Optimal viewpoint prediction is an essential task in many computer graphics applications. Unfortunately, common viewpoint qualities suffer from two major drawbacks: dependency on clean surface meshes, which are not always available, and the lack of closed-form expressions, which requires a costly search involving rendering. To overcome these limitations we propose to separate viewpoint selection from rendering through an end-to-end learning approach, whereby we reduce the influence of the mesh quality by predicting viewpoints from unstructured point clouds instead of polygonal meshes. While this makes our approach insensitive to the mesh discretization during evaluation, it only becomes possible when resolving label ambiguities that arise in this context. Therefore, we additionally propose to incorporate the label generation into the training procedure, making the label decision adaptive to the current network predictions. We show how our proposed approach allows for learning viewpoint predictions for models from different object categories and for different viewpoint qualities. Additionally, we show that prediction times are reduced from several minutes to a fraction of a second, as compared to state-of-the-art (SOTA) viewpoint quality evaluation. We will further release the code and training data, which will to our knowledge be the biggest viewpoint quality dataset available.

http://dx.doi.org