Style Transfer for CNC Machine Input Data Preprocessing

E I Nikolaev\(^1\), V V Zaharov\(^2\), N I Zaharova\(^1\)

\(^1\)North-Caucasus Federal University, 2, Kulakov Ave., Stavropol, 355029, Russia
\(^2\)Stavropol State Agrarian University, 12, Zootekhnichesky lane, Stavropol, 355017, Russia

E-mail: notdeveloper@gmail.com

Abstract. Advances in deep neural networks have led to impressive results in recent years. The new technologies such as cross-domain adaptation, reinforcement learning and generative adversarial networks have shown a real promise for industrial and real-life applications. In this paper, the results of the experimental research on designing, training and implementation of the preprocessing algorithm for the computer numerical control machine input were presented. The algorithm of neural network transfer of artistic style has demonstrated wide possibilities in the field of generating graphic content. This paper demonstrates the possibility of using a generating neural network for the synthesis of stylized images that can be used as input images for a computer numerical control machine. Thus, the proposed algorithm is pre-processing the input image. The design feature of the laser engraver does not allow styling using an arbitrary style image, so dotted or linearized binary images are used as a style. The proposed preprocessing algorithm allows synthesizing binary images reproduced by a laser engraver. At the same time, image generation is performed in one forward pass of the generating neural network.

1. Introduction
Transferring the style from one image to another image is an interesting yet difficult problem. There have been many efforts to develop efficient methods for automatic style transfer [1-5]. Recently, Gatys et al. proposed a seminal work [6]: It transfers the style of artistic images to other images using Convolutional Neural Networks (CNN). Several follow-up works improved upon this innovative approaches [7-10]. Despite the fact that this work has drawn lots of attention, one of the main disadvantage of style representation: the practical use of this technology is negligible.

In this paper, a novel application of neural style transfer by cascading it as a preprocessing algorithm for computer numerical control (CNC) machines is proposed. There are different types of CNC machines: it’s possible to build a classification based on the principle of control, the working tool, the type of material being processed. In the framework of this work, a CNC laser engraver is considered, which performs processing of wood products by burning material with a laser beam. The task of adapting the input image of photographic quality to the capabilities of the CNC machine is being solved. Modern CNC machine control systems are based on the G-code command system: it is a specialized language for industrial machines that allow formulating commands for controlling the working tool (for example, movement and power of a laser module). With this approach, the process of obtaining the resulting image \( I_e \) from the original bitmap image \( I_c \) takes a long time. In the case of using vector images \( I_c \), the speed of creating a product increases significantly. To address this problem the preprocessing algorithm based on neural style transfer method is proposed. The input bitmap image \( I_c \) is preprocessed and intermediate image \( I_m \) is generated. The \( I_m \) will be used as an input image for the physical laser engraver. The proposed method of generation \( I_m = s(I_c, I_s) \) is based on “simplify” the initial image \( I_c \). The transformation \( s = s(\cdot, I_s) \) is a stylization algorithm with a style \( I_s \). Among graphical attribute transfer algorithms, the highest efficiency is demonstrated by the neural network art style transfer methods [6]. But in the framework of this study the style image \( I_s \) effects on the efficiency of the fabrication process of the image \( I_e \) by CNC machine. As a result, the target image \( I_e \) is synthesized by a
CNC machine using an intermediate image $I_m$: $I_t = G(I_m)$, where $G(\cdot)$ is the transformation of a digital image into a physical product using the CNC machine working tool (laser module).

2. Related works
Gatys et al. observe that deep convolutional neural network is capable of extracting semantic image content from an arbitrary photograph and some appearance information from the well-known artwork [6]. Gatys et al. themselves [11] propose several slight modifications to improve their previous algorithm [6]. Although the original stylization method based on online image optimization is able to yield impressive stylized images, there are some limitations. The most concerned limitation is the efficiency issue. Another category of the stylization methods addresses the speed and computational cost issue by exploiting image generation based on offline optimization to reconstruct the stylized result, i.e., a feed-forward network is optimized over a large set of images for one or more style images. Such methods are proposed by Johnson et al. [12] and Ulyanov et al. [13] respectively. These two methods share a similar idea, which is to pre-train a feed-forward style-specific network and produce a stylized result with a single forward pass at the testing stage. They only differ in the network architecture, for which Johnson et al.’s design roughly follows the network proposed by Radford et al. [14] but with residual blocks as well as fractionally strided convolutions, and Ulyanov et al. use a multi-scale architecture as the generator network. The fast neural network transfer strategy is the most suitable method for the task addressed in this research.

In the framework of this experimental research, the problem of the intelligent control system design is also addressed. The cyber-physical systems, intelligent manufacturing systems and their role in modern industry are described in [15]. The structure of the modern intelligent manufacturing systems and the terminology of the cyber-physical systems are described in [16].

3. Input image preprocessing model
3.1. Overview
The scheme of the CNC laser engraver is depicted in Fig. 1a. A typical mode of operation of the machine is the sequential burning of points in the material. It’s possible to achieve high efficiency of the fabrication process by using a vector image as an input, but in case of using a bitmap image, the time of the burning process increases significantly. The main idea of this experimental study is to create a model for generating an intermediate image $I_m$ for the initial image $I_c$ to obtain a high efficiency of the fabrication process and preserve a similarity of the $I_c$ and $I_t$. The model takes bitmap image $I_c$ as input and produces plausible image $I_m$ via deep generative model. The $I_m$ is used as an input for the physical CNC machine tool. The construction of this CNC machine, in general, contains two axes X and Y and the laser module which moves along them. The input image is shown in Fig. 1b. CNC laser engraver generates the picture by burning material by the laser beam. Therefore, it synthesizes a binary bitmap, but sometimes it’s possible to implement a grayscale bitmap drawing by controlling laser module power.

In this research, the approach with binary input images is considered. There are two different modes of the laser engraver functioning. The first method is to control the displacement and power of the laser module using a binary input image constructed from randomly located points. Such type of input is depicted in Fig. 1c and Fig. 1f. The second one is based on using a binary image constructed from horizontal lines (Fig. 1d, Fig. 1g). This mode provides high speed of the laser engraver, as the CNC control system manipulates with line segments in the manufacturing process.
3.2. Online Neural Style Transfer

The original neural style transfer algorithm [6] tries to seek a stylized image $I_t$ that minimizes the objective:

$$ I_t = \arg \min_{I_t} \mathcal{L}_{total}; $$

$$ \mathcal{L}_{total} = \alpha \cdot \mathcal{L}_c(I_t, I_c) + \beta \cdot \mathcal{L}_s(I_t, I_s); $$

(1)

where $\mathcal{L}_c(I_t, I_c)$ compares the content representation of a given content image $I_c$ to that of the stylized image $I_t$; $\mathcal{L}_s(I_t, I_s)$ compares the Gram-based style representation derived from a style image $I_s$ to that of the stylized image $I_t$; $\alpha$ and $\beta$ are used to balance the content component and style component in the stylized result.

The content loss $\mathcal{L}_c$, defined in a form (2), represents the squared Euclidean distance between the feature representations $\mathcal{F}^l$ of the content image $I_c$ in layer $l$ and that of the stylized image $I_t$.

$$ \mathcal{L}_c(I_t, I_c) = \sum_{l \in \{l_c\}} \| \mathcal{F}^l(I_c) - \mathcal{F}^l(I_t) \|, $$

(2)

where $\{l_c\}$ denotes the set of layers which are used for the content style representation. The example of the initial image and the subset of activation maps from VGG19 [17] are depicted in Fig. 2.

The style loss $\mathcal{L}_s$ exploits Gram-based visual texture modelling technique to represent the style. Therefore, the style loss is defined by the squared Euclidean distance between the Gram-based style representations of $I_s$ and $I_t$ (3).

$$ \mathcal{L}_s(I_t, I_s) = \sum_{l \in \{l_s\}} \| \mathcal{G}(\mathcal{F}^l(I_s)) - \mathcal{G}(\mathcal{F}^l(I_t)) \|, $$

(3)

where $\mathcal{G}$ is the Gram matrix to encode the second order statistics of the set of filters; $\{l_s\}$ represents the set of layers for calculating the style loss.

The Gram matrix $\mathcal{G}$ is the inner product between the vectorised feature maps $i$ and $j$ in layer $l$:

$$ \mathcal{G}_{ij}^l = \sum_k \mathcal{F}_{ik}^l \cdot \mathcal{F}_{kj}^l, $$

(4)

where $\mathcal{F}_{ik}^l$ is the activation of the $i^{th}$ filter at position $k$ in layer $l$. 

![Figure 1. CNC machine basics: a) laser engraver scheme; b) input image; c) dotted image; d) linearized image; e) zoomed input fragment; f) zoomed dotted fragment; g) zoomed linearized fragment.](image-url)
Figure 2. Neural network image representation: initial image $I_c$ and a subset of the activation maps.

The choice of $\{l_c\}$ and $\{l_s\}$ empirically follows the principle that the usage of lower layer tends to retain low-level features (e.g., colours, simple strokes), while the usage of a higher layer generally preserves more high-level semantic content information. $L_s$ is usually computed with lower layers and $L_c$ is computed with higher layers. In framework of the current experimental research are used: $\{l_s\} = \{relu_{1,2}, relu_{2,2}, relu_{3,2}, relu_{4,2}, relu_{5,2}\}$ and $\{l_c\} = \{relu_{4,2}\}$, and VGG19 is used as a feature extractor network. To control a stylization quality the parameters $\alpha$ and $\beta$ (1) are used. The results of stylization with different stylization rate ($\alpha/\beta$) are shown in Fig. 3. The content image $I_c$ is depicted in Fig. 2. In the considered algorithm in the formulation [6] there are a large number of parameters and hyperparameters, which can be tuned to improve the quality of stylization. The crucial problem arising in the process of image stylization according to the algorithm [6] is a performance. Optimization of the objective function (1) is a long iterative process based on gradient-based methods. In order to increase the efficiency of the neural style transfer as a preprocessing stage for the CNC machine input, in this study, the generative deep model is trained.

3.3. Offline Neural Style Transfer
In the framework of this research, the modification of the networks proposed in [12, 13] is used. The architecture of the generative neural network is depicted in Fig. 4a. It contains a sequence of convolutional layers $conv$ $n \times n$ $[in, out, num]$, where $n \times n$ is a filter size, $in$ and $out$ are the input and output image sizes, and $num$ is the number of filters; residual blocks (Fig 4b); transpose convolution layers. Convolution layers contain instance normalization blocks [18].

The generative model is trained by using COCO [19] dataset, which contains 118,287 samples. The crucial task is that it is important to select the style image $I_s$ which provides intermediate image $I_m = s(I_c, I_s)$ with the most suitable structure. In the framework of this experimental research the $I_m$ is not just a stylized image, but also a map according which CNC machine tool will be moved.

The training process for the deep model depicted in Fig 4a is similar to the one in case of using online neural style transfer procedure, but rather than using per-pixel loss functions depending only on low-level pixel information, we train generative networks using perceptual loss functions that depend on high-level features from a pretrained loss network (VGG19). During training, perceptual losses measure image similarities more robustly than per-pixel losses, and at test-time, the transformation networks run in real-time. The scheme of the training pipeline is depicted in Fig. 4c. During online training, the L-BFGS-B method is used, but for generative model optimization in offline mode, Adam optimizer is used.
Figure 3. Neural Style Transfer with different stylization rate ($\alpha/\beta$) values: a) style image $I_s$; b) $\alpha/\beta = 10^{-2}$; c) $\alpha/\beta = 10^{-4}$; d) $\alpha/\beta = 10^{-7}$.

Figure 4. The Generative Network Architecture: a) neural network graph; b) residual block graph.

4. Experiments
The experiment with three different target styles was performed: a dotted style (Fig. 5a), linearized style (Fig. 5b) and vector style (Fig. 5b). The vector style aims to vectorize the input image.

Figure 5. Target styles for training generative models.
The result images are depicted in Fig. 6. The proposed algorithm of the photographic bitmap image preprocessing produces the binary image which is ready for drawing by a laser engraver. The quality of the synthesized image and the efficiency of the fabrication process depend on the initial image structure and style image structure.

![Initial image](image1)

![Dotted style](image2)

![Linearized style](image3)

![Vectorized style](image4)

**Figure 6.** Results of the preprocessing algorithm.

In Fig. 6 the vectorized style is shown for comparison purpose. The high degree of similarity between the synthesized image obtained using a vectorized style and the original image demonstrates the prospects of using a complex style for the pre-processing task. But the complexity of the reproduction of the resulting image by a laser engraver requires additional transformations, that is, lead to the need to re-apply the preprocessing algorithm.

5. Conclusion and future work

The presented experimental results can be applied not only to control the laser engraver, but also for any 2D CNC machine. The proposed preprocessing algorithm can be improved by reformulating the problem: the problem to be solved is related to the domain adaptation problem class. Solving the problem of domain adaptation can significantly improve the quality of preprocessing and, accordingly, increase the efficiency of the laser engraver. The neural network art style transfer used in this study is a special case of domain adaptation.

Another area of research is the development of the control system for generating commands for physical equipment based on reinforcement learning algorithms. A similar approach is used in [20]. The use of technologies of deep generating neural networks in conjunction with reinforcement learning methods will allow you to synthesize images using CNC machines in a human-like manner.

References

[1] Hertzmann A, Jacobs C E, Oliver N, Curless B and Salesin D H 2001 Image analogies. *In Proc. of the 28th Ann. Conf. on Computer Graphics and Interactive Tech. (ACM)* pp 327-340.

[2] Efros A A, Freeman W T 2001 Image quilting for texture synthesis and transfer. *Proc. of the 28th Ann. Conf. on Computer Graphics and Interactive Tech. (ACM)* pp 341-346.

[3] Efros A A, Leung T K 1999 Texture synthesis by non-parametric sampling. *Proc. of the Int. Conf. on Computer Vision* 2 pp 1033-1050.

[4] Shih Y, Paris S, Barnes C, Freeman W T and Durand F 2014 Style transfer for headshot portraits.
ACM Transactions on Graphics, 33(4):148.

[5] Kwatra V, Essa I, Bobick A and Kwatra N 2005 Texture optimization for example-based synthesis. ACM Transactions on Graphics, 24(3):795-802.

[6] Gatys L A, Ecker A S and Bethge M 2015 A neural algorithm of artistic style arXiv preprint arXiv:1508.06576.

[7] Li N, Wang J and Hou L 2017 Demystifying neural style transfer Proc. of the Twenty-Sixth Int. Joint Conf. on Artificial Intelligence IJCAI-17 pp 2230-2236.

[8] Li S, Xu X, Nie L and Chua T-S 2017 Laplacian-steered neural style transfer Proc. of the 2017 ACM on Multimedia Conf. (ACM) pp 1716-1724.

[9] Ruder M, Dosovitskiy A and Brox T 2016 Artistic style transfer for videos. In GCPR.

[10] Ledig C, Theis L, Huszar F and Caballero J 2016 Photo-realistic single image super-resolution using a generative adversarial network arXiv preprint arXiv:1609.04802.

[11] Gatys L A, Ecker A S, Bethge M, Hertzmann A, Shechtman E 2017 Controlling perceptual factors in neural style transfer Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, pp 3985-3993.

[12] Johnson J, Alahi A and Fei-Fei L 2016 Perceptual losses for real-time style transfer and super-resolution In European Conference on Computer Vision, pp 694-711.

[13] Ulyanov D, Lebedev V, Vedaldi A and Lempitsky V 2016 Exture networks: Feed-forward synthesis of textures and stylized images Int. Conf. on Machine Learning, pp 1349-1357.

[14] Radford A, Metz L and Chintala S 2015 Unsupervised representation learning with deep convolutional generative adversarial networks arXiv preprint arXiv:1511.06434.

[15] Chen J, Yang J, Zhou H and Xiang H 2015 CPS Modeling of CNC Machine Tool Work Processes Using an Instruction-Domain Based Approach Engineering (Beijing, China: Academy of Engineering) 1(2) pp 247-260.

[16] Lee J, Bagheri B, Kao H A 2015 A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems Manufacturing Letters 3:18-23.

[17] Simonyan K, Zisserman A 2014 Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.

[18] X. Huang and S. Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 1501-1510.

[19] Lin T, Maire M, Belongie S and Bourdev L D 2014 Microsoft COCO: Common Objects in Context. arXiv preprint arXiv:1405.0312.

[20] Ganin Y, Kulkarni T, Babuschkin I, Eslami S M and Vinyals O 2018 Synthesizing Programs for Images using Reinforced Adversarial Learning Proc. of the 35th Int. Conf. on Machine Learning (ICML) ed J Dy, A Krause (PMLR) 80 pp 1652-1661.