Predicting and Generating Wallpaper Texture with Semantic Properties

Xiaohan Feng[1], Lin Qi[1], Yanhai Gan[1], Ying Gao[1], Hui Yu[2], Junyu Dong[1]
[1]Department of Information Science and Engineering, Ocean University of China, Qingdao, China
Email: {qilin, dongjunyug}@ouc.edu.cn,{fengxiaohan, gangyanhai, gaoying}@stu.edu.cn
[2]School of Creative Technologies, University of Portsmouth, UK
Email: hui.yu@port.ac.uk

Abstract*—Humans naturally use semantic descriptions to express their visual perception of textures; this is also the fact for perception and description of wallpaper texture. Classification of wallpaper's style is mainly based on understanding of visual information. However, the complexity of real-world wallpaper images is difficult to be captured by existing datasets. Inspired by a publicly available Procedural Textures Dataset, a number of wallpaper images was collected and assembled into a wallpaper dataset. A series of psychophysical experiments was performed to further collect semantic descriptions for this dataset. Each wallpaper was labeled with 5-10 semantic descriptions. More importantly, our dataset contains complex wallpaper images with rich annotations. To our best knowledge, our dataset is the first public wallpaper dataset with semantic descriptions. We use label distribution to analysis semantic descriptions and texture characteristics. Furthermore, a texture generation method based on GAN was tested using our wallpaper dataset, which produced state-of-the-art results.

***Keywords*—Semantic descriptions, Label distribution, Texture generation**

I. Introduction

As wallpaper plays an important role to present visually pleasant effect in daily life, many people prefer to choose a delectable wallpaper to decorate their rooms. However, with the change of human aesthetic and the evolution of various wallpapers styles, consumers might also become confusing when they try to choose a satisfied one from a number of samples. Nevertheless, some people can describe which style they want using semantic descriptions, such as classic style. They naturally wish to be presented with wallpaper as described by them. In practice, this means that they need to have a method that can generate wallpaper texture according to their semantic description. To this end, we are therefore inspired to create a wallpaper dataset with semantic annotations and appeal to recent advances in Generative Adversarial Networks (GAN), so that we can train a deep network with labeled data and general wallpaper texture accordingly. The dataset includes wallpaper with diverse styles. Then we perform a set of psychophysics experiments [1] to obtain label distributions of each wallpaper images. We learn the label distribution [2] from the wallpaper dataset and predict the labels of a new one, and then generate new wallpaper that not exists in the current dataset.

One characteristic of our dataset is that it contains rich semantic annotations. Human beings are accustomed to describe textures with perceptual features, such as coarse, repetitive and blurry. Furthermore, one can easily imagine a texture according to some perceptual descriptions. When humans see a wallpaper image, they can express perceptual information by language, i.e. using semantic descriptions. We collect semantic descriptions based on texture words summarized in [6] that are frequently used in psychophysical experiments. Finally, we choose 94 semantic descriptions as labels, including 38 adjectives, 55 nouns and 1 color value.

Semantic descriptions are regarded as labels, so a multi-label learning approached is employed to learn semantic descriptions of the wallpaper. The work in [2] proposed a novel multi-label learning algorithm called label distribution learning (LDL). It covers a certain number of labels, representing the degree to which each label describes the wallpaper. We use LDL algorithms to predict the distribution of wallpaper images and calculate the distance or similarity between the real distributions and predicted distributions.

Our wallpaper dataset includes nine types of wallpaper images summarized from two online wallpaper sale websites [7][8], for a total of 1,800 wallpaper images. Each image contains 5-10 perceived semantic descriptions as labels.

Generative Adversarial Networks (GAN) [3] is a kind of deep learning model. The model produces a fairly good output through the two-player minimax game with (at least) two modules: a generative model and a discriminative model in the framework. Deep Convolutional Generative Adversarial Networks (DCGAN) [4] is a better improvement of GAN, with a major improvement in the network structure. It greatly improves the stability of GAN training and the quality of the results. We adopt DCGAN and an improved model, which was proposed in [5] and called the perception driven texture generation model, to generate new wallpaper images that do not existing in the wallpaper dataset.

II. Related Work

A. Multi-label learning algorithm

Wallpapers can have multiple semantic descriptions, which correspond to multiple labels. From the instances to the labels, we can learn a mapping. There are mainly two cases of label learning: ① Single-label learning (SLL) means that all the instances in the training set are labeled with only one label. ② Multi-label learning (MLL)[9][10] allows the training instances to be labeled with multiple labels. The differences of SLL, MLL and LDL are shown in Fig. 1. MLL can deal with more cases and is more general than SLL. MLL algorithms usually can be divided into two strategies. The first strategy is problem transformation, which is to transform the MLL task into one or more SLL tasks, e.g. binary classification [11], label ranking [12], and ensemble learning [13]. The second strategy is algorithm adaptation, which is to extend specific SLL algorithms to handle multi-label data, e.g. ML-*k*NN [14], multi-Label decision tree (ML-DT) [15], and neural networks [16].

However, neither SLL nor MLL can deal with the problem: how much does each label describe the instance? A novel label learning algorithm was proposed in [2], called label distribution learning (LDL), which means learning process on the instances labeled by label distributions.

B. Generative adversarial nets

Generative adversarial nets (GANs) were an alternative framework for training generative models and provide an alternative to maximum likelihood techniques. However, GAN is unstable to train, often leading to produce nonsensical outputs from generators. Then [16] proposed deep convolution-al generative adversarial networks (DCGANs), the nets have certain architectural constraints, and make them stable to train in most settings.

However, the patterns of wallpaper are uncontrollable if only wallpaper images are used to generate new wallpaper, the quality and resolution of wallpaper are poor. As a result, we use the joint models for perception driven texture generation [5] , which can generate texture images from perceptual descriptions, i.e. semantic descriptions. The joint deep network model essentially combines adversarial training and perceptual feature regression for texture generation. With input random noise and user-defined perceptual attributes, this model can generate high-quality textures based on human perceptual descriptions.

III. THE WALLPAPER DATASET WITH SEMANTIC LABELS

A. Image Collection

We collected wallpaper images from the website [7][8] and divided them into nine categories according to their characteristics. The wallpapers on the website are classified according to styles and elements. According to their semantic descriptions, we summarized nine categories of wallpaper images, including: vintage, post-modern, floral, geometric, European classical (or classical), fresh, striped, modern, and country-style. We collected more than 1,800 wallpapers with more 200 wallpapers in each category. However, there exists a problem, i.e. one wallpaper image appears across categories. Moreover, we performed psychophysical experiments and ask the subjects to remove the wallpapers which appear repeatedly or they feel difficult to classify. As a result, the number of wallpaper images in each category ranges from 195 to 207, for a total of 1,800. For these nine categories of wallpaper images, six of them are defined by styles and the other three are defined by the contents or elements in the wallpaper. In detail, the wallpaper categories are defined as:

- Vintage --- elegant, looks more sense of the aged, and has old feeling;
- Post-modern --- not rigidly adhere to the traditional way of logical thinking, explore innovative styling techniques, set exaggerated, deformed, and cracked patterns or combine the abstract forms of classical components with new method;
- European classical (or classical) --- with luxurious European patterns, looks elegant and neat with delicate patterns and regular arrangement;
- Fresh --- lovely, bright and light colored, gives a fresh and lovely feeling;
- Modern --- concise, trendy, fashion and personality;

The remaining three styles are described as follows:

- Floral --- only a variety of floral patterns;
- Striped --- banded, wavy stripes, Z-shaped stripes;
- Geometric --- including geometric shapes such as circles, triangles, quadrilaterals, polygons, irregular shapes, etc.;
- Country-style --- natural, including animals and plants.

B. Psychophysical experiment

1) Grouping Experiments

The purpose of grouping experiment is to select misplaced wallpaper images and then move them to the correct categories. A total of ten subjects with normal or correct to normal vision took part in this experiment, including six females and four males, aged from 22 to 26. Moreover, the environment and light condition kept same when grouping experiments were performed. The experiments were processed with a calibrated computer screen. Before the experiment, nine categories of wallpaper were placed under nine folders and renamed 1-9. The subjects were informed in advance that the wallpaper images under each folder belonged to the same category and were provided with the keywords of nine categories. Furthermore, some typical wallpapers with exacted category but not included in the dataset were show to the subjects in advance, so as to give the participants exact concepts and understandings of various types of wallpapers.

The process of grouping experiment is as follows. First, ten subjects were asked to observe wallpaper images in the No. 1 folder based on prior knowledge and their feelings, and to pick

TABLE I. EVALUATION MEASURE FOR THE DISTRIBUTION DISTANCE/SIMILARITY MEASURES.

Measure	**Formula**
Chebyshev ↓	$Dis_1\left(D,\hat{D}\right)=\max_j\left\|d_j-\hat{d}_j\right\|$
Clark ↓	$Dis_2(D,\hat{D})=\sqrt{\sum_{j=1}^{c}\frac{\left(d_j-\hat{d}_j\right)^2}{\left(d_j+\hat{d}_j\right)^2}}$

Canberra ↓	$Dis_3\left(D,\hat{D}\right)=\sum_{j=1}^{c}\frac{\left\|d_j-\hat{d}_j\right\|}{d_j+\hat{d}_j}$
Kullback-Leibler ↓	$Dis_4\left(D,\hat{D}\right)=\sum_{j=1}^{c}d_j\ln\frac{d_j}{\hat{d}_j}$
Cosine ↑	$Sim_1\left(D,\hat{D}\right)=\frac{\sum_{j=1}^{c}d_j\hat{d}_j}{\sqrt{\sum_{j=1}^{c}d_j^{\;2}}\sqrt{\sum_{j=1}^{c}\hat{d}_j^{\;2}}}$
Intersection ↑	$Sim_2\left(D,\hat{D}\right)=\sum_{j=1}^{c}\min\left(d_j,\hat{d}_j\right)$

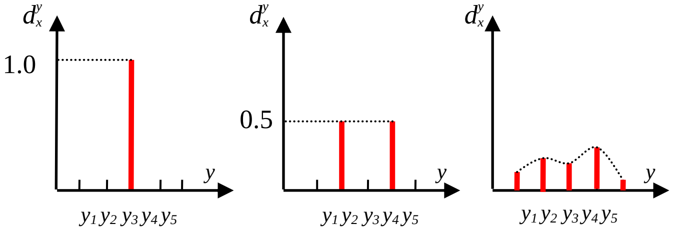

Fig. 1. Label distribution for single-label learning, multi-label learning, and the general case from left to right.

out wallpaper images that looked different from others in the same folder. Then, subjects are supposed to describe and summarize the remaining wallpaper images. Subjects can use the style words provided previously. If they think it is unreasonable, they can define new description words. Subjects performed the same operation on nine folders successively. During the experiments, if the subject thinks that one of the wallpapers does not belong to its category, he or she can create a new folder and move it to the created folder. Each subject creates a folder of its own. Considering that the long-term experiment often brings people visual fatigue, after observing a folder, subjects were forced to rest.

After observing and selecting of all the folders, the images that were picked out by the subject in his/her folder will be reclassified. If the subjects think one image cannot be reclassified, they can leave it in their folders. Thus, the wallpaper images of each folder are collected based on the experimental results of ten subjects. For the images which were excluded by more than 4 persons, we think that they do not belong to any category and then delete them. For each category of wallpaper images, more than 80% of the subjects label their categories correctly, which indicates that the category/style names we gave earlier are universally applicable. As a result, we create a wallpaper dataset with 1,800 wallpapers, including nine categories / styles, and each category containing approximately 200 images.

2) *Perceptual rating experiment*

Fig. 2. The examples of the test set from 9 styles/categories. Pictures from left to right and from top to bottom are vintage, post-modern, floral, geometric, classic, fresh, striped, modern, country-style, respectively.

Semantic descriptions are very important to wallpapers. It can describe various wallpaper images with different styles. We want to label the wallpaper dataset with semantic descriptions. We collect descriptive words firstly and then score each descriptive word through the perceptual rating experiment.

a) *Groups of wallpaper images*

The nine categories of wallpaper images are divided into three large groups, each containing three categories of wallpaper. The first group includes vintage, post-modern and floral; the second group includes geometric, European classical and fresh style; the third group includes the remaining three categories: striped, modern and country-style. Each large group contains 13 subgroups, with 45 to 47 images (we choose 15 or 16 images from the categories that included putting in this subgroup). Each time the subject performed an experiment on one subgroup of the large group. As we have three large groups, each subject needed to perform three experiments. Each subgroup was experimented by three people, which made a total of 13 subgroups, requiring 39 subjects.

b) *Experimental process*

The subjects are supposed to view wallpaper pictures one by one, and then use appropriate words to describe images they see (both nouns and adjectives are acceptable; each wallpaper needs at least three words to describe). They score each word according to their own feelings ranges from1 to 100. It means that if you feel this word can strongly describe the wallpaper they observe, the score should be high; otherwise, if one thinks this word is not appropriate with only little connection to the wallpaper, the score should be low. We stipulate that at least 3 words should be used for description, and subjects usually give 4 to 6 descriptive words and scores. The description of each

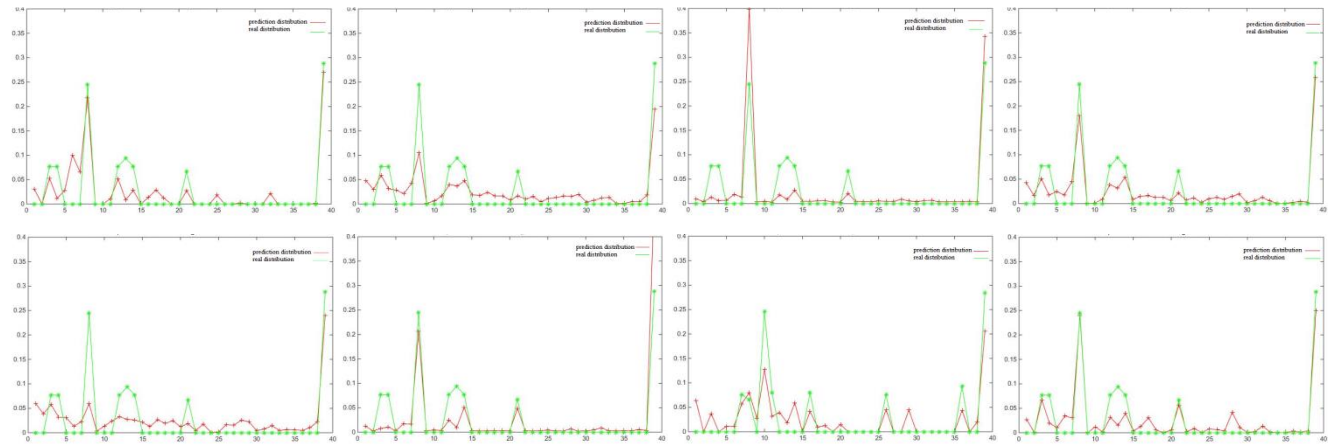

Fig. 3. The label distribution of one wallpaper (the modern wallpaper in Fig. 2). The x-axis represents the semantic labels, the y-axis indicates the value of each label. Combine each point to form the label distribution of a wallpaper. The red line is prediction distribution, and the green line is real distribution. From top to bottom, the feature of wallpaper images is extracted using Gabor filter and AlexNet, called Gabor feature and deep feature, respectively. From left to right in the top line, the LDL algorithms are AA-*k*NN, SA-IIS, SVR, and SA-BFGS, respectively. In the second line, from left to right, the LDL algorithms are AA-BP, SVR, AA-*k*NN and SA-IIS respectively.

image should be consistent with their own feelings. The requirement is that, giving the descriptive words, the image formed in the brain can be consistent with the image shown in the screen. Some descriptive words used frequently are provided just for reference only, and the subjects can use their descriptive words to describe as long as they feel appropriate.

c) Processing experimental data

We count the descriptive words and find the ones that are used frequently. Meanwhile, we omit the words that are used less or unsuited for wallpaper labels. Finally, we choose 93 words as wallpaper labels, including 38 adjectives and 55 nouns. Because color is important to the style of wallpaper, we add a color value as wallpaper label. There are 94 descriptive words in all.

Adjectives include: repetitive, floral, regular, fresh, vintage, wooden, blurry, striped, zig-zag, wavy, classical, undertint, simple, modern, lovely, worn-out, elegant, serried, exquisite, patterned, symmetrical, country-style, bright-colored, complex, bright, post-modern, colorful, messy, somber, granular, cracked, coarse, spotted, marble, stone, bent, netted, geometric.

Nouns include: triangle, leafage, rhombus, square, animal, circle, plant, bird, butterfly, dragonfly, tree, branch, letter, star, brick, plume, figure, polygon, book, bookshelf, grape, pinecone, pineapple, cherry, fish and grass, photo frame, automobile, bicycle balloon, building, cloud, airplane, mountain, arrow, mushroom, musical note, boat, horologe, dandelion, crown, button, robot, sky, vine, bowknot, seawater, water-drop, soil, skirt, snow, cotton, tyre, shoes, glasses, windmill.

The wallpapers in our dataset only have 5 to 10 descriptive words. That is because the three subjects may use the identical words to describe the same images and some words are abandoned while counting the words. We set descriptive words as wallpaper labels and calculate the value of each wallpaper label according to the scores subjects gave during experiment. For the adjective label, we add the scores of the same label, divided by 3, then normalized the value to [0, 1]. The value of adjective labels which is not used by subjects is set to 0. For noun labels, we take another strategy, the value of noun labels is 0 or 1. Once a subject uses this noun label and the noun descriptor does exist in the wallpaper for sure, the value of the noun label is 1, and 0 otherwise.

The method for calculating color values mainly from the color tone of wallpaper. In one RGB images, we treat R as warm tone, B as cool tone. We determine the wallpaper's color tone according to RGB value of each pixel in wallpaper. Compare the values of R channel and B channel, if the R-value of a pixel is larger than the B-value, the dominant color tone of this pixel is considered to be a warm color tone. If not, the dominant color tone of this pixel is considered to be cool color tone. All pixels in the wallpaper image are judged by this method. Then, count the number of pixels of R > B, and divided by the size of the entire image, leading to gain a value range [0, 1]. When the value is close to 0, it represents that this wallpaper is the colder color tone; when the color value is close to 1, it represents that this wallpaper is the warmer color tone. Because the color value is in [0, 1], this label is included in the adjective labels.

Furthermore, we visited some artists before the experiments, and we learned some questions about the wallpaper design process. The wallpaper design steps of the artists are as follows:

- Determine the theme of the wallpaper, i.e. the style of the wallpaper.
- Determine the main color tone of the wallpaper. Is it cool color tone, or warm color tone? Color has a great impact on wallpaper's styles.
- Add the content to the wallpaper: such as shading, circle, water ripples, or some other patterns.

IV. WALLPAPER LABELS PREDICTION

A novel label learning algorithm, called label distribution learning (LDL) [2], which means learning on the instances labeled by label distributions. We use x to represent instance variable, the i-th instance is denoted by x_i. And we use y to represent the label, the j-th label value is denoted by y_j. The description degree of y to x is represented by d_x^y and the label distribution of x_i is represented by $D = \{d_{x_i}^{y_1}, d_{x_i}^{y_2}, \dots, d_{x_i}^{y_c}\}$, where c is the number of possible label values. Labelling an instance x is to assign a real number d_x^y to each possible label y, representing the degree to which y describes x. Moreover, d_x^y should meet two conditions: $d_x^y \in [0, 1]$, and $\sum_y d_x^y = 1$. [2] proposed six LDL algorithms in three ways: problem transformation, algorithm adaptation, and specialized algorithm design. In order to compare these algorithms, we use six evaluation measures to compare all algorithms, TABLE I lists the formulae of the six measures. The ↑ means "the larger the better", and the ↓ means "the smaller the better".

TABLE II. A VERAGE DISTANCE&SIMILARITY MEASURES BETWEEN REAL AND PREDICTED DISTRIBUTION USING GABOR FEATURE

	Chebyshev	Clark	Canberra	Kldist	Cosine	Intersection
AA-*k*NN	0.1638	\	\	0.6688	0.8260	0.5696
SA-IIS	0.2146	5.7681	34.7082	1.1278	0.8233	0.4544
SVR	0.2207	5.7162	34.7239	0.8062	0.7570	0.5563
SA-BFGS	0.1874	5.7570	34.5336	0.9770	0.6994	0.4741

Label distribution has a data form similar to probability distribution and shares the same conditions. We can use the form of conditional probability to represent d_x^y, i.e. $d_x^y = P(y \mid x)$.Suppose $P(y \mid x)$ is a parametric model $P(y \mid x\,; \theta\,)$, where θ is the parameter vector. Given the training set S, LDL's target is to find the θ that can generate a distribution similar to the distribution given the instance X_i. As an example, the KL divergence is used as the distance measure method, then then the best parameter vector θ^* is determined by (1).

For SLL, $d_{x_i}^{y_j} = Kr\left(y_j, y(x_i)\right)$, where Kr(·,·) is the Kronecker delta function and $y(x_i)$ is the single label of x_i. Eq (1) can be changed to (2).

$$\begin{aligned}\theta^* &= \arg\min_\theta \sum_i \sum_j \left(d_{x_i}^{y_j} \ln \frac{d_{x_i}^{y_j}}{p(y_j \mid x_i;\theta)}\right) \\ &= \arg\max_\theta \sum_i \sum_j \left(d_{x_i}^{y_j} \ln p(y_j \mid x_i;\theta)\right).\end{aligned} \tag{1}$$

$$\theta^* = \arg\max_\theta \sum \ln p(y(x_i) \mid x_i;\theta). \tag{2}$$

For MLL, each instance is labeled with a label set, and then Eq (1) is changed to (3).

$$\theta^* = \arg\max_\theta \sum_i \frac{1}{|Y_i|} \sum_{y\in Y_i} \ln p(y \mid x_i;\theta). \tag{3}$$

We divided the 1800 wallpaper images into a training set (1500 images) and a test set (300 images). Each style / category was chosen randomly. In order to satisfy the condition of LDL, we only use adjective labels. Fig. 2 shows some examples of the test set. A part of results of LDL are shown in Fig.3. TABLE II and 0 list the average distance and average similarity between real and predicted label distribution of all testing samples using Gabor feature and deep feature, respectively. The special algorithms that proposed in [2] perform better than other modified algorithms.

TABLE III. A VERAGE DISTANCE&SIMILARITY MEASURES BETWEEN REAL AND PREDICTED DISTRIBUTION USING DEEP FEATURE

	Chebyshev	Clark	Canberra	Kldist	Cosine	Intersection
AA-BP	0.2236	5.7771	34.8115	1.2983	0.5970	0.4230
SVR	0.1540	5.7047	33.7765	0.4002	0.8905	0.7131
AA-*k*NN	0.1402	\	\	0.5106	0.8695	0.6356
SA-IIS	0.1232	5.6915	33.7745	0.4520	0.8964	0.6610

V. WALLPAPER GENERATION

Texture synthesis and generation have been studied for many years, and still are hot topics. Many effective methods were proposed for texture synthesis. The pixel-based approaches [17] usually synthesize a texture in scan-line order, the synthetic texture based pixels is similar with around local neighborhood pixels. Patch-based texture synthesis [18] usually synthesize a texture based patches and more effective and faster than pixel-based texture synthesis approaches. The frequently-used algorithms are image quilting [21] and graphcut textures [22].

With the development of deep learning, more novel approaches of texture generation become popular. Leon [23] proposed a new parametric texture model based on a high-performing convolutional neural network and the samples from the model are of high perceptual quality, leading to a novel texture synthesis method. The generative model is another popular method. Generative adversarial network [3] (GANs) is a recent approach to train generative models of data, which have been shown to work particularly well on image. The Spatial GAN [19] method showed the advantages of fully unsupervised GANs for texture synthesis method based on GANs. PSGAN [20] extend the structure of the input noise distribution by constructing tensors with different types of dimensions based on GANs to texture synthesis.

A. DCGAN

It is known that one of the best model for image processing is Convolutional Neural Networks (CNN) in deep learning at present. How to combine CNN and GAN, Deep Convolutional Generative Adversarial Networks (DCGAN) [4] is the better attempt in this respect. We use DCGAN to train on our

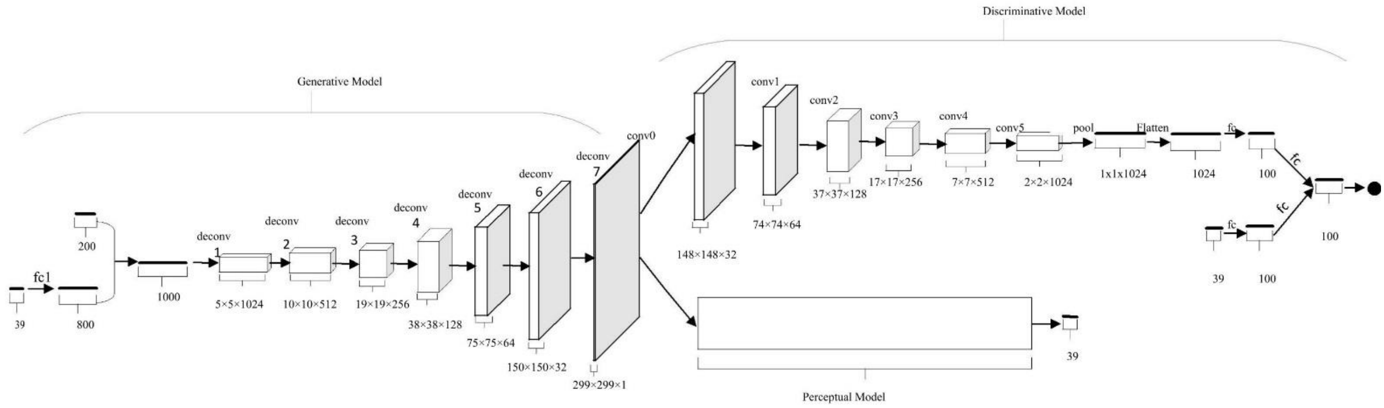

Fig. 4. The architecture of joint model for perception driven texture generation

Fig. 5. The samples generated from DCGAN without semantic labels. The images are blur with low resolution and have no evident change.

wallpaper dataset and generate new wallpaper images. Like the GAN [3], which is trained by setting a game between two models: generative model G and discriminative model D. G was trained to generate the samples which can deceive D, and the samples are intended to come from the same probability distribution as the training data (i.e. p_{data}), without having access to such data. D was trained to distinguish the samples from G rather than $p_{data.}$ D and G play the two-player min-max game with the following objective function:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))]. \tag{4}$$

where z is a noise vector from p_z, and x is a real image from the data distribution $p_{data.}$ DCGAN make some changes on the architecture of CNN. Therefore, it was more stable than GAN, enhanced the quality of samples and accelerated the speed of convergence. We only use the wallpaper examples to train the network. The generated wallpapers are show in Fig. 5. Without the semantic labels, the results are not good enough. The generated images are blur with low resolution and have no evident change.

B. Perception Driven Texture Generation

Recently, a joint deep network model was proposed in [5] to solve aforementioned questions. This network not only uses adversarial training, but also uses perceptual feature regression. Nevertheless, the input of the network requires random noise and perceptual attributions, i.e. semantic labels. In this model,

Fig. 6. The wallpaper samples generated based on the Perception Driven Texture Generation model with our 39 semantic labels (adjective labels, e.g. regular, repetitive, striped, serried, spotted etc.). The images are clearer with higher resolution than before. Each wallpaper possesses some certain perceptual attributes and have various evident changes.

the perceptual constraints were added into the generative model by modifying Inception-v3 model as perceptual model. It can give more extra information to generator and produce various textures. Fig. 4 shows the architecture of joint model for perception driven texture generation. The loss of discriminative model (D) is defined by:

$$D_{loss} = -\frac{1}{n}\sum_{i=1}^{n}(q_i \ln(D(x_i, y_i)) + (1-q_i)\ln(1-D(x_i, y_i))) \quad (5)$$

x_i is training example, y_i is the corresponding perceptual feature vector, q_i is 1 or 0, representing (x_i, y_i) is a real pair or not, n is the number of training examples. The quadratic loss of perceptual model (H) is defined as:

$$H_{loss} = \frac{1}{2n}\sum_{i=1}^{n}(H(x_i) - y_i)^2. \quad (6)$$

The loss of generative model includes the loss from discriminative model and the loss from perceptual model, which is defined as:

$$G_{loss} = G_{loss\_d} + \alpha * G_{loss\_h}. \quad (7)$$

$$G_{loss\_d} = -\frac{1}{n}\sum_{i=1}^{n}\ln(D(G(y_i, z_i), y_i)). \quad (8)$$

$$G_{loss\_h} = \frac{1}{2n}\sum_{i=1}^{n}(H(G(y_i, z_i)) - y_i)^2. \quad (9)$$

where α is a tradeoff parameter, z_i is a random noise vector. Perceptual model is preliminarily trained, and generative model and discriminative model are trained in an adversarial method. Then discriminative model makes the generator produce realistic textures, and the perceptual model makes the generated textures possess certain perceptual attributes accordingly.

We use our wallpaper dataset to generate new wallpaper images with the adjective labels. The noun labels are not adopted because its sparse values are not suitable for perceptual model training. Therefore, many generated images are background of wallpaper or containing little content in the wallpaper. The generated results are better than the results of DCGAN. Fig. 6 shows the nine generated wallpapers based on the Perception Driven Texture Generation model with 39 semantic labels (e.g. regular, repetitive, striped, serried and spotted) and possess some perceptual attributes. For instance, the wallpaper in the left top corner of Fig. 6 looks regular with strips. Furthermore, these images have higher-resolution with diverse types.

VI. CONCLUSION

In this work, we introduce a dataset of wallpaper images with semantic descriptions, and we further propose a method that can generate new wallpaper texture according to user-defined semantic depictions. Our dataset contains 1800 wallpaper images and each image has corresponding semantic labels obtained from psychophysical experiments. We also use label distribution learning (LDL) to predict the multi-labels of given new wallpaper images. Based on this dataset, we are able to train a deep network based on GAN to generate various wallpaper images according to input semantic descriptions. However, it should be noted that there are still some open questions, such as how to further increase the resolution of generated images and sharpen the edges of generated images. In future work, we will focus on processing training images with higher resolution and correspondingly generate wallpaper images with better resolution. Furthermore, since we only use adjective words as semantic labels in this study. in future work we will also try to add noun words to generate more diversiform wallpaper textures with more details and elements.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (NSFC) (No.61501417) and the Ph. D. Program Foundation of Ministry of Education of China (No.20120132110018).

REFERENCES

[1] Liu, J., Dong, J., Cai, X., Lin, Q., & Chantler, M."Visual Perception of Procedural Textures: Identifying Perceptual Dimensions and Predicting Generation Models." Plos One10.6(2015): e0130335.

[2] Geng, Xin. "Label Distribution Learning." IEEE Transactions onKnowledge & Data Engineering 28.7(2014):1734-1748.

[3] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., & Ozair, S., et al. "Generative adversarial nets." International Conference on Neural Information Processing Systems MIT Press, 2014:2672-2680.

[4] Radford, Alec, L. Metz, and S. Chintala. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks."Computer Science (2015).

[5] Yanhai Gan, Huifang Chi, Ying Gao, Jun Liu, Guoqiang Zhong, Junyu Dong. “Perception Driven Texture Generation.” ICME, 2017.

[6] Bhushan, Nalini, A. R. Rao, and G. L. Lohse. "The texture lexicon: Understanding the categorization of visual texture terms and their relationship to texture images." Cognitive Science 21.2(1997):219-246.

[7] https://www.wallpaperfromthe70s.com/

[8] https://www.houzz.com/photos/wallpaper

[9] Tsoumakas, Grigorios, I. Katakis, and D. Taniar. "Multi-Label Classification: An Overview." International Journal of Data Warehousing & Mining 3.3(2008):1-13.

[10] Zhang, Min Ling, and Z. H. Zhou. "A Review on Multi-Label Learning Algorithms." IEEE Transactions on Knowledge & Data Engineering26.8(2014):1819-1837.

[11] Read, J., Pfahringer, B., Holmes, G., & Frank, E. "Classifier chains for multi-label classification. "Machine Learning 85.3(2011):333-359.

[12] Hüllermeier, E., Fürnkranz, J., Cheng, W., & Brinker, K. "Label ranking by learning pairwise preferences." Artificial Intelligence 172.16(2008):1897-1916.

[13] Li, Ping, H. Li, and M. Wu. "Multi-label ensemble based on variable pairwise constraint projection." Information Sciences 222.3(2013):269-281.

[14] Zhang, Min Ling, and Z. H. Zhou. "Multi-label learning by instance differentiation." National Conference on Artificial Intelligence AAAI Press, 2007:669-674.

[15] Read, J., Bifet, A., Holmes, G., & Pfahringer, B."Scalable and efficient multi-label classification for evolving data streams." Machine Learning 88.1-2(2012):243-272.

[16] Zhang, Min Ling, and Z. H. Zhou. "Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization." IEEE Transactions on Knowledge & Data Engineering 18.10(2006):1338-1351.

[17] Paget, R, and I. D. Longstaff. "Texture synthesis via a noncausal nonparametric multiscale Markov random field." IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society7.6(1998):925.

[18] Efros, Alexei A, and T. K. Leung. "Texture synthesis by non-parametric sampling." The Proceedings of the Seventh IEEE International Conference on Computer Vision IEEE, 2002:1033.

[19] Jetchev, Nikolay, U. Bergmann, and R. Vollgraf. "Texture Synthesis with Spatial Generative Adversarial Networks." (2017).

[20] Bergmann, Urs, N. Jetchev, and R. Vollgraf. "Learning Texture Manifolds with the Periodic Spatial GAN." (2017).

[21] Efros, Alexei A, and W. T. Freeman. "Image quilting for texture synthesis and transfer." Proc. SIGGRAPH 2001 (2001):341-346.

[22] Kwatra, V., Essa, I., Turk, G., & Bobick, A."Graphcut textures: image and video synthesis using graph cuts." ACM SIGGRAPH ACM, 2003:277-286.

[23] Gatys, Leon A, A. S. Ecker, and M. Bethge. "Texture Synthesis Using Convolutional Neural Networks." Febs Letters 70.1(2015):51-55.