Perceptual-Based CNN Model for Watercolor Mixing Prediction

Ya-Bo Huang
National Taiwan University
pixiyon@cmlab.csie.ntu.edu.tw

Mei-Yun Chen
National Taiwan University
meiyun@cmlab.csie.ntu.edu.tw

Ming Ouhyoung
National Taiwan University
ming@csie.ntu.edu.tw

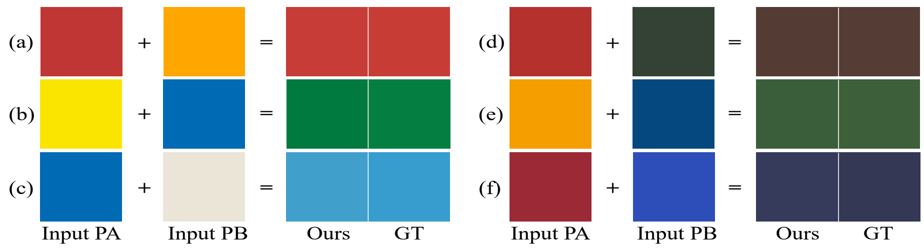

Figure 1: Part of color mixing results (a)~(f), showing different cases in our test set. With two input pigments PA and PB, our model can predict the results (Ours, left) which are almost indistinguishable to the ground truth (GT, right) of the real mixed pigment.

ABSTRACT

In the poster, we propose a model to predict the mixture of watercolor pigments using convolutional neural networks (CNN). With a watercolor dataset, we train our model to minimize the loss function of sRGB differences. In metric of color difference ΔE_{Lab}, our model achieves 88.7 % of data that $\Delta E_{Lab} < 5$ on the test set, which means the difference can not easily be detected by human eye. In addition, an interesting phenomenon is found; Even if the reflectance curve of the predicted color is not as smooth as the ground truth curve, the RGB color is still close to the ground truth.

CCS CONCEPTS

• **Computing methodologies** → *Non-photorealistic rendering*;

KEYWORDS

pigment mixing, convolutional neuron network, machine learning, color matching function

ACM Reference Format:

Ya-Bo Huang, Mei-Yun Chen, and Ming Ouhyoung. 2018. Perceptual-Based CNN Model for Watercolor Mixing Prediction. In *Proceedings of SIGGRAPH '18 Posters*. ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3230744.3230785

1 INTRODUCTION

Pigment mixing plays a crucial part in a digital painting system. A good model of color mixing can be beneficial to render realistic color for painting. Previous works on the color mixing [Haase and Meyer 1992] and [Baxter et al. 2004] use Kubelka-Munk model [Kubelka 1948] to predicts the reflectance of given mixture of pigments . However, limitations of KM model have been reported for translucent media such as watercolor [Edström 2007].

Therefore, we propose a feed-forward network with a perceptual-based loss function to predict a realistic color mixing of the watercolor pigments. In the work, we use the reflectance and transmittance of two watercolor pigments as well as the reflectance of a white paper as the input; the output is the reflectance of a mixed pigment. Lastly, we use standard RGB (sRGB) with CIE 1931 2° observer and standard illuminant D65 to show RGB colors of predicted results. Besides, a metric of color difference metric ΔE_{Lab} is chosen to evaluate the accuracy of our result.

2 OUR APPROACH

Our goal is to generate mixed color reflectance from two input pigments and to minimize the perceptual difference between the ground-truth color and our result.

2.1 Dataset

We use a dataset of watercolor pigments from [Chen et al. 2018]. The dataset contains 1043 mixed pigment, with 13 primary pigments in 12 different quantities (thickness). Each mixture is mixed by two primary pigments. In the dataset, the primary pigment is denoted by its transmittance and reflectance spectrum; the mixture and paper is only represented by the refletance. Each spectra is a 41-dimension vector. Thus, our input has 205 features, namely the transmittance and reflectance of two primary pigments as well as the reflectance of paper; then, the output is a mixed pigment reflectance with 41 dimensions.

To match the size of the input layer, input features are resized into (41,5) for CNN model.

2.2 CNN Model

Our model is an alternative ResNet18 [He et al. 2016] , which is bulit with different filter sizes, pooling and output layers to generate a reflectance spectrum. We found that the modified ResNet can

SIGGRAPH '18 Posters, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5817-0/18/08.
https://doi.org/10.1145/3230744.3230785

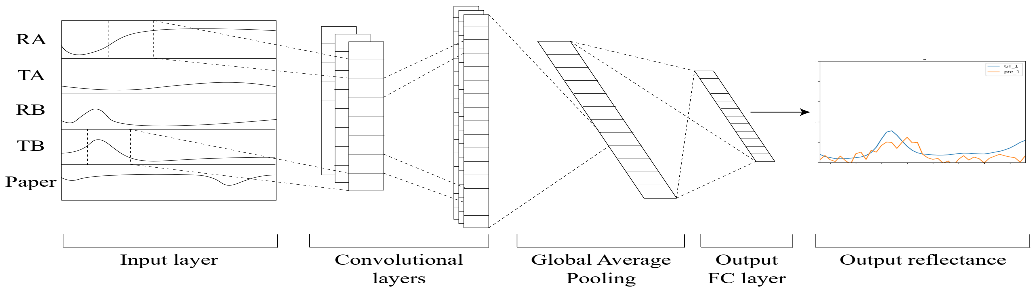

Figure 2: An illustration of the architecture of our model. We use a alternative ResNet18 to extract multi-level features from input reflectance and transmittance (RA, RB, TA, TB and Paper). A global average pooling layer is connected to the last conv. layer to reduce the number of parameters in our model. Fully connected layers after the poolong layer perform regression to minimize the perceptual loss between output and ground truth.

achieve a good performance in color mixing problem and adopted it into our work. The filter sizes are (9,1) for the first two ResBlocks and (3,1) for the last two ResBlocks. The filter numbers are {64, 128, 256, 512} from the first to the last ResBlock. Those changes made our model be able to extract multi-level features from input reflectance and transmittance. To minimize overfitting, a global average pooling layer is connected to last convolutional layer to reducing the total number of parameters in our model. After the poolong layer, the fully connected layers perform regression to fit the output reflectance and minimize the perceptual loss between reflectance and ground truth.

2.3 Loss Function

In order to minimize the perceptual differences in color, we choose the distance of tristimulus value distance, which corresponds to three kinds of cone cells in human eyes, as our loss function for training. The output reflectance spectrum and ground truth is converted into tristimulus values. Then, we compute Euclidean distance between output reflectance and ground truth. In praticle, we choose the distance in linear sRGB colorspace as loss function (Eq. 1).

$$Loss_{sRGB} = \sqrt{(\Delta R)^2 + (\Delta G)^2 + (\Delta B)^2} \tag{1}$$

There are 2 reasons to choose sRGB as loss function. (1). sRGB colorspace is widely used on monitors, printers, and the Internet. (2). Linear sRGB transformation is differentiable so it can be optimized by stochastic gradient descent (SGD).

In addition, the Lab distance (ΔE_{Lab}) (Eq. 2) is chosen as the evaluation metric.

$$\Delta E_{Lab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{2}$$

Since Lab transformation is not differentiable therefore it is not eligible for loss function.

3 RESULT

Part of our results are shown in Fig. 1. Our model achieves mean ΔE_{Lab} of 2.29 and $\Delta E_{Lab} < 5$ of 88.7% on test set. The cases where $\Delta E_{Lab} > 5$ means that we can easily notice a difference between the prediction and the ground truth.

Another finding in output reflectance is shown in Fig. 3. We found that the output spectra only focus on the part corresponding to visible and primary colors, and ignore the others. Therefore, the output spectrum in Fig. 3 act as color matching functions to blue and a little green.

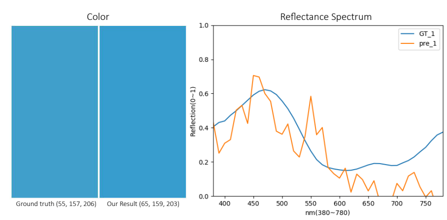

Figure 3: An illustration of the output reflectance of our model. The curve between visible wavelengths (380 to 750 nm) mainly focus on the short-term average, while the curve beyond 750nm can be totally ignored (being invisible to human eye).

ACKNOWLEDGMENTS

This project was partially supported by Meidatek Inc. under Grant No.:˜MTKC-2018-0167 , Ministry of Science and Technology(MOST), Taiwan under Grant No.:˜106-3114-E-002-012 and˜105-2221-E-002-128-MY2

REFERENCES

William Baxter, Jeremy Wendt, and Ming C Lin. 2004. IMPaSTo: a realistic, interactive model for paint. In *Proc. of the 3rd international symposium on Non-photorealistic animation and rendering*. ACM, 45–148.

Mei-Yun Chen, Ci-Syuan Yang, and Ming Ouhyoung. 2018. A Smart Palette for Helping Novice Painters to Mix Physical Watercolor Pigments. In *Proc. of EuroGraphics 2018, Posters,*, Eakta Jain and Jirí Kosinka (Eds.). The Eurographics Association, April 2018. https://doi.org/10.2312/egp.20181008

Per Edström. 2007. Examination of the revised Kubelka-Munk theory: considerations of modeling strategies. *JOSA A* 24, 2 (2007), 548–556.

Chet S Haase and Gary W Meyer. 1992. Modeling pigmented materials for realistic image synthesis. *ACM Transactions on Graphics (TOG)* 11, 4 (1992), 305–335.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proc. of the IEEE conference on computer vision and pattern recognition*. 770–778.

Paul Kubelka. 1948. New contributions to the optics of intensely light-scattering materials. Part I. *Josa* 38, 5 (1948), 448–457.