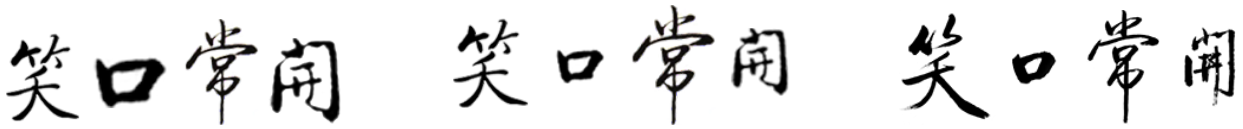
Optimal Character Composing for Chinese Calligraphic Artwork

Yiping Meng[1,2] Fan Tang[1,2] Weiming Dong[1] Xiaopeng Zhang[1]
[1]NLPR-LIAMA, Institute of Automation, Chinese Academy of Sciences [2]University of Chinese Academy of Sciences



(a) Characters are in the same font size (b) Our result (c) Calligrapher's artwork

Figure 1: *Our method can optimally adjust the sizes of characters to create a Chinese calligraphic artwork with style similar to that of the original calligrapher. Compared with the result (a) with characters in the same font size, our result (b) is more similar as the composing style of a real calligraphy (c) and always outperform those in traditional font size. In Chinese the four characters mean "Grinning all the time".*

Keywords: Chinese calligraphic, Markov process

Concepts: •**Computing methodologies** → **Image processing;**

1 Introduction

Chinese Calligraphy is a form of calligraphy widely practised in China and revered in the Chinese cultural sphere, including countries such as Japan, Korea, Vietnam, etc.. The variation of characters in shape, size and shade is the key skill in creation of Chinese calligraphic artworks. Characters should be carefully designed and composed together to create an excellent artwork. Previous works usually focus on synthesizing new calligraphic characters of a specific style [Xia et al. 2013], reproducing the writing processes of Chinese calligraphic artworks [Yang et al. 2014], or animating the construction process of Chinese ink-wash paintings [Tang et al. 2014], instead of creating and optimizing new calligraphic artworks. In this work, we develop a data-driven framework to optimally compose multiple characters picked from a database to form a Chinese Calligraphic artwork (see Figure 1). Our system can automatically generate a calligraphy in the style of a specific calligrapher. Currently we focus on the adjustment of character sizes in an artwork and formulate the composing of characters as a Markov process. The understanding of art varies with people, so our objective is finding a *reasonable* solution that is plausible to an artist's creating style, instead of searching for the *best* composing result. To the best of our knowledge, our work is the first attempt to optimize the creating process of Chinese calligraphic artworks.

2 Technical Approach and Future Work

Database To simulate the creating style of a calligrapher, we first build a database that contains multiple calligraphic artworks and also single characters written by the calligrapher. In this work, we focus on the simulation of "Qi Gong" style (Qi Gong was one of the most famous calligraphers in China). We downloaded 500 images of Qi Gong's calligraphic artworks from internet and segmented the characters in each image by following their bounding boxes. We then labelled every character with its original text and also its script type. The calligraphic artworks and their corresponding characters were organized together to construct the database.

(a) Character (b) Outer contour (c) Feature

Figure 2: *Extracting feature vector from a character.*

Feature extraction For each character in the database, we first extract its outer contour by Canny detector. We then use the orthocenter of the outer contour as the original point and uniformly launch 120 rays from it (i.e., the angle between two rays is 3°). For each ray, we get the line segment between the origin point and the intersection of the ray and the outer contour, and use the length of the segment as a property of the character. Thus, we get a 120-Dimensional vector $\vec{f_c}(l_1, l_2, \ldots, f_{120})$ and then use this vector as the feature of a character, which describes both the size and the script structure of the character (see Figure 2).

Learning strategy There are two factors which affect the character sizes in a calligraphic artwork: the intrinsic size of a character, and the layout and structure when different characters are composed together to form an artwork. The intrinsic size of a character means that in an artwork some characters are smaller than others, due to the personal custom of the calligrapher. For traditional printed characters, the widths of characters in a font library are usually same, so the font size is specified by the width of of each character. However, this definition method cannot depict the intrinsic sizes of different characters in calligraphic artworks. Thus, different with the definition of traditional font size, we define the font size of a calligraphic artwork as the average width of all characters in it.

Besides the intrinsic character size, the size of a character is also affected by its neighbours. We simply this affection to a first-order Markov process by assuming the size of a character is only affected by the previously added character. For all the characters in an artwork, we first perform zero mean processing to each dimension of their features, and then use the sum of all dimensions of the new feature of a character C_i as its state $X(C_i)$ (the size in the calligraphic artwork). For simplicity, we uniformly sampling all the state values to 20 states and construct the state space S. Each calligraphic handwriting can be expressed as a sequence of state values in S. We denote $F_{uv}(u, v \in S)$ as frequency that state S_u converts

SA '16, December 05-08, 2016, , Macao
ISBN: 978-1-4503-4540-8/16/12
DOI: http://dx.doi.org/10.1145/3005274.3005292

to S_v, thus the transition probability p_{uv} from S_u to S_v is:

$$p_{uv} = \frac{F_{uv}}{\sum_{t \in S} F_{ut}}$$

Supposing the state of the ith character $X(C_i)$ is S_u, according to the first-order assumption, the probability that $X(C_{i+1})$ squals S_v can be formulated as:

$$P(X(C_{i+1}) = S_v) = P(X(C_{i+1}) = S_v | X(C_i) = S_u) = p_{uv}$$

The transition probability Matrix $\mathcal{P}_{20*20} = (p_{uv})$ reveals the variety of characters in the same artwork and can be generated by traversing all the artworks in the database and recording the state transfer between characters.

Calligraphy generation Given the texts and the font size, we first pick the corresponding characters and specify the size of each character following our definition of calligraphic font size. Once the first character's size is assigned, we compute its state and then predict the latter character states one by one based on the probability transfer matrix P. We then utilize the states to further adjust the character sizes to get the final calligraphic artwork.

In the future work, we will add more features to represent the characteristics of a character in a calligraphic artwork. Furthermore, we will try to figure out whether different artists adjust character sizes in a different way, and whether there is a common composing rule that is used by all artists.

Acknowledgements

This work was supported by National Natural Science Foundation of China under nos. 61672520 and 61372184, by Beijing Natural Science Foundation under No. 4162056, and by CASIA-Tencent BestImage joint research project. Weiming Dong is the corresponding author.

References

TANG, F., MENG, Y., DONG, W., MEI, X., HUANG, F., AND ZHANG, X. 2014. Animated construction of ink-wash paintings. In *SIGGRAPH Asia 2014 Posters*, ACM, New York, NY, USA, SA '14, 23:1–23:1.

XIA, Y., WU, J., GAO, P., LIN, Y., AND MAO, T. 2013. Ontology-based model for chinese calligraphy synthesis. *Computer Graphics Forum 32*, 7, 11–20.

YANG, L.-J., XU, T.-C., LI, X.-S., AND WU, E.-H. 2014. Feature-oriented writing process reproduction of chinese calligraphic artwork. In *SIGGRAPH Asia 2014 Technical Briefs*, ACM, New York, NY, USA, SA '14, 5:1–5:4.