

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

佐藤 信¹

概要: 本稿では、深層生成モデルにより学習した潜在空間を用いた和紙のテクスチャの補間手法を提案する。深層生成モデルである DCGAN を用いて和紙のテクスチャの学習をおこない、繊細なテクスチャの学習が可能であることを示す。そして、学習したモデルの生成器の潜在空間を用いて、学習のためのサンプル画像として用いたテクスチャに類似のテクスチャを生成することにより、テクスチャの補間をおこなう。提案手法は、少数の和紙のテクスチャ画像から、多数の類似のテクスチャパッチを生成することが可能であることから、和の素材を生成するための手法として適しているばかりではなく、多数の類似なテクスチャを必要とするような実時間または対話的な応用に適しているといえる。

Washi texture Interpolation Using Latent Spaces Learned with Deep Generative Model

MAKOTO SATOH¹

Abstract: This paper presents an interpolation method using latent spaces learned with a deep generative model. The practicability of learning delicate textures is validated by learning Washi textures using DCGAN, one of the deep generative models. Then Washi textures are interpolated by generating texture images similar to sample texture images used for training. Using the method many similar texture image patches can be generated from a few Washi texture images. For this reason, the proposed method is suitable for being used in real-time or interactive applications which need many similar texture images, as well as for being used for generating japanese traditional materials.

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のテクスチャの潜在空間を用いたテクスチャの補間手法を提案する。提案手法の特徴を次に示す。

- 和紙のテクスチャの繊細な特徴を学習するために深層生成モデルのひとつである DCGAN(Deep Convolutional Generative Adversarial Network) [5] を用いる。
- 学習した和紙の潜在空間を用いて、学習のためのサンプル画像として用いた和紙のテクスチャに類似のテクスチャを生成することにより、徐々に変化する和紙の補間テクスチャを生成する。生成されるテクスチャに

類似性を基準とした順序付けをおこなうことにより、可能な範囲で滑らかに変化する補間テクスチャを生成する。

提案手法の目的は、和紙の繊細な質感を表現した落ち着いた和の素材を生成することである。そして、少数のテクスチャ画像から多数の類似のテクスチャパッチを生成することが可能であることから、多数の類似なテクスチャを必要とするような実時間または対話的な応用に適しているといえる。また、テクスチャの繊細な特徴の学習を必要とする応用分野は、植物の生育の検査などのように、多数存在することから、そのような応用のための基礎研究としても提案手法は重要であるといえる。

これ以降の構成について、簡単に説明する。2 節では、関連研究について述べる。そして、3 節では、深層生成モ

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

デルにより学習した和紙の潜在空間を用いた類似性に基づく和紙テクスチャの補間手法を提案する。実験結果および検討について4節で述べる。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。

2. 関連研究との比較

2.1 深層生成モデル

深層学習は、コンピュータ科学および工学分野の専門家はもちろん、広範な分野の専門家に関心をもつ学問分野となっている。例えば、[4]のような科学の広い分野の専門家を対象とする専門誌においても、深層学習の紹介記事が掲載されている。

生成モデルは、機械学習のモデルのひとつであり、[2]において深層学習が発表される以前から存在した。そして、深層学習が登場してからは、深層学習を用いた生成モデルである深層生成モデルについての研究が活発におこなわれた。しかし、生成されるデータの品質は、深層学習の登場以前と比較すると格段に向上したものの、さらに品質を向上させるために研究が必要であるという状況であった。このような状況の中で、ほぼ同時期に、VAE(Variational Auto-Encoder) [3] および GAN(Generative Adversarial Network) [1] が発表された。VAEでは変分下限により最適化をおこなうのに対して、GANでは生成器および識別機の損失によりミニマックス解を求めるというように、VAEとGANとは異なる動作原理による手法である。これらの手法を用いると、高品質のデータの生成が可能であることがあるため、多くの研究者の注目を集めることとなった。そして現在でも、多くの関連モデルおよび応用例が発表され続けている。

本稿の提案手法では、GANの関連モデルである、DCGAN(Deep Convolutional Generative Adversarial Network) [5] を用いて、和紙のテクスチャを学習し、それに類似のテクスチャを生成する。これまでのDCGANの適用研究では、コントラストの鮮明な画像を学習した例が多いといえる。提案手法では、和紙の繊細で複雑な繰り返しパターンを学習する。

2.2 テクスチャの類似性

提案手法では、類似性に基づきテクスチャの順序付けをおこなうことにより補間テクスチャを生成する。類似性の基準としてMSSIM(Mean Structural Similarity) [9] を用いる。画像の品質を評価するために多数の基準[7]が提案されているが、SSIM [8] [9] [10] は、画像の輝度(luminance)、コントラスト(contrast) および構造(structure) に基づいた画像の類似性の評価基準である。画像の品質の基準とする参照画像と比較することにより画像品質の評価基準として用いることも可能である。SSIMを使用した例には、映像データの要約を生成するために、SSIMによる類似性の値を用いる手法[6]などがある。

Algorithm 1 Learning Washi Textures

```
Prepare a Washi texture training image set  $W$ .
Prepare a DCGAN model to train the Washi texture images.
Prepare an empty set  $M$  to save trained models.
Set trained model saving epochs to  $E$ .
Set the maximum number of epochs for training to  $N_T$ .
 $n \leftarrow 0$ 
while  $n < N_T$  do
  Train the model for one epoch using the training set  $W$ .
  if  $n$  exists in  $E$  then
    Save current model in trained model set  $M$ .
  end if
   $n \leftarrow n + 1$ 
end while
```

3. 深層生成モデルにより学習した潜在空間を用いた類似性に基づく和紙テクスチャの補間

3.1 深層生成モデルによる和紙テクスチャの学習

深層生成モデルを用いた和紙のテクスチャの学習手順をアルゴリズム1に示す。学習に用いた深層生成モデルは、DCGAN(Deep Convolutional Generative Adversarial Network)である。DCGANの識別機(discriminator)の入力に与えるサンプル画像(real image)として、和紙のテクスチャ画像を用いる。DCGANの訓練が進むにつれて、サンプルとして与えた和紙のテクスチャに類似のテクスチャをもつ画像(fake image)が、DCGANの生成器(generator)の出力として生成されるようになる。このように訓練が進むにつれて、和紙のテクスチャの潜在空間を表現するモデルが学習されることになる。ここでは、3.2節で述べる和紙のテクスチャの潜在空間を用いたテクスチャの補間に用いるために、予め設定したエポックまたは繰り返しにおいて学習モデルを保存する。

3.2 和紙テクスチャの潜在空間を用いた類似性に基づくテクスチャ補間

アルゴリズム2,3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た類似性に基づくテクスチャの補間手法を示す。アルゴリズム2では、乱数などを用いて潜在空間上の2点を選択する。そして、それらの点を結ぶ直線上の点を、アルゴリズム3の関数を用いて探索する。アルゴリズム3では、潜在空間上の探索区間および類似度の探索区間を適応的に変化させながら、潜在空間上の点を再帰的に探索する。そして、類似度についての条件を満たす点を選択する。再帰的な処理が終了すると、アルゴリズム2に戻り、選択した潜在空間上の点からテクスチャを生成する。なお、テクスチャの生成には、アルゴリズム1を用いて学習したDCGANの学習モデルを用いる。

Algorithm 2 Washi Texture Interpolation Launcher

Prepare a Washi texture model D trained with DCGAN.
 Select latent space points which are interpolated: \mathbf{p}_1 and \mathbf{p}_2 .
 Generate texture images from \mathbf{p}_1 and \mathbf{p}_2 using D : T_1 and T_2 respectively, T_1 is used as a reference texture image.
 Compute SSIMs between T_1 and T_1 , and between T_1 and T_2 : C_1 and C_2 respectively.
 Set the number of SSIM intervals on which the textures are interpolated based: N_S .
 Set the SSIM intervals \mathbf{S} and interpolation ratios in latent space \mathbf{R} :

$$\mathbf{S} = \{S_i\} = \{[S_{1_i}, S_{r_i}]\}, \mathbf{R} = \{R_i\}$$
 such that $S_{1_i} \leftarrow C_1, S_{r_{N_S}} \leftarrow C_2, R_i \leftarrow -1, S_{1_i} \geq S_{r_i}, S_{r_j} = S_{1_{j+1}} \quad (1 \leq i \leq N_S, 1 \leq j < N_S)$.
 Set current latent space searching interval (ratios): $\mathbf{L}_c = [L_{1_c}, L_{r_c}] \leftarrow [0, 1]$.
 Set current SSIM searching interval: $\mathbf{S}_c = [S_{1_c}, S_{r_c}] \leftarrow [C_1, C_2]$.
 Note: \mathbf{S} and \mathbf{R} ; global scope and static allocation, \mathbf{L}_c and \mathbf{S}_c ; local scope and automatic allocation.

Set the maximum number of iterations at each recursive level of **function** `WashiTextureInterpolation`: N_I .
 Call **function** `WashiTextureInterpolation` ($D, \mathbf{L}_c, \mathbf{S}_c, T_1, \mathbf{p}_1, \mathbf{p}_2, N_I$)

Generate interpolating points $\mathbf{p}_{I_i} \leftarrow (\mathbf{p}_2 - \mathbf{p}_1) \times R_i + \mathbf{p}_1 \quad (1 \leq i \leq N_S)$.
 Generate interpolating texture images from \mathbf{p}_{I_i} using D .

Algorithm 3 function WashiTextureInterpolation ($D, \mathbf{L}_c, \mathbf{S}_c, T_1, \mathbf{p}_1, \mathbf{p}_2, N_I$)

$n \leftarrow 1$
while $n \leq N_I$ **do**
 if $S_{1_c} \leq S_{r_c}$ **then**
 return
 end if
 Compute current interpolation ratio: $L_c \leftarrow (L_{1_c} + L_{r_c})/2$.
 Generate current latent space point: $\mathbf{p}_c \leftarrow (\mathbf{p}_2 - \mathbf{p}_1) \times L_c + \mathbf{p}_1$.
 Generate current interpolating texture image from \mathbf{p}_c using D , and compute its SSIM C_c using the reference image T_1 .
 if $C_c \geq S_{r_c}$ **and** $C_c \leq S_{1_c}$ **then**
 Find the index of the SSIM interval in which C_c is contained: I_c .
 Set R_{I_c} to L_c .
 if $I_c > 1$ **then**
 Prepare new current latent space searching interval: $\mathbf{L}_n \leftarrow [L_{1_c}, L_c]$.
 Prepare new current SSIM searching interval: $\mathbf{S}_n \leftarrow [S_{1_c}, S_{r_{I_c-1}}]$.
 Call **function** `WashiTextureInterpolation` ($D, \mathbf{L}_n, \mathbf{S}_n, T_1, \mathbf{p}_1, \mathbf{p}_2, N_I$)
 end if
 if $I_c < (\text{size of } \mathbf{S})$ **then**
 Prepare new current latent space searching interval: $\mathbf{L}_n \leftarrow [L_c, L_{r_c}]$.
 Prepare new current SSIM searching interval: $\mathbf{S}_n \leftarrow [S_{1_{I_c+1}}, S_{r_c}]$.
 Call **function** `WashiTextureInterpolation` ($D, \mathbf{L}_n, \mathbf{S}_n, T_1, \mathbf{p}_1, \mathbf{p}_2, N_I$)
 end if
 else if $C_c < S_{r_c}$ **then**
 Modify the current latent space searching interval: $\mathbf{L}_c \leftarrow [L_{1_c}, L_c]$.
 else if $C_c > S_{1_c}$ **then**
 Modify the current latent space searching interval: $\mathbf{L}_c \leftarrow [L_c, L_{r_c}]$.
 end if
 $n \leftarrow n + 1$
 end while

Note: \mathbf{S} and \mathbf{R} ; global scope and static allocation, function arguments, \mathbf{L}_n and \mathbf{S}_n ; local scope and automatic allocation.

探索に用いる SSIM による類似度には、潜在空間上の点から生成されるテクスチャと参照テクスチャとの MSSIM の値を用いる。参照テクスチャには、アルゴリズム 2 において生成した潜在空間上の 2 点のうち的一方から生成したテクスチャを用いる。

このようにテクスチャの類似性を考慮することにより、可能な範囲で、テクスチャの特徴の変化が単調であり特徴

の偏りが少ない、均一な補間テクスチャを生成することが可能となる。

提案アルゴリズムは、潜在空間上の点の座標値の変化に対して、その点から生成されるテクスチャの類似度が単調または滑らかに変化する場合に効率的で効果的な手法である。4 節に示すように、深層生成モデルである DCGAN により和紙のテクスチャを学習した場合には、生成されるテ

クスチャの類似度が比較的滑らかに変化する潜在空間を学習できるが、その類似度の変化は均一ではない。そこで、そのような比較的滑らかであるが偏りのある潜在空間を用いて可能な範囲で類似度が均一に変化する補間テキストを生成する場合に提案の探索手法が有効となる。

4. 実験結果と検討

4.1 深層生成モデルによる和紙テキストの学習

3節において述べた手法を用いて和紙のテキストの学習をおこなない、学習した潜在空間から補間テキストの生成をおこなった。補間テキストの生成では、アルゴリズム 2,3 (3.2 節) を用いて、テキストの類似性を考慮した補間をおこなった。

和紙のテキストの学習には、深層生成モデルのひとつである DCGAN を用いた。4種類の和紙のテキスト画像から、乱数を用いて画像領域を選択することにより、各和紙について 10,000 のテキストパッチを生成した。このテキストパッチを、DCGAN の識別機 (discriminator) の訓練のためのサンプル画像 (real image) として用いた。また、DCGAN の生成器 (generator) の入力には、 $[-1, 1]$ の一様乱数を与えた。一様乱数を用いたのは、学習した生成器に $[-1, 1]$ の値を入力することにより、和紙に類似のテキストを生成するためである。訓練に用いた和紙のテキスト画像の例を、図 1 に示す。

4.2 和紙テキストの生成と補間

4.2.1 潜在空間からの和紙テキストの生成

4.1 節において学習をおこなった DCGAN の生成器を用いて和紙に類似のテキストを生成することにより、学習モデルの確認をおこなった。 $[-1, 1]$ の一様乱数を生成器の各次元の入力として生成したテキストを、図 2 に示す。

4.2.2 潜在空間を用いた和紙テキストの補間

DCGAN により学習した和紙テキストの潜在空間を用いて、和紙テキストの補間をおこなった。始めに、一様乱数により潜在空間上の 2 点を選択し、それらを等間隔に補間する点列を生成した。そして、各補間点を DCGAN の生成器の入力としてテキストを生成した。図 3 に、生成されたテキストを示す。各行の青色の枠で囲まれたテキストが、乱数により選択した 2 点から生成したテキストである。それらの間に並ぶテキストが、潜在空間上の各補間点から生成したテキストである。潜在空間上の点列の順序と図 3 に示すテキストの順序とは対応している。

4.2.3 潜在空間のズームインによる和紙テキストの生成

図 4 は、潜在空間をズームインすることにより生成した和紙のテキストである。図 4 の第 1 行は、図 3 と同様である。図 4 の第 2 行に示すテキストの生成では、始めに、第 1 行の画像を生成するために用いた潜在空間上の点列か

ら 2 点を選択し、その選択した点を補間する点列からテキストを生成した。具体的には、第 1 行では、赤色および茶色の枠で囲むテキストに対応する潜在空間上の 2 点が乱数により選択された。そして、黄色および緑色の枠で囲むテキストに対応する潜在空間上の 2 点を選択した。第 2 行では、第 1 行で選択した黄色および緑色の枠で囲むテキストに対応する潜在空間上の 2 点の補間点を生成し、それらの各補間点から和紙のテキストを生成した。残りの行についても、同様である。

4.3 MSSIM による類似性に基づいた和紙テキストの補間

3.2 節で述べたアルゴリズム 2,3 を用いたテキストの類似性に基づいた和紙テキストの補間を、図 5 に示す。各行の両端のテキストは、乱数を用いて選択した潜在空間上の 2 点から生成したテキストである。この 2 点を端点として補間点を生成した。図 5 の第 1 行は、潜在空間上で等間隔に並ぶ各補間点から生成した補間テキストであり、第 2 行は、アルゴリズム 2,3 により生成した補間テキストである。MSSIM による類似性の計算では、各行の左端のテキストを参照テキストとして用いた。

表 1 に、図 5 に示すテキストの生成で用いた潜在空間上の補間位置および SSIM の値を示す。補間の端点間の距離に対する、一方の補間の端点（各行の左のテキストに対応）と各補間点との距離の割合により、補間位置を示している。

なお、図 5 の各列と表 1 の各行とは対応している。

4.4 検討

DCGAN の訓練に用いた和紙のテキスト画像 (図 1) と訓練した生成器から生成したテキスト (図 2) とを比較することにより、訓練に用いた和紙のテキストに類似のテキストを生成可能であることが分かる。

図 3 からは、和紙の潜在空間上の 2 点を補間する各点から生成されるテキストは、繊細な和紙の特徴を捉えた補間テキストであることが分かる。

図 4 では、和紙の潜在空間を徐々に拡大するにあわせて、繊細な和紙の特徴の微妙な変化を詳細に捉えたテキストを生成できることが分かる。

図 5 および表 1 からは、3.2 節で述べたアルゴリズム 2,3 を用いることにより、潜在空間での等間隔の補間点を用いた場合と比較して、テキストの類似性の偏りが少ない補間テキストを生成することが可能であることが分かる。

これらにより、深層生成モデルである DCGAN を用いると、学習により和紙の繊細なテキストを生成することが可能であるといえる。また、提案アルゴリズムを用いると、テキストの類似性の偏りが少ない補間テキストを生成することが可能であることが分かる。

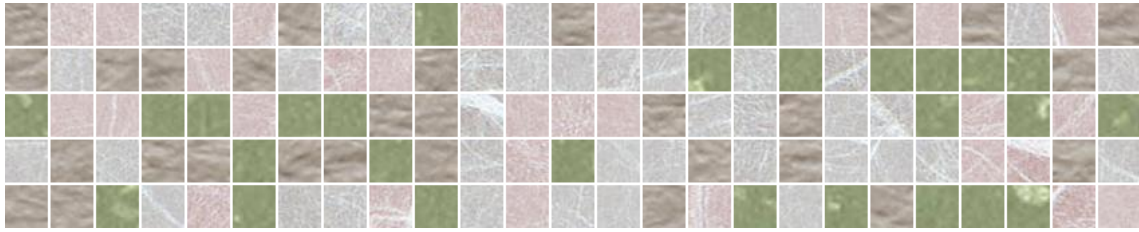


図 1 DCGAN の訓練に用いた和紙テクスチャ画像の例

Fig. 1 Examples of Washi texture images used to train DCGAN.

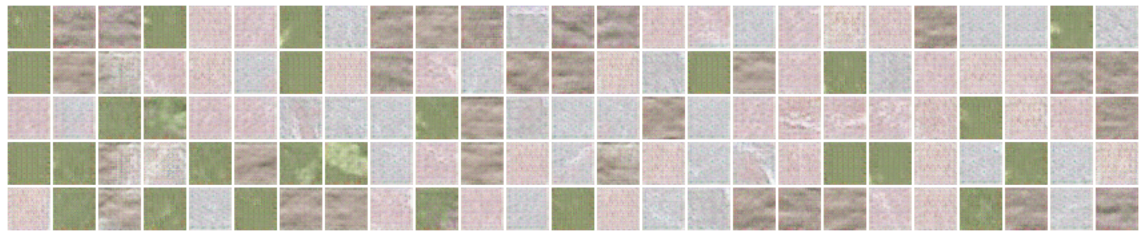


図 2 ランダムに抽出した潜在空間上の点から生成した和紙テクスチャ

Fig. 2 Washi texture images generated from randomly-sampled latent space points.

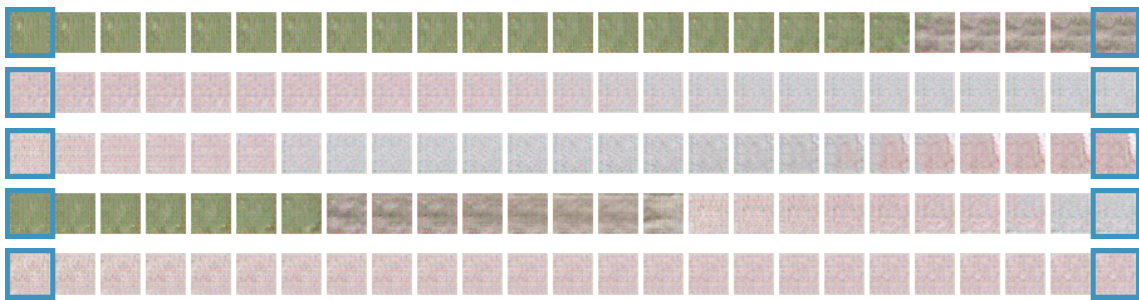


図 3 学習した潜在空間を用いた和紙テクスチャ補間

Fig. 3 Washi texture interpolation using learned latent space.

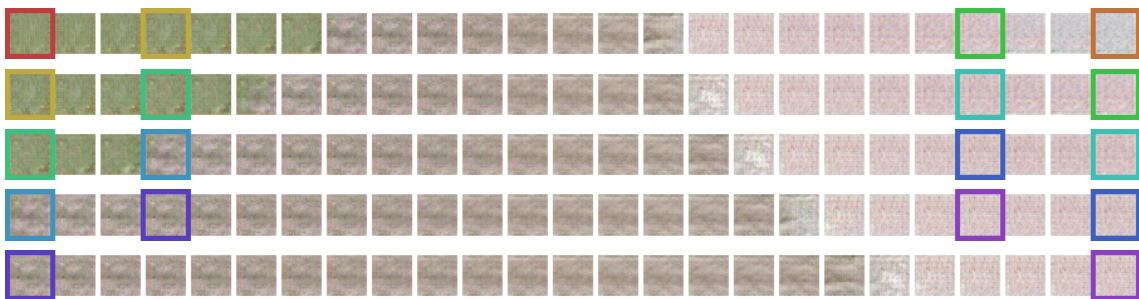


図 4 潜在空間のズームインによる和紙テクスチャの生成

Fig. 4 Generating Washi texture images by zooming in of latent space.

5. おわりに

深層生成モデルにより和紙の繊細なテクスチャを学習することが可能であることを示した。潜在空間から補間テクスチャを生成することにより、テクスチャの特徴の変化を滑らかに表現する潜在空間を学習していることを示した。

そして、テクスチャの特徴に基づいて補間テクスチャを

生成するためのアルゴリズムを提案した。潜在空間上での等間隔の補間点を用いる手法と提案手法とを比較することにより、提案手法を用いることにより特徴の偏りの少ない補間テクスチャを生成することが可能であることを示した。

今後の課題としては、多種類の和紙のテクスチャの学習、アプリケーションへの応用などがある。また、和紙以外の繊細なテクスチャの学習についても重要な課題である。



図 5 形状類似性に基づく和紙テクスチャの補間

Fig. 5 Washi texture interpolation based on shape similarity: textures at both ends in each row (blue edge) were generated from randomly-sampled latent space points. Textures between the textures at the both ends were generated from interpolating points between the randomly-sampled points. Top row: equal-spacing interpolation. Bottom row: interpolation based on shape similarity using MSSIM (algorithm 2, 3 in section 3.2). Each column of figure 5 corresponds to each row of table 1.

表 1 潜在空間から生成したテクスチャの類似度 (MSSIM) に関する補間手法の比較

Table 1 Comparison in terms of the similarities (MSSIM) of texture images generated from latent space between interpolation methods: Interpolation ratio; the ratio of distance between one end point for interpolation and a interpolating point to distance between the end points for interpolation, and shape similarity; MSSIM between a reference texture image and the texture image generated from the interpolating point in latent space, are shown. A texture image generated from the one end point was used as the reference texture image. Equal-spacing interpolation method and interpolation method based on texture image similarity (MSSIM) are compared. Each column of figure 5 corresponds to each row of table 1.

interpolation ratios		similarities (MSSIMs)	
equal spacing	algorithm 2, 3	equal spacing	algorithm 2, 3
0.000	0.000	1.000	1.000
0.053	0.125	0.991	0.951
0.105	0.188	0.964	0.913
0.158	0.250	0.928	0.874
0.211	0.273	0.903	0.824
0.263	0.277	0.856	0.775
0.316	0.279	0.470	0.731
0.368	0.281	0.498	0.675
0.421	0.283	0.521	0.612
0.474	0.500	0.541	0.540
0.526	0.563	0.537	0.519
0.579	0.584	0.501	0.458
0.632	0.586	0.308	0.410
0.684	0.588	0.281	0.354
0.737	0.625	0.259	0.310
0.789	0.750	0.238	0.255
0.842	0.906	0.220	0.178
0.895	0.922	0.197	0.139
0.947	0.938	0.085	0.102
1.000	1.000	0.054	0.054

参考文献

- [1] 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).
- [2] Hinton, G. E. and Salakhutdinov, R. R.: Reducing the Dimensionality of Data with Neural Networks, *Science*, Vol. 313, No. 5786, pp. 504–507 (2006).
- [3] Kingma, D. P. and Welling, M.: Auto-Encoding Variational Bayes, *ArXiv e-prints* (2013).
- [4] LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, *Nature*, Vol. 521, No. 7553, pp. 436–444 (2015).
- [5] 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).
- [6] Sandhu, S. K. and Agarwal, A.: Summarizing Videos by Key Frame Extraction Using SSIM and Other Visual Features, *Proceedings of the Sixth International Conference on Computer and Communication Technology 2015, ICCCT '15*, New York, NY, USA, ACM, pp. 209–213 (online), DOI: 10.1145/2818567.2818607 (2015).
- [7] Thung, K. H. and Raveendran, P.: A survey of image quality measures, *2009 International Conference for Technical Postgraduates (TECHPOS)*, pp. 1–4 (online), DOI: 10.1109/TECHPOS.2009.5412098 (2009).
- [8] Wang, Z. and Bovik, A. C.: A universal image quality index, *IEEE Signal Processing Letters*, Vol. 9, No. 3, pp. 81–84 (online), DOI: 10.1109/97.995823 (2002).
- [9] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P.: Image Quality Assessment: From Error Visibility to Structural Similarity, *Trans. Img. Proc.*, Vol. 13, No. 4, pp. 600–612 (online), DOI: 10.1109/TIP.2003.819861 (2004).
- [10] Wang, Z., C. Bovik, A. and Simoncelli, E.: *Structural Approaches to Image Quality Assessment*, pp. 961–974 (2005).