Sketch2Normal: Deep Networks for Normal Map Generation

Wanchao Su,[1] Xin Yang,[1,2] Hongbo Fu[1]

[1]School of Creative Media, City University of Hong Kong
[2]Dalian University of Technology

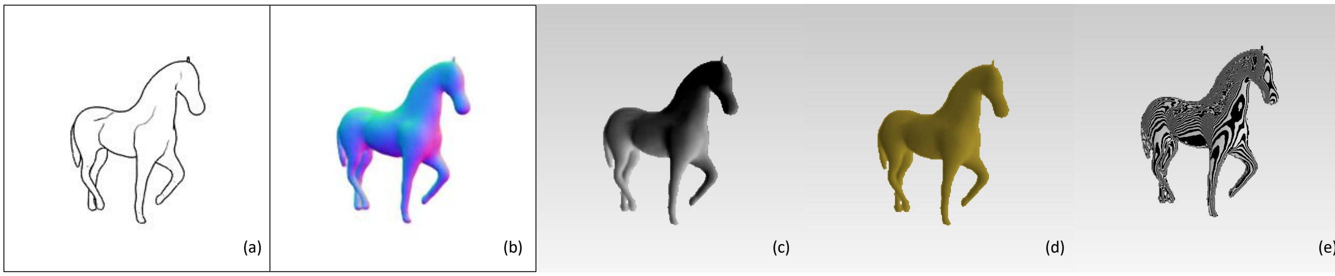

Figure 1: (a) The input sketch in our system. (b) The generated normal map. (c) The normal map with Phong illumination. (d) Render the normal map with dimple shader. (e) Rendering with reflexion lines.

KEYWORDS

Sketch, Normal Map, Generative Adversarial Network

ACM Reference Format:
Wanchao Su,[1] Xin Yang,[1,2] Hongbo Fu[1] [1]School of Creative Media, City University of Hong Kong [2]Dalian University of Technology . 2017. Sketch2Normal: Deep Networks for Normal Map Generation. In *Proceedings of SIGGRAPH Asia 2017 Poster, Bangkok, Thailand, November 2017 (SIGGRAPH Asia 2017),* 2 pages.
https://doi.org/10.475/123_4

1 INTRODUCTION

Normal maps are of great importance for many 2D graphics applications such as surface editing, re-lighting, texture mapping and 2D shading etc. Automatically inferring normal map is highly desirable for graphics designers. Many researchers have investigated the inference of normal map from intuitive and flexiable line drawing based on traditional geometric methods or deep networks while our proposed method shows more robustness and provides more plausible results.

We present a normal map generating system, using the sketch as input to produce a normal map. We treat the sketch-to-normal map generation problem as an image translation problem, utilizing a conditional GAN-based framework to "translate" a sketch image into a normal map image. Since there are ambiguities in the sketch-normal pairs, our system incorporates the user interactivity as the additional guidance in generating normal map according to the design. Users of our system can choose to add point normal to reduce the ambiguity of sketch and enlarge the design choices of the normal maps.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGGRAPH Asia 2017, November 2017, Bangkok, Thailand
© 2017 Copyright held by the owner/author(s).
ACM ISBN 123-4567-24-567/17/06…$15.00
https://doi.org/10.475/123_4

2 METHOD AND IMPLEMENTATION

We adapt our discriminator D and generator G from those in pix2pix [3]. Using the conditional Wasserstein GAN (WGAN), the objective function for D can be interpreted as identifying pixel-wise "realness". G here can treat the input as a guidance for generating a sample, which gives more information than a random noise. We use a mixed loss function here as:

$$L = L_{WGAN} + \lambda_{L_1} L_{L_1} + \lambda_{mask} L_{mask}. \quad (1)$$

Similar to ideepcolor [4], our system enables users directly assign point normals to specific locations in the input sketch. The selected point has the normal information (assigned RGB color) and the value of the corresponding location in the binary mask is set to one. In Eq. (2), we first differentiate the generated image and real image using L1 norm, then pixel-wisely product (noted as $\odot$) the residual image with input mask, and finally get the average of the result with respect to the number of input pixels.

$$L_{mask} = \frac{1}{\sum Mask} E[\|Real - Generated\|_1 \odot Mask]. \quad (2)$$

To pass more sketch information to the normal map, we incorporate the U-Net architecture in generator G. In order to get the corresponding sketch image and normal map, we propose to use a non-photorealistic rendering method [1] to generate suggestive contours as sketches in our system. To simulate the user additional input, we adopt the method in ideepcolor [4]. For each image, we use a geometric distribution to generate the number of input points. Each point location is sampled from a 2D Gaussian distribution of the normal map area (the color is not white).

3 EXPERIMENTS AND RESULTS

The main reference work for comparison in the experiments is pix2pix[3]. The qualitative results are illustrated in Fig. 2. We also

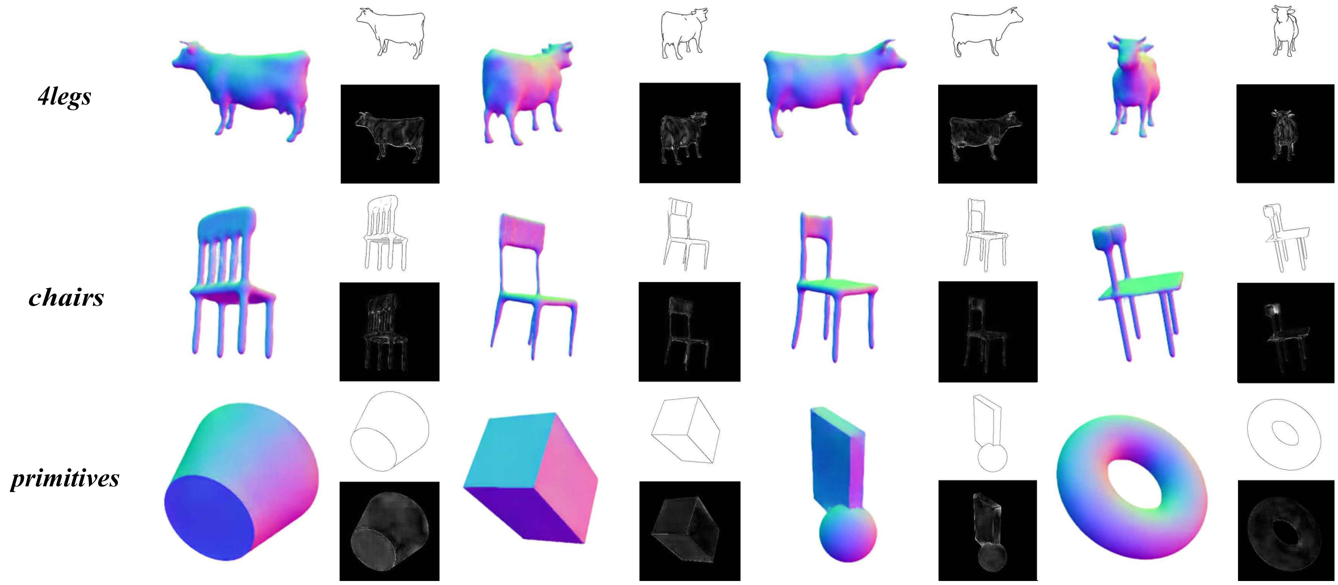

Figure 2: Different categories of results generated with our method. The loss images are visualized (in black) based on pixel-wise angualr loss value comparing to the ground truth, the corresponding inputs are illustrated above the loss map.

compare the results in a quantitative way: Measuring the difference between the result and the ground truth in manners of L1 distance, L2 distance and angular difference. The comparison details are presented in Table 1, and we also visualize the loss in Fig. 2.

Our method uses Wasserstein distance instead of the original GAN loss, enabling better results than pix2pix [3]. The WGAN setting can prevent gradient vanishing effectively, giving reliable gradient information to update the generator G. For the method WGAN-GP, there is no batch normalization in the implementation, which adds the instability in the training process.

Table 1: Different methods comparison. WGAN-GP is the same settings as our method except the basic GAN loss is the Wasserstein GAN loss with gradient penalty [2].

Loss Type	pix2pix	WGAN-GP	Ours
Angular	436.076	281.978	194.043
L1	3190.279	2174.009	1433.284
L2	611.189	304.980	102.521

In Table 1, we can see that our method outperforms the other two, showing significant lower loss values. Converting to degree value, our method achieves loss at pixel-wised maximum 24° and 0.17° on average. In Fig. 2, we show different results generated with our method. To further test the system, we use several hand drawn sketches to test the system, and the results are shown in Fig. 3. For completely new sketches, our system can provide plausible results.

4 CONCLUSION

In this paper, we introduce a novel method inferring normal maps from sketches with the conditional WGAN and user guidance. Our method shows superb performances than the previous approaches in both qualitative and quantitative way. Moverover, we rendered the normal maps under various shaders demonstrated in Fig. 1.

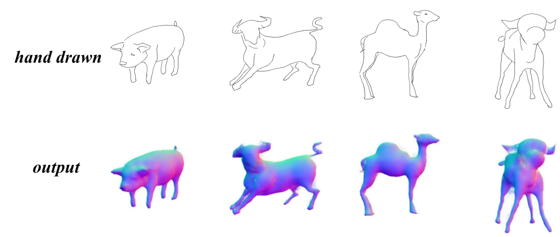

Figure 3: Test of hand drawn sketches in our system. Our system can infer reasonable normal maps with complete new sketch inputs.

ACKNOWLEDGEMENT

This work was partially funded by the RGC of Hong Kong SAR (Project No. CityU11204014), National Natural Science Foundation of China (Project No. 61632006), the Hong Kong Scholars Program, and supported by NVIDIA Corporation with the donation of the Titan Xp GPU.

REFERENCES

[1] Doug DeCarlo, Adam Finkelstein, Szymon Rusinkiewicz, and Anthony Santella. 2003. Suggestive contours for conveying shape. *ACM Transactions on Graphics (TOG)* 22, 3 (2003), 848–855.

[2] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. 2017. Improved training of wasserstein gans. *arXiv preprint arXiv:1704.00028* (2017).

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2016. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004* (2016).

[4] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. 2017. Real-Time User-Guided Image Colorization with Learned Deep Priors. *ACM Transactions on Graphics (TOG)* 9, 4 (2017).