Applied Science and Innovative Research
ISSN 2474-4972 (Print) ISSN 2474-4980 (Online)
Vol. 8, No. 3, 2024
www.scholink.org/ojs/index.php/asir

Original Paper

Exploration of Stable Diffusion in Architectural Design

Ximin Nie*

Department of Architecture and Civil Engineering, Southwest Jiaotong University Hope College, Chengdu, Sichuan 610400, China

* Ximin Nie, Department of Architecture and Civil Engineering, Southwest Jiaotong University Hope College, Chengdu, Sichuan 610400, China

Received: May 12, 2024 Accepted: June 1, 2024 Online Published: June 5, 2024

doi:10.22158/asir.v8n3p1 URL: http://doi.org/10.22158/asir.v8n3p1

Abstract

With the rapid development of artificial intelligence technology, its application in the field of architectural design has gradually become a research hotspot. In particular, Stable Diffusion technology shows its unique advantages and potential in the production of architectural renderings, which has attracted wide attention. In this paper, the application of Stable Diffusion technology in architectural creation is deeply discussed, aiming at providing new perspectives and ideas for architectural design and promoting the innovation and development of architectural design process.

Keywords

Artificial intelligence, Stable Diffusion, Architectural design

1. Introduction

With the progress of science and technology and the development of society, the field of architectural design constantly pursues innovation and breakthrough, trying to improve the design efficiency at the same time, better show the aesthetic value of architecture and the characteristics of The Times. In this process, architectural AI creative expression plays an extremely important role as an intuitive display of design concepts. Traditional architectural design relies on the expertise and experience of designers, but this process is often time-consuming, costly, and difficult to quickly adapt to the needs of design changes. With the rise of artificial intelligence technology, especially the development of advanced technology such as Stable Diffusion, it provides new and diverse possibilities for the creation of architectural design process. These technologies can generate high-quality images in a short time, not only improve production efficiency, but also greatly reduce costs, but also provide a broader space for architectural design innovation.

2. Introduction to Stable Diffusion Technology

2.1 Features and Advantages of Stable Diffusion Technology

At present, the mainstream AI drawing software has Stable Diffusion and Midjourney. They can realize the drawing of architectural renderings, the difference is that the Stable Diffusion function is more powerful after the continuous update, there are many plug-ins in one, such as the controlnet plug-in often used in the field of architectural design.

The working principle of Stable Diffusion in the field of architectural design effect drawing generation is composed of three parts: Language mode (i.e., language model). It is the core part of the whole Stable Diffusion), Diffusion model (diffusion model) and decoder (decoder).

To sum up, as one of the most advanced artificial intelligence image generation technologies, the core feature of Stable Diffusion technology is that it can generate high quality and high fidelity images according to text description. This technology is able to interpret complex text information and transform it into visual images, showing great creativity and flexibility. Its advantages are reflected in several aspects: First, Stable Diffusion technology has a strong adaptive ability, which can quickly generate the image corresponding to the theme and style after receiving different text input. This means that designers can easily explore and experiment with different design ideas by tweaking text descriptions. Secondly, the technology supports the control of architectural details, and designers can refine many details of architectural images through accurate text and picture descriptions, so as to generate highly customized architectural design results. Finally, compared with the traditional image generation technology, Stable Diffusion can complete the image generation task with higher efficiency, which significantly improves the efficiency of the architectural design process.

2.2 Working Principle of Stable Diffusion

The working principle of Stable Diffusion is more complicated. Here, the working principle of Stable Diffusion can be compared to the process of picture clear - picture fuzzy - picture clear. The process from picture clear to picture fuzzy is the process of increasing noise. Then through calculation, the picture state from fuzzy to clear process, this process is the process of removing noise.

3. Application Method

3.1 Model

AI can carry out deep learning through super many materials, therefore, we can train the algorithm, let the AI learn the characteristics of a certain class or several classes of picture information (such as keywords, effects and other information) that the drawing author is inclined to, and the data packet after the training is completed, which is the AI model. It can be said that the model is the core soul of the entire AI architectural creation, which can automatically extract a variety of information (such as text, pictures, and information in the picture) to recombine, and finally output the architectural rendering.

To generate a rendering through Stable Diffusion, the first step is to understand the concept of the model. In AI, there are two broad categories of models commonly used:

The first type is Checkpoint (i.e. large model of complete data), the Checkpoint large model manages the most critical model of the overall effect of the rendering, which can help us control the style of the rendering. Such as the large tone of the architectural effect of the rainy night scene, you can choose the corresponding large model to control.

After selecting the style of the effect image, select the corresponding model in the Stable Diffusion model in the web UI.

The second category is Lora, Embeding, Hypernetwork (that is, the extended model used with the large model). In the extended model, the lora model is commonly used for the rendering of architectural renderings (Hypernetwork's rendering accuracy is low; Embeding model is difficult to change the overall style of the effect diagram, the drawing limitation is high, so the architectural effect diagram is rarely used in the generation.

Select the desired Lora model from the Lora button in the webui (at this time, the corresponding lora model will appear in the Vincennes Diagram dialog box), and the Lora model is selected successfully. In addition, in the Vincennes chart dialog box Lora model weight, according to their own needs to control the value (we commonly used architectural renderings, the number here is between 0 and 1, through research is not difficult to find that after skilled use, the commonly used weight value can be fixed between 0.55 and 0.7). After choosing the Lora model, we can write some trigger keywords, so that we can more accurately generate the effect we want.

In the concept of the model, it is worth noting that each AI creator can train the model indefinitely according to his own needs, simply speaking, if the more raw material training (raw material here refers to the useful material after screening, is a direction of the material, rather than a variety of directions of the material), the more rich the information, The more accurate the direction of the image generated by the subsequent AI is, which is the result that the designer hopes for, which can greatly improve the efficiency of the work. However, there is a problem in training large models, which is that the training time will be very long and it will occupy a lot of computing power. Therefore, among many AI creators, Lora extended model will be trained, which can greatly improve the efficiency of training, and the generated results of architectural renderings are also more in line with the designer's expectations.

3.2 Vincent Diagram

Vincennes diagram is one of the important methods in the drawing method of architectural renderings. Usually, architectural renderings are commonly produced in the way of Tu-sheng diagram with controlnet plug-in, and Tu-sheng diagram is commonly used for local redrawing operations.

In the Vincencene Diagram generating building renderings, enter a positive prompt for the renderings in the Vincencene Diagram dialog box in the webui (a positive prompt is a rendering word that you want to appear in the renderings, usually an adjective), and a reverse prompt (a reverse prompt is a word that you don't want to appear in the renderings). In the forward prompt word, enter two parts in

the prompt, one part is the content prompt words, such as public buildings, sunny days, modern style, glass curtain wall, trees, etc., these words are the words that control the effect of the renderings. The other part is the standard prompt words, such as high quality, 4k image quality, realistic style, high precision, high definition, unparalleled architectural photography words, these words are describing the effect of the image quality words.

In prompt, English input must be used, Chinese input cannot be used. Therefore, we can enter Chinese directly through the plug-in (in the advanced version of the webui, there is already a built-in translation plug-in, click the one-click translation button below the prompt dialog box to translate the prompt word); You can also use AI translation software to translate the Chinese and input it into the prompt dialog. In the process of writing prompt words, each prompt word is separated by English commas, and the weight of many prompt words is also corresponding to their positions, such as white building, glass curtain wall, sunscape renderings of these prompt words, white building occupies a greater weight, that is, the more restrictive the prompt words are. If we want to change the weather in the architectural renderings, we can locally modify the prompt words, such as changing the above white buildings, glass curtain walls, and day renderings to "rainy day" or "rainy day renderings" and other related words, so that we can directly generate the desired effect.

Through the prompt words of Vincennes diagram to generate the effect map, the range of "drawing card" is infinite, often can not get the effect map results we want. Therefore, we will attach the generation through Vincennes diagrams with important controlnet plug-ins. With the controlnet plugin, you can realize the uncontrollable problem of building renderings, fixing the composition, Angle, building outline and other information, so that Stable Diffusion only plays the part of it efficiently "created". It is worth noting that both controlnet and Lora models serve large models, and Lora is the characteristic style of control renderings; controlnet is the general "outline" of the control rendering, and it is indispensable for the generation of architectural renderings. In the role of controlnet, it also includes detailed control factors such as composition, contour, precision control, material division, and depth.

In prompt, after entering the forward and reverse prompts, drag the target reference of the renderings into the Single image dialog box in the controlnet unit in the webui, then select the control type from the control type below the webui (in building renderings, Commonly used types are Lineart (line draft), SoftEdge (soft edge) and Scribble/Sketch (Scribble/ sketch), so that the processing type of the preprocessor is automatically loaded. In each type of preprocessing, there are a variety of processing standards (usually we choose the line class), and the corresponding model can be automatically generated. Finally click controlnet Enable, Perfect Pixel mode, Allow preview three buttons. Then, adjust the rendering weight, the default parameter of the weight is 0 (this weight will greatly affect the architectural details, exaggeration, rationality, etc. generated by the rendering diagram, so it is recommended that the weight should not be too large or too small, usually choose an intermediate value); Then adjust the boot intervention time (this parameter is a parameter when controlnet is active,

if set to 0.3, it means that controlnet starts at 30% of the picture drawing process); Then adjust the boot termination time (this parameter is when controlnet terminates the function of the parameter, if set to 0.7, it means that controlnet begins to end control from 70% of the picture drawing process), in these two parameters, the boot intervention time is usually set to 0. Set the boot termination time to a value between 0.6 and 0.9. In control mode, you can choose according to your own habits, usually choosing more prompt word options, so that we can get both Stable Diffusion created content, but also limited by controlnet "outline". This rendering is then preprocessed, resulting in a rendering with only line drafts (this rendering of line drafts is a type commonly used in architectural renderings). Finally, you get the final image with controlnet control that combines the forward and reverse prompts.

3.3 Graphs Generate Graphs

In the drawing of architectural renderings, Tuson drawing and Vincennes drawing mentioned above are the two main directions of architectural renderings.

In the method of Tu-sheng diagram, firstly, the work flow of Tu-sheng diagram adds prompt words to the reference diagram, and finally generates the final architectural painting through Stable Diffusion calculation. In addition, we also need to understand that the limitations of Tucson drawings are relatively strong, and basically the architectural renderings generated by Tucson drawings will finally be drawn from the original drawings.

Drag the selected architectural effect into the picture generation box in the Stable Diffusion, and then we can input the keywords of the effect diagram we want in the prompt words (here do not write, because mentioned above: the limitations of the diagram is relatively strong, basically through the diagram generated by the effect diagram will finally draw the original diagram to draw this rule). Then we can adjust the graph parameters, such as resize, zoom mode, redraw amplitude, etc. (The parameters here will not be discussed in detail in this article. Because data preparation and preprocessing is an important step in the application of Stable Diffusion technology. At this stage, a large number of high-quality architectural design images need to be collected as data sets, which will serve as the basis for model training. The quality of image data directly affects the fidelity and aesthetics of the final architectural renderings. Therefore, the construction of data sets needs to ensure the diversity and coverage of images, as well as the accuracy of image annotation). Finally, click the Generate button to get the architectural design drawing by Stable Diffusion.

3.4 Generate HD Images

Due to the characteristics of Stable Diffusion technology, when we generate architectural renderings, the precision is often not very high in the process of regeneration into pictures. Therefore, after we generate the final version of the architectural renderings we want, we also need to do a step, which is the high-definition amplification of the picture, that is, the enlargement of the high-definition picture.

In Stable Diffusion, there are usually three ways to enlarge the building effect picture, which are: Tiled Diffusion (it can be used in Vincennes, Tu-sen), Ultimate SD upscale (it's only used in Tu-sen), and High-Resolution repair (it's only used in Vincennes).

Through high resolution repair means: in the webui, input the corresponding keywords in the forward prompt word, set the size of the picture to 512*512, click generate, you can generate a 512*512 building effect map. Enlarged images are blurry due to pixel size limitations. At this time, you can find the picture in the "boundless image browse", send it to the Vincennes diagram (the sent picture will keep the parameters and the number of random seeds unchanged), then check the high resolution repair, and then set the magnification (that is, 512*512*3=1536*1536 precision photos), the amplification algorithm is selected R-ESRGAN 4X+, The number of iteration steps can be selected 20, the redrawing range is not recommended to choose too large (generally choose below 0.5), and finally click Generate to get a high-precision picture. High-resolution restoration means, in fact, through the process of repeated drawing with low resolution pictures, high-precision architectural renderings are obtained through continuous drawing.

The second way to improve accuracy is Ultimate SD upscale. The Ultimate SD upscale is very convenient for high-precision rendering, and you can directly choose the scripts that are encoded in the advanced Web user interface. The first choice is to set the zoom (you can also enter the method value manually). The zoom algorithm is still R-ESRGAN 4X+, the drawing type can be Chess (which is usually better), and then select the Half tile offset pass +interse in Seams fix. Finally, we send the diagram to the graph by the above method, set the relevant parameters, and you can get a 2k precision architectural effect diagram.

The third method is Tiled Diffusion, this method is also through block processing to improve the accuracy of the picture, the picture is processed in blocks, and finally the various blocks are put together to get a final architectural effect map. In higher versions of WeBuIs, there are Tiled Diffusion and Tiled VAE in the dialog boxes of the Graph and Vincennes diagram, both of which need to be set up, because Tiled Diffusion takes care of expanding the picture and Tiled VAE takes care of encoding. After entering the corresponding keywords in the forward and reverse prompt words, the size of the image can be set slightly higher, such as 1024*512, etc., click to generate an architectural effect drawing (this drawing is blurred when enlarged), and send the drawing to the Picture-generation dialog box in the infinity image browser, and select a lower value (such as 0.3) for redrawing amplitude. Select MultiDiffusion scheme in Tiled Diffusion, select R-ESRGAN 4X+ magnification algorithm, often select 4 times magnification, and finally click Generate to get the same high-definition image.

Tiled VAE is the encoded parameter in the calculation process of the graphics card. If the video memory is not large, the block size value can be appropriately reduced.

4. Conclusion

This paper discusses the application of Stable Diffusion technology in the process of architectural design, and highlights its remarkable advantages in creation efficiency and image quality. As an innovative tool, this technology opens up new possibilities for architectural design and rendering, improves the flexibility and efficiency of design, and also provides designers with more abundant and

high-quality visual display and creative thinking direction. For designers engaged in architectural design and related fields, actively exploring and applying Stable Diffusion technology can improve work efficiency and reduce cost to a certain extent.

References

Chai, X. W. (2024). Application and Benefit analysis of Intelligent Design Software in Architectural design process. *China Construction*, *2024*(06).

He, W. Y., & Yang, L. S. (2023). Exploration of Generative Artificial Intelligence in Architectural design: A case study of Xiaoku AI Cloud. *Architectural Journal*, *2023*(10).

Lei, Y., Liu, L., & Hong, G. (2021). Application of virtual reality technology in architectural renderings. *Housing and Real Estate*, *2021*(22): 97-99.

Li, D., & Liu, J. T. (2019). Research on Architectural design and Optimization based on Al and Virtual Reality technology. *Science and Technology Innovation and Application*, *14*(14).

Li, J. Y., Zhao, H. Y., Wu, J. Y., Dai, S. Y., & Xu, H. H. (2023). Discussion and Exploration of small rural Building Design based on AI-Assisted Building Design technology -- A case study of Stable Diffusion. *Building a digital Education: 2023 National Conference on Teaching and Research of Digital Technology in Architecture*, 2023.091754.