Arbitrary Style Transfer of Facial Image Based on Feed-forward Network and Its Application in Aesthetic QR Code

Shanqing Zhang[1], Shengqi Su[1], Li Li[1], Jianfeng Lu[1], Ching-Chun Chang[2]

[1] Hangzhou Dianzi University, School of Computer Science and Technology, China
[2] Department of Computer Science, University of Warwick, UK
sqzhang@hdu.edu.cn, sushengqi1994@163.com, lili2008@hdu.edu.cn, jflu@hdu.edu.cn,
ching-chun.chang@warwick.ac.uk

Abstract

QR code has become essential in daily-life because of the popularity of mobile devices. The visual effect of a conventional QR code is not ideal. Consequently, many good aesthetic algorithms have been proposed. However, both the decoding rate and visual effect of a QR code cannot be guaranteed simultaneously when facial image serves as the background. We propose an arbitrary style transfer of facial image based on feed-forward network as a preprocessing algorithm for an aesthetic QR code. The deep characteristics of content image and style image are unified in the same layer of convolutional neural networks in our style transfer network. Styles are changed. The result of style transfer is restricted with semantic segmentation result, color uniform regularization of facial image and repeating restriction similarity constraints. Experimental results show that both the decoding rate and visual effect of a QR code are guaranteed when our method is used in background preprocessing.

Keywords: Arbitrary style transfer, Feed-forward network, Facial image, Aesthetic QR code

1 Introduction

With the economic development in recent years, smart mobile devices have be-come more and more popular. The QR code has the advantages of being produced conveniently, having large storage capacity, strong error correction capability and anti-rotation capability. Consequently, QR code is used extensively as an interface between smart mobile devices and the Internet. Today, QR codes have become an integral part of life. An ordinary QR code with visual-unpleasant appearance consists of monotonic black/white encoding modules, which cannot be recognized by human eyes. Recently, visual optimization of QR codes has attracted extensive attention in academia.

Consequently, many good aesthetic algorithms have been proposed. However, most aesthetic QR algorithms are focused on improving the similarities to the original background, but neglect low decoding rate. We found that most aesthetic QR algorithms are suitable for flat texture or uniformly colored anime style images. We use the scanning error analysis method [1] to analyze the effects of different background images. In Figure 1, the red parts indicate errors and the green parts indicate correctly-identified values. There are fewer errors in the sampling results matrix of the simple texture image. According to scanning principles, if there are errors in the matrix of sampling results, they need to be corrected by performing data correction. When the percentage of the generated errors exceeds a certain threshold, error correction may fail, which results in a scanning failure. Preprocessing the background is therefore used to increase the decoding rate. Common preprocessing algorithms enhance contrast, remove texture, etc.. The visual effect is impacted although satisfactory decoding rate is obtained. It is critical to obtain suitable preprocessing algorithms for facial image with a good visual effect and decoding rate.

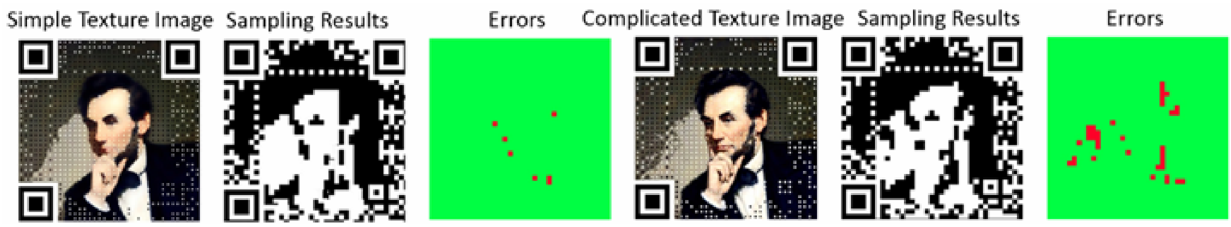

Figure 1. Error analysis of the intermediate results of scanning different QR codes using the ZXing library

*Corresponding Author: Li Li; E-mail: lili2008@hdu.edu.cn
DOI: 10.3966/160792642020032102019

We introduce convolutional neural networks (CNN) based style transfer into preprocessing. Style transfer removes details from facial image, thereby improving the decoding rate, achieving both reality and aesthetics. However, the current style transfer cannot be used in facial image because style distortion and non-unified image contents. The corresponding CNN is required for every style image training in conventional image stylization network.

We referenced the network structure of Chen et al. [2]. The image contents and deep characteristics of style images are unified into one layer of CNN, style change is performed, which resulted in faster speed of arbitrary style transfer of images. We further introduced semantic segmentation result to mitigate the impact on texture and color by the background. Color uniform regularization of facial image was introduced and the maximum times in any patch was con-strained, greatly improving the quality of facial style images. The experimental results showed that our method yields both a high decoding rate and good visual effect.

2 Related Work

In this section, we review some techniques primarily concerning two topics, aesthetic QR code and image style transfer.

2.1 Aesthetic QR Code

Recently, more and more researches have focused on aesthetic QR code. Current aesthetic algorithms include two categories.

The first category uses the error correction encoding mechanism of RS code [3-5]. It can highlight some important areas of image so that the region of interest (ROI) can be completely displayed. The experimental results show that there are almost no noise pixels on ROI. The main drawback is that the replacement area is limited by error correction capacity. In addition, a processed image is covered with a lot of the black and white points and the visual effect of salient regions is poor. The results are shown in Figure 2(a).

The second category uses the XOR characteristic of RS code [6-8] to beautify the QR code. The results are shown in Figure 2(b) and Figure 2(c). In our previous work, we proposed a novel algorithm based on the XOR mechanism of hybrid basis vector matrices and background image synthetic strategy [9]. The hybrid basis vector matrices include the reverse basis vector matrix (RBVM) and positive basis vector matrix (PBVM). First, RBVM and PBVM are established by Gauss Jordan elimination method according to the characteristics of the RS code. Then, the modification of the parity area of the QR code can be performed with the XOR operation of RBVM. And the XOR operation of PBVM is used to change the data area of the QR code. So a QR code can be modified to be very close to the background image without affecting the error correction ability. Finally, in order to further decrease the difference between a QR code and the background image, a new synthesis strategy is adopted to achieve better aesthetic effect. The experimental results show that this method yield better visual effect without sacrificing recognition rate. The results are shown in Figure 2(d).

(a) The result of [4] (b) The result of [6] (c) The result of [7] (d) The result of [9]

Figure 2. The aesthetic QR codes of different algorithms

At present, it is difficult for the existing QR code beautification algorithms to address both the decoding rate and the aesthetics of a QR code, especially for facial images.

2.2 Style Transfer

Recently, image style transfer has become a hot topic in the AI field. It is related to texture synthesis. Texture synthesis methods are divided into two categories, i.e., parametric [10-12] and non-parametric [13-15] method. Para-metric methods of texture synthesis aim to represent textures through proper statistical models. Non-parametric methods aim to build a perfect image rather than build rich models to understand textures. Li and Wand [16] showed that representations of image content and style were separable by various CNN convolutional layers. Moreover, representation provided the possibility for image decoupling and recombining. Gatys et al. [17] formulated style transfer as an optimization problem that combines texture synthesis with content reconstruction. By matching the global statistics of deep features, the parametric methods can be used to preserve the content of image and the overall visual effect of an artwork.

Based on the existing image style transfer approaches with CNN, we propose an arbitrary style transfer method of facial image based on single-layer activations in a pre-rained CNN. Similar to Chen's work [2], we reconstruct the complete best matching activation target in the activation space. More importantly, a visually appealing style transfer can be achieved by using our method, which is able to directly adapt to the facial area. We adjust the network of the style image transfer according to the special requirement of facial image as the content image. First, unlike the existing style transfer methods [16-17], our method does not use pixel-level loss, instead, our method uses loss on the activations. Second, semantic segmentation result and repetitive restriction can help achieve arbitrary style transfer during semantic matching between images. Finally, we use a global constraint based on YUV color space of an image to avoid the color-overmatch problem, which greatly improves the photorealism of the transfer results.

3 A New Method of Style Transfer

The network architecture of our method is shown in Figure 3. The inputs of feed-forward network are facial image (i.e., content image) and style image. DeepLab [18] is used to obtain semantic segmentation result before content images are inputted into the network. Every patch activated by content images is re-placed by the corresponding patch of style transfer images with global balance and semantic segmentation result. Facial image with high fidelity is acquired through a uniform style transfer is obtained. The VGG-19 [19] model is used in this study.

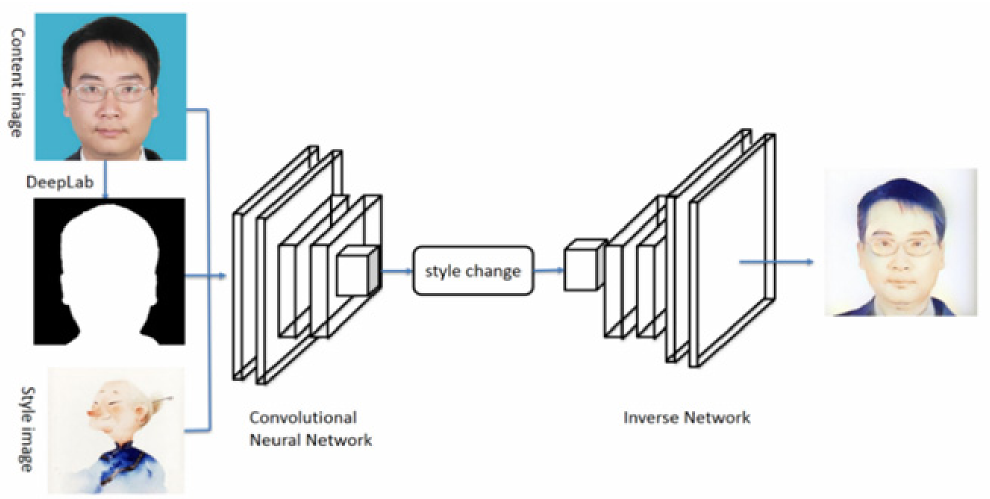

Figure 3. Network architecture of arbitrary style transfer of a facial image based on feed-forward network

3.1 The Activation Based on Semantic Segmentation Result

In neural algorithm of style transfer, CNN activation is used as the feature of an image. Experimental results show that the interference on facial image from background is the primary factor contributing to infidelity in the process of style transfer. We defined an activation function based on semantic segmentation to solve this problem.

DeepLab [18] is used to generate semantic segmentation result for a facial image. The content image is divided into face and background. The activation of content image with semantic segmentation result is the input of arbitrary style transfer network. The enhanced activation of content image $F_{seg}(C)$ is

$$F_{seg}(C)=F(C)J(C). \quad \textbf{(1)}$$

where, $F(C)$ is the activation of a layer CNN of content image C, $J(C)$ is the semantic segmentation result of the corresponding content image. Downsampling operation matches the content image activation size of every CNN layer.

3.2 Non-parametric Image Style Tran Sfer Network

In order to obtain better visual effects, we adopt a non-parametric algorithm of style transfer, and its comparison with the results of the parametric style transfer network is shown in Figure 4. Let C and S denote content image and style image of a style transfer, respectively. $F(\bullet)$ is the full convolution operation for the pre-trained CNN. An image is mapped to an intermediate activation space in CNN from its original image space. The image style change is performed on the activation space $F(S)$ and the enhanced content image activation $F_{seg}(C)$ with semantic segmentation. The change process is as follows:

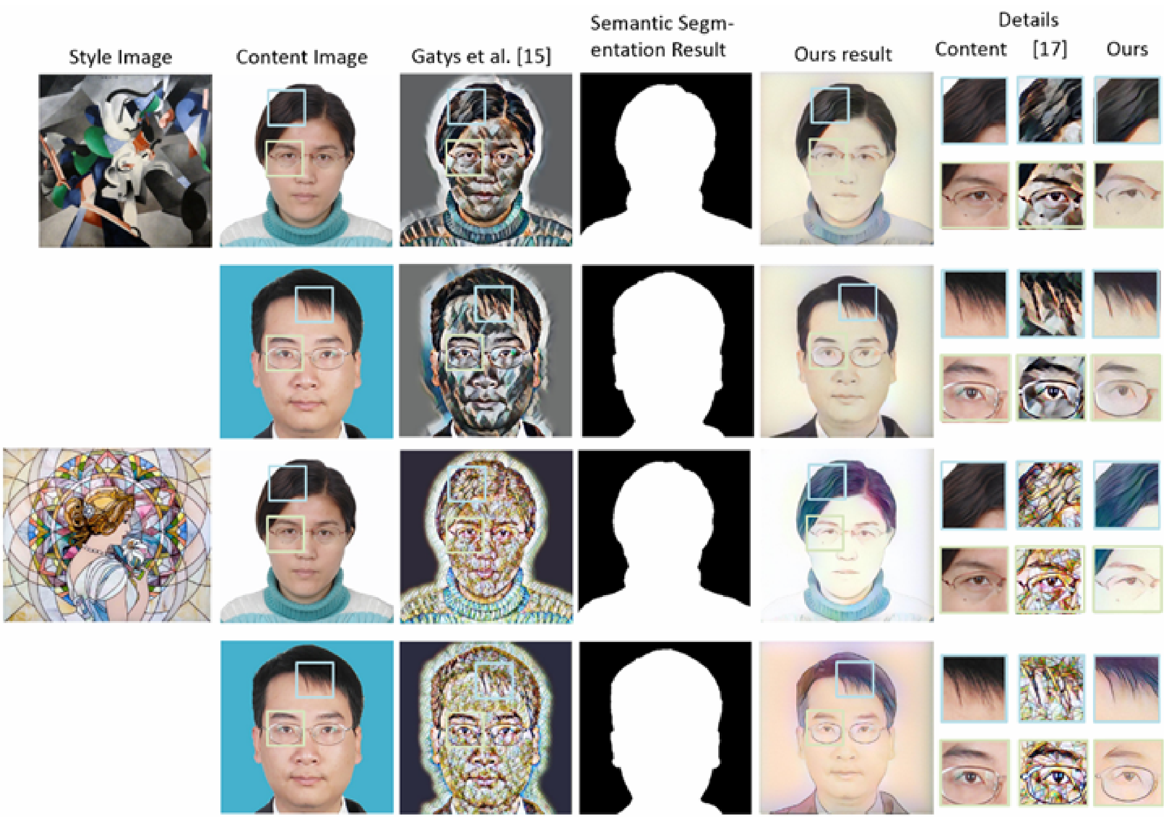

Figure 4. The contrast of the parameter style transfer network [17] with our non-parametric feed-forward network style

(1) Pre-trained VGG-19 network is used to generate the corresponding activations of a content image C and style image S. Patches with size $r \times r$ are extracted from the activations. The patches are marked as $f_{seg\_i}(C)$, $i \in n_C$ and $f_j(S)$, $j \in n_S$, respectively. The size of the content image is $h \times w$, n_C and n_S represent the number of extracted patches of the content image and style image from the current layer. The extracted patches should have sufficient overlap and include all activated channels.

(2) The activation of the content image with semantic segmentation result is correlated to the best matching patch of the style image.

$$f_i^{st}(C,S) = \underset{f_j(S), j=1,\dots,n_S}{\arg\max} \frac{\langle f_{seg\_i}(C), f_j(S) \rangle}{\| f_{seg\_i}(C) \| \cdot \| f_j(S) \|}. \quad (2)$$

(3) The activation patch of the content image $f_{seg\_i}(C)$ is replaced by the best matching activation patch of the style image $f_i^{st}(C,S)$.

(4) The full content activation is re-constructed and is represented by $F^{st}(C,S)$.

3.3 Repeating Restriction Similarity Constraints

The enhanced activation patch of the content image with semantic segmentation result is correlated to the best matching patch of the style image, and the full target activation is reconstructed. A part of the activation patch of the style image is used repeatedly. It adversely affects the uniform transfer of the image style. It cannot produce rich style transfer of the image. Inspired by [2-20], we introduced the repeating restriction parameter to limit the usage of activation patch of the style image.

As its name implies, repeating restriction similarity constraints addresses similarity of image area. It is completed in Step 2 of the non-parametric image style transfer in Section 3.2. When the best matching style patch with a patch of con-tent image is determined, a count matrix with initial zero is introduced. When an activation patch of the style image is used, the corresponding matrix element is incremented. The patch cannot be used any more after it reaches the threshold and the next new best matching style patch is used. The style expands more uniformly and image style transfer is richer when the threshold is 5.

3.4 The Color Uniform Regularization of Facial Image

Image realistic regularization [21] uses an affine function to widen color difference. It is not suitable for facial image. And regularization in [21] on RGB space requires heavy computation for style transfer. In order to solve these problems, we propose the color uniform regularization of facial image. We directly employ the semantic segmentation result from DeepLab in Section 3.1, with the scope to local transformation being controlled. Local transformation is performed in channel Y on YUV space of the content image. The style transfer is faster. Importantly, we decrease the range of the brightness field to weaken the impact from the brightness difference in the facial area. The results are shown in Figure 5.

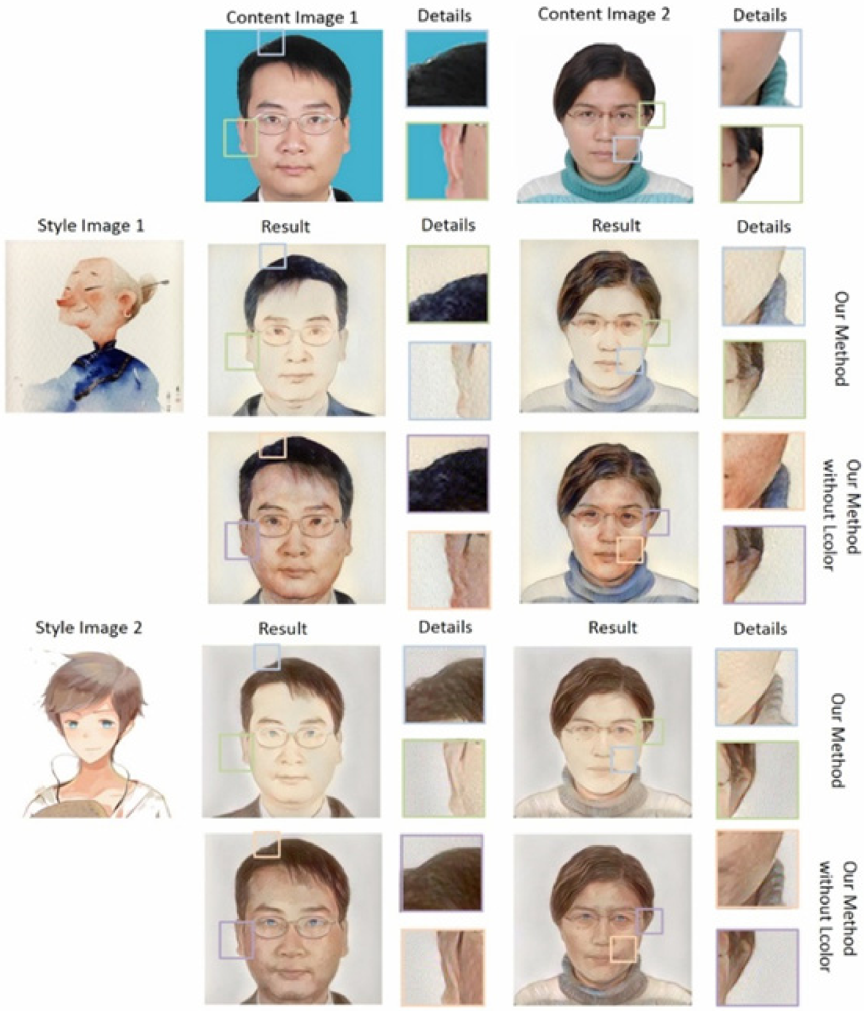

Figure 5. The contrast of our feed-forward network with the same network without color uniform regularization

Color uniform regularization of facial image is defined as follows:

$$L_{color}\left(I,C\right)=V_Y\left(I\right)^T M_Y\left(C\right)V_Y\left(I\right), \tag{3}$$

where content image C has N pixels. $V_Y\left(I\right)$ the vectorized version $\left(N\times 1\right)$ of the style transfer image I in channel Y. The matrix $M_Y\left(C\right)$ $\left(N\times N\right)$ is a representation of a standard linear system which can minimize a locally affine function of the content image C in Y channel [21].

3.5 Loss Function

Different from [16-17], we use the squared-error in our loss function. The enhanced image activation $F_{seg}\left(C\right)$ with semantic segmentation result is redefined. The full activation is reconstructed under the repeating restriction similarity constraints. We pioneer to calculate the total loss in YUV space and add the color uniform regularization of facial image in the loss function. Our loss function is defined as follows:

$$\begin{aligned} I_{style}\left(C,S\right)=&\arg\min_{I\in\mathbb{R}^{h\times w\times d}}\left\|F\left(I\right)-F^{st}\left(C,S\right)\right\|_F^2 \\ &+\alpha L_{TV}\left(I\right)-\beta L_{color}\left(I,C\right), \end{aligned} \tag{4}$$

where $\left\|\bullet\right\|_F$ is the Frobenius norm, L_{color} is the color uniform regularization of the facial image, $L_{TV}\left(\bullet\right)$ is the total variation regularization term [22] that is widely used in image deblurring;

$$\begin{aligned} L_{TV}\left(I\right)=&\sum_{i=1}^{h-1}\sum_{j=1}^{w}\sum_{k=1}^{d}\left(I_{i+1,j,k}-I_{i,j,k}\right)^2 \\ &+\sum_{i=1}^{h}\sum_{j=1}^{w-1}\sum_{k=1}^{d}\left(I_{i,j+1,k}-I_{i,j,k}\right)^2, \end{aligned} \tag{5}$$

where $F\left(\bullet\right)$ is activation from the pre-trained CNN. Therefore, the final solution of (4) can be obtained by subgradient-based optimization. Based on [2], an inverse network can be trained to approximate an optimum of the loss function in (4) for any activations. We define the optimal inverse function as follows:

$$\arg\inf_{g} \mathrm{E}_H\left[\left\|F\left(g\left(H\right)\right)-H\right\|_F^2 + \alpha L_{TV}\left(g\left(H\right)\right)-\beta L_{color}\left(H\right)\right], \quad \textbf{(6)}$$

where g represents a deterministic function and H is a random variable corresponding to the target activation. The total variation regularization term is defined in the previous sections. And the color uniformity regularization of facial image uses the following function in the reverse network:

$$L_{color}\left(H\right)=V_Y\left(g\left(H\right)\right)^T M_Y\left(H\right)V_Y\left(g\left(H\right)\right). \quad \textbf{(7)}$$

4 The Implementation of Image Style Transfer

A pre-trained VGG-19 is used as the feature extractor of our feed-forward net-work. The training of the inverse network [2] is referenced, Microsoft COCO (MSCOCO) [23] is employed, two training periods are done. The open sources in DeepLab v3+ [18] of encoder-decoder are used. We use the Torch7 framework to implement our method. All results in this paper are generated between 3 and 8 seconds on NVIDIA GeForce GTX 1060 6GB GPU.

Target layer. The effect of different layers of VGG-19 for our method is shown in Figure 6. The deeper layer of VGG-19 is selected, the more obvious is the texture of the style transfer image. A better style transfer result is in layer Relu3_1, the original image is preserved and texture of style image is added. More image style transfer experiments are performed in layer Relu 3_1. According to the naming convention of VGG-19, "Relu X_1" refers to the first ReLU layer after the (X-1)$_{th}$ maxpooling layer.

Style-matching tuning. Image style is directly affected by the patch size and patch stride of convolution in CNN. More texture of styled image is preserved whereas more special structures of the content image are lost with big patch size and patch stride (Figure 7 and Figure 8). We therefore set patch size as 3×3 and patch stride as 1.

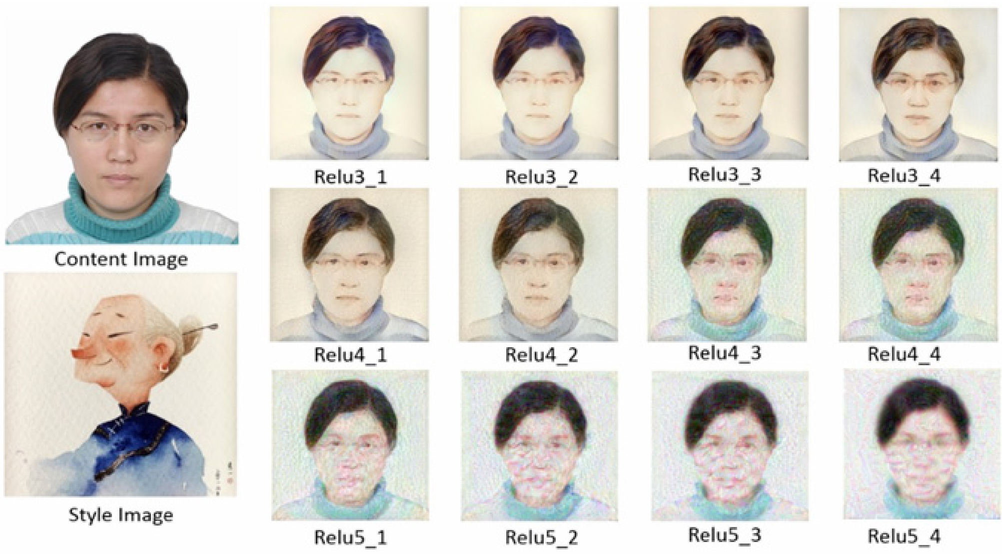

Figure 6. The effect of arbitrary style transfer network results in different layers of VGG-19

Figure 7. The effect of arbitrary style transfer network results in different patch size

Figure 8. The effect of arbitrary style transfer network results in different patch stride

Computation time. The time needed for style transfer with related work is shown in Table 1. Compared with parameter style transfer [16-17], only one layer of pre-training VGG-19 network is used in our study with less time in every iteration and fast speed. Our style transfer effect is slightly better (Figure 11) than [2], but it takes more time. Our total loss function is added color field transfer of visual model, and the style transfer is performed with semantic segmentation result and repeated restriction similarity constraints.

5 Experimental Results of Image Style Transfer

Experimental results support the effectiveness of our method. Parameters are adjusted as described above. Patch size is 3×3, patch stride is 1, $\alpha=10^{-6}$ and $\beta=10^{-2}$.

The effectiveness of style transfer of facial image. Our purpose is to obtain more artistic style transfer images for facial images via aesthetic QR code. The style transfer should be both artistic and realistic, and a high decoding rate of style transfer QR code should be achieved. The ideal style transfer results are achieved for different facial images in different style images in our network (Figure 9). Details of facial images are well preserved by the new network. Compared with [2] and [20], some results are shown in Figure 10. Our network obviously improves the effect of facial image style transfer.

Figure 9. The results for different facial images in different style images using our network

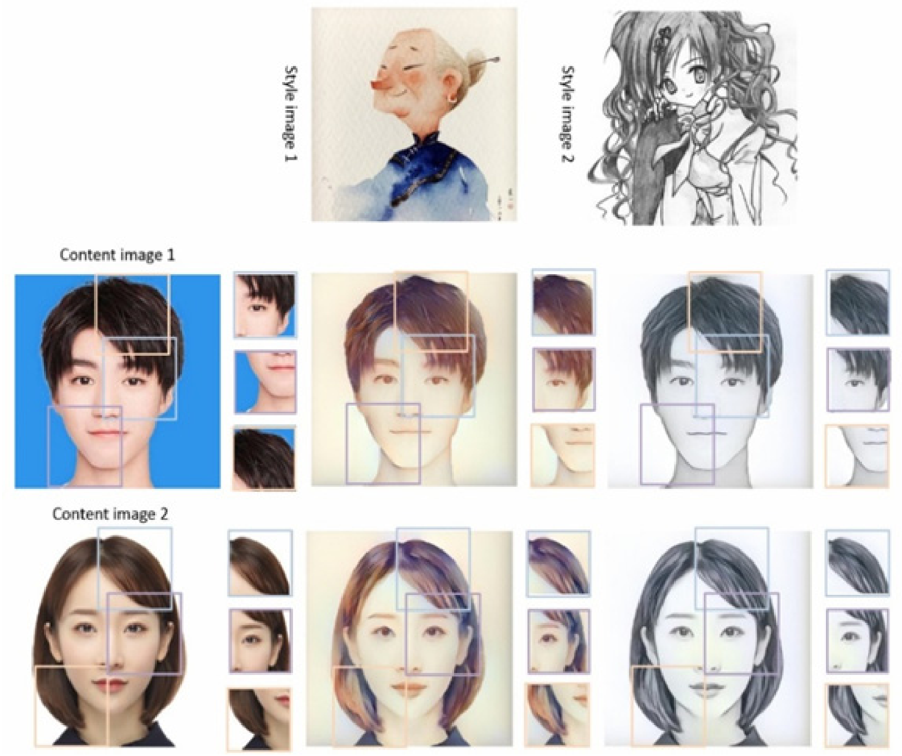

Figure 10. Details of facial images (such as hair, facial features, and outlines) using our method

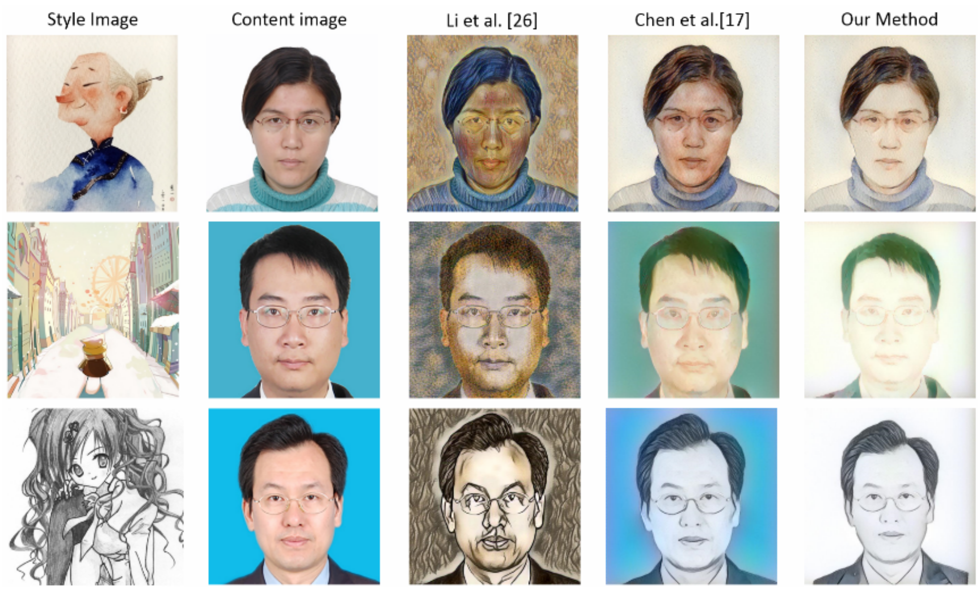

Figure 11. The results generated by different style transfer method

Adaptability to multi-style transfer. We propose a feed-forward network for arbitrary style transfer for facial image in order to obtain high fidelity facial image contents and artistic facial style. We need only one network, which supports arbitrary style transfer of facial image contents and avoids different networks for style image training in previous studies.

Table 1. Different methods stylize 555*555 image average computation time on NVIDIA GeForce GTX 1060 6GB GPU

Method	N. Iters.	Time/Iter. (s)	Total (s)
Gatys et al. [17]	800	0.208	116.4
Chen and Mark [2]	1	2.132	2.13
Li and Wand [16]	200	0.988	197.6
Our method	1	4.256	4.26

6 Aesthetic QR Code Method Based on Arbitrary Style Transfer Network of Facial Image

6.1 RS Code Encoding Mechanism

QR code uses RS code's error correction. RS code has the features of XOR operation in that a new RS code is obtained with XOR operation of two different RS codes [8]. In [9], we proposed a new algorithm for aesthetic QR code by using the XOR characteristics of RS code with PBVM and RBVM to reduce as much as possible the visual differences between the aesthetic QR code and the background image.

6.2 Aesthetic QR Code Generation

Combining [9], we propose a new aesthetic QR algorithm with the arbitrary style transfer network of facial image in Figure 12. The details include:

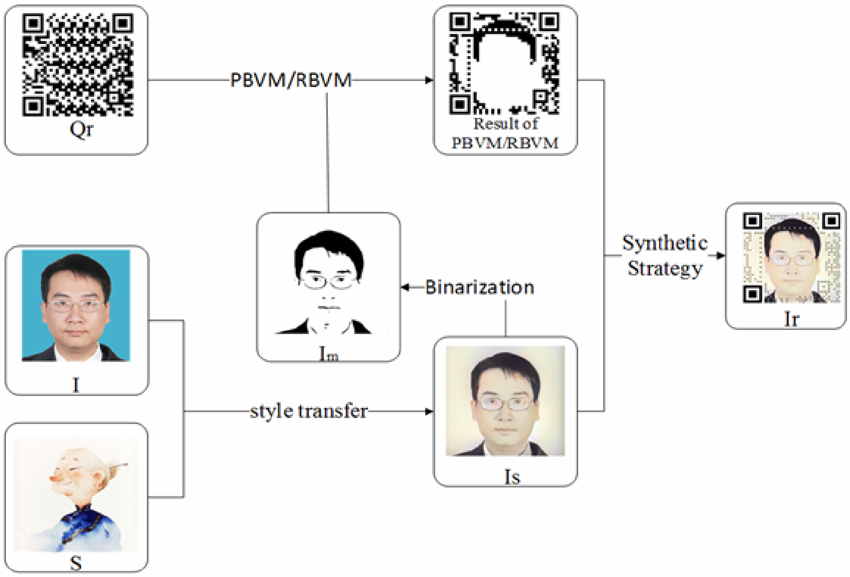

Figure 12. Flowchart of an aesthetic QR algorithm with arbitrary style transfer network

(1) Generate the background image using our style transfer network.

(2) The XOR characteristics of RS code with PBVM and RBVM is applied to beautify the QR code, and reduce the visual differences between the aesthetic QR code and background image as much as possible.

(3) Embed QR code into the background image.

6.3 Experimental Results

Visual effect. A collection of 100 images with different styles serves as a dataset in our experiment. All images in the dataset are used as background images to generate aesthetic QR codes. The selected results generated by our method are shown in Figure 13. The QR version is 5 and the error correction level is L. QR code with stylized image background has more artistic and diverse visual effect. No-tice that the additional points that need to be added to the face area are significantly less in the style transfer QR code.

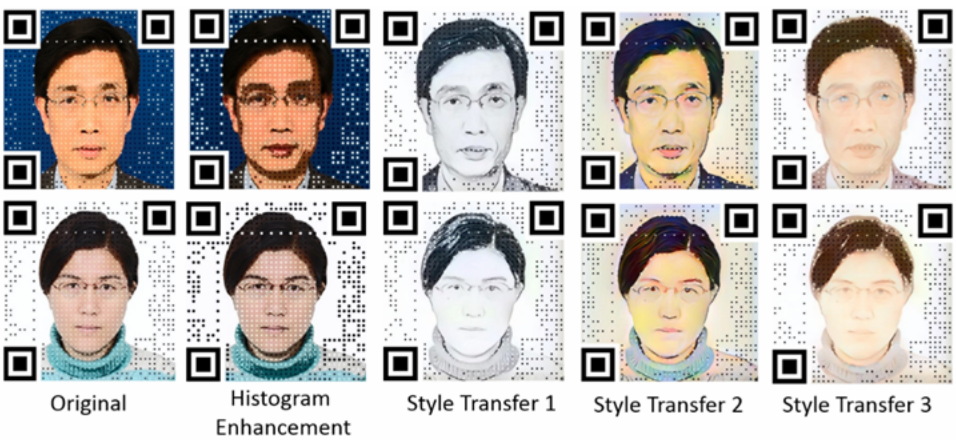

Figure 13. The generation of aesthetic QR codes with different background image

Scanning robustness. Scanning robustness of the QR codes was evaluated from various aspects, such as scanning angle variation, scale variation, brightness variation, and coverage. A collection of 100 images with different styles serves as a dataset in our experiment. And the same device (iPhone 8) is used for scanning and decoding tests. It is evaluated whether the decoding of the QR code is successful or not. The results are shown in Figure 14.

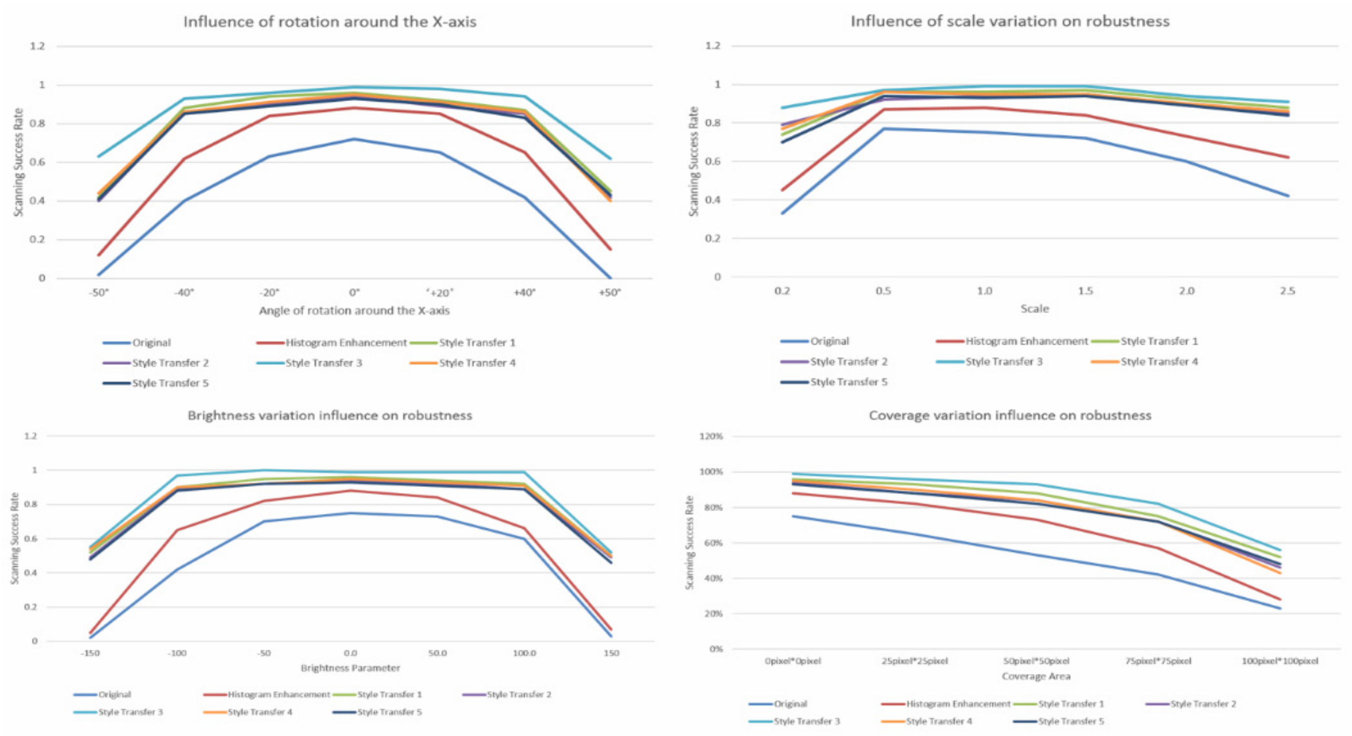

Figure 14. The results of scanning robustness

Experimental results show that our method not only ensures code usability, but also has the effect of QR code beautification.

7 Conclusion

In this study, we propose an arbitrary style transfer network of facial image based on feed-forward network in order to obtain high fidelity images for aesthetic QR codes. The network is combined with visual model and semantic segmentation. Global brightness balance is used, good visual effect is obtained, and the decoding rate of aesthetic QR code is guaranteed. However, for a small portion of the facial images, it is hard to control color distribution. We will study this further in the future.

Acknowledgments

This work was mainly supported by National Natural Science Foundation of China (No. 61370218, No. GG19F020033), Public Welfare Technology and Industry Project of Zhejiang Provincial Science Technology Department (No. 2016C31081, No. LGG18F020013, No. LGG19F020016).

References

[1] M. Xu, Q. Li, J. Niu, X. Liu, W. Xu, P. Lv, B. Zhou, ART-UP: A Novel Method for Generating Scanning-robust Aesthetic QR Codes, *Journal of Latex Class Files*, Vol. 14, No. 8, pp. 1-14, August, 2018.

[2] T. Q. Chen, M. Schmidt, Fast Patch-based Style Transfer of Arbitrary Style, *Constructive Machine Learning Workshop (NIPS 2016)*, Barcelona, Spain, 2016, pp. 1-5.

[3] S. Ono, K. Morinaga, S. Nakayama, Two-dimensional Barcode Decoration Based on Real-coded Genetic Algorithm, *2008 IEEE Congress on Evolutionary Computation (IEEE CEC)*, Hong Kong, China, 2008, pp. 1068-1073.

[4] D. Samretwit, T. Wakahara, Measurement of Reading Characteristics of Multiplexed Image in QR Code, *2011 Third International Conference on Intelligent Networking and Collaborative Systems (2011 INCOS)*, Fukuoka, Japan, 2011, pp. 552-557.

[5] L. Li, J. Qiu, J. Lu, C. C. Chang, An Aesthetic QR Code Solution Based on Error Correction Mechanism, *Journal of Systems and Software*, Vol. 116, No. C, pp. 85-94, June, 2016.

[6] K. Fujita, M. Kuribayashi, M. Morii, Expansion of Image Displayable Area in Design QR Code and Its Applications, *Forum on Information Technology*, Vol. 10, No. 4, pp. 517-520, September, 2011.

[7] S. S. Lin, M. C. Hu, C. H. Lee, T. Y. Lee, Efficient QR Code Beautification with High Quality Visual Content, *IEEE Transactions on Multimedia*, Vol. 17, No. 9, pp. 1515-1524, September, 2015.

[8] R. Cox, *QArt Codes*, https://research.swtch.com/qart.

[9] L. Li, B. Wang, J. Lu, S. Zhang, C. C. Chang, A New Aesthetic QR Code Algorithm Based on Salient Region Detection and SPBVM, *Journal of Internet Technology*, Vol. 20, No. 3, pp. 935-946, May, 2019.

[10] J. S. De Bonet, Multiresolution Sampling Procedure for Analysis and Synthesis of Texture Images, *the 24th Annual Conference on Computer Graphics and interactive techniques*, Los Angeles, CA, USA, 1997, pp. 361-368.

[11] D. J. Heeger, J. R. Bergen, Pyramid-based texture analysis/synthesis, *the 22nd Annual Conference on Computer Graphics and Interactive Techniques*. Los Angeles, CA, USA, 1995, pp. 229-238.

[12] J. Portilla, E. P. Simoncelli, A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients, *International Journal of Computer Vision*, Vol. 40, No. 1, pp. 49-70, October, 2000.

[13] N. Ashikhmin, Fast Texture Transfer, *IEEE Computer Graphics and Applications*, Vol. 23, No. 4, pp. 38-43, July, 2003.

[14] A. A. Efros, W. T. Freeman, Image quilting for texture Synthesis and Transfer, *the 28th Annual Conference on Computer Graphics and Interactive Techniques*, Los Angeles, California, USA, 2001, pp. 341-346.

[15] V. Kwatra, A. Schödl, I. Essa, G. Turk, A. Bobick, Graphcut Textures: Image and Video Synthesis Using Graph Cuts, *ACM Transactions on Graphics (ToG)*, Vol. 22, No. 3, pp. 277-286, July, 2003.

[16] C. Li, M. Wand, Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis, *the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas, NV, USA, 2016, pp. 2479-2486.

[17] L. A. Gatys, A. S. Ecker, M. Bethge, *A Neural Algorithm of Artistic Style*, https://arxiv.org/abs/1508.06576.

[18] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation, *the European Conference on Computer Vision (ECCV)*, Munich, Germany, 2018, pp. 833-851.

[19] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks For Large-Scale Image Recognition, *3rd International Conference on Learning Representations (ICLR 2015)*, San Diego, CA, USA, 2015, pp. 1-14.

[20] S. Gu, C. Chen, J. Liao, L. Yuan, Arbitrary Style Transfer with Deep Feature Reshuffle, *the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, USA, 2018, pp. 8222-8231.

[21] F. Luan, S. Paris, E. Shechtman, K. Bala, Deep Photo Style Transfer, *the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, Honolulu, HI, USA, 2017, pp. 6997-7005.

[22] J. Johnson, A. Alahi, F.-F. Li, Perceptual Losses for Real-time Style Transfer and Super-resolution, *European Conference on Computer Vision (ECCV 2016)*, Amsterdam, Netherlands, 2016, pp. 694-711.

[23] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, C. L. Zitnick, Microsoft Coco: Common Objects in Context, *European Conference on Computer Vision (ECCV 2014)*, Zurich, Switzerland, 2014, pp. 740-755.

Biographies

Shanqing Zhang received his B.S. and M.S. degrees in Mathematics department in 1992 and 1995, respectively, from Lanzhou University. And he obtained his Ph.D. from Computer Science School of East China Normal University in 2004. He has worked at Hangzhou Dianzi University since 2004. From 2006, he is an associate Professor of Hangzhou Dianzi University. His research interests are focus on image processing and computer vision.

Shengqi Su received the B.S. degree in Network Engineering from Jiangxi Normal University, Nanchang, China, in 2016. She is currently pursuing the M.S. degree in computer science and technology in Hangzhou Dianzi University. Her research interests include deep learning, image processing, and steganography.

Li Li received the B.S. and M.S. degrees in Mathematics from Dalian University of Technology, Dalian, China, in 1994 and 1997 respectively, and received Ph.D. degree in computer science department from Zhejiang University in 2004. She has been working in School of Computer Science and Technology in Hangzhou Dianzi University since 2004, where she is an Professor. Her main research interests are broadly in data hiding, image/video/3D mesh watermarking, QR code, and image/video processing.

Jianfeng Lu received his B.S. and M.S. degrees from the department of Mechanical Engineering in 1998 and 2001 respectively, and his Ph.D. degree in computer science department from the Zhejiang University in 2005. He works in the department of computer science in Hangzhou Dianzi University from 2005. His main research interests are in the areas of digital video and image processing and scientific visualization.

Ching-Chun Chang is a Ph.D. student under the supervision of Prof. Chang-Tsun Li in the Department of Computer Science, University of Warwick, UK. He received the BBA degree in information management from National Central University, Taiwan, in 2015. He was granted the Marie-Curie fellowship in 2017. He engaged in a short-term scientific mission supported by European cooperation in science and technology actions at the Faculty of Computer Science, Otto von Guericke University Magdeburg, Germany, in 2016. He participated in a research and innovation staexchange scheme supported by Marie Skłodowska-Curie actions at the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, USA, during 2017. He was a visiting scholar at the School of Computer and Mathematics, Charles Sturt University, Australia in 2018 and at the School of Information Technology, Deakin University, Australia in 2019. His research interests include digital watermarking, steganography, secret sharing, applied cryptography, digital forensics, multimedia security, image processing, machine learning, and computational linguistics.