國立臺灣大學電機資訊學院資訊工程學系

碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

互動式風格轉換在虛擬藝廊的應用

Applications of Interactive Style Transfer to Virtual Gallery

簡莘洳

Hsin-Ju Chien

指導教授: 李明穗 博士

洪一平 博士

Advisor: Ming-Sui Lee Ph.D.

Yi-Ping Hung Ph.D.

中華民國 111 年 9 月

September, 2022

doi:10.6342/NTU202203319

摘要

近年来，由於疫情的爆發，許多活動都改為線上的方式舉辦以避免疫情擴散，如虛擬藝廊。然而，在目前的虛擬藝廊中，無論是策展人或是觀賞者皆只能對藝術品進行有限的操作，為此我們希望透過風格轉換來讓策展者能夠藉此技術發展更多元的體驗形式，並讓觀賞者透過風格轉換的過程增加對於藝術品的認識，加深與藝術作品的連結。在本篇論文中，我們提出了一個互動式風格轉換虛擬藝廊系統，在系統設計階段，我們比較不同風格轉換方法產生的結果，並根據文中提及的考量因素選擇其中兩種方法整合至我們的系統。接著，使用者可以透過使用者介面對虛擬藝廊中的畫框、卡紙、圖片、以及 3D 藝術品進行風格轉換。最後，我們進行了一項使用者研究來評估本系統的有效性。結果顯示，在虛擬藝廊中加入風格轉換不只能夠增添虛擬藝廊體驗的豐富性，對於提升藝術品的美感也有一定的幫助，對於觀賞者而言，不僅能在觀賞過程中增添趣味性，更能藉由選擇不同風格的過程，加深對於藝術品的印象。

關鍵字：風格轉換、虛擬藝廊、使用者系統互動

doi:10.6342/NTU202203319

Abstract

In recent years, due to the outbreak of the epidemic, many events have been held online to avoid spreading illness, such as the virtual gallery. However, in the current virtual galleries, both curators and viewers can only do limited manipulation of the artwork. Therefore, we hope that through style transfer, curators can use this technology to develop more diversified forms of experience, and viewers can increase their understanding of the artworks and deepen their connection with the artworks through the process of style transfer. In this thesis, we propose an interactive style transfer virtual gallery system. In the system design phase, we compare the results of different style transfer methods and select two of them to integrate into our system based on our considerations. Then, users can perform style transfer on pictures, paper jams, frames, and 3D artworks in the virtual gallery through the user interface. Finally, we conducted a user study to evaluate the effectiveness of this system. The results show that adding style transfer to the virtual gallery not only increases the richness of the virtual gallery experience, but also enhances the aesthetics

doi:10.6342/NTU202203319

of the artwork. For the viewers, it not only makes the viewing process more interesting, but also deepens the impression of the artwork through the process of choosing different styles.

Keywords: Style Transfer, Virtual Gallery, User System Interaction

doi:10.6342/NTU202203319

Contents

doi:10.6342/NTU202203319

doi:10.6342/NTU202203319

List of Figures

doi:10.6342/NTU202203319

doi:10.6342/NTU202203319

List of Tables

doi:10.6342/NTU202203319

Chapter 1 Introduction

In recent years, due to the outbreak of the pandemic, many activities have been conducted online to avoid spreading illness, such as virtual galleries, online concerts, etc. Virtual galleries are different from traditional physical galleries in many ways. For instance, viewers can visit online at any time and place and can visit multiple times without restrictions, artworks do not need to worry about being destroyed by humans, and artworks can be moved and scaled at will in the virtual space, which is more flexible than a physical gallery and does not need to employ a tremendous amount of manpower and consume additional cost. In addition, different ways of visiting the virtual gallery will also bring different experiences to the users. There are many platforms that provide experience in multiple ways. For example, OnCyber [1] can be accessed with a web browser or in VR mode, MomentX [2] can be accessed in VR mode but requires the purchase of a specific NFT, and GOXR[3] can be accessed with a mobile app or in VR mode.

However, in the current virtual galleries, both curators and viewers can only do limited manipulation of the artworks, such as basic zooming and moving, etc. Considering that the artworks in the virtual gallery are all digitalized, we can make more use of this information to further develop new interactive experiences. Therefore, we hope that through style transfer, curators can use this technology to develop more diversified forms of experience, and viewers can increase their understanding of the artworks and deepen their

doi:10.6342/NTU202203319

connection with the artworks through the process of style transfer.

Style transfer is a technique for synthesizing new images, where the resulting image maintains the structure of the content image but has the texture and color of the style image. And Style transfer can be divided into two categories: Optimization-based and Model-based. The former requires iterative optimization of the initialized image, which is more time-consuming, while the latter is to train a network to know how to fuse the content feature and style feature to do the transformation through datasets, and the result can be obtained by a forward pass at test time.

In this thesis, we propose an interactive style transfer virtual gallery system. We compare the results of different style transfer methods and select two of them to integrate into our system based on our considerations. Thus, users can perform style transfer on pictures, paper jams, frames, 3D artworks in the virtual gallery through the user interface. Finally, we conducted a user study to evaluate the effectiveness of this system. The results show that adding style transfer to the virtual gallery not only increases the richness of the virtual gallery experience, but also enhances the aesthetics of the artwork. For the viewers, it not only makes the viewing process more interesting, but also deepens the impression of the artwork through the process of choosing different styles.

doi:10.6342/NTU202203319

Chapter 2 Related Work

Our work compared the results of different style transfer algorithms. The previous works can be categorized into three main categories: image style transfer, video style transfer, and 3D style transfer.

2.1 Image Style Transfer

The image style transfer can be traced back as far as the algorithm proposed by Gatys et al. [4] They demonstrated that the high-dimensional features extracted from the image through the DNN model that has been well-trained to classify the image in the ImageNet dataset can represent the overall structure of an image, and the correlation between features can be used to capture the style of an image. They design loss function based on those two findings and optimize the original image by multiple iterations, and thus the result image can maintain the structure of the content image and have the characteristics of the style image.

However, the method by Gatys et al.[4] is extremely slow due to its optimization process. In order to reduce the computational cost, Johnson et al.[5] proposed a feed-forward neural network, and thus the result can be generated through a single forward pass at test time which perform more efficiently than the optimization-based method. But the limita-

doi:10.6342/NTU202203319

tion of feed-forward method is that a model can only transfer one fixed style. Dumoulin et al.[6] addressed this problem by proposing a conditional instance normalization layer, they found that they can achieve multi-style transfer by learning a set of affine transformation parameters, and based on the method of Dumoulin et al.[6], Huang et al.[7] proposed to match the mean and variance of content feature to the style feature extracted from the convolutional neural network, which can achieve arbitrary style transfer. WCT[8] also proposed a similar concept that is matching the content feature covariance matrix to the style feature covariance matrix through whitening and coloring transforms.

The previous arbitrary style transfer approaches are holistic transformations that dismissed local style patterns. Park et al.[9] introduced self-attention mechanism to learn the semantic correlation between the content features and the style features for generating more local style details. Liu et al.[10] indicates that the low-level patterns of image also have an impact on the style transfer result. Therefore, they concatenate the low-level features with the high-level features, and thus the model can pay more attention to the shallow features during training. In order to make the transferred results more artistic and realistic, Chen et al.[11] use GAN to learn human-aware style information. They train a discriminator to distinguish transferred result is fake or not, and train a generator which is dedicated to deceiving the discriminator.

2.2 Video Style Transfer

Compared with image style transfer, the goal of video style transfer is to focus on not only generating pleasing stylization results but also consider the consistency of transferred results between adjacent frames of the video, so as to avoid flickering problem. Ruder et

doi:10.6342/NTU202203319

al.[12] introduce temporal constraints by calculating optical flow to generate consistent stylized video. However, the generated result depends on the accuracy of optical flow algorithm. In order to generate consistent result without using optical flow estimation, Li et al.[13] proposed a linear transformation model, which is mainly to learn a transformation matrix to replace whitening and coloring transforms. Since the linear transformation can preserve content affinity at test time, the stylized videos are more stable and consistent. Deng et al.[14] calculate per-channel correlation across content and style representations, the generated video satisfies the constraint that the flicker of the video is not perceived by the human eye as proposed by Wang et al.[15]. While adding the vivid style pattern on the results and maintaining the consistency of the generated videos.

2.3 3D Style Transfer

3D style transfer requires more consideration of the three-dimensional information of 3D objects than image style transfer, otherwise transformed result will have variety visual artifacts.

Kato et al.[16] proposed to use a 3D rendering process to render an image of the 3D object from a random angle as the input of the CNN model to perform style transfer. Mordvintsev et al.[17] also use the 3D rendering process to turn 3D objects into 2D RGB images, but unlike the method of Kato et al.[16], they do not modify the geometry of the 3D object. The texture of the 3D object is optimized by backpropagating the gradient of the loss function proposed by Gatys et al.[4].

There are also some methods that do not use differential renderer for style transfer, such as Hsu.[18], which use the method of Gatys et al.[4] to transferred the texture of the

doi:10.6342/NTU202203319

3D object first, and calculate the normal map and displacement map based on the generated texture to make the shading more consistent with the transferred texture.

In addition, not only transferring the textures of 3D objects, there are also some studies that focus on transferring geometric style between 3D objects. For example, Segu et al.[19] proposed a generative model with disentanglement of content and style representation for synthesizing new 3D shapes. The method proposed by Yin et al.[20] combines geometry style transfer network and texture style transfer network, and following a differential Renderer to create 2D images of the 3D object as the input of CNN model to synthesize a variety of 3D objects.

doi:10.6342/NTU202203319

Chapter 3 System Design and Implementation

In this chapter, we describe the system design considerations in detail, including the selection of style transfer methods, and user interface design, and the details of the implementation and analysis of the results is described in Section 3.5.

3.1 Object Type

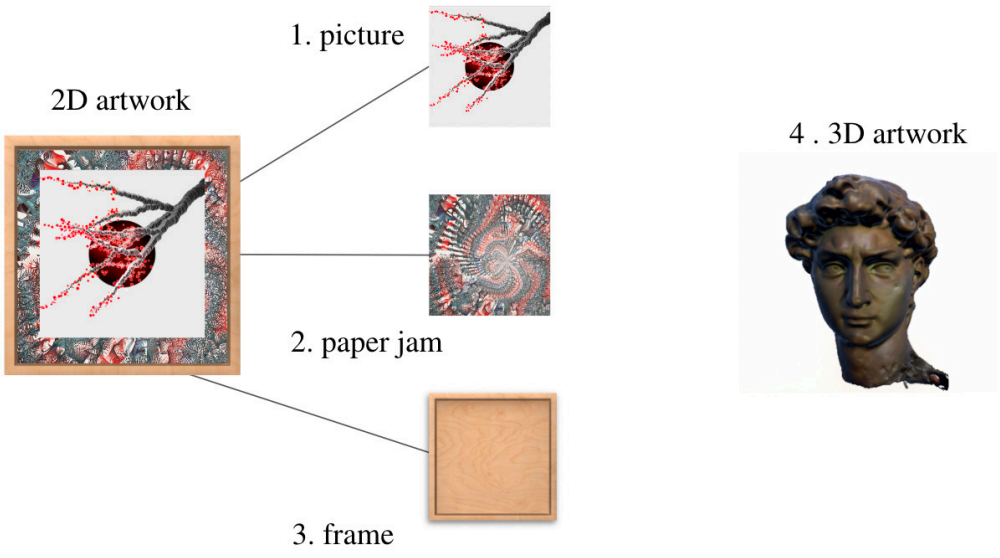
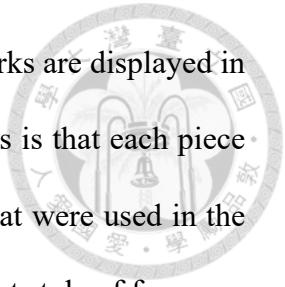


Figure 3.1: Four types of objects.

doi:10.6342/NTU202203319



In the virtual gallery, a wide range of 2D artworks and 3D artworks are displayed in this space for the viewers to appreciate. The charm of these artworks is that each piece contains the ideas that the creator wanted to convey, the emotions that were used in the creation, and the means of artistic expression. In addition, the different style of frame on the 2D artwork will bring different emotions to the viewer. Among the many categories of frames, we choose the common Chinese painting frame as the basic frame in our virtual gallery, and the paper jam in it plays the role of embellishment and accent, which will also have a certain influence on the overall artwork.

Therefore, as shown in Figure 3.1, there are four types of objects that can be transferred in our virtual gallery, including pictures, paper jams, frames, and 3D artworks. In our system, the viewer can use his or her own sense of aesthetics to select the right style for style transfer, which enhances the interaction and connection between the viewer and the artworks.

3.2 Model Selection

In this section, we will analyze and compare the results of different style transfer methods. The comparison is divided into two parts: style transfer for 2D image and style transfer for 3D object.

3.2.1 Qualitative Comparison of Stylized 2D Image

We compare optimization-based method[4], and six model-based method, including AdaIN[7], Linear[13], SANet[9], MCCNet[14], AdaAttN[10], IEST[11]. Linear[13] and MCCNet[14] is video style transfer method, but it can also be applied to images. The

doi:10.6342/NTU202203319

results of the compared methods are generated by conducting their publicly available code and default configurations.

Since the picture is the most important part of the artwork, we don't want the transferred picture to be unnecessarily distorted and have unwanted visual artifacts, losing the meaning of the original picture. About the paper jam, which is mainly used to set off the picture. Due to the paper jam being placed under the picture, we don't want the result to be too messy or have too much texture, which might degrade the overall visual quality. Therefore, our main considerations for choosing the method to perform style transfer on pictures and paper jams are artifact-free, content structure preservation, and overall visual quality. And the vividness of style patterns is the secondary considerations. In Figure. 3.2, we show the stylization results generated by the compared methods.

Unlike the model-based methods which need to be concerned with the problem of generalization, Gatys is able to generate the pleasant results (e.g., the 1^{st} and 7^{th} rows) due to optimization-process which is focus on transferring the specified characteristic of style image onto the content image. However, the results sometimes encounter the problem of bad local minimum (e.g., the 5^{th} and 9^{th} rows). Because of the over-simplified feature transferred calculation in AdaIN, it is likely to generate the crack pattern (e.g., the 2^{nd} , 4^{th} , 7^{th} –9^{th} rows), which have a large impact on the visual quality. In addition, content detail lost in some results (e.g., all the leaves are missing in the result of the 3^{th} row). Moreover, due to the trade-off between content image and style image is not that well, the results sometimes introduce some colors that are not present in the style image. (e.g., the green color of the lake in the result of the 8^{th} rows). Although the transferred result of the Linear is clean, it tends to only transfer the low-level texture pattern like colors. In addition, it often generates unwanted haloing artifacts around the object (e.g., the 8^{th} and 9^{th} rows).

doi:10.6342/NTU202203319

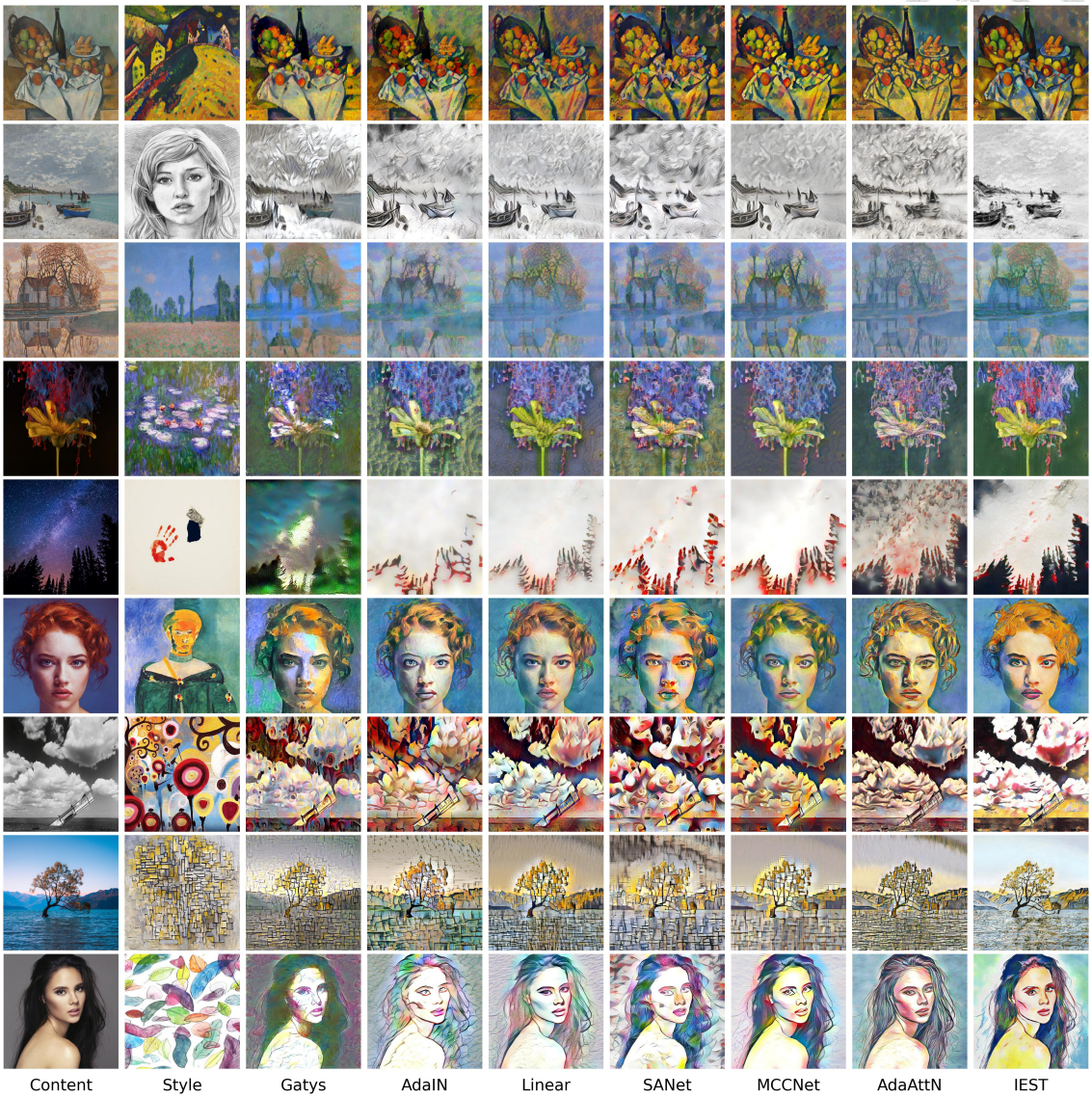
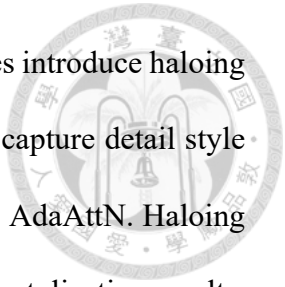


Figure 3.2: Results from compared style transfer methods.

SANet can faithfully transfer adequate local stroke patterns (e.g., the circular style pattern at the 7^{th} row and the feathers textured pattern on the hair at the 9^{th} row), and is enabled to capture the content detail (e.g., the stars of the content image can be observed in the transferred image at the 5^{th} rows). But some results have unpleasant content distortion (e.g., the 2^{nd} rows). In addition, when the background of the content image only has similar color, the stylization result will produce repetitive style textures (e.g., the 4^{th} , 6^{th}, 8^{th} , and 9^{th} rows). The visual artifact of repetitive patterns makes the results look messy. MCCNet lost some content detail. In addition, the colors of the generated result

doi:10.6342/NTU202203319



are disharmonious (e.g., the 6^{th} and 9^{th} row), and the results sometimes introduce haloing artifacts (e.g., the 8^{th} and 9^{th} rows). Moreover, their method cannot capture detail style patterns. The content structure is well maintained in the results of the AdaAttN. Haloing artifact which is easily produced by many methods is not present in the stylization results. However, the colors of the results have lower saturation and the lines of the object contour are too sharp. By learning human-aware style information, IEST is superior to the other compared style transfer methods in terms of visual quality. Although the style patterns in the results are less than SANet, the overall color distribution is more harmonious. In addition, the content structure of the transferred results is well-preserved, and no haloing artifacts and crack pattern problems.

Based on the above analysis of the four aforementioned considerations, the artifact-free methods are AdaAttN and IEST, while the methods with well content preservation are also AdaAttN and IEST. As for the overall visual quality, the overall color tone of the results generated by IEST is the most harmonious and realistic, while the style patterns are more transferred by SANet. In summary, we finally chose IEST as the model to be used in our system to perform style transfer on pictures and paper jams.

3.2.2 Qualitative Comparison of Stylized 3D Object

There is still less paper study on transferring the texture of the 3D object. Thus, we apply 2D image style transfer method to the texture of the 3D object for comparison. The compared method is the same as subsection 3.2.1 and we add 3D style transfer method[17] into comparison.

The main considerations in choosing a model for transferring 3D artwork are artifact-

doi:10.6342/NTU202203319

free, the richness of the style patterns, and the overall visual quality. Artifact-free means the result won't have unwanted colors or retains the content image color. Since we perform style transfer on the texture of the object without changing the shape of the object itself, the slight distortion of the texture of the object will not affect our ability to recognize the object. Thus we believe that increasing the richness of the texture is more important than preserving the content structure in terms of enhancing the aesthetics. In Figure 3.3, we show the stylization results generated by the compared methods.

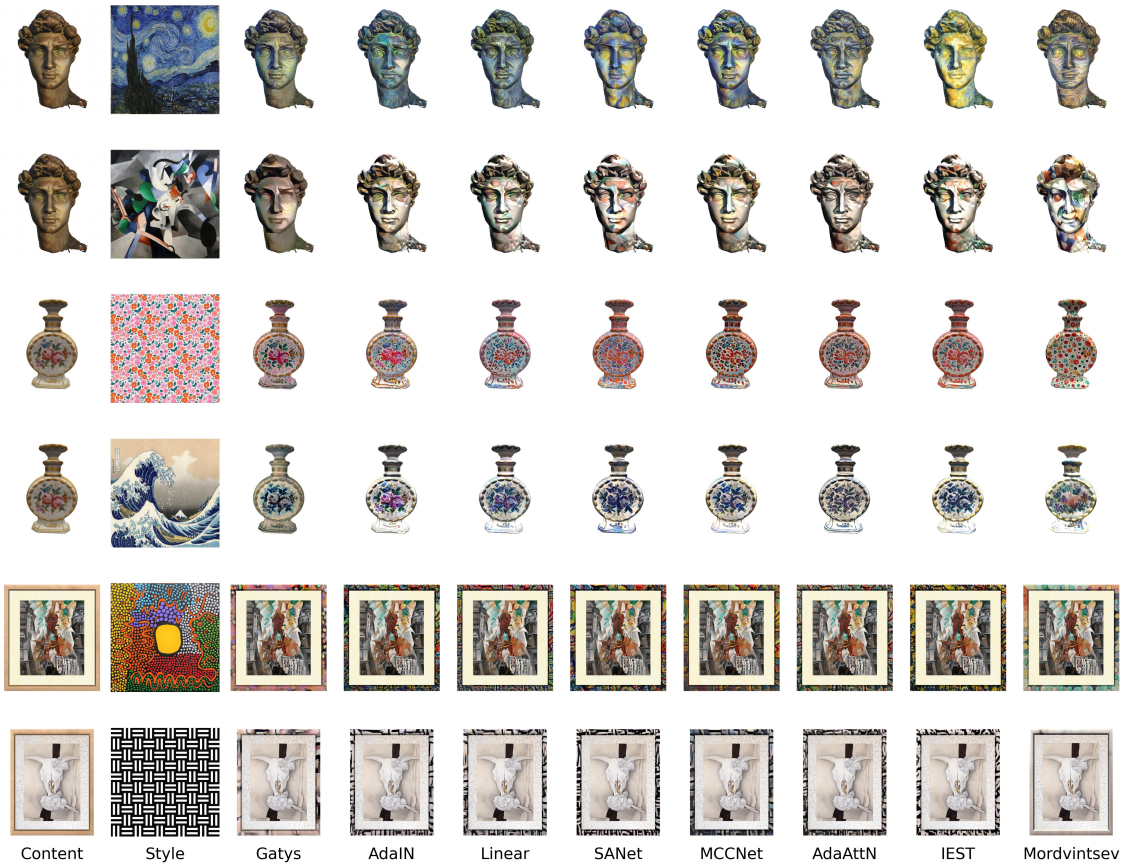


Figure 3.3: Examples of using compared style transfer method to stylize the textures of the 3D models.

Gatys sometimes fails to capture the stroke details (e.g.,1^{st} and 2^{nd} rows). In addition, we observed that Gatys is easier to suffer from getting stuck in the local minima if the texture of the object is simple (e.g., 5^{th} and 6^{th} rows). The results generated from AdaIN tends to present disharmonious colors and the method sometimes retains the color

doi:10.6342/NTU202203319

of content images (e.g., 3^{rd}). Linear is likely to produce unwanted colors which don't exist in the style image (e.g., 3^{rd} row), and the color of the content image retains (e.g., 6^{th} row). SANet can generate the most style patterns (e.g., 3^{rd} row) and detailed stroke of the style image can be transferred (e.g., 1^{st} and 2^{nd} rows) that makes the 3D object look more appealing and artistic. MCCNet cannot capture the detailed texture patterns, and tends to generate unseen colors (e.g., 3^{rd} and 6^{th} rows). AdaAttN and IEST can generate harmonious colors with well-perserved content structure. However, the style patterns are rarely seen in the stylization results and sometimes present messy patterns (the 6^{th} row). Although Mordvintsev considers the underlying UV-mapping of the 3D objects, it also encounters the problem of bad local minima due to the optimization process (e.g., 2^{nd}, 5^{th} and 6^{th}). In addition, this method sometimes cannot preserve the content structure well (e.g., the flower pattern of the 3D object is missing at 3^{rd} row).

Based on the above analysis of the four aforementioned considerations, the artifact-free methods are SANet, AdaAttN, and IEST, while the method that can transfer vivid style patterns and detail stroke is SANet. As for visual quality, SANet, AdaAttN, and IEST can generate more harmonious colors, while AdaAttN can preserve content structure well. As we focus more on producing results with a rich style pattern, we finally chose SANet as the model to be used in our system to perform style transfer on 3D artwork.

3.3 Style Selection

In order to perform the style transfer, our system provide 6 categories from which users can select a image as the style image. We called the 6 categories "Style Library". As shown in Figure 3.4(a) The 6 categories are Impressionism, Art burt, Ink Painting

doi:10.6342/NTU202203319

Cubism, Pattern, Pop Art. Each category contains three images that match the category.

In consideration of the diversity and flexibility of choosing a style image, we provide a feature that allows the user to select a picture from his or her NFT wallet as a style image. As shown in Figure 3.4(b), the NFTs in the user's wallet will be shown on the panel. Even if the user does not have an NFT wallet or no NFT can still choose style image from the style library for style transfer.

(a) Style Library panel (b) My NFT panel

Figure 3.4: Screenshot of two selection panels.

3.4 User Interface Design

A user interface is the means by which a user interacts with a system. Oppermann [21] stressed that a user-friendly interface is a central issue for the usability of a software product. In order to achieve intuitive operation and ease of use, in our system, as shown in Figure 3.5, the user interface panel is divided into two parts: a display area, which is on the left half of the panel, displays the object user is selected; an operation area, which is on the right half of the panel, allows users to select style image and adjust style intensity for style transfer. In the operation area, we use buttons to let users intuitively click on the style categories or a single style image, as shown in Figure 3.5, and use a slider to let users

doi:10.6342/NTU202203319

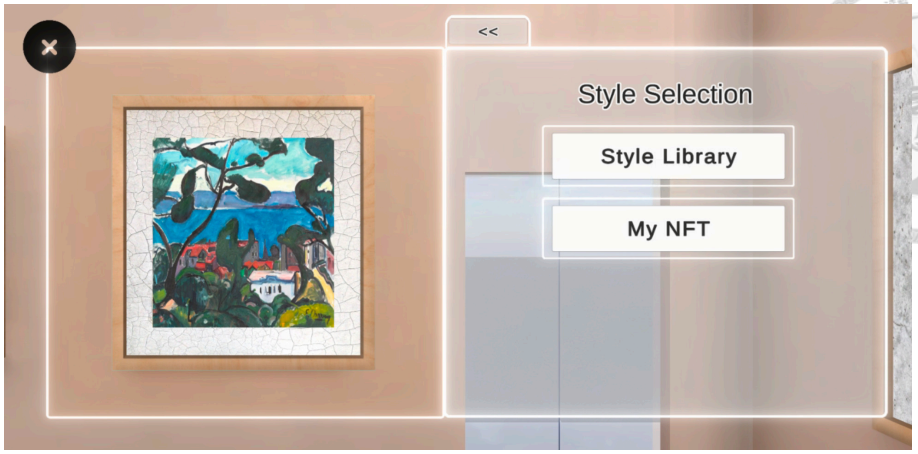


Figure 3.5: Screenshot of the user interface.

easily and effectively adjust the style intensity, as shown in Figure 3.6.

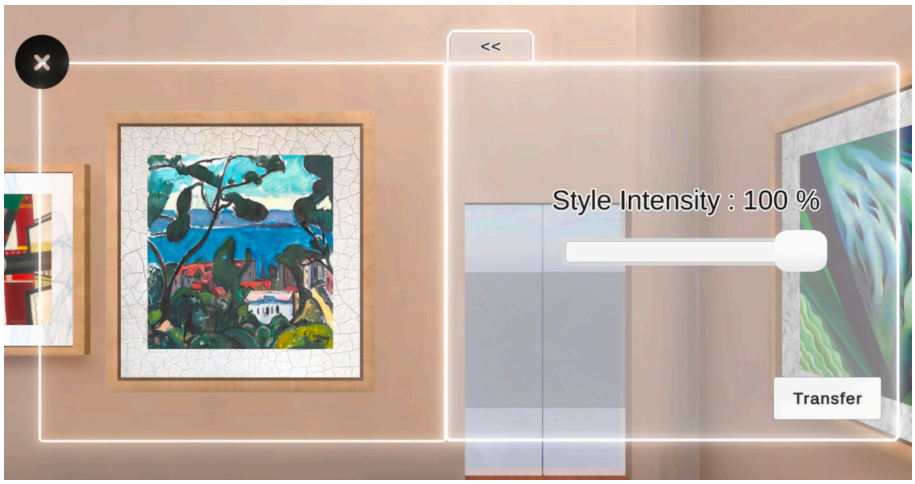
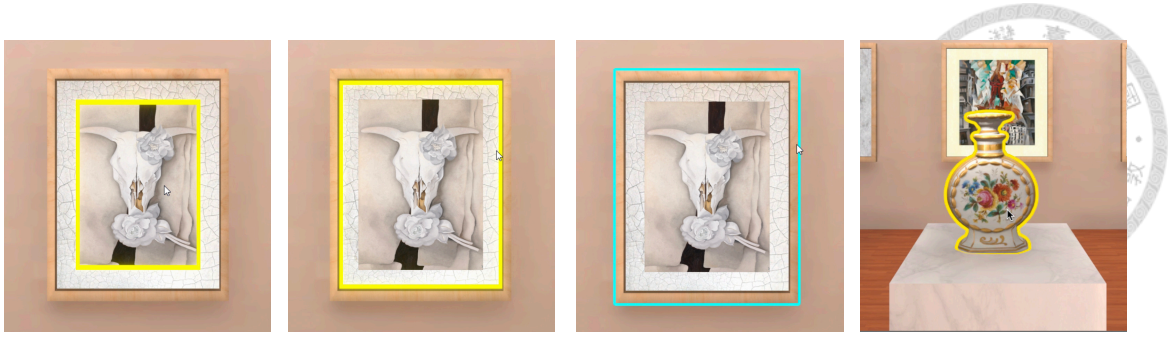


Figure 3.6: Screenshot of the user interface with slider component.

In order to let the user know exactly what he is hovering over and to indicate how many objects can do style transfer in the virtual gallery, we introduced visual feedback. That is, when the mouse moves to an object that can perform style transfer, the object will render an outline on the edge of it, as shown in Figure 3.7, which indicates the outline effect of hovering over a picture, paper jam, frame, and 3D artwork respectively.

In order to eliminate the unnatural feeling of switching from pre-transformation to post-transformation when the transfer is finished, we introduce a dissolve effect to the transformation process to make the style change process look smoother. In addition, the

doi:10.6342/NTU202203319



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

Figure 3.7: Screenshot of outline effect on different object when hovering over the object. (a) hovering over the picture. (b) hovering over the paper jam. (c) hovering over the frame. (d) hovering over the 3D artwork.

dissolve effect will only work on the object that the user is currently selected.

3.5 Implementation

In this section, we will describe our implementation in detail and discuss some issues about the transferred results in our virtual gallery.

3.5.1 Virtual Gallery in Desktop

The application is developed with Unity 3D 2020.3.15f2. User predominantly experiences virtual gallery from a first-person perspective with the camera at a fixed height and capable of rotating 360°. User interaction is done with mouse and keyboard. User interacts with objects and manipulates user interface through a mouse, and user is enabled to use WASD keys or arrow keys to move in the virtual environment. Figure 3.8 presents the screenshot of our application.

doi:10.6342/NTU202203319

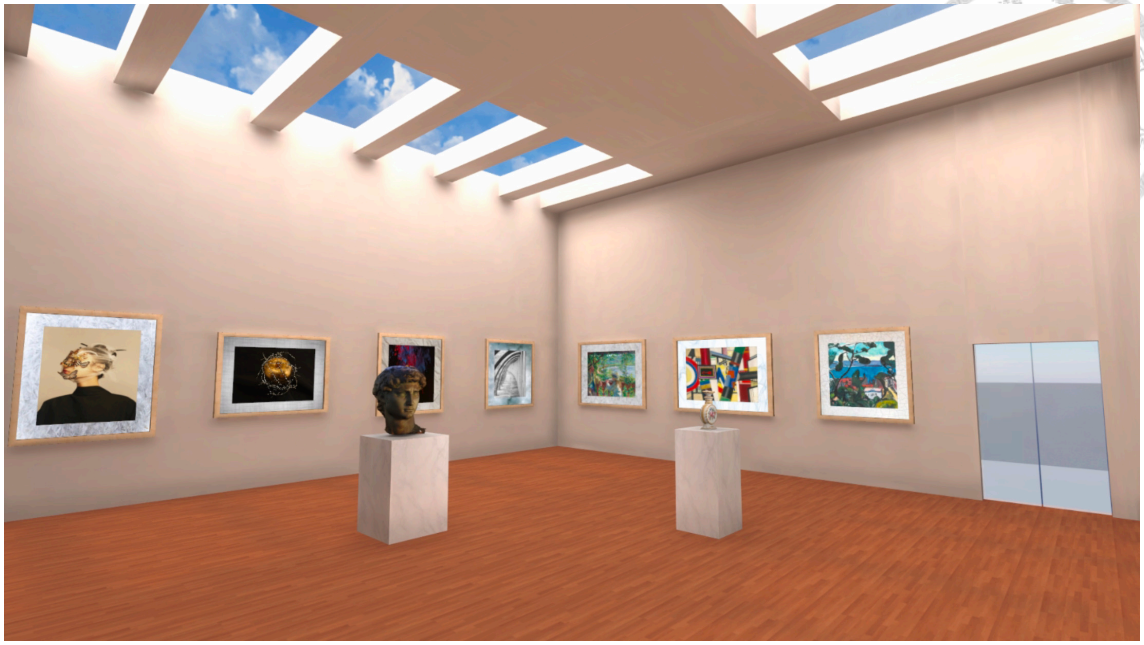


Figure 3.8: Sceenshot of virtual gallery.

3.5.2 Data Flow

In Figure 3.9 , we show how data flows through the system. The user can click the interested object for style transfer in the system by mouse. After that user needs to select a style image by clicking the button and adjusts the style intensity by dragging the slider. After the user presses the transfer button, we send four kinds of data to our server through TCP socket.The four kinds of data are style transfer method name, selected object texture, style image, and style intensity ranges from zero to one. Then server execute style transfer code to inference the transferred output. Finally, the result will be sent back to unity and update the style of the selected object.

3.5.3 Transferred Results

In this subsection, we show the transferred results of different objects and analyze the impact of the transferred results using different style images on the original artwork.

doi:10.6342/NTU202203319

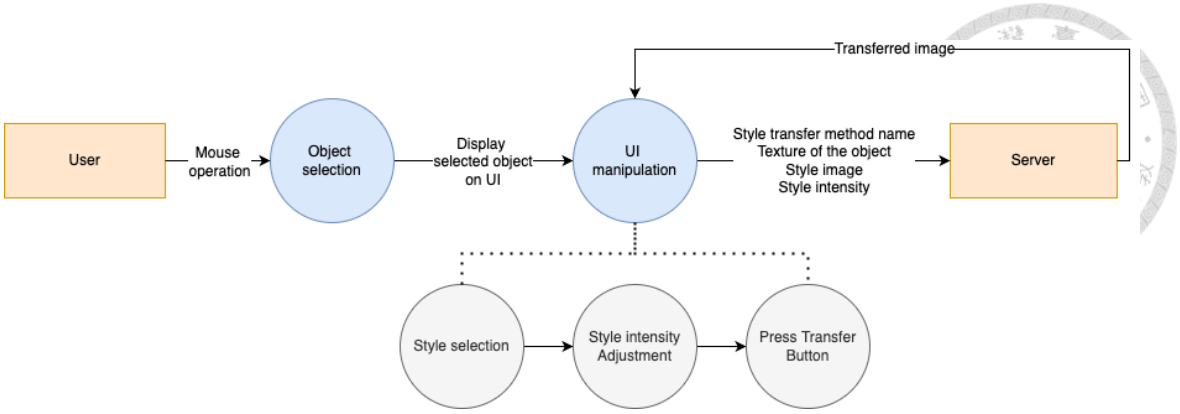


Figure 3.9: The data flow diagram.

Results Transferred by Picture

Figure 3.10 shows the result of the style transfer performed on the frame using the corresponding picture. We observe that compared to the original plain frame, the stylized frame seems to be designed especially for this artwork because it has similar color tone and style elements with the picture. Thus we can consider that the transferred frame can enhance the overall aesthetic of the artwork. Figure 3.11 shows the result of using the

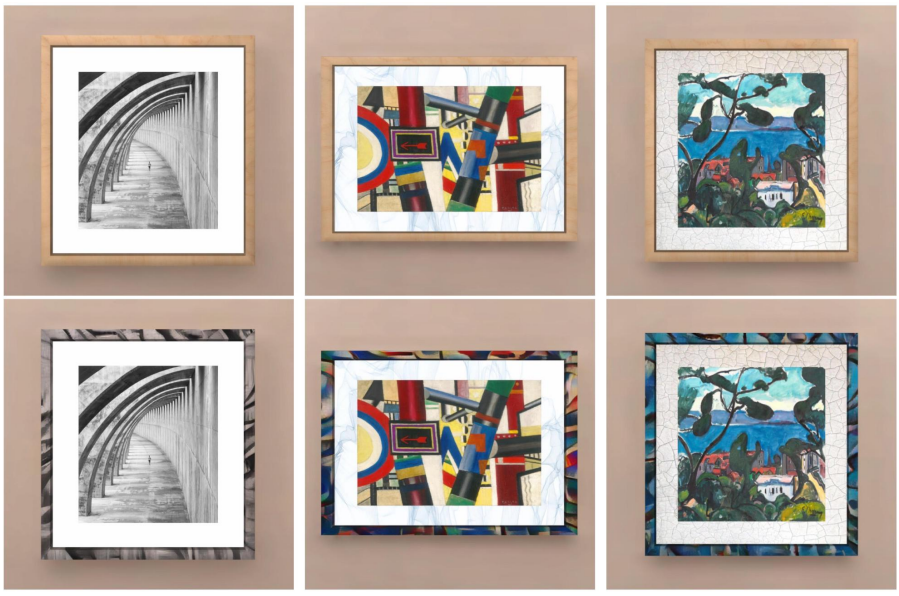


Figure 3.10: Results in the bottom row are frame transferred results of the top row using the picture as the style image.

corresponding picture to do style transfer on the paper jam. We observed that the transferred paper jam can present the effect of extension. However, because the paper jam and

doi:10.6342/NTU202203319

the picture are not like the frame and the picture, there is a paper jam as a buffer, if the style is too strong can make the entire artwork look messy, so we thought that the result with low style intensity can add richness to the artwork without taking away the flair of the original. Figure 3.12 shows the results of using the corresponding picture to do style transfer on the frame and the paper jam. By combining the advantages of both mentioned before, the transferred result has added more attractions to the artwork.

Figure 3.11: Results in the bottom row are paper jam transferred results of the top row using the picture as the style image.

Figure 3.12: Results in the bottom row are frame and paper jam transferred results of the top row using the picture as the style image.

doi:10.6342/NTU202203319

Results Transferred by Arbitrary style

Figure 3.13 shows the transferred result of the frame by using a style similar to the color of the picture. We believe that even if we do not use the picture as the style, we can still add visual quality to the overall artwork by choosing the right style. Figure 3.14 shows the transferred result of the paper jam by using an arbitrary style image, we observed that the use of different styles brings a different mood to the picture, and because of the proportion issue, the impact is greater than that of the frame. Figure 3.15 shows the transferred result of the picture and frame by using an arbitrary style image, we observed that performing style transfer on the picture will bring a new aspect of the picture to the viewers, which is interesting and match the same style of the frame will enhance the consistency of the whole artwork.

Figure 3.13: Results in the bottom row are frame transferred results of the top row using the right bottom image as the style image.

doi:10.6342/NTU202203319

Figure 3.14: Results in the bottom row are frame transferred results of the top row using arbitrary style images.

Figure 3.15: Results in the bottom row are picture and frame transferred results of the top row using arbitrary style image.

doi:10.6342/NTU202203319

Transferred Results of 3D Artworks.

Figure 3.16 and Figure 3.17 shows the transferred results of the two 3D artworks respectively. We observe that the transferred result are more vivid and artistic due to the addition of stylistic textures.

Figure 3.16: Results of bottle using different style images for style transfer.

Figure 3.17: Results of sculpture using different style images for style transfer.

doi:10.6342/NTU202203319

Chapter 4 System Evaluation

4.1 User Study Design

In order to evaluate the effectiveness of our system, we conduct a within-subjects user study based on measuring user satisfaction towards transferred results and evaluating usability of the system by questionnaires and user interviews.

4.1.1 Scene Setup

As shown in Figure 4.1(a-d), a total of fourteen 2D artworks are hung on the wall in the scene, each containing a picture, paper jam, and frame. All frames have the same wood texture. In addition, there are two 3D artworks on the display in the scene, and Figure 4.1(e)(f) shows the 3D scene viewed from two different viewpoints.

4.1.2 Tasks

In order to clearly identify the participant's perception of different objects performing style transfer, we divided the task into four parts according to the transformed objects, including frame transformation, paper jam transformation, arbitrary transformation and 3D artwork transformation, where the arbitrary transformation task meant that the user

doi:10.6342/NTU202203319

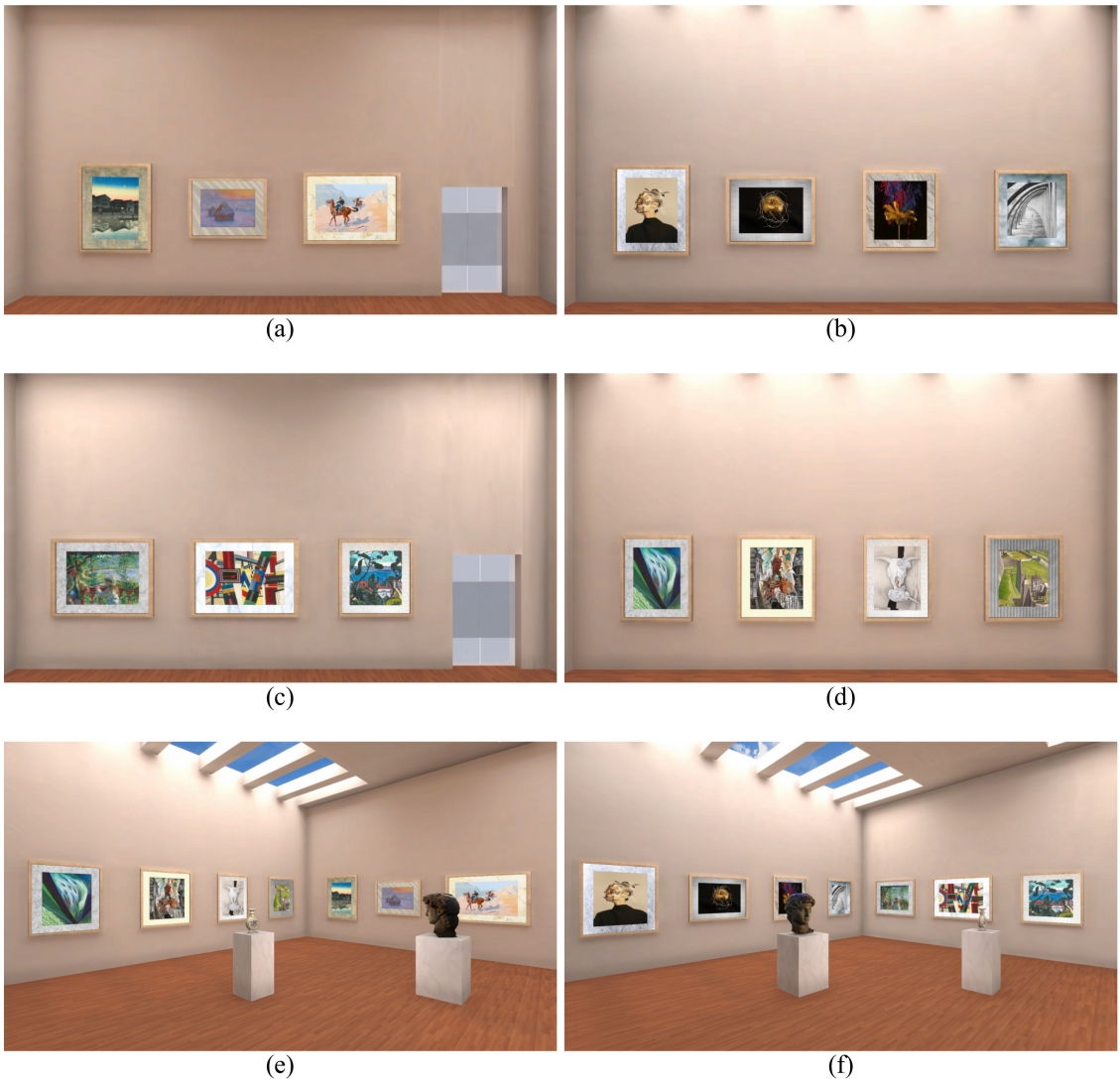


Figure 4.1: Screenshots of the scene.

could perform style transfer on the picture, paper jam and frame at will. We asked participants to use their own sense of aesthetics to transform the objects and compare the overall feeling of the transformed objects to the original ones using the questionnaire. For frame transformation and paper jam transformation, we verbally guide the participants to do two actions: First, use the picture as the style image for style transfer; Second, select style images excluding the picture for style transfer.And prompt him to adjust the style intensity to produce different results. As for arbitrary transformation and 3D artwork transformation, participants can choose any style image for style transfer.

doi:10.6342/NTU202203319

4.1.3 VR Mode

After the four style transfer tasks are completed, users can wear VR headsets to watch the virtual gallery in an immersive way. Teleportation can be used to move around the virtual gallery.

To increase the interactivity of the artwork in virtual reality, we proposed three visual effect of transformation process: Instant transformation, Fast transformation, Slow transformation. Instant transformation means that when the participant enters a 10-meter area in front of the 2D artwork, the artwork will immediately convert to the transferred result, and when the participant leaves the area, the artwork will return back to its original appearance. The transformation process is divided into four stages, and four times are used to control the switching between the stages, including the time to display the original artwork, the time to change to the transferred result, the time to stay in the transferred result and the time to return back to the original artwork.Corresponding to these four time periods, fast transformation is 4 seconds, 1 second, 2 seconds and 1 second, and slow transformation is 2 seconds, 0.5 seconds, 1.5 seconds and 0.5 seconds. We asks user to experience the three visual effects by using teleportation to approach the three transformed artworks respectively.

4.2 Procedure

At the beginning of the experiment, each participant is briefed with the purpose of the experiment, and they will be briefly introduced to the method of moving in the virtual gallery and how to click on the objects to do the style transfer.

doi:10.6342/NTU202203319

After understanding the operation, the participant started to perform the four transformation tasks. We asked the participant to select the one they liked the most from the 14 artworks for the frame transformation task. After finishing the task, the participant filled out the questionnaire about satisfaction. The questionnaire statements is shown in Table 4.1(Q1-Q3,Q6). And we asked for the reason why they liked the transferred result with that style. For the next task of paper transformation, we asked the participant to choose another artwork to do the style transfer, the questionnaire statements are the same as the frame transformation task, as shown in Table 4.1(Q1-Q3,Q6). For the arbitrary transformation, the participant was asked to choose another favorite artwork for style transfer, the questionnaire statements are shown in Table 4.1(Q4,Q6), and the participant was asked what the overall matching points were.For the 3D artwork transformation, we asked the participant to choose one of the two 3D artworks for style transfer, and the questionnaire statements are shown in Table 4.1(Q5,Q6). After the four transformation tasks were completed, we asked the participant to put on the VR headset to watch the virtual gallery in an immersive way, and we asked the participant to approach the transferred artworks to experience three different visual effects of the transformation process. Finally,the participant filled out custom user satisfaction questionnaire as shown in Table 4.1(Q7-Q15), system usability questionnaire (SUS)[22], ranked the objects in order of preference for style transfer and ranked the visual effect of transformation process in order of preference.

doi:10.6342/NTU202203319

Table 4.1: Statements for user satisfaction.

Statements
Q1: I thought the transferred result using picture is better than the original.
Q2: I thought the transferred result using my selected style is better than the original.
Q3: I thought the transferred result using picture is better than using my selected style.
Q4: I thought the transferred result is better than the original.
Q5: I thought the transferred 3D artwork is better than the original.
Q6: I thought this system makes it easy for me to find the style I want.
Q7: I thought the application is running smoothly.
Q8: I thought that the beauty of the artwork can be enhanced through style transfer.
Q9: I thought that style transfer enriches the virtual gallery experience.
Q10: I thought the process of style transfer deepened my impression of the artwork.
Q11: I liked to select style images from the Style Library for style transfer.
Q12: I liked to select style images from My NFT for style transfer.
Q13: I thought the number of user interface panels is too much.
Q14: I thought the elements of the user interface is not clear.
Q15: I preferred to appreciate the artwork in virtual reality rather than in the desktop version.

4.3 Experimental Results

We use 5-point Likert scale to assess the statements Q1-Q15 about user satisfaction. Five-point means ”strongly agree”, three-point means ”neutral”, and one-point means ” strongly disagree”.

Participants

In this evaluation, 10 participants were recruited for the experiment, 5 males and 5 females with ages from 20 to 25 (M = 23.2, STD = 1.69). All of the participants has experience in VR, and 5 participants are VR developers. 5 participants have experience visiting virtual galley.

doi:10.6342/NTU202203319

Frame Transformation

As shown in Figure 5.1(a), the average score of Q1 and Q2 is 4.1 and 4 respectively, we can conclude that most of the participants consider that the result of style transfer is better than the original one, The preference for two transferred results is not much different. 90% of the participant agree to the Q4, which means that it is easy for them to find the right style.

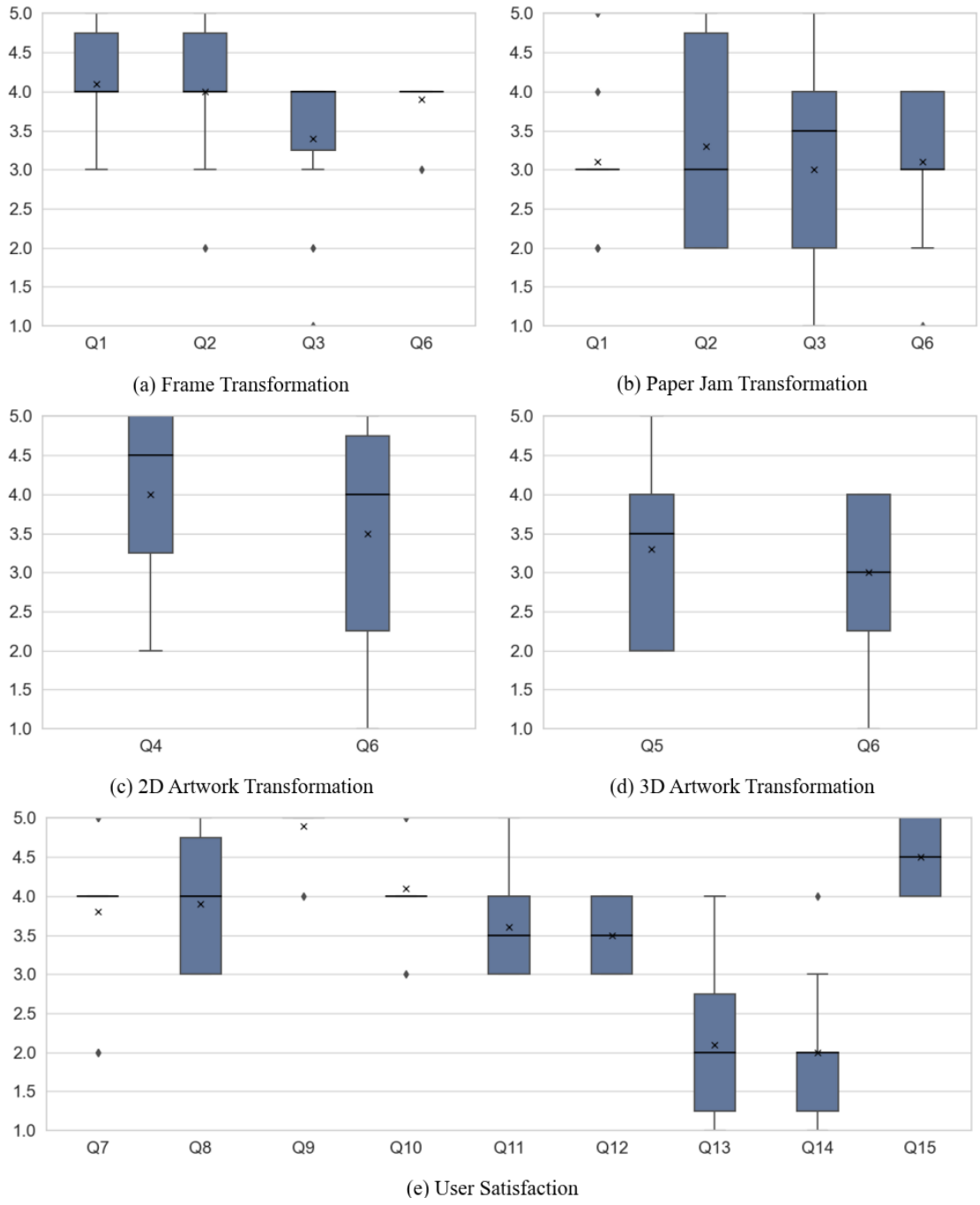
Paper Jam Transformation

As shown in Fig.5.1(b), the average score of Q1 and Q2 is 3.1 and 3.2, we can conclude that there was no significant preference between the transferred result and the original one.Although the average score of Q2 is slightly higher than Q1, it has higher variance.Based on the score of Q4, participants need to spend certain amount of time choosing the right style.

Arbitrary Transformation

As shown in Fig.5.1 (c).The average score of Q5 is 4, which means that most participant consider that the transferred result is better than the original one. Note that some of the them didn't perform style transfer on the picture. Based on the score of Q4, participants need to spend a certain amount of time choosing the right style due to the need of transferring three objects.

doi:10.6342/NTU202203319



(a) Frame Transformation

(b) Paper Jam Transformation

(c) 2D Artwork Transformation

(d) 3D Artwork Transformation

(e) User Satisfaction

Figure 4.2: Questionnaire results for user satisfaction.

3D Artwork Transformation

As shown in Fig.5.1 (c).The average score of Q6 is 3.3, and the number of participants who like the transferred 3D artwork is slightly higher than that of participants who do not like it. Base on the score of Q4, half of the participants need to spend more time to find

doi:10.6342/NTU202203319

the right style, and one participant strongly disagree to the Q4.

User Satisfaction Questionnaire

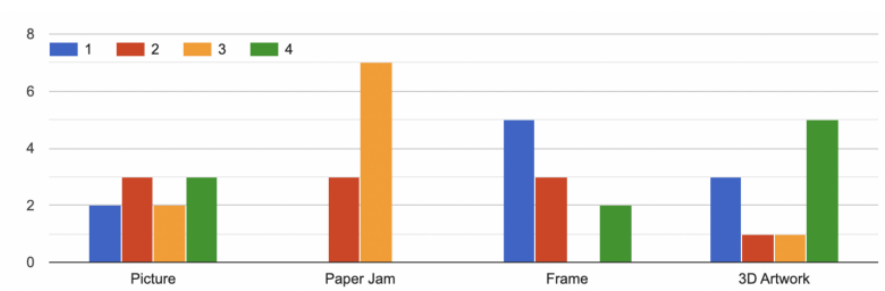
As shown in Figure 5.1 (d), we received high scores in Q7-Q10 and Q15. It is important to notice that 90% of the participants strongly agreed to the Q9 for fun level, and most of the participants agreed to the Q10. Based on the result of Q11 and Q12, we observed that participants have no preference on where to choose the style images. About the question regarding UI, although we got high score in Q13 and Q14, some participants said that the UI need to be improved. In addition, all participants preferred to appreciate the artwork in VR mode.

Preference Questionnaire

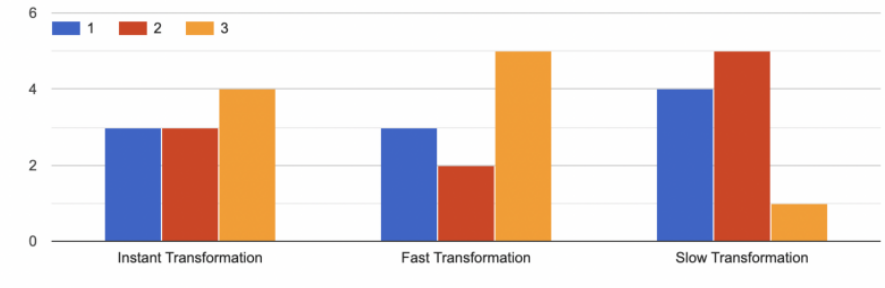
Participants ranked the objects they like to do style transfer based on their preferences from 1 to 4, with 1 being the most preferred. The result is shown in Figure 5.2(a) where the y-axis represents the number of people. We found that most of the participants like to perform style transfer on the frame most. 3D artwork ranked the lowest, with five participants giving it fourth place.

Participants ranked the visual effect of transformation process based on their preferences from 1 to 3, with 1 being the most preferred. The result is shown in Figure 5.2(b) where the y-axis represents the number of people. Most of the participants like slow transformation the most. The participants who like it consider that the speed of the slow transformation is just enough to clearly see the transformation process. Fast transformation ranked the lowest, with five participants giving it third place.

doi:10.6342/NTU202203319



(a)



(b)

Figure 4.3: Rank score of user preference.

System Usability Scale

We use System Usability Scale (SUS) for measuring the usability of our system. As shown in Figure 5.3, we got a score of 79.5 which is above the average SUS score 68.

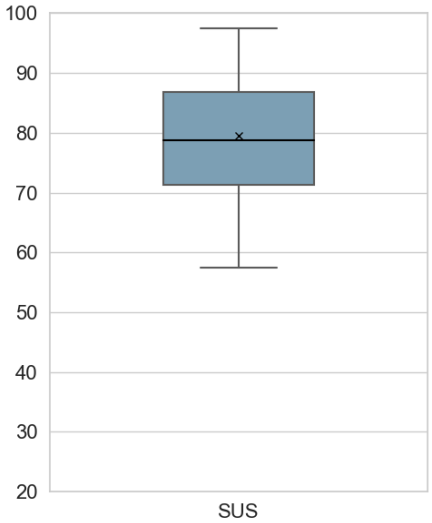


Figure 4.4: Result of System Usability Scale.

doi:10.6342/NTU202203319

Chapter 5 Discussion

In this chapter, we summarize the positive feedback and negative feedback of performing style transfer on different objects based on the experimental results and user interviews.

Frame Transformation

The frame is the object that participants like to perform style transfer most. According to the participants' feedback, we observed that most of the participants indicated that the color and style patterns of the transferred frame were consistent with the picture. Therefore, the overall result after style transfer is more harmonious and more artistic. Some participants also mentioned that if choosing a suitable style to transfer frame, it will not only increase the richness of the picture, but also highlight the content of the picture, which can be regarded as an artwork overall.

About negative feedback, one participant indicated that the frame has little effect on the overall visual effect because the frame accounts for a small proportion of the whole artwork.

doi:10.6342/NTU202203319

Paper Jam Transformation

Because the paper jam is stacked behind the picture and there is no obvious separation between the picture and the paper jam, most participants who prefer to stylize the paper jam choose a style image that is very different from the original picture because it can improve the contrast with the picture, and only one participant prefers to use the picture as style image for transformation because she said the transferred result extends the style of the picture and have more space for additional information, but even so, the overall number of participant who preferred to transform paper jams was less than those who preferred to transform frames. Most of the participants indicate that if the transferred result has more patterns or strong colors would make the artwork look messy and reduce the overall aesthetics.

Arbitrary Transformation

About positive feedback, one participant said that transforming different pictures into a particular artist style could give the pictures new life. With the virtual gallery display, it has the feeling of paying tribute to a particular artist's exhibition. Another participant said that using a style image with less color in it can degrade the saturation of the picture which brings more realistic feel, and the transferred paper jam increases a connection to the picture, enriching the overall storytelling of the artwork.

About the negative feedback, many participants said that if the pictures were transformed, it would affect the meaning that the original pictures were intended to convey, and it would be easy to create a cluttered feeling if the participant did not match the style

doi:10.6342/NTU202203319

of the three objects well.

3D Artwork Transformation

In preference questionnaire, 3D artwork was ranked lower than the other objects. Some participants said only certain styles are suitable for 3D artwork and sometimes the transferred result brings the feeling of inconsistency. The participants who like the transferred result said that the style transfer can add more colors to the original monotonous 3D artwork and enhance the overall look.

Visual effect of transformation process

Regarding the speed of the transformation process, participants who like slow transformation the most said that it is able to watch the process of transformation clearly. Two participants suggested that average the speed of slow transformation and fast transformation will be better. And the other two participants suggested that transforming slowly when approaching the artwork and staying with the transferred result until the user leaves the area because they consider that this method will not interrupt the process of appreciating the artwork and can also increase the naturalness of the experience.

doi:10.6342/NTU202203319

Chapter 6 Conclusion and Future Work

6.1 Conclusion

The results show that adding a style transfer to the virtual gallery can make the experience more interesting and deepen the viewer's impression of the artwork. Several of the participants gave the artwork a new story when viewing the transformed artwork, effectively increasing the connection between the viewer and the artwork. From our experimental results, we found that the frame transformation was more effective than other objects in improving the overall aesthetics of the artwork and was more acceptable to the viewers. On the other hand, the paper jam transformation and 3D artwork transformation have little effect on enhancing the aesthetics.

6.2 Future Work

Based on the response of the participants, the user interface should be further improved to make it more intuitive and easy to use. For instance, show how many images are in a category, and place the transferred button close to the slider.

doi:10.6342/NTU202203319

We can further develop the style transfer method in three directions. First, we found that the current style transfer method for the texture of 3D object still lacks development, a style transfer method considering the underlying UV-mapping of the 3D model and being able to transfer the vividness of style patterns at the same time is needed. Second, we can develop a method for changing the geometric of the frame to match the style of the picture. Third, develop a method to allow the color and stroke of the style image to be adjusted separately, and thus users can have more dimensions to adjust freely and produce more diverse and rich results.

In addition, considering that users may not be sensitive to color, they may feel that the result is not ideal but they do not know how to adjust it. To address this problem, we can develop a recommendation system to recommend the style and intensity that looks best with the current artwork according to color or texture, and thus users can get attractive results more easily. Moreover, we would like to recruit participants from different backgrounds, such as people with artistic backgrounds or professionals, to explore more about the views of different identities towards performing style transfer in the virtual gallery.

doi:10.6342/NTU202203319

References

[1] Cyber. Oncyber, 2021.

[2] MomentX-Network. Momentx, 2021.

[3] XRSPACE. Goxr, 2021.

[4] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2414–2423, 2016.

[5] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, Computer Vision – ECCV 2016, pages 694–711, Cham, 2016. Springer International Publishing.

[6] Vincent Dumoulin, Jonathon Shlens, and Manjunath Kudlur. A learned representation for artistic style, 2016.

[7] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Oct 2017.

[8] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, and Ming-Hsuan Yang.

doi:10.6342/NTU202203319

Universal style transfer via feature transforms. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc., 2017.

[9] Dae Young Park and Kwang Hee Lee. Arbitrary style transfer with style-attentional networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[10] Songhua Liu, Tianwei Lin, Dongliang He, Fu Li, Meiling Wang, Xin Li, Zhengxing Sun, Qian Li, and Errui Ding. Adaattn: Revisit attention mechanism in arbitrary neural style transfer. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 6649–6658, October 2021.

[11] Haibo Chen, lei zhao, Zhizhong Wang, Huiming Zhang, Zhiwen Zuo, Ailin Li, Wei Xing, and Dongming Lu. Artistic style transfer with internal-external learning and contrastive learning. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, Advances in Neural Information Processing Systems, volume 34, pages 26561–26573. Curran Associates, Inc., 2021.

[12] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. Artistic style transfer for videos. In Bodo Rosenhahn and Bjoern Andres, editors, Pattern Recognition, pages 26–36, Cham, 2016. Springer International Publishing.

[13] Xueting Li, Sifei Liu, Jan Kautz, and Ming-Hsuan Yang. Learning linear transformations for fast image and video style transfer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[14] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. Artistic style transfer for

doi:10.6342/NTU202203319

videos. In Bodo Rosenhahn and Bjoern Andres, editors, Pattern Recognition, pages 26–36, Cham, 2016. Springer International Publishing.

[15] Wenjing Wang, Jizheng Xu, Li Zhang, Yue Wang, and Jiaying Liu. Consistent video style transfer via compound regularization. Proceedings of the AAAI Conference on Artificial Intelligence, 34(07):12233–12240, Apr. 2020.

[16] Hiroharu Kato, Yoshitaka Ushiku, and Tatsuya Harada. Neural 3d mesh renderer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

[17] Alexander Mordvintsev, Nicola Pezzotti, Ludwig Schubert, and Chris Olah. Differentiable image parameterizations. Distill, 3, 07 2018.

[18] Jhih-Hong Hsu. Style transfer of 3d scenes, 2019.

[19] Mattia Segu, Margarita Grinvald, Roland Siegwart, and Federico Tombari. 3dsnet: Unsupervised shape-to-shape 3d style transfer, 2020.

[20] Kangxue Yin, Jun Gao, Maria Shugrina, Sameh Khamis, and Sanja Fidler. 3dstylenet: Creating 3d shapes with geometric and texture style variations. In Proceedings of International Conference on Computer Vision (ICCV), 2021.

[21] Reinhard Oppermann. User-interface Design, pages 233–248. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.

[22] John Brooke. Sus: A quick and dirty usability scale, 1996.

doi:10.6342/NTU202203319