

論文内容の要旨

博士論文題目 CNN-based Scene Modeling: From Depth Estimation to 3D Reconstruction

氏名 Niu Zhaofeng

(論文内容の要旨)

With the rapid improvement of image sensors and computer vision technologies, scene modeling has become more and more popular. For applications like augmented reality (AR) and virtual reality (VR), quick and precise 3D reconstruction of the real world, which describes actual objects in data format that can be used for displaying and computing, is necessary. Therefore, researchers have paid lots of attention to develop an efficient yet accurate scene modeling method and have achieved many

encouraging results. However, most of these methods are based on expensive depth sensors and are vulnerable to depth noises, which largely limit their application areas. In the dissertation, the author focuses on the design and implementation of CNN-based scene modeling, in order to get rid of expensive depth sensors and to make a more robust reconstruction method.

Firstly, the author focuses on developing a new monocular depth estimation method for easing the task of depth acquisition. Monocular depth estimation is an essential technique for tasks like 3D reconstruction. Although many works have emerged in recent years, they can be improved by better utilizing the multi-scale information of the input images, which is proved to be one of the keys in generating high-quality depth estimations. The author proposes a new monocular depth estimation method named HMA-Depth, in which the author follows the encoder-decoder scheme and combines several techniques such as skip connections and the atrous spatial pyramid pooling. To obtain more precise local information from the image while keeping a good understanding of the global context, a hierarchical multi-scale attention module is adopted and its outputs are combined to generate the final output that is with both good details and good overall accuracy. Experimental results on two commonly-used datasets prove that HMA-Depth can outperform the existing approaches.

Then the author attempts to develop a new 3D reconstruction method that is robust to depth noises. The truncated signed distance function (TSDF) fusion is one of the key operations in the 3D reconstruction process. However, existing TSDF fusion methods usually suffer from the inevitable sensor noises. The author proposes a new TSDF fusion network, named DFusion, to minimize the influences from the two most common sensor noises, i. e., depth noises and pose noises. To the best of the author's knowledge, this is the first depth fusion for resolving both depth noises and pose noises.

DFusion consists of a fusion module, which fuses depth maps together, as well as a following denoising module, which removes both depth noises and pose noises for TSDF volumes. To utilize the 3D structural information, 3D convolutional layers are used in the encoder and decoder parts of the denoising module. Also, a specially-designed loss function is adopted to improve the fusion performance in object and surface regions. The experiments are conducted on a synthetic dataset as well as a realscene dataset. The results prove that the proposed method outperforms existing methods.

氏 名	Niu Zhaofeng
-----	--------------

(論文審査結果の要旨)

令和4年8月30日に本博士論文の最終審査を行った。その結果、本博士論文は、提出者が独立した研究者として研究活動を続けていくための十分な素養を備えていることを示すものと認める。

Niu Zhaofengさんは、本論文において、コンピュータビジョン分野の幾何学的問題を機械学習に基づく方法で解くことを試みた。RGB画像から三次元シーンを復元する問題は、この分野の伝統的な問題であり、さまざまな手法が提案されてきた。この論文では、この問題を、RGB画像から奥行き推定を行い、奥行き情報を含むRGB-D画像を生成する問題と、一つのシーンを異なる視点から観測することで得られた複数のRGB-D画像を統合することで三次元シーンの立体形状を復元する問題に分割し、それぞれにおいて、新たな方法を提案し、その有効性を確認した。

前者に関しては、入力画像からマルチスケール画像を生成し、それぞれのスケールごとに奥行き推定とアテンション領域の抽出を行い、それらを統合する方式を提案し、実験により、その方法が従来方法より高精度に奥行き推定ができることを示した。この問題はコンピュータビジョン分野の多くの研究者が取り組んでいる問題であり、その中で、従来手法よりよい性能を持つ手法を提案できたことは評価できる。

後者に関しては、高精度な復元が可能な従来手法が存在したが、それがカメラの位置姿勢推定誤差を考慮していなかった点に注目し、カメラの位置姿勢推定に誤差が含まれる場合に、その従来手法の性能が劣化することを発見した上で、その問題を解決する改良法を提案した。三次元シーンの立体形状復元問題では、カメラの位置姿勢推定に誤差が含まれることは普通であり、その条件においてより効果的な手法を提案できたことは、高く評価できる。

以上の成果から、本論文は、博士(工学)の学位論文として、メディア情報学分野における十分な学術的価値を有することが認められると判断した。