An Interactive Neural Network-Based System for Confined Stylization of Product Design

Man-Kit, Tang

Department of Computing

The Hong Kong Polytechnic University, Hong Kong SAR

man-kit-jacky.tang@connect.polyu.hk

Abstract. Deep learning technology has significantly improved image style transfer. The present techniques, however, do not explore much on enabling users to control the confined transfer region of an image, such as separating particular clothing from a figure. To accomplish the goal, this research suggests an interactive image stylization technique. The Grab Cut Algorithm was employed to extract additional image from the content image, the product designer can simply drag a rectangle around the desired product image. Then, a distance loss function was introduced to preserve the shape of the product during stylization. A GUI application that implements the suggested stylization strategy was created. The experimental result shows that the original clothing shape can be well reserved and it demonstrated that our suggested approach is feasible to design initial product images.

Keywords: Image Style Transfer, Fashion Design, Convolutional Neural Network, Product Visualization, Image Segmentation

1. Introduction

Image Style Transfer is a computer vision technique to recompose the content of an image in the style of another. The problem of image style transfer via deep neural networks has received numerous contributions due to the recent rapid growth of machine learning. In 2015, Gatys et al.
[1] introduced a revolutionary method to transfer an artistic visual style of a style image into a photorealistic image using a Convolutional Neural Network (CNN).The fundamental concept is to use the VGG convolutional network to segregate and reconstruct the content and style images.(Fig. 1).[1]'s work has served as a basis to support the succeeding research and development of image style transfer.

Fig. 1. Example of Image Style Transfer [1]

Apart from Image Stylization, Convolutional neural networks have recently been effective in a number of Computer vision applications, including texture generation, image segmentation, and object recognition. In the area of fashion, this has supported numerous research projects and innovations, including clothing classification, clothing parsing and recommendation. This paper focuses on the self-control of clothing style in confined position during the product design phases. According to our approach, the fashion designer can select a photorealistic clothing image as content image, and find an art picture that they appreciate as a style image. The stylization algorithm is then adopted to generate another unique clothing design, which combines the style of the style image and the clothing shape of content image, while the shape of the original clothes was preserved.

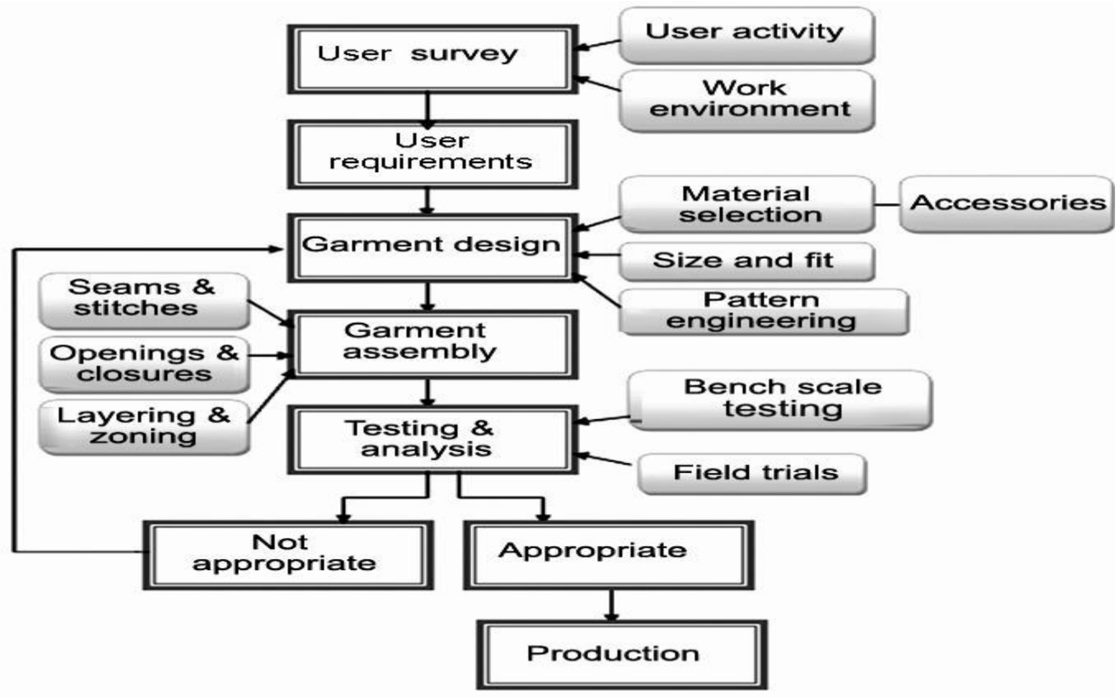


Fig. 2. Flow Chart of Garment Design Process[2]

Our developed GUI application will be served as a fashion visualization tool for the professional designer during the initial garment design phase. While there are already fashion visualization tools existing in the market like Digital Fashion Pro and CLO3D, the major difference between them and our application is obvious: our application can stylize the specific region of the clothes with arbitrary style image provided by designer while the common visualization tools can only provide their limited choices of color and pattern for clothes stylization.

The stylized clothes which have another color, texture or pattern will be visualized by our application. As shown in Fig. 2., the fashion designer can determine whether it is feasible to proceed to the next phases by considering different factors, like the production cost of the stylized clothe, clothes's material selection, customer survey and requirements. If it is not ideal to produce such a stylized cloth, the designer can give up or find another style image to produce a more appropriate product image. Otherwise, the designer may proceed to the next design phase or production phase.

2. Related Work

Image style transfer had been studied by Computer Vision researchers. At the beginning, as image style transfer can be considered as a process of image texture synthesis, [3] had proposed a texture modelling methods to model the texture by using the image statistics from the texture's sample and the summary statistical attribute. In 2015, the first neural algorithm of artistic style transfer, proposed by Gatys et al. [1], can separate and recombine the content and style of images. They created a model based on convolutional neural networks(CNN), extracting and storing visual features using a pretrained VGG network. When the new image matches both the style features of the style image and the content features of the content image, the total loss, which is a linear combination of the content and the style loss, is updated. [4] significantly increased the image style transfer speed by including perceptual loss functions for training feed-forward networks. Although the aforementioned methods produce remarkable style transfer results, but it will cause poor content detail retention during style transfer. In contrast to the earlier research studies, this work aims to apply neural style transfer to the fashion industry, giving the fashion designer the ability to create their own distinctive product in accordance with their preference. If the normal style transfer is used for fashionable product image, the whole generated image will be stylized. Instead of a simple image without any person or background, it is expected that our proposed method can process content image with complicated background. As a result, this study focuses on how users can control the confined transfer position of a content image while matching the style of the style image and the shape of the product inside the content image.

3. Proposed Work

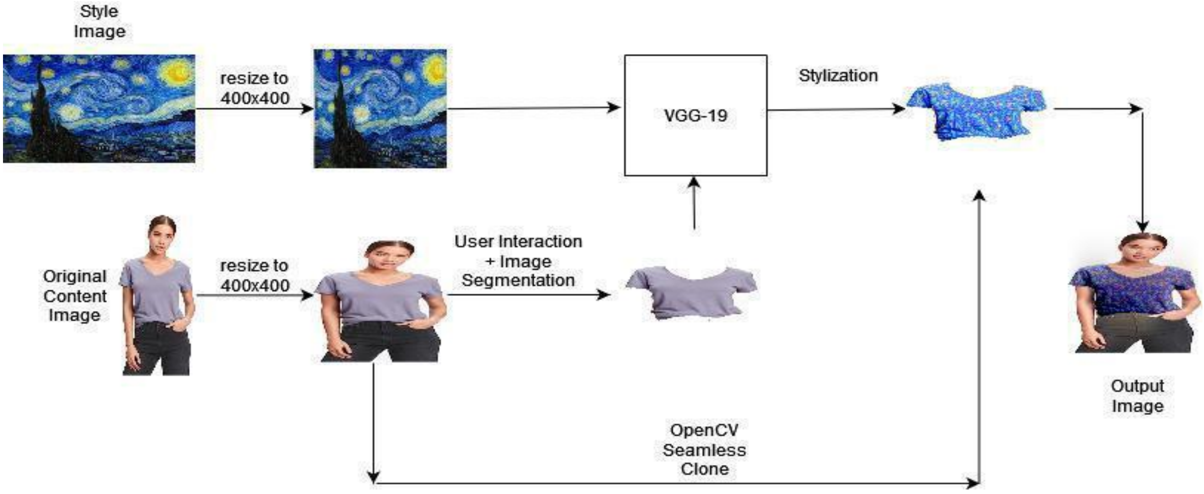
(Fig. 3) shows the overall process of image style transfer method as proposed in this paper.The proposed method was implemented by Python 3 and a GUI application[5] was primarily written. The following operations will be performed by the application:

(1) The dimension of the selected content image and style image will be resized into 400x400 to fit the layout of the GUI application.

(2) The user will be required to drag a rectangle around the desired product of the content image to obtain an extracted image. This can be achieved by using GrabCut Algorithm[6].

(3) The features of the segmented image and style image will be

Fig. 3. The Proposed Framework for Localized Image Stylization



extracted and stored using a pretrained VGG-19 network. Apart from the original content loss and style loss as proposed by [1], the additional distance transform loss [7] had also been introduced in this model for the calculation of total loss during neural style transfer. This will help to stylize the only pre-segmented region.

(4) The stylized segmented image will be cloned[8] into the resized content image to generate an output image with non-altering background.

3.1 GrabCut Algorithm

Based on the content image's color, grayscale and texture, it can be partitioned into multiple pieces called segments by Image Segmentation Algorithm. After segmentation, the features of the image exhibit disparities across diverse places yet show similarities in the same areas(Fig. 4). This study uses the GrabCut algorithm [6] to handle segmentation in order to give users the ability to precisely control the transfer position of an image.

Fig. 4. Example of Grab Cut Algorithm [5]

3.2 Stylization for Confined Region

The fundamental idea behind neural style transfer is to take the content and style information from the input image and combine them using a pretrained convolutional neural network (CNN).The input images used in this study include a content image with a complicated background, an extracted image that GrabCut Algorithm extracts from the content image, and a style image.

The VGGNet[1] generates a feature map from the filter responses to each layer for a given input image.An input image's content can be viewed as being represented by feature maps on particular layers. A feature space that is made to capture texture information is utilized to extract the style representation of an input image. It is made up of feature correlations that the Gram matrix provides in several layers.Given is the Gram matrix ,

$$G_{ij}^{l} = \sum_{k} F_{ik}^{l} F_{jk}^{l} \qquad (1)$$

where F_{ik}^{l} and F_{jk}^{l} refer to feature maps i and j in layer l.

For stylization, a content image $\vec{P}$ first inputs to the VGGNet and its feature maps in selected layer are stored as the content representation P^{l} on l^{th} layer. Next, a style image $\vec{a}$ passes through the network. The sum of Gram matrices on every layer are computed and stored as style representation A^{L} of a style image.Then, the image to be generated , which is initialized as the content image, passes through the network. Using its feature maps, the content representation F^{l} and the style representation G^{L} of the generated image are computed on same layers as the respective representations.With content and style representations, loss functions used for generating images can be calculated. The content loss Lcontent is calculated as shown in Eq. (2)

$$L_{\text{content}}(\vec{p}, \vec{x}, \vec{l}) = \frac{1}{2} \sum_{ij} (F_{ij}^{l} - P_{ij}^{l})^{2} \qquad (2)$$

Style loss L_{style} can be calculated as shown in Eq. (3)

$$L_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^{L} w_l E_l, \qquad (3)$$

where

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{ij} (G_{ij}^l - A_{ij}^l)^2 \qquad (4)$$

w_l are weighting factors for the contribution of layer l to style loss, N_l is the number of filters in layer l, M_l is the dimensions of layer l . To stylize the confined region of the content image, [7] proposed a new loss function using distance transform of the input images. The distance transform assigns a value—the distance to the closest pixel that is silhouette—to each pixel in a binary image. The distance transform image D has same dimensions as original image, but its pixel values are the values of distance transform. For every pixel d_{ij} of the distance transform image D, emphasis with power n would look like,

$$d_{ij} = \begin{cases} 0 & \text{if inside of a silhouette} \\ d_{ij}^n & \text{otherwise} \end{cases} \qquad (5)$$

The distance transform loss $L_{distance}$ would be

$$L_{\text{distance}} = \frac{1}{2} (\vec{p} \circ D_{\text{content}}^n - \vec{x} \circ D_{\text{content}}^n)^2 \qquad (6)$$

By simply adding $L_{distance}$ to the total loss with weighting factor γ, we can obtain the following total loss function for model optimization.:

$$L_{total} = \alpha L_{content} + \beta L_{style} + \gamma L_{distance} \qquad (7)$$

4. Experimental Results

(a) (b) (c) (d) (e)

Fig. 5. :The results with zebra pattern and watercolour paintings as style images. (a) Style Image. (b) Content Image. (c) Extracted Image. (d)The method proposed by Gatys et al. [1] (e) Our Method

In Figure 5, we randomly picked four pictures from Internet as content images and style images. The shirts in Figure 5(c) was extracted by our application[5] using the GrabCut Algorithm[6]. We compare our method with the stylization method proposed by Gatys et al. [1]. It was found that only our method keeps the background clean.In [1]'s method, the style from the style image is transferred to the whole content image rather than only clothes.

5. Conclusion and Future Work

In this work, an interactive stylization approach was specifically suggested for product images. New cloth designs can be created by merging appropriate style photos, making it simpler for designer to create the initial product image as they want. The designer would determine to further proceed to the next design phase or manufacturing phase if the stylized product image was considered to be reasonable for manufacturing. Additionally, a GUI application had been primarily developed[5].Although the stylization method adopted by this paper was based on [1] and seems initial, the experimental results shows that it was still effective on creating fashionable cloth image. As there is currently not much existing research on confined region stylization, the stylization method used in this paper may be naïve. Given the simplicity of our approach, future researches may examine more sophisticated stylization algorithms, and experiments to compare them to other algorithms.

Acknowledgements

The author thanks the general support of the volunteer Yik Ki Chan for his editing work on the video demonstration of the GUI application.

References

1. L. A. Gatys, A. S. Ecker & M. Bethge, "A Neural Algorithm of Artistic Style," arXiv:1508.06576v2 [cs.CV], Sep. 2015.
2. Gupta, Deepti. (2011). Design and engineering of functional clothing. Indian Journal of Fibre and Textile Research. 36. 327-335.
3. David J. Heeger and James R. Bergen. 1995. Pyramid-based texture analysis/synthesis. In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (SIGGRAPH '95). Association for Computing Machinery, New York, NY, USA, 229–238. https://doi.org/10.1145/218380.218446
4. Johnson, J., Alahi, A., Fei-Fei, L. (2016). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science(), vol 9906. Springer, Cham. https://doi.org/10.1007/978-3-319-46475-6_43
5. Man-Kit,T. GitHub. Retrieved February 11, 2023, from https://github.com/jackyt1010/An-Interactive-Neural-Network-Based-System-for-Contained-Stylization-of-Product-Design
6. Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. 2004. "GrabCut": interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. 23, 3 (August 2004), 309–314. https://doi.org/10.1145/1015706.1015720
7. Atarsaikhan, Gantugs & Iwana, Brian & Uchida, Seiichi. (2018). Contained Neural Style Transfer for Decorated Logo Generation. 317-322. 10.1109/DAS.2018.78.
8. Masayuki Tanaka, Ryo Kamio, and Masatoshi Okutomi. 2012. Seamless image cloning by a closed form solution of a modified Poisson problem. In SIGGRAPH Asia 2012 Posters (SA '12). Association for Computing Machinery, New York, NY, USA, Article 15, 1. https://doi.org/10.1145/2407156.2407173
9. BOYKOV, Y., AND JOLLY, M.-P. 2001. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *Proc. IEEE Int. Conf. on Computer Vision*, CD–ROM