Waterloo Exploration Database: New Challenges for Image Quality Assessment Models

Abstract:

The great content diversity of real-world digital images poses a grand challenge to image quality assessment (IQA) models, which are traditionally designed and validated on a handful of commonly used IQA databases with very limited content variation. To test the generalization capability and to facilitate the wide usage of IQA techniques in real-world applications, we establish a large-scale database named the Waterloo Exploration Database, which in its current state contains 4744 pristine natural images and 94 880 distorted images created from them. Instead of collecting the mean opinion score for each image via subjective testing, which is extremely difficult if not impossible, we present three alternative test criteria to evaluate the performance of IQA models, namely, the pristine/distorted image discriminability test, the listwise ranking consistency test, and the pairwise preference consistency test (P-test). We compare 20 well-known IQA models using the proposed criteria, which not only provide a stronger test in a more challenging testing environment for existing models, but also demonstrate the additional benefits of using the proposed database. For example, in the P-test, even for the best performing no-reference IQA model, more than 6 million failure cases against the model are “discovered” automatically out of over 1 billion test pairs. Furthermore, we discuss how the new database may be exploited using innovative approaches in the future, to reveal the weaknesses of existing IQA models, to provide insights on how to improve the models, and to shed light on how the next-generation IQA models may be developed. The database and codes are made publicly available at: https://ece.uwaterloo.ca/~k29ma/exploration/.