

Pipes of AI – Machine Learning Assisted 3D Modeling Design

Chuan Liu¹, Jiaqi Shen², Yue Ren³, and Hao Zheng²(⊠)

¹ Zhejiang University of Technology, Hangzhou, China ² University of Pennsylvania, Philadelphia, USA zhhao@design.upenn.edu ³ University College London, London, UK

Abstract. Style transfer is a design technique that is based on Artificial Intelligence and Machine Learning, which is an innovative way to generate new images with the intervention of style images. The output image will carry the characteristic of style image and maintain the content of the input image. However, the design technique is employed in generating 2D images, which has a limited range in practical use. Thus, the goal of the project is to utilize style transfer as a toolset for architectural design and find out the possibility for a 3D modeling design. To implement style transfer into the research, floor plans of different heights are selected from a given design boundary and set as the content images, while a framework of a truss structure is set as the style image. Transferred images are obtained after processing the style transfer neural network, then the geometric images are translated into floor plans for new structure design. After the selection of the tilt angle and the degree of density, vertical components that connecting two adjacent layers are generated to be the pillars of the structure. At this stage, 2D style transferred images are successfully transformed into 3D geometries, which can be applied to the architectural design processes. Generally speaking, style transfer is an intelligent design tool that provides architects with a variety of choices of idea-generating. It has the potential to inspire architects at an early stage of design with not only 2D but also 3D format.

Keywords: AI in design · Neural Networks · Style transfer

1 Principle of CNN

Convolutional Neural Networks (CNN) is a class of deep neural networks that are most commonly used to analyze visual images. Figure 1 shows the image that analyzes how an input image is processed by CNN. Firstly, the computer reads the image as pixels and represents it as a matrix, which will then be processed by the convolutional layer. This layer uses a set of learnable filters that convolve across the width and height of the input file and compute the dot product to give the activation map. Different filters that detect different features are convolved in the input file and output a set of activation maps that are passed to the next layer in the CNN. The pooling layer between the convolutional layers can be found in the CNN architecture. This layer substantially reduces the number of parameters and calculations in the network and controls overfitting by gradually reducing the size of the network [7].



Fig. 1. Introduction of CNN structure.

The next layer is the fully connected layer, in which the neurons are fully connected to all activation of the previous layer. Therefore, their activation can be calculated by matrix multiplication followed by offset. This is the final stage of the CNN network [2, 3]. In general, CNN uses relatively less preprocessing than other image classification algorithms. This means that the network learns manually designed filters in traditional algorithms. This independence from prior knowledge and human effort in feature design is a major advantage.

1.1 Principle and Applications of Style Transfer

CNN-based Style Transfer is a type of algorithm for processing digital images or video, using the look or visual style of another image. In the paper *A Neural Algorithm of Artistic Style*, Gatys introduces a Style Transfer Network to create high-quality artistic images [5]. Moreover, it has been successfully applied in key areas such as object and face generation. The style transfer process assumes an input image and a sample style image. The input image is fed through the CNN and the network activation is sampled at each convolutional layer of the VGG-19 architecture [6]. The content image is then obtained as a resulting output sample. The style image is then fed through the same CNN and the network activation is sampled. These activations are encoded as a matrix representation to represent the "style" of a given style image (Fig. 2). The goal of Style Transfer is to synthesize an output image that shows the content of a content image to which a style image style is applied.

To fully understand Style Transfer, Ostagram, an online style transfer tool, is used to generate images. A pair of images is selected as the content image and style image for the experiment, then switch the role of the image to another for another set of experiments. It's evident that the output image presents the geometry of its content image and the style of the style image (Fig. 3).

1.2 Project Goal

Style Transfer allows us to create new images with high perceived quality, combining the content of any photo with the look of many well-known artworks. Gatys [4] and



Fig. 2. Style deconstruction of the Style Transfer Neural Network.



Fig. 3. Example of style transfer.

Mordatch [8] mention the versatility of the Neural Network. For example, a map is generated in light of the features of the target image, such as image density, to constrain the texture synthesis process. Besides, image analogy is applied to transfer textures from already stylized images to target images.

Despite the remarkable results of Style Transfer, design techniques are only used to design images to be generated at the 2D level, such as graphic design [1]. In order to broaden its scope of use, it is necessary to design 3D levels in order to apply the generated results to architectural design. In order to achieve the project goal, 3D geometry is required as the result of the Style Transfer. With the completed 3D geometry, further designs like 3D structures or buildings can be used.

In order to put this idea into practice, it is necessary to answer questions about the nature of creativity and the criteria for evaluating creativity. Can style transfer create a novel sensibility, can we as humans perceive and understand it? In order to answer these questions, the project objectives presented here need to solve the problem not only from the aesthetic aspect - the idea that style transfer can produce fascinating 2D images - and from the point of view of emphasizing practical value: considering about the practicality of image-generated content and the usability of converting it to a 3D format.

2 2D Image Representation of 3D Volume

Since Style Transfer processes images on a 2D scale, the 3D model is needed to be converted into 2D images at first. An existing project is selected as the carrier of the design boundary. The project is a theater design, which is diverse in interior spaces, has a variety of floor plans of different heights (Fig. 4). By adjusting the position of section plan in Sketchup, a series of floor plans are gained according to the height increment of the building. Then, the ten-floor plans are converted into images that only carry the information of the outline, which is the boundary between the interior and exterior space of every floor plan (Fig. 5). Black and white colors are used to distinguish the interior and exterior space. At the stage, two sets of color-filling methods are applied as one set black to interior space and white to exterior space while one set the opposite. The methods will bring different results for output images since the color is an influential factor of Style Transfer. Thus, the most suitable result will be selected and be employed in further design.



Fig. 4. Origin model of the content images.



Fig. 5. The geometry of floor plans.

In general, by translating 3D volume into a 2D image, content for the input image is obtained as the basic element of the Style Transfer Neural Network. Later on, the style image will be imported into the network with the content images to start the generating process.

2.1 The Effect of Style Weight in Style Transfer

After the content images are obtained, a style image should be imported so that the generating phase can be activated. For style image, facade images and landscape images

are firstly employed to the Neural Style Transfer Neural Network. The result of the output images turns out not available for further design because of its indistinguishable geometry. To eliminate the chance of an unavailable output image, style images that contain a clearer outline and distinctive color contrast are applied. Thus, a framework of a truss structure is set as the style image (Fig. 6).



Fig. 6. Style image.

Before importing style image and content image into the network, several input parameters are needed to be confirmed. These parameters including "Style Weight", "Content Weight" and "tv Weight". There will be different effects presented by the output images by adjustments of every input parameter separately. With the awareness, a content image is imported into the neural network while adjusting a single parameter regularly to receive a series of output images. From the three sets of output images, it is easier to tell the difference between every set of output images with the influence of every parameter. At the same time, the regular pattern is easier to be observed of every set so that we can decide which input parameter is more suitable for further design.

After comparing each set of output images, we noticed that "Style Weight" is the input parameter that brings the most ideal generating result. Figure 7 shows the images that exported as the output images while adjusting the value of only style weight. From the images, the difference is obvious between every single image and the surprising found is the regular pattern of the set of images. With the increase of the value on style weight, the part of generated texture in the black area, which represents interior space, is more evident as well as taking up more percentage of the black area. Those phenomena are not shown in the other two sets of experiments with the adjustment of "tv Weight" and "Content Weight". Thus, "Style Weight" is selected as the input parameter for design.



Fig. 7. Output images generated by adjusting style weight.

With the conclusion above, firstly, "tv Weight" and "Content Weight" is given a specific value in the design. For "Style Weigh", a range from 500 to 5000 is given to the content images since the range shows the most ideal case for the contents of new generating images. For instance, the "Style Weight" value of the 1st-floor plan is 500, and

that of the 2nd-floor plan is 1000. With the set of rules, values are distributed for every content image so that the contents of output images are well-regulated and persuasive as well (Fig. 8).



Fig. 8. Transferred images generated by adjusting style weight.

2.2 Transformation of Image to Geometry

After the processing phase of Style Transfer, output images are obtained for further design. At the stage, the major concern is how to utilize the images in architectural design, since the 2D image is unavailable for geometry-oriented modeling software. Thus, 2D images need to be converted into geometry to fit it in architectural language. Rhinoceros, which is a geometry-based modeling software, is applied to accomplish the conversion.

To translate the image into geometry, a PDF version of the output images is imported to Rhino. After the import phase is completed, the geometry like polyline is used for geometric generation. There are several rules at the editing phase, the first is to delete the area that is useless for the coming design. The area in white color, which represents the exterior part of the floor plan, is defined as the nugatory space of the design.

After eliminating the white area in every image, there is another set of rules for further edition. Since the generated texture in the black area presents as intersected line segments from the top perspective, the polyline making tool is utilized to translate the texture into geometry. By connecting all the line segments together with polylines, a geometry version of the image is completed (Fig. 9). The rest of the output images will be edited according to the same set of rules until all the geometric layers are obtained (Fig. 10).



Fig. 9. Transferred image translation.



Fig. 10. Translation of all the geometric layers.

2.3 Algorithm Analysis of Geometry Generation Between Adjacent Layers

The 3D model will be accomplished after all the geometric layers are stacked together according to the original sequence of the floor plans. To make the model applicable for architectural design, pillars should be added between the layers so that the load condition of the structure is reasonable.

There are lots of possibilities for the geometry of the pillars, since the only rule a line need to obey is started from an intersection point and end on another one at the upper layer. However, if it is the only rule for the pillar's geometry, the design is likely unavailable for forming an architecture model. Realizing the problem of connecting uncertainty, Grasshopper, a plug-in based on Rhinoceros, is applied to explore the best way of connection from an algorithmic perspective.

After the script is completed, the information of two adjacent layers is imported into Grasshopper. Then, there are two input parameters, which decide the number and tilt angle of lines, needed to be modified for generating the vertical geometry. The input parameter which decides the line's number is being modified at first since a determined value of the parameter is a prerequisite for further adjustment like a tilt angle. A number slider, which domain is from 0–1, is employed to adjust the number of lines between the layers. In the experiment, the domain is divided into five groups of subdomains, which are 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1. Then the median is selected from every subdomain to be the value of the input parameter for each set of experiments.

After the number of geometry is fixed, a panel is set for the domain of the tilt angle. To determine which domain is the most suitable, another five sets of experiments are conducted according to the same rule of the former experiment. A domain from 0-10 is divided into five subdomains, which are 0-2, 2-4, 4-6, 6-8, and 8-10. Then the subdomains are processed by the same kind of rule. In general, the vertical geometry is determined based on the generated lines, which are baked according to the selected values of two input parameters. With the vertical and horizontal geometry, the structure will be further designed in a more aesthetic way.

3 Result of Section Plans

To translate the geometry into architectural language, piping, a design tool in Rhinoceros, is used to convert geometry into a 3D format. By adjusting the radius of both horizontal and vertical geometry, the optimal format is generated. Figure 11 shows the image of two section plans, one is the view from the right and the other one is from the front. The views are selected according to the structural complexity, which presents the distribution and relationship of vertical geometry between layers.



Fig. 11. Front facade (left) and right facade (right).

Moreover, the human scale is considered when designing the height between the adjacent layers so as to meet the requirement of use. The height of the 1st floor is 8 meters, which is two times higher than the rest of the floor heights which are 4 meters. Public spaces can be positioned at the area where has a more open vertical space with fewer pillars while private spaces the opposite.

3.1 Result of Perspective View

The vertical and horizontal pipes are presented in different textures so as to provide a more comfortable space experience for visitors. Vertical pipes are designed as pillars,

thus grey is used with a hint of transparency to mimic the format of pillars. Horizontal pipes are designed as floors since there is no given requirement for the texture, white is used to distinguish themselves with vertical pipes.

Figure 12 shows the image that depicts the relationship between the building and its surroundings. The original building is evolved into several variants, which are achieved by stretch, compression, and superimposition of the original format. This step helps to expedite the design process, which translates the original building into a series of new formats so as to fit in different circumstances of a city.



Fig. 12. Panorama.

4 Conclusion

Generally speaking, 2D geometry is successfully converted into a 3D format that is available for architectural design. Since 2D images are the basis for 2D geometry, Style Transfer, a neural network to generate 2D images, is of great significance for the whole design. From the project, Style Transfer not only successfully generates new images that inspire architects for further 3D model design but also provides a series of choices by adjusting input parameters so that architects can select out the one based on sensibility and aesthetics.

All of this brings to an envisagement that if Style Transfer can be employed as a post-human approach to architecture. The neural network used in the project successfully blends the style into the content image. Therefore, it is possible to generate a new style in light of two input images that are of distinctive design styles of two architects. It would be more interesting if the new style can be applied in architectural language so that Style Transfer will proceed into a creating level.

Future application in architectural design with Style Transfer involves in the creation of architectural images as well as vectorized data, for example the iterative optimization of building data with similar loss functions. Apart from application in architecture, it is in the prospect of being utilized in fields like graphic design and animation design. Thus, it is highly possible that Style Transfer Neural Network will be an indispensable tool for designers at the creating stage in the coming future.

References

- 1. del Campo, M., Manninger, S., Sanche, M., Wang, L.: The church of ai-an examination of architecture in a posthuman design ecology. In: CAADRIA 2019 (2019)
- 2. Fukushima, K.: Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybern. **36**(4),193–202 (1980)
- 3. Fukushima, K.: Neocognitron. Scholarpedia 2(1), 1717 (2007)
- 4. Gatys, L.A., Ecker, A.S., Bethge, M.: A neural algorithm of artistic style.arXiv preprint arXiv: 1508.06576 (2015)
- Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2414–2423 (2016)
- Kümmerer, M., Wallis, T.S.A., Bethge, M.: Deepgaze ii: reading fixations from deep features trained on object recognition. arXiv preprint arXiv:1610.01563 (2016)
- 7. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436-444 (2015)
- 8. Mordatch, I., Abbeel, P.: Emergence of grounded compositional language in multi-agent populations. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

