

Img2Avathar: A Novel Hybrid Ensemble Approach For 3D Object Reconstruction from Multi-View Monocular RGB images

Ajith Balakrishnan ¹ Dr. Sreeja S²

APJ Abdul Kalam Technological University
ajith.b06@wipro.com, sreejasat81@gmail.com

Abstract

Generating 3D models from multi-view 2D RGB images has got high attention in the past few years. Some of the recent successful methods (e.g., 3D-R2N2) uses Recurrent (RNN) networks for fusing the 2D images extracted by shared architecture of encoder and decoder network. Due to the permutation variant property of RNN these methods does not delivers good results. Also these methods fails with more number of inputs due to long term memory loss of RNN network. Other widely used methods like pooling captures only first order information's and ignores valuable features. The method called as **Attsets** with unique training algorithm called as **JTSO** provides much better results with permutation invariant property. Attsets replaces RNN module with Attnet and outputs an improved reconstruction of the objects in the form of 3D occupancy grid. To improve the reconstructed output we propose an extension to the architecture Attsets which is called as **Img2Avathar**. we present a novel network called as **Refiner** inspired from the architecture called **3D U-Net** and an improved algorithm to train the network which is called as **STSO-JTSO**. In this paper we evaluated various properties of our network with large public data sets. The proposed method executes much faster than 3D-R2N2 in terms of inference. The proposed network outperforms state-of-the-art methods for single and multi view reconstruction and creates 3D models for the objects wherever traditional SFM/SLAM methods fails due to viewpoints separated by large margin.

Keywords

Multi view 3D model reconstruction, auto-encoder, Attsets, Img2Avathar, STSO-JTSO, 3D U-Net, 3D-R2N2