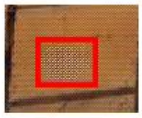
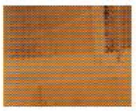
User study on the results of style transfer

Style transfer is a hot topic in the fields of multimedia, image processing and computer graphics. The ideal style transfer should transfer image with the style of style image while maintaining the content consistency with original image. We have prepared a series of results from baseline method and our method. Please carefully compare the results following specific standards below as your opinions will be used to evaluate the methods. There are four comparison items, each comparison item involves four comparison groups. Thank you for your cooperation.

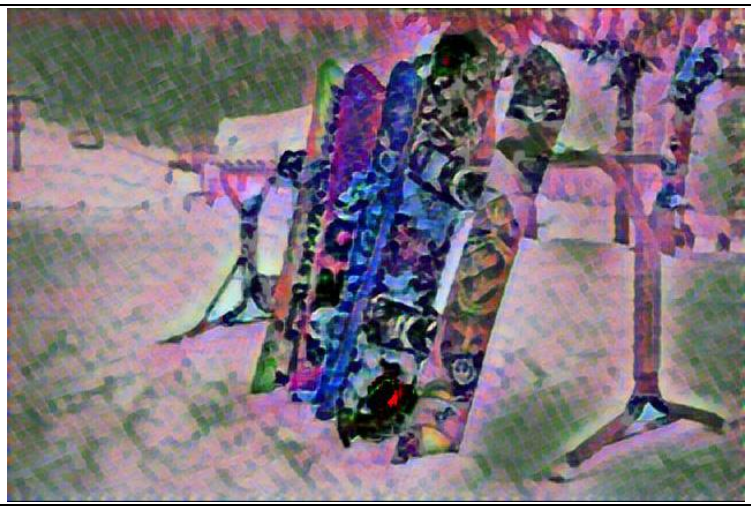
1.Image quality

Compare item	Image quality
Compare standard	Please follow standards below to compare results, and give your preference in the lower right column. Note: When comparing this item, please enlarge the image to compare as the noise will be more obvious after zoomed in Standards: (1) There is no discordant texture (noise) at details. (2) The picture does not have checkerboard effects. (3) There is no line-like noise in the images We have provided some noise examples for your reference, including but not limited to these.

	Mark √ to the prefer ones in the last column
example s	 checkerboard noise  checkerboard noise  global line-like noise
Group 1	Preference
A	
B	
Group 2	Preference

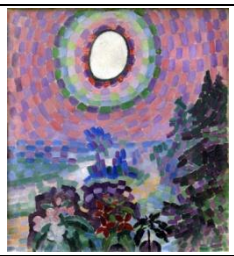
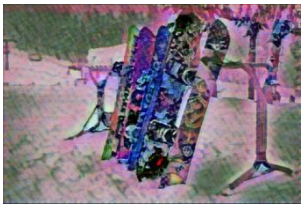
A		
B		
Group 3		Preference

A		
B		
Group 4		Preference
A		

B		

2. Saliency order preservation

Compare item	Saliency order preservation
Compare standard	Please follow the standards below to compare images in the table, the content images and the stylized image according to the style image, and select the images that you would prefer in each group. Standards: (1) The style images and the stylized images in the table have the same emphasis on important objects. For example, the core color of the style image is transformed to the core elements of the result image. The non-core parts have the colors and textures that correspond to non-core areas in the style image. (2) Core elements can be highlighted by the color and textures distribution. Mark √ to the prefer ones in the last column

Style image		Content image	Stylized image	
Group 1				Preference
A				
B				
Group 2				Preference
A				
B				
Group 3				Preference
A				
B				
Group 4				Preference

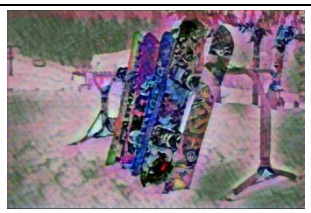
A				
B				

3. Image content preservation

Compare item	Image content preservation
Compare standard	Please compare the content images in the table with the stylized image according to the following standards, and select the result that you would prefer in each group. 1. The content image and the stylized image have the same content according to visual judgement. We provided some examples for reference. Mark √ to the prefer ones in the last column
examples	 This is an example of content being destroyed. The 3rd and 4th images from left to right are generated according to the content of 2nd image. The style image of the 3rd image is the 1st image. However, the human eyes in the 3rd image are

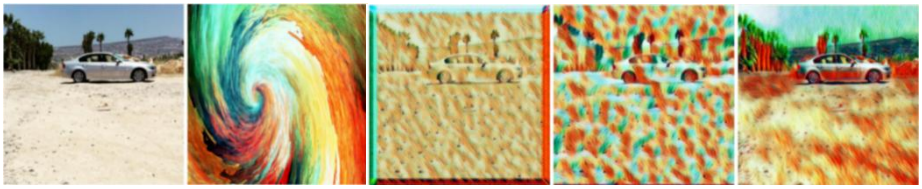
	enlarged, which makes the portrait looks more like the person in the 1st image, rather than the person in the 2nd image. Although the 4th image is also generated according to the content of the 2nd image, it is almost impossible to recognize the original content from the 4th image. Therefore, the preference of the 3rd ,4th images will be relatively low and the preference of 4th image should be lower than that of the 3rd image.

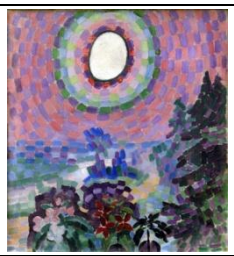
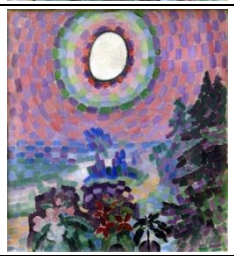
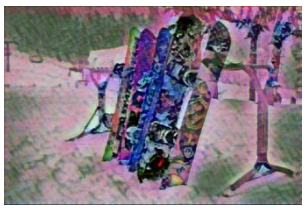
Content image		Stylized image	
Group 1			Preference
A			
B			
Group 2			Preference
A			
B			
Group 3			Preference

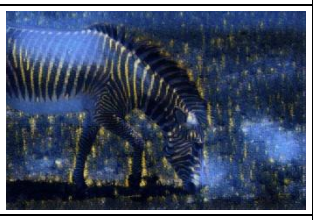
A			
B			
Group 4			Preference
A			
B			

4. Style similarity

Compare item		Style similarity
Compare standard	Please follow standards to compare the images in the table, the content images and the stylized images according to the style image, and select the result that you would prefer in each group. 1.The style images and stylized images have the same style. 2.The stylized images and the content images have the same content. 3. The stylized images are visually pleasing. We have provided corresponding examples for reference. Mark √ to the prefer ones in the last column	

examples	 Examples of relatively good and bad results are shown. The 1st image is the content image. The 2nd image is the style image. The 3rd, 4th, and 5th images show examples of poor and visually pleasing stylized results. The corresponding preference should increase in turn.

	Style image	Content image	Stylized image	
Group 1				Preference
A				
B				
Group 2				Preference
A				
B				

Group 3				Preference
A				
B				
Group 4				Preference
A				
B				

Results

Image quality	
Group 1	Number of Preference
Baseline(A)	17
our method(B)	46

Group 2	
Baseline(A)	28
our method(B)	35
Group 3	
Baseline(A)	26
our method(B)	37
Group 4	
Baseline(A)	35
our method(B)	28
Sum of groups	
Baseline	106
our method	146

Saliency order preservation	
Group 1	Number of Preference
Baseline(A)	23
our method(B)	40
Group 2	
Baseline(A)	23
our method(B)	40
Group 3	
Baseline(A)	23

our method(B)	40
Group 4	
Baseline(A)	36
our method(B)	27
Sum of groups	
Baseline	105
our method	147

Image content preservation	
Group 1	Number of Preference
Baseline(A)	31
our method(B)	32
Group 2	
Baseline(A)	22
our method(B)	41
Group 3	
Baseline(A)	21
our method(B)	42
Group 4	
Baseline(A)	42
our method(B)	21
Sum of groups	

Baseline	116
our method	136

Style similarity	
Group 1	Number of Preference
Baseline(A)	40
our method(B)	23
Group 2	
Baseline(A)	34
our method(B)	29
Group 3	
Baseline(A)	22
our method(B)	41
Group 4	
Baseline(A)	21
our method(B)	42
Sum of groups	
Baseline	117
our method	135