AI Painting: An Aesthetic Painting Generation System

Cunjun Zhang†
Tsinghua University
cunjunzhang@tsinghua.edu.cn

Kehua Lei†
Tsinghua University
kehualei@tsinghua.edu.cn

Jia Jia*
Tsinghua University
jjia@tsinghua.edu.cn

Yihui Ma
Tsinghua University
mayihui12@foxmail.com

Zhiyuan Hu
Tsinghua University
zy-hu16@mails.tsinghua.edu.cn

ABSTRACT

There are many great works done in image generation. However, it is still an open problem how to generate a painting, which is meeting the aesthetic rules in specific style. Therefore, in this paper, we propose a demonstration to generate a specific painting based on users' input. In the system called AI Painting, we generate an original image from content text, transfer the image into a specific aesthetic effect, simulate the image into specific artistic genre, and illustrate the painting process.

CCS CONCEPTS

• **Human-centered computing** → *Graphical user interfaces*;

KEYWORDS

Painting Content Generation, Aesthetic Effect Modification, Artistic Effect Simulation, Painting Process Illustration

ACM Reference Format:
Cunjun Zhang[2], Kehua Lei[2], Jia Jia, Yihui Ma, and Zhiyuan Hu. 2018. AI Painting: An Aesthetic Painting Generation System. In *2018 ACM Multimedia Conference (MM '18), October 22–26, 2018, Seoul, Republic of Korea.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3240508.3241386

1 INTRODUCTION

In this work, we are interested in painting generation with specific artistic genre, based on content text given by users. We build up an input-output system named AI painting. There are three parts of our system input, including content (object or scene described by natural language), aesthetic effect word (such as joyful, depressed) and artistic genre (e.g. Impressionism, Suprematism, Chinese ink-wash painting). The output is defined as an automatically generated artwork in a dynamic painting process.

Great progress had been made in the field of image generation before. Traditional solution is capturing the detailed visual information in attribute representation[2] and encoding in a vector[11]. Recently, Deep Recurrent Attentive Writer(DRAW) has been used in realistic image generation[4]. When it comes to aesthetic impression, researchers have tried to build a image space bridging color features and fashion words[9]. For style transfer, most traditional textual transfer researches are non-parametric algorithms[1]. It is a remarkable breakthrough that convolutional neural networks are used to transfer a image in style of another image[3].

In this paper,we are focused on 3 key challenges:

- propose a novel framework to generate images as real paintings with illustration of drawing process
- make the painting more natural to aesthetic impression
- illustrate drawing process approaching real process

*Corresponding author.
†Joint first author: Cunjun Zhang, Kehua Lei (According to the alphabetical order of first name)

MM '18, October 22–26, 2018, Seoul, Republic of Korea

ACM ISBN 978-1-4503-5665-7/18/10.
https://doi.org/10.1145/3240508.3241386

2 DEMONSTRATION

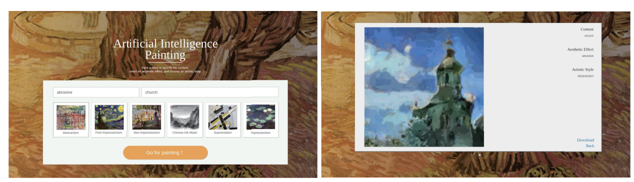

Figure 1: User interface of AI Painting

There are three parts in our input user interface: 1) a select box for users to choose aesthetic effect word; 2) an input box for users to describe painting content by text; 3) six cards representing six different artistic genre for users to select. After several seconds for system running, the generated painting will be provided to users with a short video of the painting process illustration. Furthermore, users can download the final painting for other application.

3 SYSTEM FRAMEWORK

The proposed system generates a piece of painting automatically, according to the content, aesthetic effect word and artistic genre given by users, which is depicted in Figure 2. To train the proposed system, we collected paintings with six different artistic genre to construct the dataset. The main workflow consists of 4 steps: 1) generating an image based on the content by a Stacked Generative Adversarial Networks (StackGAN++) [13] module; 2) modifying the image into the specific aesthetic effect based on Image Aesthetic Space by a Bimodal Deep Autoencoder with Cross Edges(BDA-CE)[8] module; 3) transferring the image into the specific artistic genre by neural style transfer and brush stroke enhancement; 4) illustrating the painting process dynamically by a short video.

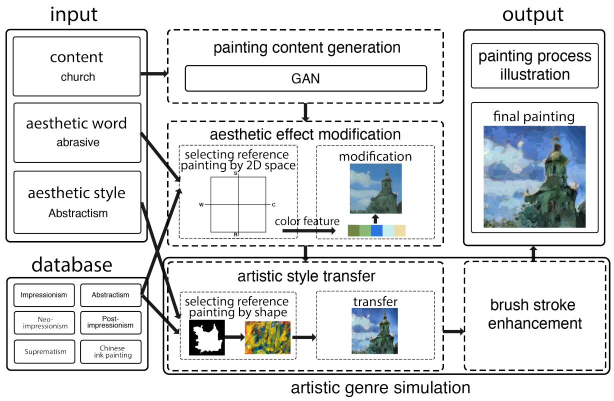

Figure 2: System Framework of AI Painting

3.1 Datasets

To better learn paintings' traits in different artistic genres, we construct a database made up of about 4000 paintings belonged to six different artistic genres (Impressionism, Neo-impressionism, Post-impressionism, Abstractism, Suprematism and Chinese-ink-wash). From this dataset, we extract a set of features to describe their color pattern including five-color combination, saturation and its contrast, brightness and its contrast, warm or cool color and clear or dull color[12]. Besides, we extract figures from each painting by salient region detection technique.

3.2 Painting content generation

To generate the content given by users, we apply StackGAN++ modules in our system. In this paper, we choose six kinds of items as examples, including flower, church, cat, dog, Eiffel Tower, and Mount Fuji. For each item, a StackGAN++ model was trained with a corresponding dataset, which is open source or crawled from Flickr.

3.3 Aesthetic effect modification

To transfer our painting into a specific aesthetic effect, we manage to bridge aesthetic effect words and color pattern features by introducing a two-dimensional Image Aesthetic Space (IAS). With a pre-trained Bimodal Deep Autoencoder with Cross Edges (BDA-CE) module, we mapped all the 4000 paintings in our dataset to the IAS. When receiving aesthetic effect and artistic genre by user, our system pick up one painting from the same artistic genre in our dataset, whose coordination is closest to that of aesthetic effect word in IAS. Finally, according to the color pattern of painting picked up before, we do the color modification[10] of our target image to give the image specific aesthetic effect.

3.4 Artistic genre simulation

There are two steps in the artistic genre simulation part: artistic style transfer and brush stroke enhancement.

Artistic style transfer. We choose a neural style transfer method to make our target image more like a painting of the artistic genre chosen by users. Using this method, we need to find a reference painting. To better the transfer effect, we compare the shape features and find the most similar one from the specific artistic genre. We encode the target image and the reference image, use an adaptive instance normalization (AdaIN) layer[6] to transfer style in the feature space, and decode the AdaIN output into image spaces.

Brush stroke enhancement. In this step, we divide our six categories painting into two typical painting materials: oriental ancient ink painting for Chinese-ink-wash and western classical oil painting for other genres. For oil painting, the painting was drawn by a series of brush stroke in different size[5].The painting first renders with large brush strokes and covered by smaller one again and again. Those brush strokes are long and curved, whose colors extracted from the target painting. For ink wash painting, we firstly detect significant edges of the target painting. Then, we use an edge-preserving energy minimization model[7] to propagate the pixel values from the edges to the blank regions.

3.5 Painting process illustration

We present the painting process in two ways based on types of materials. For oil painting, we firstly draw the outline of the painting, then color it layer by layer from large stroke to small stroke. As for ink painting, we divide the painting into different stroke parts, and demonstrate them one by one from main objects to the background.

4 USER STUDY

We invited 10 students (5 males, 5 females) from the Academy of Arts & Design and prepared 30 different paintings of various content text, aesthetic effect word and artistic style. We showed the paintings and their corresponding information to those participants. Participants were asked to use a 10-point scale('10' means the most satisfied) to evaluate three things:

- how do our painting looks like a real painting? The mean of the results we get is 9.22(standard deviation=0.28).So we can infer that our approach have a excellent performance in generating something like real paintings.
- how beautiful is our painting? The mean of the results we get is 8.39(standard deviation=0.59). It indicates that most participants take our paintings as beautiful art works.
- how is our uses' appreciation of our painting consistent with the aesthetic effect word? The mean of the results we get is 7.25(standard deviation=0.83). We can imply that our paintings are able to leave a specific aesthetic impression on participants roughly.

5 CONCLUSION

In this paper, we make a proposal of AI painting system, which can generate a painting based on content text, aesthetic effect word and artistic genre given by users. The painting generated by our system can be further used in various field. Our system can provide designers with decorating paintings used in posters and slides. Besides, the painting generated by our AI painting, can also be used in education field, such as inspiring imagination of children.

6 ACKNOWLEDGE

This work is supported by National Key Research and Development Plan (2016YFB1001200),the Innovation Method Fund of China (2016IM010200).

REFERENCES

[1] Alexei A Efros and William T Freeman. 2001. Image quilting for texture synthesis and transfer. In *Conference on Computer Graphics and Interactive Techniques*. 341–346.

[2] A. Farhadi, I. Endres, D. Hoiem, and D. Forsyth. 2009. Describing objects by their attributes. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. 1778–1785.

[3] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2016. Image Style Transfer Using Convolutional Neural Networks. In *Computer Vision and Pattern Recognition*. 2414–2423.

[4] Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Jimenez Rezende, and Daan Wierstra. 2015. DRAW: a recurrent neural network for image generation. *Computer Science* (2015), 1462–1471.

[5] Aaron Hertzmann. 1998. Painterly rendering with curved brush strokes of multiple sizes. In *Conference on Computer Graphics and Interactive Techniques*. 453–460.

[6] Xun Huang and Serge Belongie. 2017. Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. (2017).

[7] Lingyu Liang and Lianwen Jin. 2013. Image-Based Rendering for Ink Painting. (2013), 3950–3954.

[8] Yihui Ma, Jia Jia, Yufan Hou, Yaohua Bu, and Wentao Han. 2018. Understanding the Aesthetic Styles of Social Images. In *Acoustics, Speech and Signal Processing (ICASSP), 2018 IEEE International Conference on*. IEEE.

[9] Yihui Ma, Jia Jia, Suping Zhou, Jingtian Fu, Yejun Liu, and Zijian Tong. 2017. Towards Better Understanding the Clothing Fashion Styles: A Multimodal Deep Learning Approach. (2017).

[10] Attila Neumann. 2005. Color style transfer techniques using hue, lightness and saturation histogram matching. In *Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*. 111–122.

[11] Devi Parikh and Kristen Grauman. 2011. Relative attributes. *Iccv* 6669, 5 (2011), 503–510.

[12] Xiaohui Wang, Jia Jia, Jiaming Yin, and Lianhong Cai. 2014. Interpretable aesthetic features for affective image classification. In *IEEE International Conference on Image Processing*. 3230–3234.

[13] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, and Dimitris Metaxas. 2017. StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks. *arXiv preprint arXiv:1710.10916* (2017).