

深層学習した和紙テクスチャの潜在空間を用いた 繊細な動的 Small Multiple Textures の生成

佐藤 信¹

概要: 本稿では、深層生成モデルにより学習した和紙テクスチャの潜在空間を用いて、繊細な動的 Small Multiple Textures を生成する。提案手法では始めに、テクスチャの類似度の変化のパターンを用意する。次に、深層生成モデルにより学習した和紙テクスチャの潜在空間の探索をおこない、与えられた類似度の変化のパターンにあわせて変化する複数のテクスチャを生成する。潜在空間の探索では、深層生成モデルにより学習した潜在空間をテクスチャの類似度に基づき探索するための既提案手法を用いる。環境音楽にあうような繊細に変化する類似度のパターンを与えることにより、繊細な動的 Small Multiple Textures を生成可能であることを示す。提案手法は、落ち着いた雰囲気のを表現するコンテンツを制作するために適した手法である。

Generating Nuanced Dynamic Small Multiple Textures with Deep Learned Washi Texture Latent Spaces

MAKOTO SATOH¹

Abstract: This paper presents a method for generating nuanced dynamic Small Multiple Textures using latent spaces of Washi textures learned with a deep generative model. In the method, first, the variation patterns of the similarity of the textures are prepared. Next, the latent spaces of the Washi textures learned with the deep generative model are searched, then multiple textures varying with the given similarity variation patterns are generated. For searching the latent spaces, a proposed method for searching latent spaces is utilized, searching latent spaces learned with deep generative models based on the similarity of textures. It is shown that nuanced dynamic Small Multiple Textures can be generated using the nuanced varying similarity patterns matched with ambiente music. The method is suitable for creating artistic contents representing relaxing japanese traditional ambiente.

1. はじめに

本稿では、深層生成モデルにより学習した和紙テクスチャの潜在空間を用いて、繊細な動的 Small Multiple Textures を生成する。特徴は、次のとおりである。

- 深層生成モデルを用いて生成した滑らかに類似度が変化する補間テクスチャ列を、量子化した類似度の変化のパターンに写像する。
- 動的な繊細なテクスチャの調節が可能である。

類似度が滑らかに変化するパターンの生成には、コンピュータ・グラフィックスのための流体シミュレーションを用いる。

新たな手法を提案することにより、コンテンツを制作するための素材の選択の自由度を大きくすることが、本研究の目的である。提案手法は、落ち着いた雰囲気のを表現するコンテンツを制作するために適した手法である。

これ以降の構成について、簡単に説明する。2節では、関連研究との比較をおこなう。そして、3節では、繊細な動的 Small Multiple Textures の生成手法を提案する。4節では実装を用いた実験結果を示し検討をおこなう。そして最後に、5節で本稿のまとめと今後について述べる。

¹ 岩手大学
Iwate University, Ueda, Iwate 020-8551, Japan

2. 関連研究との比較

2.1 聴覚情報と視覚情報との関連付け

聴覚情報と視覚情報とを関連付けることは、長年にわたり研究者の興味をひきつけてきたテーマであり、数多くの研究が存在する [4], [6], [12]. また、聴覚情報のなかでもある種類の音には聴くものをリラックスさせる効果があることは、日常的な経験から広く認められているといえる。例えば、よく眠るため、または、集中力を高めるためなどに音楽が効果的である場合がある。

そのような音の効果について、[1] では心理学的に研究がおこなわれている。一方、環境音楽（アンビエント音楽）に代表される、リラックスするため、または、落ち着いて考えるための音楽が発表され始めたのは、1970年代後半からである。そして1980年代からは、環境音楽と映像とを融合した作品の制作が試みられ続けている。それらの作品で用いられる映像は、実写したビデオ映像を計算機により処理したものが多くといえる。

本稿は、自然の風を間接的にイメージする風鈴の音を、和紙の繊細なテクスチャに関連付けるための手法 [8], [9], [10] の関連研究である。ゆったりと変化する落ち着いた雰囲気音楽にあわせることを目的として、微調節が可能な繊細な動的なテクスチャを生成するための手法を提案する。

2.2 深層生成モデル

機械学習における生成モデルとは、学習データに類似のデータを生成することが可能な学習モデルであり、深層学習が発表される以前から存在した。しかし、それらにより生成されるデータの品質は十分とはいえなかった。

深層学習 (deep learning) [7] が発表されると、それまでの生成モデルに深層学習を適用する手法が発表されたが、それらにより生成される画像はそれまでのものと比較すると格段に高品質であるが、コンテンツの制作に用いるためには十分とはいえなかった。

そして、VAE (Variational Auto-Encoder) [5] および GAN (Generative Adversarial Network) [3] などの、深層生成モデル (deep generative model) が発表され、学習した潜在空間から学習データに類似なデータを高品質に生成することが可能であることが示された。そのことから、これらの学習モデルは多くの研究者の注目を集め、多くの関連モデルおよび応用例が発表され続けている。

本稿では、GAN の関連モデルである DCGAN (Deep Convolutional Generative Adversarial Network) [11] を用いて学習した、和紙の繊細なテクスチャの学習モデルを用いる。学習モデルの生成器から和紙のテクスチャに類似な繊細なテクスチャを生成し、それらを用いて動的なテクスチャを生成する。生成では、テクスチャの視覚的な類似度を基準として用いることにより写像をおこなう。

Algorithm 1 Generating Nuanced Textures

Step 1

Prepare a Washi texture model \mathcal{M} trained with DCGAN.
Generate texture images \mathcal{T} using the generator of \mathcal{M} .
Generate a interpolating texture images \mathcal{I} from \mathcal{T} .

Step 2

Create a pixel pattern \mathcal{P} with a fluid simulation method.

Step 3

Generate quantified pixel pattern indexes \mathcal{Q} from \mathcal{P} .

Step 4

Generate nuanced textures by means of mapping \mathcal{I} to \mathcal{Q} .

3. 繊細な動的テクスチャの生成

3.1 提案手法の概要

アルゴリズム 1 に、提案手法を示す。Step 1 では、深層生成モデルにより和紙のテクスチャを学習した学習モデル \mathcal{M} の生成器を用いて、和紙に類似なテクスチャをもつ画像 \mathcal{T} を生成する。そして、 \mathcal{T} の中から意図にあわせて画像を選択し、[10] での手法により、それらを補間するテクスチャ画像列 \mathcal{I} を生成する。 \mathcal{I} は、テクスチャの類似度が滑らかに変化する画像列である。Step 2 では、流体の視覚的なシミュレーション手法 [13] により、滑らかに変化するピクセル・パターン \mathcal{P} を生成する。Step 3 では、学習によりピクセル・パターン \mathcal{P} を量子化し、学習したカラーパレットのインデックスにより構成されるピクセル・パターン・インデックス \mathcal{Q} を生成する (量子化については、3.2 節を参照)。Step 4 では、ピクセル・パターン・インデックス \mathcal{Q} と補間テクスチャ画像列 \mathcal{I} に含まれるテクスチャ画像のインデックスとを対応付けることにより、動的な Small Multiple Textures を生成する。

3.2 ピクセル・パターンの量子化

3.1 節 Step 3 でのピクセル・パターン \mathcal{P} の量子化では、自己組織化マップ [14] を用いた学習手法 [2] により、画像量子化をおこなう。ここでは以下の手順により、量子化関数 $f(\mathbf{p}_{x,y})$ の学習をおこなう。

始めに、Step 2 での流体シミュレーションにより得られるフレームの中から学習に用いるためのフレームを選択する。そして、選択したフレームのピクセル・パターン \mathcal{P} に含まれる各ピクセルの値 $\mathbf{p}_{x,y}$ を学習データとして、以下の関数を学習する。

$$i = f(\mathbf{p}_{x,y})$$

ここで、関数の出力 i は、 $\mathbf{p}_{x,y}$ に対応するパレットの要素のインデックスである。

次に、学習した量子化関数に、Step 2 での流体シミュ

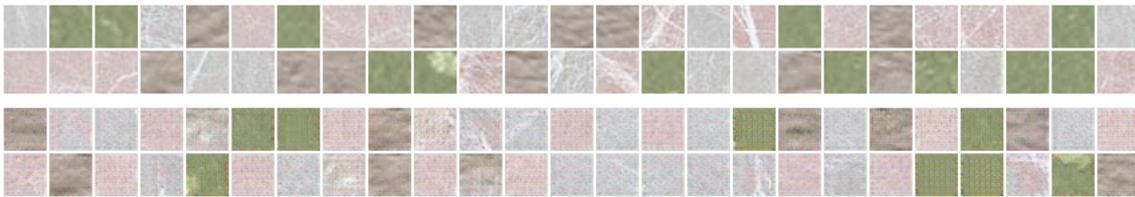


図 1 DCGAN の訓練に用いた和紙のテクスチャ画像（上半分），および，学習した潜在空間から生成したテクスチャ画像（下半分）の例

Fig. 1 Examples of Washi texture images used to train DCGAN (upper half), and texture images generated from the learned latent space (lower half).

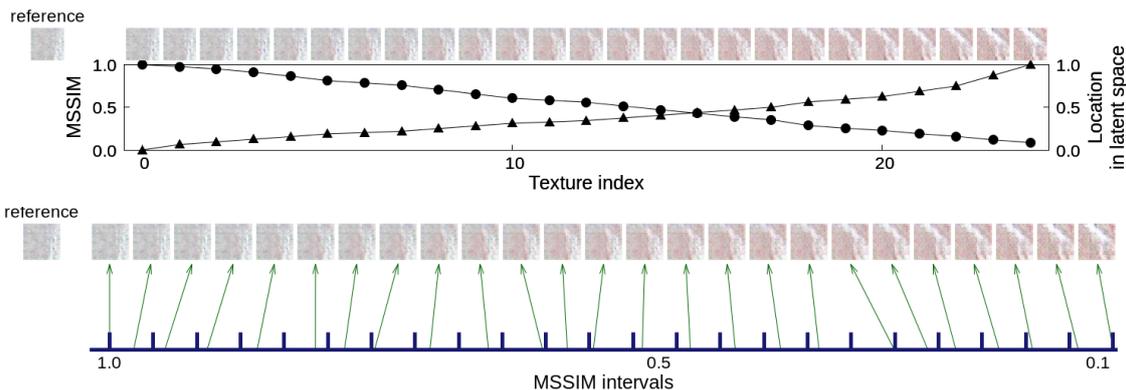


図 2 潜在空間の探索による補間テクスチャの生成

Fig. 2 Generating interpolating textures with the latent space search method described in [10]; The first row shows the MSSIMs of the interpolated and interpolating textures (circles), and the locations of the latent space points from which the textures was generated (triangles). The locations are normalized with respect to the distance between the latent space points from which interpolated textures was generated. The second row shows the similarity searching intervals (MSSIM intervals) used in latent space search (blue scale), and the similarity (MSSIMs) of the generated textures (green arrows).

レーションにより得られる各フレームのピクセル・パターン P を入力する。そして、その出力 i を用いて量子化したピクセル・インデックス・パターン Q を生成する。

4. 実験と結果の検討

4.1 和紙テクスチャの学習

実験では、[8] での和紙テクスチャの学習モデルを用いた。そこでの学習に用いた和紙のテクスチャ画像の例を、図 1(上半分) に示す。なお、深層生成モデルとして DCGAN を用いた。

学習モデルの生成器を用いてテクスチャ画像を生成することにより、学習モデルの確認をおこなった。図 1(下半分) は、乱数により選択した和紙テクスチャの潜在空間上の点から生成したテクスチャ画像である。

4.2 補間テクスチャ画像列の生成

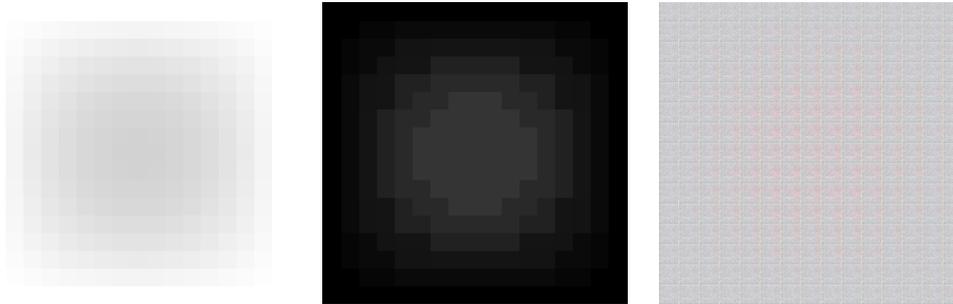
図 2 は、潜在空間の探索手法 [10] により、和紙テクスチャの学習モデルから生成した補間テクスチャ画像列であ

る。テクスチャ列の両端のテクスチャ（質感）を補間するように画像を生成している。両端のテクスチャ画像は、乱数により生成したテクスチャ画像の中から選択した。テクスチャの類似度の基準には MSSIM を用いた。MSSIM 値を計算するための参照画像には、補間テクスチャ画像列の左側の画像を用いた。

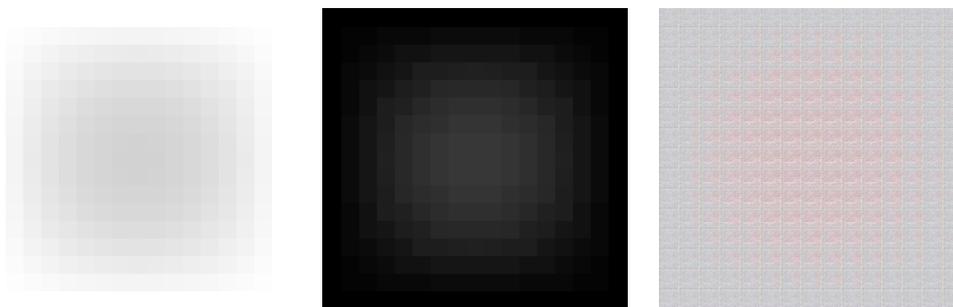
図 2 の第 1 行のグラフの丸印は、参照テクスチャと補間テクスチャ列の各テクスチャとの類似度 (MSSIM 値) である。比較される 2 つのテクスチャが全く同じである場合には、MSSIM 値が 1.0 となる。テクスチャの類似度が小さくなるほど、MSSIM 値は小さくなる。グラフの三角印は、各テクスチャの生成に用いた潜在空間上の点の位置である。ここでは潜在空間上での位置を、潜在空間上での補間の端点を結ぶ線分上での一方の端点からの正規化距離として示している。なお、潜在空間上での補間の端点とは、補間されるテクスチャに対応する潜在空間上の 2 点のことである。潜在空間上での補間の端点のそれぞれの点から、補間テクスチャ列の両端のテクスチャが生成される。



(a) Generated Washi textures.



(b) 10 level quantification.



(c) 20 level quantification.

図 3 繊細なテクスチャの生成

Fig. 3 Generating nuanced Small Multiple Textures: Washi textures generated with a learned deep generative model (a), and nuanced Small Multiple Textures generated with 10 level quantification (b) and with 20 level quantification (c). In (b) and (c), a pixel pattern (left), its quantified pixel pattern (center) and generated Small Multiple Textures (right) are shown.

第 2 行のグラフの青色の目盛りは、探索において用いた類似度探索区間 ([10] を参照) である。[10] の探索手法では、類似度探索区間に設定された類似度の変化の条件を満たすような補間テクスチャ列を生成するための潜在空間上の補間点列を探索する。グラフの緑色の矢印は、補間テクスチャ画像列に含まれる各テクスチャ画像の類似度とその画像を生成した類似度探索区間上での位置とを対応付けている。

4.3 繊細なテクスチャの生成

アルゴリズム 1 に示す提案手法の実装をおこない、動作の確認をおこなった。和紙テクスチャの学習モデルには、4.1 節でのモデルを用いた。

図 3 に、アルゴリズム 1 の各ステップにおいて得られた結果を示す。図 3 (a) は、Step 1 で生成された補間テクスチャ画像列 \mathcal{I} である。生成方法については、4.2 節と同様である。図 3 (b) の左の図は Step 2 での流体シミュレ

ションにより作成したピクセル・パターン \mathcal{P} である。図ではピクセルを拡大している。図 3 (b) の中央の図では、Step 3 でのピクセル・パターン \mathcal{P} の量子化において用いられるカラーパレットの学習状態の確認をおこなっている。確認の手順は、次のとおりである。始めに、ピクセル・パターン \mathcal{P} に含まれる各ピクセルについて、量子化関数を用いて、そのピクセルの量子化のためのカラーパレットのインデックスを求める。そして、求めたカラーパレットのインデックスをピクセル・パターンにあわせて配置することにより、ピクセル・パターン・インデックス \mathcal{Q} を生成する。ここで、ピクセル・パターン・インデックスに格納されているインデックスに対応するカラーパレットの値を用いてピクセル・パターン \mathcal{P} を表示したものが、図 3 (b) の中央の図である。この図は、ピクセル・パターン \mathcal{P} を学習により量子化した結果であるが、提案手法では、量子化したピクセル・パターンを直接的に用いるのではなく、次に述べるように、量子化のために用いるカラーパレット

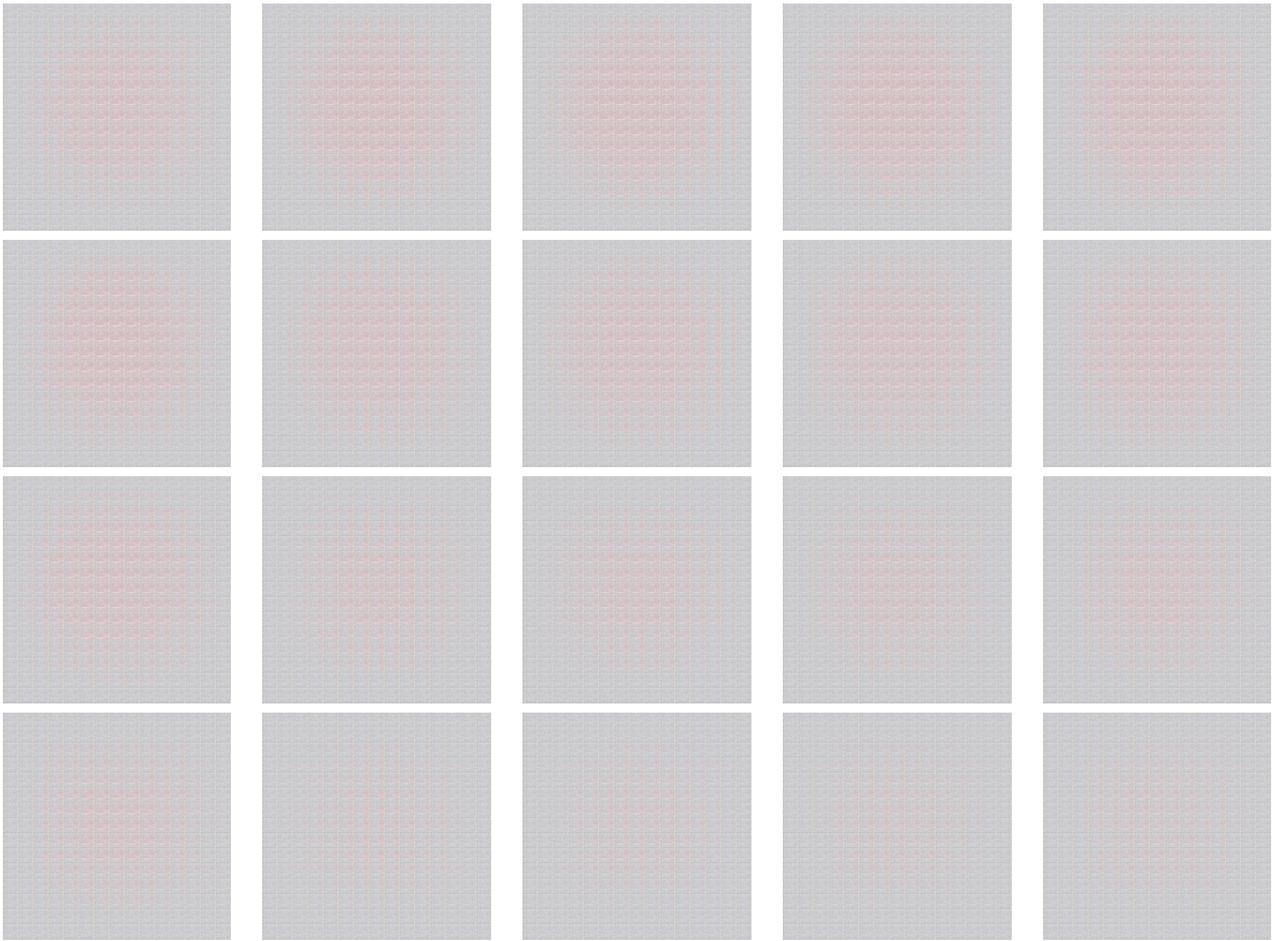


図 4 動的な Small Multiple Textures

Fig. 4 Dynamic Small Multiple Textures: the texture sequence starts at upper left texture, and ends at lower right texture.

のインデックスを格納することにより生成したピクセル・パターン・インデックス Q を用いる。図 3 (b) の中央の図は、量子化関数の学習状態を視覚的に確認するためのものである。図 3 (b) の右の図は、Step 4 において生成される Small Multiple Textures である。Step 1 において生成された補間テクスチャ画像列 I に含まれるテクスチャ画像を、ピクセル・パターン・インデックス Q に対応するように配置することにより生成した。自己組織化マップによるピクセル・パターンの量子化レベルは、図 3 (b) では 10 である。Step 4 での対応付けには、図 3 (a) の左半分のテクスチャ画像を用いた。

図 3 (c) は、量子化レベルが 20 であり、Step 4 での対応付けに図 3 (a) の全テクスチャ画像を用いた。それ以外は、図 3 (b) と同様である。

4.4 動的な Small Multiple Textures

図 4 に、動的な Small Multiple Textures を生成した例を示す。テクスチャ列は、左上のテクスチャから始まり右

下のテクスチャで終わる。流体シミュレーションの連続したフレームからサンプリングしたピクセル・パターンを用いて、アルゴリズム 1 により生成した。Step 3 でのピクセル・パターンの量子化のために用いる自己組織化マップの学習では、テクスチャ列の始めから 2 番目のテクスチャに対応するピクセル・パターンを訓練データとして用いた。量子化レベルは 20 とした。そして、その学習済み量子化関数を用いて、全ピクセル・パターンについてピクセル・パターン・インデックス Q を生成した。

4.5 検討

図 1 からは、深層生成モデルの訓練に用いた和紙のテクスチャ画像と訓練した生成器から生成したテクスチャ画像とを比較することにより、深層生成モデルである DCGAN を用いて訓練した生成器から、和紙の繊細な特徴を捉えたテクスチャを生成可能であることが分かる。

図 2 では、深層生成モデルの潜在空間の探索手法 [10] を用いることにより、類似度が滑らかに変化する和紙の補間

テクスチャ画像列を生成可能であることが分かる。また、類似度探索区間上での探索位置(矢印)、および、潜在空間上での探索位置(三角印)からは、深層生成モデルを用いることにより、滑らかに変化する和紙テクスチャの潜在空間を学習できていることが分かる。

図3からは、和紙の繊細なテクスチャ(質感)を捉えた、Small Multiple Texturesを生成可能であることが分かる。図3(b)の左と中央の図を比較することにより、ピクセル・パターンに含まれるピクセルの中で特徴が類似なもの(値が近いもの)は、近い値を学習したカラーパレットの要素に対応付けられることが分かる。これについては、[2]と同様である。図3(b)の中央と右の図を比較することにより、本稿の提案手法では、前述のような性質をもつように設計したピクセル・パターン・インデックスと滑らかに類似度が変化するテクスチャ列に含まれるテクスチャとを対応付けることにより、繊細なテクスチャ・パターンを生成することを可能にしていることが分かる。そして、図3(b)と(c)とを比較することにより、ピクセル・パターンの量子化レベルを変化させることにより、生成されるテクスチャのパターンの微調節が可能であることが分かる。

図4からは、繊細な変化をする動的な Small Multiple Texturesを生成可能であることが分かる。なお、生成したテクスチャを用いてアニメーションを作成することにより、これらのテクスチャだけを用いて滑らかに変化するアニメーションを制作することが可能であることを確認した。

5. おわりに

深層生成モデルにより学習した和紙のテクスチャの潜在空間を用いて、繊細な動的 Small Multiple Texturesを生成するための手法を提案した。特徴は、深層生成モデルを用いて生成した滑らかに類似度が変化する補間テクスチャ列を、量子化した類似度の変化のパターンに写像する点である。

提案手法を実装することにより、繊細な変化をする動的なテクスチャを生成可能であることを示した。特に、和紙の繊細な特徴を捉えたままで、生成されるテクスチャを微調節することが可能であることを確認できた。提案手法は、落ち着いた雰囲気のを表現するコンテンツを制作するために適した手法であるといえる。

今後の課題としては、多種類のテクスチャへの適用、生成したテクスチャと音によるコンテンツの制作、および、潜在空間の探索のための手法の改良などがある。

参考文献

- [1] Brewer, J. F.: Healing sounds, *Complementary Therapies in Nursing and Midwifery*, Vol. 4, No. 1, pp. 7–12 (online), DOI: [https://doi.org/10.1016/S1353-6117\(98\)80006-1](https://doi.org/10.1016/S1353-6117(98)80006-1) (1998).
- [2] Dekker, A. H.: Kohonen neural networks for optimal

- colour quantization, *Network: Computation in Neural Systems*, Vol. 5, No. 3, pp. 351–367 (online), DOI: 10.1088/0954-898X.5.3.003 (1994).
- [3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative Adversarial Nets, *Advances in Neural Information Processing Systems 27* (Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. and Weinberger, K. Q., eds.), Curran Associates, Inc., pp. 2672–2680 (online), available from <http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf> (2014).
- [4] Kim, J., Ananthanarayan, S. and Yeh, T.: Seen Music: Ambient Music Data Visualization for Children with Hearing Impairments, *Proceedings of the 14th International Conference on Interaction Design and Children, IDC '15*, New York, NY, USA, ACM, pp. 426–429 (online), DOI: 10.1145/2771839.2771870 (2015).
- [5] Kingma, D. P. and Welling, M.: Auto-Encoding Variational Bayes, *ArXiv e-prints* (2013).
- [6] Kosugi, N.: Misual: Music Visualization Based on Acoustic Data, *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services, iiWAS '10*, New York, NY, USA, ACM, pp. 609–616 (online), DOI: 10.1145/1967486.1967581 (2010).
- [7] LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, *Nature*, Vol. 521, No. 7553, pp. 436–444 (2015).
- [8] 佐藤 信: 深層生成モデルにより学習した潜在空間を用いた和紙テクスチャの補間, 情報処理学会研究報告, Vol. 2018-CG-169, No. 2, pp. 1–6 (2018).
- [9] 佐藤 信: 動的な Small Multiple Textures の生成のための和紙テクスチャの潜在空間への風鈴音の写像, 情報処理学会研究報告, Vol. 2018-AVM-102, No. 1, pp. 1–6 (2018).
- [10] 佐藤 信: 風鈴音にあわせた動的な和紙テクスチャの生成のための潜在空間の探索区間のガンマ補正, 情報処理学会研究報告, Vol. 2019-CG-173, No. 8, pp. 1–8 (2019).
- [11] Radford, A., Metz, L. and Chintala, S.: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, *CoRR*, Vol. abs1511.06434 (online), available from <http://arxiv.org/abs/1511.06434> (2015).
- [12] Smith, S. M. and Williams, G. N.: A Visualization of Music, *Proceedings of the 8th Conference on Visualization '97, VIS '97*, Los Alamitos, CA, USA, IEEE Computer Society Press, pp. 499–ff. (online), available from <http://dl.acm.org/citation.cfm?id=266989.267131> (1997).
- [13] Stam, J.: Stable Fluids, *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99*, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co., pp. 121–128 (online), DOI: 10.1145/311535.311548 (1999).
- [14] T. コホネン: 自己組織化マップ, 丸善出版 (2016).