスチャ画像とランダムに初期化した基底画像と重みによる再構成画像のグラム行列の差が最小化するように勾配降下法を用いて基底画像と重みを得る. また, 使用したテクスチャ画像はPertex [17]から選定し, 基底画像枚数は4とした.

次に, グラム行列を用いて基底抽出を行った結果を図7に示す. 結果より, 再構成画像はテクスチャ画像に比べ, 模様が崩れた画像になっている. 一方, 図3の画像類似度指標を用いた基底抽出手法による再構成画像に比べ, 複数の画像が重なったような再構成画像になっていない. そのため, 画像としての崩れを許容すれば多数のテクスチャ提示を行うための手法として応用できる可能性がある.

6 まとめ

本稿では多数のテクスチャ提示が可能なマテリアルディスプレイのために, 基底素材と重みの線形和によってテクスチャ画像が再構成でき, 全ての値が非負

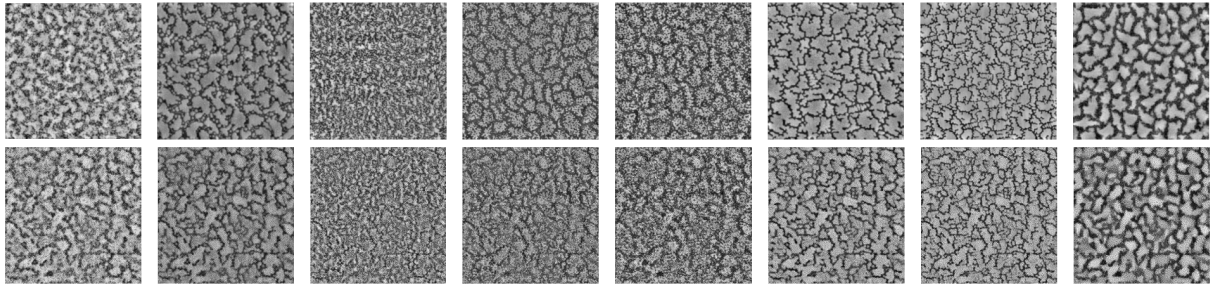


図7 グラム行列を用いた基底抽出. 上段がテクスチャ画像, 下段が再構成画像
 Fig.7 Basis extraction using a gram matrix. Upper panel: texture image, lower panel: reconstructed image

の制約のもと画像類似度指標を用いた基底抽出手法を提案した. 検証から, 既存手法に比べ, 提案手法が多くのテクスチャ画像を再構成できることを確認した. また, 本手法は多数のテクスチャ提示を行うためのスクリーン配置素材の作成に用いることができる可能性を示した.

参考文献

- [1] Matthias B Hullin, Ivo Ihrke, Wolfgang Heidrich, Tim Weyrich, Gerwin Damberg, and Martin Fuchs. State of the Art in Computational Fabrication and Display of Material Appearance. In *Eurographics Annual Conference (STAR)*, 2013.
- [2] Ray Asahina, Takashi Nomoto, Takatoshi Yoshida, and Yoshihiro Watanabe. Realistic 3D Swept-Volumetric Display with Hidden-Surface Removal Using Physical Materials. *IEEE VR*, pp. 113–121, 2021.
- [3] Tangjitruamboon Tananan, 彭浩倫, 渡辺義浩. ポリユメトリックディスプレイのための素材融合による色操作に関する検証. 映像情報メディア学会 情報ディスプレイ研究会, 2021.
- [4] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *nature*, Vol. 401, No. 6755, pp. 788–791, 1999.
- [5] Yoichi Ochiai, Alexis Oyama, Takayuki Hoshi, and Jun Rekimoto. Poppable display: A display that enables popping, breaking, and tearing interactions with people. In *2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE)*, pp. 124–128. IEEE, 2013.
- [6] Leo Miyashita, Kota Ishihara, Yoshihiro Watanabe, and Masatoshi Ishikawa. Zoematrope: a system for physical material design. *ACM Transactions on Graphics*, Vol. 35, No. 4, pp. 1–11, 2016.
- [7] Hayase Nishi, Daisuke Iwai, and Kosuke Sato. 3D texture representation in projection mapping onto a surface with micro-vibration. In *SIGGRAPH Asia 2024 Posters*, pp. 1–2. 2024.
- [8] George H Dunteman. *Principal components analysis*, Vol. 69. Sage, 1989.
- [9] Stéphane G Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on signal processing*, Vol. 41, No. 12, pp. 3397–3415, 1993.
- [10] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, Vol. 313, No. 5786, pp. 504–507, 2006.
- [11] Kaixuan Zhang, Zhaochen Shi, Jana Zujovic, Huib De Ridder, Rene Van Egmond, David L Neuhoff, and Thrasymvolos N Pappas. Training and testing texture similarity metrics for structurally lossless compression. *IEEE Transactions on Image Processing*, Vol. 33, pp. 1614–1626, 2024.
- [12] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, Vol. 13, No. 4, pp. 600–612, 2004.
- [13] Keyan Ding, Kede Ma, Shiqi Wang, and Eero P Simoncelli. Image quality assessment: Unifying structure and texture similarity. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 44, No. 5, pp. 2567–2581, 2020.
- [14] Leon Gatys, Alexander S Ecker, and Matthias Bethge. Texture synthesis using convolutional neural networks. *Advances in neural information processing systems*, Vol. 28, , 2015.
- [15] Poly Haven. Poly Haven, 2020年7月24日. 2024年11月8日閲覧.
- [16] Rephael Lacoste. textures.com, 2013年10月27日. 2024年11月8日閲覧.
- [17] Fraser Halley, et al. *Perceptually relevant browsing environments for large texture databases*. PhD thesis, Citeseer, 2012.