Generating Colorized 3D Object Model with 2D Image Views

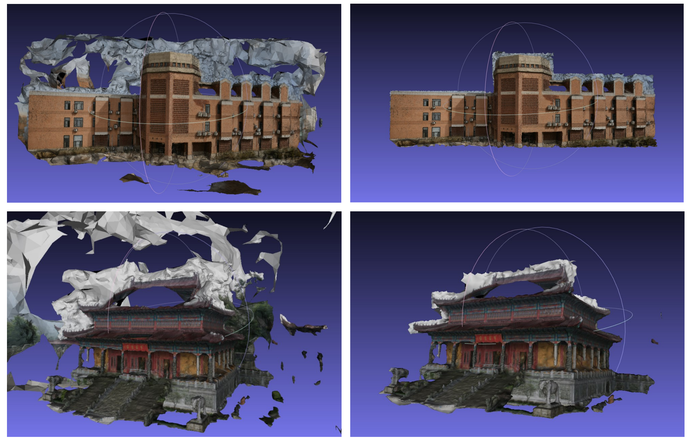
Tuan-Anh VU* Yingshu CHEN* Wishsmitha Samadhi MENDIS* Sai-Kit Yeung

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Introduction

It is difficult to automatically generate a 3D model with different types of lighting and shading to effectively develop virtual scenes for various industries, e.g. gaming, cinematography, etc. To make it efficient and feasible for variant illumination production within the same scenes, we propose a methodology to generate colorized 3D object models from 2D images.

1. realizes an automatically 3D reconstruction process via only a bunch of images of views, with alternative shading indicating different times in a day.
2. provides an easy way for game designers, movie producers and animators to transfer style (change the time of the scene from day to night without complex 3D rendering process).
3. eliminates useless background, foreground and restores detail contribute to high-quality 3D model reconstruction (example below).



Data Sources and Pre-processing

Input images: we grab video frames from a computer graphic architecture animation fly-through video as synthesized building images, and obtain images from [1].

Colorization reference images: we employ the frames in the time-lapse videos from Shih et al. [2] as the time target image and we only use one video for experiment in this paper. Other style reference images are searched from public domains (e.g. Pexels, Flickr, Google images, etc. and all rights reserved) manually with personal preference.

Data pre-processing: To get sufficient experiments within a short time, only 40 or so images were selected for colorization and 3D reconstruction for every single architecture. And as the machine limitation (8GB of system memory), we resize all selected images into images with 700 pixels as maximum width or height for colorization.

Methodology and experimental results

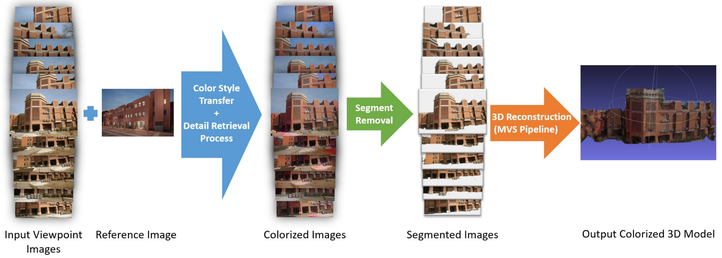


Image Color Style Transfer: The basic pipeline of color style transfer follows the Deep Photo Style Transfer algorithm which proposed by Luan et al. in [3]. As Luan's method cannot process images under large size with basic machine equipment. To assure the resolution of final reconstructed model, we post-process all colorized images under basic small size (up to 700 pixels as width or height) into a larger size (e.g. 1920px x 1280px). We utilize a photorealistic style transfer method proposed by Mechrez et al. [4] to further make the colorized image more realistic.

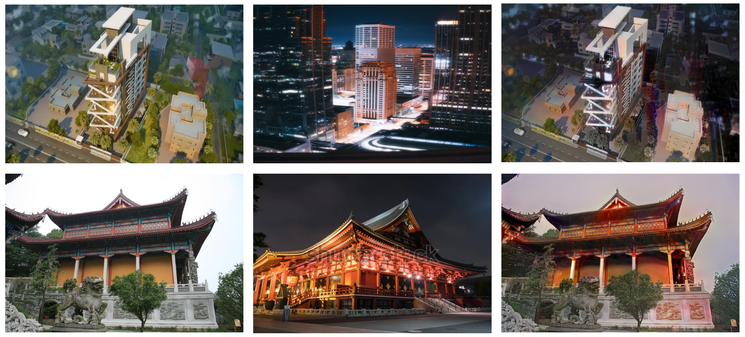
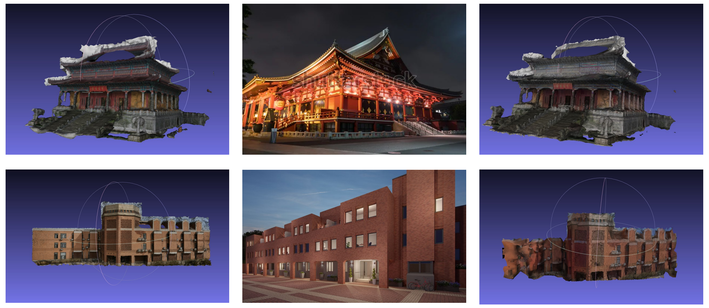


Image Segmentation: we used the novel model ResNet-50-dilated [5] for Encoder and Pyramid Pooling Module with loss optimization from PSPNet [6] for Decoder. The reason to split the model into decoder and encoder is that encoders are usually modified directly from classification networks, and decoders consist of final convolutions and up-sampling.

3D Reconstruction: we used photogrammetry based Multi-View Stereo (MVS) approach [7] for 3D reconstruction with the input is the segmented images. The main pipeline includes camera alignment, feature extraction and image matching, structure from motion, depth map generation, meshing and texturing using the colorized images.

According to the results, there is a slight loss of features in the colorized image compared to the original image. That is mainly due to the reduction of the resolution of viewpoint images after the colorization process. When the input image resolution is higher, the more features are extracted from the images and the reconstructed 3D model is much more detailed.



Conclusion

The colorization method we utilize now is promising as it can generate a plausible and usable 2D colorized image for later texturing use. The implemented flow in our methodology is capable of generating 3D models with desired color style transfer with a set of viewpoint images. However, as discussed in the above section, the details and the features of the reconstructed model might reduce compared to the original model due to the loss of features and resolution due to colorization. In the future, this problem can be addressed by introducing a new texturing method to the original 3D model itself based on the reference images for the color transfer.

References

[1] 3d reconstruction dataset. Data retrieved from Robot Vision Group at the National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences (CASIA), http://vision.ia.ac.cn/data

[2] Y. Shih, S. Paris, F. Durand, and W. T. Freeman. Data-driven hallucination of different times of day from a single outdoor photo. ACM Transactions on Graphics (TOG), 32(6):200, 2013.

[3] F. Luan, S. Paris, E. Shechtman, and K. Bala. Deep photo style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[4] R. Mechrez, E. Shechtman, and L. Zelnik-Manor. Photorealistic style transfer with screened poisson equation. Procedings of the British Machine Vision Conference 2017, 2017.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. CoRR, abs/1512.03385, 2015.

[6] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.

[7] M. Goesele, B. Curless, and S. M. Seitz. Multi-view stereo revisited. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006.