Adapting ISO 20462 Softcopy Quality Ruler Method for on-line Image Quality Studies
Peter D. Burns, Jonathan B. Phillips¹ and Don Williams²
Burns Digital Imaging, ¹NVIDIA Corporation and ²Image Science Associates

ABSTRACT
The ISO 20462 method for subjective image quality evaluation relies on a set of reference images, which are calibrated in terms of known absolute quality differences. The original reference images, Standard Reference Stimuli (SRS), were in the form of photographic prints, to be viewed under controlled lighting and fixed viewing conditions. This method was then extended for scenes presented on computer monitors as Digital Reference Stimuli (DRS). This softcopy ruler method was developed as part of the Camera Phone Image Quality (CPIQ) Initiative and has now been adopted as an updated ISO 20462 method (ISO 20462-3:2012). This extended method of using the softcopy rulers was validated by CPIQ using two systems of laboratory lighting and display. That effort, while successful, required significant effort and resources to accomplish. Chief among these was gathering a sufficient number of qualified viewers who could commit to viewing images on a narrowly defined schedule at a limited number of laboratories. In our study, we investigate whether and to what extent the ISO 20462 softcopy ruler method can be adapted to Internet-based subjective evaluations. Our objective is to develop and test a method that uses a commercial online survey service. We describe several limitations to be overcome, including image file size, and a static interface rather than one allowing dynamic updating of the reference image. The method that we developed and tested uses reference anchor images rather than the slider-selected reference. However, our anchors were drawn from the ISO 20462 set, and therefore were taken as calibrated reference images, albeit viewed under uncontrolled conditions. We describe the verification study that was completed using Survey Monkey®, and compare results with the corresponding Softcopy Ruler data. Similar results were obtained for observer ratings and their scene-dependency. We conclude that crowdsourcing is useful, particularly when common non-laboratory image viewing is the intent. When calibrated subjective image quality measures are needed, our adapted method should be considered an efficient alternative to the ISO 20462 standard, provided that common reference images are used.