# Abstract

Being able to accurately predict human perceived quality of multimedia content (e.g., images and videos) through computer algorithms is an important and hot research topic, since it directly leads to several advantages: storage capacity and bandwidth savings for large organizations, and possibility to offer higher quality services when the communication channel capacity is limited. This PhD thesis focuses on using machine learning algorithms and advanced statistical methods to find new ways of addressing the problem.

Human viewers can easily distinguish between an image with a pleasant perceptual quality and another one whose quality is corrupted by artifacts introduced for instance by a codec. This task that is straightforward for a human due to his daily experiences and background, is absolutely not easy for an algorithm. For this reason, in the literature, human opinion scores on the perceptual quality of a multimedia content are considered as ground truth data. Any quality prediction algorithm (QPA) is then designed to predict the so-called Mean Opinion Score (MOS) of a multimedia content, i.e., the average of the scores that a group of human viewers would express if they were asked to watch that content and rate its perceptual quality on a given numerical scale.

Several QPAs aiming at MOS prediction have been proposed in the literature but the state-of-the-art in media quality assessment is still suffering some crucial limitations. Some of these limitations can be summarized as follows: i) none of the existing QPAs is accurate in all application scenarios, and thus there is still large room for their improvement; ii) QPAs, as estimators of the MOS, do not provide a complete measure of quality-of-experience (QoE). In fact, QPAs simply measure the quality perception of an "average viewer" and only marginally account for the individual expectation and uncertainty that characterize different users of the same service. For instance, QPAs do not provide answers to the following questions: what is the percentage of final users that would not be satisfied by the perceptual quality of a given video sequence? What are the characteristics of the final users that are expected to not be satisfied? The answers to these two questions are obviously of primary importance for any company that produces and markets multimedia content. There is therefore a crucial need to go beyond the MOS as suggested by recent publications in the literature.

This thesis contributes to advancing the state-of-the-art by proposing some solutions to cope with the aforementioned state-of-the-art limitations. Such proposals can be summarized under three main items.

- **The perceptual quality as a random variable**: A large number of stochastic influence factors, e.g. the subject's emotional state and the way each subject interprets the quality scale, that are not under the control of the test designer, influence the scores of a human subject during a subjective test. To jointly account for the inaccuracy of existing QPAs and these stochastic influence factors, both the prediction of a QPAs and the MOS are defined as random variables. A probabilistic approach is proposed to find the quality range to which the perceptual quality of a given processed video sequence (PVS) is expected to belong with a user specified probability.

- **Measuring the reliability of a MOS prediction**: since existing QPAs are not always accurate, integrating their prediction with an index that informs on how reliable the predicted quality score is, acquires a significant importance. Two different approaches to cope with this issue are proposed in this thesis. The first one is based on machine learning and it aims at predicting the intrinsic ability of a PVS to confuse human viewers and hence the QPAs as they are trained to predict human ratings. The second approach is instead based on the level of disagreement between many different QPAs when used to assess the quality of a given PVS. In fact, an index to measure the disagreement of QPAs is proposed and it is shown that such an index allow to distinguish between the cases in which QPAs are accurate and when they are not.

- **Artificial Intelligence-based observers**: finally this thesis presents a more complete approach to quality assessment. Unlike traditional approaches, the proposed one allows to fully consider individual expectations when automatically assessing the perceptual quality of multimedia content. Instead of predicting the MOS for each content under evaluation, as it is traditionally done in the literature, a different direction is proposed. An artificial neural network is trained to mimic an individual observer in terms of quality perception. Such a neural network can then be considered as a "virtual observer", and it is called an artificial intelligence-based observer (AIO). A large number of AIOs can then be trained, each representing an actual observer with well known characteristics and expectations. The advantage of this approach is to be able to more accurately model the distribution of the opinion scores that form the quality prediction as well as the uncertainty that intrinsically characterizes human viewers. The proposed approach therefore allows to: i) predict the percentage of viewers that would not appreciate the perceptual quality of a given processed content; ii) make inference on the characteristic of the unsatisfied viewers; iii) perform the simulation of subjective tests.