

階層型確率的な主成分分析モデルによるテクスチャの生成

鈴木 藍雅^{1,2,a)} 庄野 逸^{3,b)}

概要: 多様なテクスチャを表現できるモデルの作成は、自然画像の微視的なモデリングとして重要な課題である。Portilla と Simoncelli は、ヒトの視覚メカニズムに基づいたテクスチャ特徴量、およびその特徴量を元にしたテクスチャの生成モデルを提案し、このモデルによって多様なテクスチャが生成できることを示している。一方で、この Portilla らのモデルは、自然画像に現れる典型的なテクスチャにおいては、特徴量の成分同士に関連性をもつ。本研究では、Portilla らの提案した特徴量のうち、テクスチャの生成に寄与する潜在的な構造因子を抽出するモデルの提案を行う。このモデルは、特徴量のグループ構造に着目し、階層的な主成分分析を行うことで有効な特徴量の縮約表現を獲得するものである。実験では、提案したモデルを用い、テクスチャデータセットから特徴量の縮約表現を獲得し、低次元の表現によって、テクスチャ画像の生成が行えることを示した。

A Generative Model of Textures Using Hierarchical Probabilistic Principal Component Analysis

AIGA SUZUKI^{1,2,a)} HAYARU SHOUNO^{3,b)}

1. はじめに

我々の視覚的刺激をもたらすあらゆる風景は、微視的には多種多様なテクスチャのパッチワークにより構成されているといえる。生物学的な知見として、人間の物体の視覚的認知には、その形状 (shape) に加えて、それらのテクスチャ構造も大きく寄与していることが知られている [1]。このことから、コンピュータビジョンにおける自然画像のモデリングにおいて、多様なテクスチャを表現できるテクスチャ構造のモデル化は、重要な課題であるといえる。

近年における主要なテクスチャの生成モデルとしては、Markov 確率場 (Markov Random Field; MRF) を用いたものが挙げられる [3], [4], [5], [8], [9]。これらの手法に共通する考え方は、テクスチャ画像は 2 次元格子空間上の確率場と

して構造化されているという仮定である。MRF による手法は、この仮定によって認知的な質感の等化 (perceptually equivalence) に関して考慮されていないという問題が存在する。

質感の等化とは、我々があるテクスチャ A と B を視覚的に認知した時に、「A と B は同じ模様である」と認識するメカニズムのことである。これは自然画像のモデリングにおいては自然な要請であるが、格子空間上の画素値間のモデリングという観点から、陽に得られる結果ではない。

質感の等価を取り入れたテクスチャのモデリングとして、Portilla と Simoncelli による特徴量 (Portilla-Simoncelli statistics, 以下 PSS とする) ベースの手法がある [6], [7]。この手法は、各種のテクスチャは PSS によって表現できると仮定し、特徴量マッチングによる最適化問題として、その生成系を構成するものである。PSS は多様なテクスチャが表現できることが示されている一方で、統計的な拘束条件を導入している原理上、PSS はその要素間に関連を持つと考えられる。さらに、自然画像に現れるような典型的なテクスチャは、ある限定された構造を持つと考えられるため、より小さい次元によって記述可能であると予測できる。

¹ 筑波大学システム情報工学研究科

〒 305-8577 茨城県つくば市天王台 1-1-1

² 産業技術総合研究所 人工知能研究センター

〒 305-8560 茨城県つくば市梅園 1-1-1

³ 電気通信大学情報理工学研究所

〒 182-8585 東京都調布市調布ヶ丘 1-5-1

^{a)} ai-suzuki@aist.go.jp

^{b)} shouno@uec.ac.jp

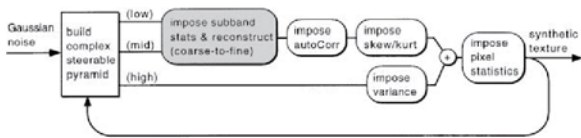


図1 PSSによるテクスチャ生成の流れ図. [7]より引用. Gauss ノイズから PSS が所望の値に近づくよう再帰的に繰り返すによる最適化を行う.

本稿では、テクスチャの生成に寄与する潜在的な構造因子を確率的な主成分分析 (Probabilistic Principal Component Analysis; PPCA) を用いて、獲得する手法を提案する. また得られた構造因子を用いて、少数のパラメータによって、多様なテクスチャを生成するモデルについての検討を行う.

2. PSS の概要

本節では、提案手法において重要な概念となる Portilla, Simoncelli らにより提案されたテクスチャ特徴である PSS について、簡単に述べる.

はじめに入力されたテクスチャ画像を Steerable filter pyramid [10] と呼ばれる、マルチスケール・方位選択性を持つ画像分解手法を用いて分解する. この分解画像を再構成することで、入力されたテクスチャの特定のスケール・方位の成分を抽出することができる. PSS の構築においては、テクスチャのスケール・方位ごとの再構成画像から、テクスチャの構造を決定する 10 グループの基準を考える. PSS はこれらのグループごとに、相関などの種々の要約統計量を算出し、得られた統計量をひとつのベクトルとしてまとめることでテクスチャ特徴量とする.

PSS から実際にテクスチャを生成する際は、図1に示すように、Gauss ノイズの画像から繰り返すによる最適化によって所望の画像に近づけてゆく画像を構成する. このアルゴリズムはおよそ 50 回の繰り返すによって、概ね所望の PSS に収束することが実験的に示されている [7].

PSS はピクセルの近傍画素を定義するパラメータ K , 画像分解における方位数 N , スケール数 M により制御され、その次元数が決定される. 以後は [7] で一貫して用いられている $N = 4, K = 4, M = 7$ を実験において用いる. またこの時の PSS は $D = 1784$ 次元となる.

3. 問題設定

3.1 実験に用いたテクスチャデータ

実験に用いるテクスチャデータベースには、Uppsala University より提供されている Kylberg Texture Dataset v. 1.0 を用いた. これは 28 種類の様々なテクスチャ画像を含むデータセットであり、各クラスごとに 1920 枚のユニークなパッチ画像が含まれている. 各パッチは 576×576 [px] の 256 階調のグレースケール画像であり、パッチごとに平均の輝度値が 127 となるよう正規化されている. 実験の際には、各パッチごとに 1920 枚のうちの 1720 枚をモデルの

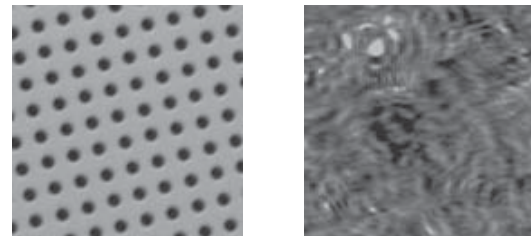


図2 単純な Linear-PPCA の適用による Ceiling クラスの再構成画像. (左): 元画像 (右): 1000 次元による再構成画像; 元画像の特徴を全く表現できていないことが分かる.

学習データ, 残りの 200 枚をモデルの性能検証用のテストデータとして用いた. また, モデルへの入力には, 画像の解像度が高すぎるため, 128×128 [px] ヘダウンサンプリングし用いた.

3.2 モデルの性能評価

モデルによるテクスチャ生成の評価尺度には, 2 つのテクスチャ間の類似度の指標である Texture Similarity Score (TSS) [3] を用いる. 生成されたテクスチャパッチ \mathbf{s} と比較対象のテクスチャ \mathbf{x} が与えられた元での TSS はそれらの Normalized Cross Correlation (NCC) として

$$\text{TSS}(\mathbf{s}, \mathbf{x}) \triangleq \max \left\{ \frac{\mathbf{x}_{(1)}^T \mathbf{s}}{\|\mathbf{x}_{(1)}\| \|\mathbf{s}\|}, \dots, \frac{\mathbf{x}_{(I)}^T \mathbf{s}}{\|\mathbf{x}_{(I)}\| \|\mathbf{s}\|} \right\} \quad (1)$$

と定義される. ここで $\mathbf{x}_{(i)}$ は \mathbf{s} と同じサイズの \mathbf{x} 上のパッチで, I は取りうるすべての場合の数である. 実験における TSS の計算は [3] と同様に, 19×19 [px] の領域を \mathbf{s} からランダムに抽出し求める.

4. PPCA による PSS の階層的な次元削減

本研究では次元削減の手法に, Gauss 性のノイズをもつデータの分布が最大となるような, 新たな正規直交軸を与える確率的な主成分分析 (Probabilistic Principal Component Analysis; PPCA) [11] を用いた. PPCA は通常の決定論的な主成分分析 (Principal Component Analysis; PCA) とほとんどの同様が結果を得られるが, 次元が極めて大きくなった際の最適化が容易であることや, 情報量基準による適切な削減次元の決定ができることが利点として挙げられる [2].

4.1 予備実験:単純な PPCA による次元削減の試み

予備実験として, はじめに単純な Linear-PPCA によって直接の次元削減を試みた. しかし, 図2に示すように, 元の PSS の構造を十分に保持するであろう, 高次元への変換においても十分な結果が得られなかった. これは単純な PPCA の適用によっては, PSS の潜在構造を十分に表現できないためであると考えられる.

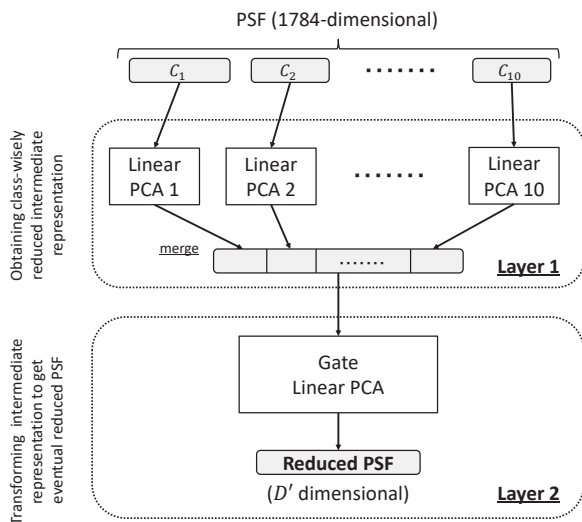


図3 階層型 PPCA モデルの概要図。はじめに PSS を各グループ毎に分割し、それぞれの構造を中間層の Linear-PPCA を用いて学習する。それらから得られた中間表現に対して再度 PPCA を適用し、次元削減を行うモデルである。

4.2 階層型 PPCA モデルによる PSS の次元削減

2 節で述べたように、PSS は画像分解から得られた各種の要約統計量を並べてベクトル化したものである。すなわち、元の PSS は複数の異なる特徴量を束ねたものであると言える。この点に着目し、特徴量のグループ構造を考慮した、階層的な PPCA による PSS の次元削減を行うモデルを提案する。提案する階層型 PPCA モデルの概要を図 3 に示す。

まず 2 節で述べた PSS の 10 種類の要約統計量ごとに、異なる Linear-PPCA を用いて、グループごとの構造を学習する。その後、各 Linear-PPCA から得られた縮約表現をベクトル化し、モデルの中間表現とする。最後に、中間表現を入力とする Linear-PPCA を用いて、各グループ間の相互作用を加味した構造の学習を行い、最終的な PSS の縮約表現を得る。

5. 実験

5.1 適切な次元数の決定

はじめに、中間層 (図 3, Layer 1) の各 PPCA における適切な次元数を決定する。出力層 (図 3, Layer 2) の累積寄与率を 1、すなわち次元削減を行わないようモデルを固定し、中間層の次元選択の際の累積寄与率 r を変化させた時のテストデータに対する TSS の変化をみる。この時の TSS の変化に関するグラフを図 4 に示す。

これをみると、TSS は中間層の累積寄与率に対して、おおむね単調に増加していることがわかる。また、0.99999999 以降は累積寄与率の変化が 32-bit 浮動小数点数の計算機イプシロンを下回るほど小さな値であったため、累積寄与率は上記の $r = 0.99999999$ を用いた。また、この時の PSS の中間表現の次元は 965 次元であった。

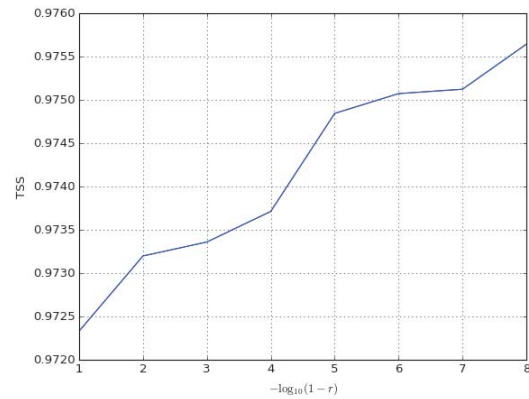


図 4 中間層の累積寄与率に対する TSS の変化。(縦軸) TSS, (横軸) $-\log_{10}(1-r)$.

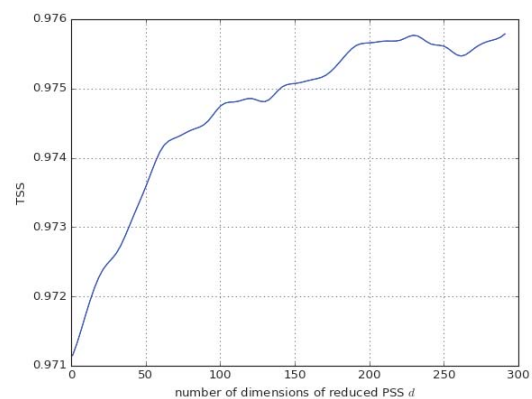


図 5 出力層の出力次元に対する TSS の変化。(縦軸) TSS, (横軸) 出力層の次元。

次に、出力層の適切な次元数を決定する。中間層の累積寄与率を、先に求めた値に固定し、出力層の出力次元を変化させた際のテストデータに対する TSS の変化をみる。この時の TSS の変化に関するグラフを図 5 に示す。

これをみると、変動は大きいもののおよそ 200 次元を超えたあたりで TSS の向上は飽和しているように見える。そのため、最終的に用いる PSS の縮約表現の次元は 200 次元とした。

5.2 視覚評価

提案手法を用いて縮約した PSS を用いて、テキストの生成が行えることを示す。はじめに、前節で定めたパラメータを用いて、テストデータのテキストから得られた PSS を構築したモデルで $1784 \rightarrow 200 \rightarrow 1784$ 次元へ再構成した。そうして得られた PSS を用いて、テキストの生成を行い、それらを元の入力画像と比較し視覚評価を行う。図 6 に示す。

この結果から細かなパターンに関しては、入力と近いテキストを生成できていることがわかる (図 6 上 ceiling1, cushion1, blanket1 クラス)。一方で大きなパターン構造を

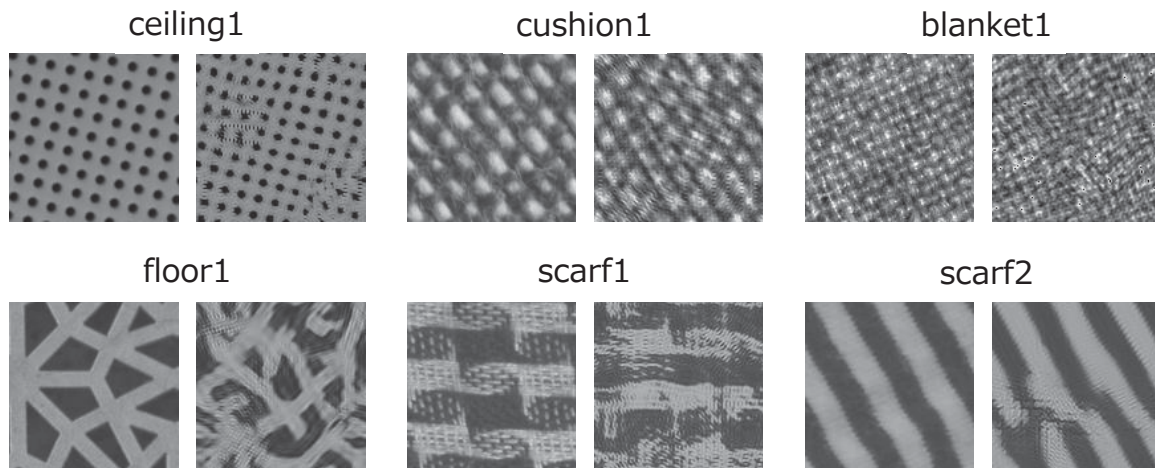


図6 200次元の縮約されたPSSからの再構成結果。それぞれのクラスごとに左側が元の画像、右側が再構成画像。クラス名はKylberg Texture Datasetにおけるクラス名と同じ。クラスによりばらつきはあるが、おおむね元のクラスの特徴を表現できていることがわかる。

もつクラスにおいては、おおむね画像の傾向は表現できているものの、テクスチャの連続するべき構造をうまく再現できない例が見受けられた。(図6下 floor1, scarf1, scarf2 クラス)。これは、提案したモデルによる縮約表現が画像の低周波成分の情報をうまく再現出来ていないためであると考えられる。また、再構成結果のいくつかの例において、直交 Wavelet 基底を用いた画像の再構成において見られるようなアーティファクトが生じていた(図6 ceiling1 クラス左上, floor1 クラス左上, scarf2 クラス中央下)。これは SF pyramid からの画像再構成アルゴリズムに起因するものと考えられる。

6. まとめと今後の課題

本稿では、高次元のテクスチャ特徴量から、有効な縮約表現を得るための階層的な PPCA モデルの提案を行った。実験では、テクスチャデータセットに対して、その縮約表現の解析を行った。その結果、もとの特徴量の 1/10 程度の 200 次元の縮約表現によって、もとのテクスチャの構造を表現し、テクスチャの生成が行えることを示した。

今後の課題として、大量の自然画像データセットを用い、実際に自然画像に現れるテクスチャ構造に対して、有効なモデルを構築し、検討することが挙げられる。また、各 PPCA において適切な次元数の決定や、用いる PPCA のカーネル化など、よいよ縮約表現を得るためのアーキテクチャの最適化を行う。

参考文献

[1] Adelson, E. H.: On seeing stuff: the perception of materials by humans and machines, *Photonics West 2001-electronic imaging*, International Society for Optics and Photonics, pp. 1–12 (2001).
[2] Bishop, C. M.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New

York, Inc., Secaucus, NJ, USA (2006).
[3] Heess, N., Williams, C. K. and Hinton, G. E.: Learning Generative Texture Models with extended Fields-of-Experts., *BMVC*, pp. 1–11 (2009).
[4] Kivinen, J. J. and Williams, C. K. I.: Multiple Texture Boltzmann Machines, *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS 2012)*, Vol. 22, pp. 638–646 (online), available from <http://homepages.inf.ed.ac.uk/s0960152/papers/MTBM-AISTATS12.pdf> (2012).
[5] Luo, H., Carrier, P. L., Courville, A. and Bengio, Y.: Texture Modeling with Convolutional Spike-and-Slab RBMs and Deep Extensions, *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 31, pp. 415–423 (online), available from <http://arxiv.org/abs/1211.5687> (2012).
[6] Portilla, J., Navarro, R., Nestares, O. and Taberner, A.: Texture synthesis-by-analysis method based on a multiscale early-vision model, *Optical Engineering*, Vol. 35, No. 8, pp. 2403–2417 (1996).
[7] Portilla, J. and Simoncelli, E. P.: Parametric texture model based on joint statistics of complex wavelet coefficients, *International Journal of Computer Vision*, Vol. 40, No. 1, pp. 49–71 (2000).
[8] Ranzato, M., Mnih, V. and Hinton, G.: Generating more realistic images using gated MRF's, *Nips*, pp. 1–9 (2010).
[9] Ranzato, M., Mnih, V., Susskind, J. M. and Hinton, G. E.: Modeling natural images using gated MRFs, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 9, pp. 2206–2222 (2013).
[10] Simoncelli, E. P. and Freeman, W. T.: The steerable pyramid: A flexible architecture for multi-scale derivative computation, *Image Processing, 1995. Proceedings., International Conference on*, Vol. 3, IEEE, pp. 444–447 (1995).
[11] Tipping, M. E. and Bishop, C.: Probabilistic Principal Component Analysis, *Journal of the Royal Statistical Society, Series B*, Vol. 21/3, pp. 611–622 (1999).