Context-Aware Text-Based Binary Image Stylization and Synthesis

Shuai Yang, Jiaying Liu, Senior Member, IEEE, Wenhan Yang, Student Member, IEEE, and Zongming Guo, Member, IEEE,

As supplementary material of our paper, we present the following contents:

- Visual comparison with state-of-the-art style transfer methods. (Figs. 1-3)
- Synthesized nine pieces of visual-textual presentations. (Fig. 4)
- Illustration of text image stylization on twelve fonts and twelve languages. (Figs. 5-6)
- Texture rendering for Zodiac symbols and emoji icons. (Figs. 7-8)
- Summary of parameters for user control. (Table I, Figs. 9-13)
- More stroke based image stylization results. (Fig. 14)
- Guideline for picking up the background photo.

I. COMPARISON OF STYLE TRANSFER METHODS



Fig. 1: Visual comparison of text stylization. The top row shows source style images. For each result group, the first one is the input target text. Other images are results by supervised (the upper row) and unsupervised (the lower row) methods. For supervised methods, the structure guidance map extracted by our method is directly given as input. *Image credits: Unsplash users Aaron Burden, Ashim D'Silva, Yanguang Lan.*



S' (flame)

S' (coral reef)

S' (wall)



Fig. 2: Visual comparison of text stylization. The top row shows source style images. For each result group, the first one is the input target text. Other images are results by supervised (the upper row) and unsupervised (the lower row) methods. For supervised methods, the structure guidance map extracted by our method is directly given as input. *Image credits: Unsplash users Anna Popovic, Ishan@seefromthesky.*



 $S' \ (spume)$

 $S' \ (neon)$



Fig. 3: Visual comparison of text stylization. The top row shows source style images. For each result group, the first one is the input target text. Other images are results by supervised (the upper row) and unsupervised (the lower row) methods. For supervised methods, the structure guidance map extracted by our method is directly given as input. Image credits: Unsplash users Cassie Matias, Lex Guerra, Christian Nielsen.

II. VISUAL-TEXTUAL PRESENTATION SYNTHESIS







Fig. 4: Visual-textual presentation synthesis. For each result group, three images in the upper row are I, S' and T, respectively. The lower one is our result. *Image credits: Unsplash users Math, NASA, Christian Nielsen, Yanguang Lan, Tim Stief, Henning Witzel, Tim Gouw, Thomas Kelley.*

III. GENERATING STYLISH TEXT IN DIFFERENT FONTS



Fig. 5: Performance on different languages. Image credits: Unsplash users Anna Popovic, Aaron Burden, Steinar Engeland, Yanguang Lan.

IV. GENERATING STYLISH TEXT IN DIFFERENT LANGUAGES



Fig. 6: Performance on different fonts. Image credits: Unsplash users Ashim D'Silva, Grant McCurdy, Yanguang Lan.

V. SYMBOL AND ICON RENDERING



S' (water)

S' (star)





Fig. 7: Rendering rippling/starry Zodiac symbols. Image credits: Unsplash users Raphaël Biscaldi, Christian Nielsen.





Fig. 8: Rendering emoji icons with different styles. Image credits: Unsplash users Ashim D'Silva, Anna Popovic, Cassie Matias.

VI. PARAMETERS FOR USER CONTROL

Parameters	Description	Default value	Low value	High value
L	number of levels in structure transfer	5	more rigid text contour	more stylish text contour
λ_1	weight of the distribution term in texture transfer	0.1	more blurring text contour	cleaner text contour
λ_2	weight of the psycho-visual term in texture transfer	0.001	more texture details	more diversified textures
λ_3	weight of the saliency term in texture transfer	0.05	less salient foreground text	more salient foreground text
λ_4	weight of the aesthetics cost in position estimation	0.5	more seamless embedding	more centered text
colorFlag	flag indicating whether to perform color transfer	FALSE	-	-
positionFlag	flag indicating whether to estimate text position	TRUE	-	-

TABLE I: Parameters for user control



Fig. 9: Effects of the number L of image pyramid layers. From left to right: our legibility-preserving structure transfer result using L = 3, 7, 11, respectively.



Fig. 10: Effects of the weight λ_1 of the distribution term in texture transfer. By incorporating distance term, our method better characterizes the texture patterns along the text outline.



Fig. 11: Effects of the weight λ_2 of the psycho-visual term in texture transfer. A low λ_2 causes texture over-repetitiveness while a large λ_2 tends to blur texture details. Cropped regions are zoomed for better comparison.



Fig. 12: Effects of the weight λ_3 of the saliency term in texture transfer. The saliency term makes the foreground character more prominent and the background cleaner, thereby enhancing the legibility.



Fig. 13: Effects of the weight λ_4 of the aesthetics cost in position estimation. The white dotted lines are added to facilitate observation and comparison.

VII. MORE STROKE BASED IMAGE STYLIZATION RESULTS



Fig. 14: More stroke based image stylization results. For each result group, two images in the upper row are S' and T, respectively. The lower one is our result. *Image credits: Unsplash users Steinar Engeland, Andreas Gücklhorn, Matt Jones, Greg Shield, Glen Carrie.*

VIII. BACKGROUND PHOTO GUIDELINE

- It is recommended that the background image has a flat area of sufficient size to place text.
- It is recommended that the background image has certain semantic relevance to the style images, such as blue sky and white clouds.
- It is recommended that the background image and the style image do not differ too much in color. If the background image and the style image have totally different color distributions, more sophisticated color transfer methods or user interactions may be required to perform color matching.

REFERENCES

- [1] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," in Proc. Conf. Computer Graphics and Interactive Techniques, 2001, pp. 327-340.
- [2] A. J. Champandard, "Semantic style transfer and turning two-bit doodles into fine artworks," 2016, arXiv:1603.01768.
 [3] S. Yang, J. Liu, Z. Lian, and Z. Guo, "Awesome typography: Statistics-based text effects transfer," in *IEEE International* Conference on Computer Vision and Pattern Recognition, 2017.
- [4] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in Proc. ACM Conf. Computer Graphics and Interactive Techniques, 2001, pp. 341-346.
- [5] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2016.
- [6] C. Li and M. Wand, "Combining markov random fields and convolutional neural networks for image synthesis," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2016.