Automated 3D Reconstruction of the Retina Surface from Monocular Fundus Images

A. Du[1], B. M. Williams[1]

[1] *School of Computing and Communications, Lancaster University, LA1 4WA.*

Introduction

The accurate diagnosis of certain diseases, such as Glaucoma, requires an understanding of the three-dimensional (3D) surface of the optic disc and surrounding area. Recently, considerable attention has focused on developing automated methods for disease diagnosis and the interrogation of medical images in ophthalmology, which can also benefit from three-dimensional information that is not provided by common modalities such as fundus imaging. While technologies such as optical coherence tomography (OCT) can provide 3D surface information, it remains expensive and so prohibitive for many settings, including developing nations and primary care. 3D reconstruction from 2D images is a challenging task [1] and traditionally requires multiple viewpoints. While stereoscopic fundus images can be achieved, this is not always the case, particularly with emerging modalities such as mobile fundus imaging, which holds great potential for supporting developing nations. Reconstruction from single images is an even more challenging task that has been tackling by recent work in other domains, such as hand reconstruction for pose estimation, supporting the idea that depth information is implicitly encoded in and can be extracted from images. In this work, we aim to estimate the 3D surface information from monocular images of the retina, such as those achievable from lower-cost devices such as fundus cameras and ultimately mobile fundus cameras. We develop a dataset comprising images of the retina and corresponding depth information that we derive from collocated OCT images, that we segment automatically. We develop an approach for determining the correspondence between the photographic and surface depth information and evaluate this on a testing set or retinal images.

Methods

A significant challenge for this work is the availability of corresponding retinal image and surface information data. We develop a dataset of 32 patient eyes, including OCT information and corresponding images of the fundus, including the location of the slices. We extract surface height information by training a encoder-decoder segmentation approach on a larger dataset of OCT images to determine the front surface of the nerve fiber layer. We establish a world coordinate system, integrating the retina image and its corresponding surface height information. Based on this dataset, we modify a U-Net [2] segmentation network to a regularised regression task to predict the pixel-level depth value of each pixel. We develop an energy functional comprising L1 and L2 loss to train the network [3], and optimise the lr value and drop strategy.

Results & Discussion

We evaluate our approach on a separate subset of 12 fundus images with depth maps determined by our segmentation approach. Our surface reconstruction yields good results in estimating the depth at each pixel as demonstrated in Figure 1. We evaluate our results numerically using pointwise mean absolute distance, achieving a value of 17.24 for the example shown, which is largely impacted by the difficulty in reconstructing the boundary information.

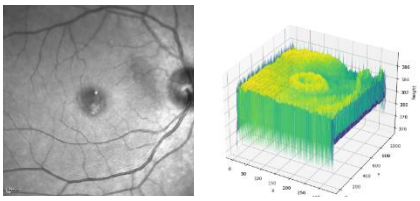

Figure 1: fundus image (left) and surface reconstruction (right).

Conclusion

In this work, we have developed a novel dataset including retinal images and corresponding depth values achieved from OCT images, and a method of depth estimation for 3D surface reconstruction. Further, we have demonstrated the feasibility of extracting depth information from monocular fundus images for 3D reconstruction of the retinal surface. While more work is needed to improve this proof-of-concept in terms of noise reduction and boundary values, this step forward is an exciting move towards achieving accurate 3D surface reconstruction of the retina from low-cost fundus imaging devices, which has potential to improve the automated diagnosis and analysis of diseases such as Glaucoma.

References

[1] Alhashim and Wonka. High quality monocular depth estimation via transfer learning. arXiv, 2018.
[2] Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. MICCAI, 2015.
[3] Qian et al. Sgas: Sequential greedy architecture search. CVPR, 2020.