Oil Painting Style Transfer for Images Using Both Traditional and Deep Learning Methods

Chuan Sun  
University of KansasLawrence, KS  
chuansun@ku.edu

*Abstract*—In this paper, .

Keywords—Style transfer, F

# Introduction

With the rapid development of image processing and computer vision, the analysis on image styles has been making decent progress. Image styles are characterized by the artists’ painting custom and effect, which in practice correspond to a bunch of painting styles such as watercolor, sketch, oil painting, sand painting and so on. Transferring the style of an image to another is one of the main focus and practical application amidst the domain, which aims to take an arbitrary image as input, and generate an output with the same content but in the desired painting style.

A number of recent works have been conducted on analyzing the style of an image and trying to convert it to another. Traditional ways tend to utilize image processing techniques including image sharpening, edge detection, segmentation, reconstruction, etc. [1], to compute style features mathematically and apply the extracted features onto new images.

The work by Haeberli [2] proposed a method to simulate image tile mosaics using Voronoi diagram and distance measurement, which generates a mosaic-like filter style on images while preserving its contents. Fan Hui, Chen Zheng and Li Jinjiang designed a sand style transfer algorithm [1] using fuzzy enhancement and texture synthesis. These methodologies make full use of mathematical tools such as gradient, differential, probabilistic modeling, Manhattan distance, etc. to simulate the desired style patterns to be applied.

On the other hand, the flourishing development on neural network and deep learning provides a different approach for style comprehension and transformation. A lot of research work has been done on representing image content and style using deep learning techniques such as CNNs and VGG networks. For instance, the work by Leon A. Gatys separates the representation of image content from its style through convolutional neural networks [3]. Justin Johnson replaced the traditional pixel-level loss with perceptual loss defined on high level features [4] for the neural network doing image transformation.

Deep learning methodologies enables better generalization on feature extraction and representation through a stack of neural layers such as normalization, convolution, pooling, or even long short term memory for preserving long sequence information. While at the same time, they require real target tags to do supervised training in most cases.

In this paper, we explore the methodologies of transferring the style of an image to that of an oil painting. To achieve a good blend between oil painting style and image content, the oil brush patterns are analyzed. Then both traditional method and deep learning networks are applied to simulate the oil painting style and compared to see their pros and cons.

# Method

The oil painting style contains a characteristic visual appearance pattern originated from the oil brush, which differentiates itself from other painting styles. We start from constructing a transformation to simulate this characteristic pattern using traditional image processing techniques.

## Traditional Approach

Since the idea comes from producing oil painting styles by simulating every physical stroke of oil paintings, we first need to determine the locations, magnitudes and directions of all potential brush strokes to be put on the canvas. Based on the knowledge learned during the course, edge detection algorithms provide a superb solution for all these three requirements. Gradients of the image are calculated based on formula 1.



The gradients on both horizontal and vertical axis determine the magnitude and the size of the stroke, denoted as

2

In our method, the pattern of every stroke is approximated by an eclipse, and the magnitude of every gradient determines the length and size of the eclipse. The direction of the eclipse is determined by

3

We evenly distribute every stroke across the entire width and height of the canvas on a predefined brush width, and add some random noise to the stroke order so that every stroke looks different and randomized in both size and location.

To approach more resemblance of manually crafted paintings, the original image is first slightly blurred using a medium filter to smooth out sharpness from real camera-taken images. Then the stroke orders parameterized by brush size are generated with random noise, determining the oil brush stroke locations. Finally different eclipses are constructed on these locations based on edge magnitude and orientation.

## Deep Learning Approach

The traditional method turns out to work quite well on transferring image into an oil painting style, while at the same time having limitations as well, especially in terms of style generalization. In this project, the deep learning methodology is further added to explore the effect of neural network layers on style representation and loss optimization during the style transfer process.

This proceeds our work from unsupervised computation into supervised learning domain, where a desired style example will be required as input, and the model is trained to generates an output image in the style of this input.

Convolution Neural Networks are proficient at understanding image data through layers of feature extraction and pooling summarization. When applying such a network onto our research, we need to correspond this advantage to our problems, and make convolutional layers extract the representation of both image contents and image styles. With the model stacking more and more convolutional layers, it tends to extract higher and higher level of information, from image surface patterns to edges and segments finally to combined objects. In order to separate image style from content, we consider the lower layers as representation of styles, and higher layers as representation of contents.

### (1) Model Architecture

The structure and weights of VGG-19 model is utilized here to stack a bunch of convolutional layers together, which consists of 5 model blocks, with each block containing different number of convolutional layers and a max pooling layer, depicted in Fig 1.

A close up of text on a white background

Description automatically generated

1. Architecture of VGG-19.

Each convolution layer applies a number of feature filters to extract different features from the input image, and the model summarizes the extracted feature using max pooling layers in a hierarchical manner. The pretrained ImageNet weights enable strong capability in object recognition and feature extraction. Among this architecture, the first convolution layer in every block is extracted as our style layers, and the last convolution layer serves as our content layer. Style layers are stacked with the content layers to form our deep model.

### (2) Loss Function

Calculating loss has been crucial for the model evolving towards its best solution. For this project, the model should keep a decent balance between style transformation and content preservation, so the loss should take both content loss and style loss into consideration.

When the original content image and the model final output are fed into the network, the model output for filter at layer are denoted as and . The content loss is retrieved from the difference between model output and the target at each layer, which is denoted by

4

This content loss ensures that the model output does not end up far away from the content input. In order to transfer the input image toward the desired style input, the style loss is defined in a similar way, with the only difference that style representation is not calculated from model output, instead the inner correlation between all filters inside a layer is used:

5

where i, j represents any two filters inside a layer, N is the number of filters inside layer , M is the size of the image(width \* height), and denotes the inner product between two filter map in layer fed with the input and respectively.

The total style loss is the sum of all layer losses. To summarize, the style loss does not focus on the feature outcome generated by layers, but on the style resemblance between each two filter maps. Larger inner correlation between two filters indicates higher level of style resemblance, thus ensuring that the model output contains a similar style to the style input.

### (3) Gradient Descent

The model trains for a predefined number of iterations. For each iteration, the model computes total loss and the gradient towards minimizing total loss using Adadelta optimizer. Calculated gradients are then applied onto the input content image for gradual evolving. The best model with lowest total loss is preserved as the final output.

# Experiment

The oil painting transfer is conducted in both traditional and deep learning methods. Five images representative in 5 domains(landscape, architecture, natural wonder, people, object portrait) are selected as content input image. For deep learning methodology, authentic oil paintings are selected as style input.

## Traditional Approach

Fig 2 shows two samples of original content images and the corresponding outputs by the traditional method.

The constructed oil brush strokes achieve a decent effect of imitating authentic oil painting styles, with the brush size tuning to a proper ratio compared to real strokes. Random noise added to brush order locations help randomize brush strokes in large monotonous color areas, so as to make it vivid enough.

A picture containing photo, building, front, man

Description automatically generated

1. Original input images and their corresponding outputs.

The effect of different edge detection algorithms on the output are also analyzed.

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

Then different type of filters with varying parameters were applied to the pattern image to demonstrate their corresponding filtering effect. And finally we tuned a notch filter and a lowpass filter to alleviate the noises in the space image.

##### References

1. M. Souden, J. Benesty and S. Affes, "On Optimal Frequency-Domain Multichannel Linear Filtering for Noise Reduction," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 2, pp. 260-276, Feb. 2010.
2. Soo-Chang Pei and Chien-Cheng Tseng, "Two dimensional IIR digital notch filter design," in IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 41, no. 3, pp. 227-231, March 1994.