Glare Detection and Inpainting in Facial Imaging Using Deep Learning Methods

Dennis Chao

A-TOP AI Innovation, 2024 A-TOP Summer Internship

Abstract

3D Facial imaging is used in various fields, such as human-computer interaction, security, animation, and healthcare (eg., plastic surgery simulations). A-TOP's 3D face reconstruction is based on Structure-from-Motion (SfM) and Multi-View Stereo (MVS) technique. However, this kind of reconstruction method requires continuous image capture from different perspective and may result in glare spots on the nose region due to light reflections and skin conditions. One problem identified is that these glare spots are incompatible with the reconstruction pipeline, because they would lead to broken parts when creating the 3D shape of the face, especially serious on the nose region. The main objective is to develop a solution that can automatically detect glare spots on the nose and repair them with the realistic skin color.

Introduction

Image inpainting is a process used to fill in or reconstruct missing or damaged parts of an image in field of Computer Vision. Traditionally, OpenCV[1][2] uses a non-linear partial differential equation that propagates information from the surrounding pixels into the missing or damaged regions. However, one of its biggest limitations is that the inpainting color will always be affected by the surrounding color. It may result in inaccuracy when filling in a region that should not be in the same color as its surrounding. As a result, some advanced inpainting methods that uses deep learning methods, such as Large Mask model (LaMa)[3], are able to create better inpainting results. LaMa uses a combination of Fast Fourier convolutions (FCC) and attention mechanisms to effectively handle both local and global image contexts. This allows it to generate plausible content for large missing areas while maintaining coherence with the surrounding image.

Methodology

Our proposed method consists of two parts: glare spots detection and LaMa inpainting. Glare spots detection first loads a pre-trained face parsing model. Once the image is inputted, it then process the image using the model and parses the nose region. To detect glare spots, we first converts the original image to gray scale, and then identifies glare spots by applying a brightness threshold and combining this with the nose mask to focus only on glare spots on the nose. The LaMa inpainting process the image through multiple layers of FFCs and attention modules. After going through the process, the output is a completed image with the masked region filled in.

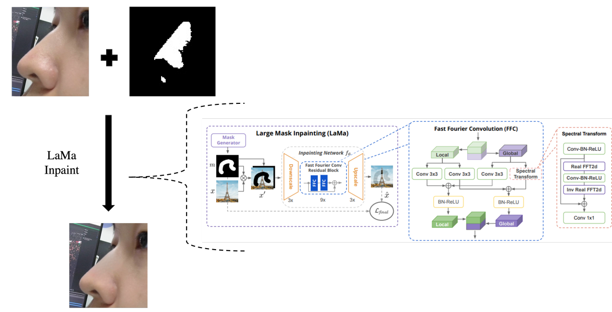

Figure 1. An overall workflow for our proposed method.

Results

The resulting images are shown as below. From left to right: The first picture is OpenCV's NS Result, the second picture is OpenCV's TELEA Result, and the third picture is LaMa's Inpaint Result. It's obvious that the results from OpenCV is greatly affected by the background color while the result using the LaMa model maintains the overall features of the image.

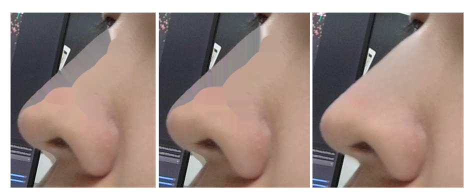

Figure 2. Comparison between OpenCV's inpaint algorithms and our solution.(The results are OpenCV's TELEA, OpenCV's NS, and our proposed method from left to right respectively.)

References

[1] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.

[2] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, pp. I–I, IEEE, 2001.

[3] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," *arXiv preprint arXiv:2109.07161*, 2021.