Report on the Evaluation of Current and Future Image Compression Technologies

Picture Coding Symposium, Cairns, Australia

June 2015

Edited by Fernando Pereira with contributions by Ralf Schaefer, Touradj Ebrahimi, Joern Ostermann, Ed Delp

This document reports the conclusions, comments and recommendations resulting from the final panel discussion which happened at the Session on Evaluation of Current and Future Image Compression Technologies held at the Picture Coding Symposium, Cairns, Australia, in June 2015.

The Session on Evaluation of Current and Future Image Compression Technologies comprised three parts:

- Presentation of the proposed coding solutions, notably:
    - ***New coding solution*** labelled as Dalaa from Mozilla Foundation presented by Nathan Egge
    - ***JPEG2000 compliant solution with perceptual optimization*** from Univ. New South Wales presented by David Taubman
    - ***JPEG-compliant solution that creates images with lower rate and minimal induced distortion*** from University of Southern California presented by Antonio Ortega
- Presentation of the objective and subjective quality assessment by Touradj Ebrahimi on behalf of the QUALINET consortium; the quality assessment included the proposed solutions as well as many standard solutions, e.g. JPEG, JPEG 2000, JPEG XR, HEVC Intra
- Panel discussion moderated by Fernando Pereira with Ed Delp (Purdue University, US), Ralf Schaefer (HHI, Germany) and Joern Ostermann (Leibniz Universität Hannover, Germany) as panelists.

This document has annexes containing the documents which were submitted with the proposed coding solutions as well as the four presentations made; this material is also available at http://www.pcs2015.org/.

The discussion following the presentations led to the following conclusions and remarks:

- It is important to keep in mind that codec implementations and not codecs in abstract are being assessed and compared.
- It was great that software crosschecking has been performed.
- The confidence interval bars on the subjective tests were relatively large such that it is not possible to draw too many conclusions. An analysis of the reasons for this would be helpful. A possibility is to increase the number of subjects involved.
- It was not easy to fully understand the results based on the presented charts. Some further charts would be helpful, e.g. mean results for the complete set of pictures, and

ranking of the coding methods over all pictures e.g. for lowest, middle and highest data rates.

- PNSR results were as had been expected. The relatively consistent results for PSNR and the other objective quality measures was also expected. Some codecs show huge PSNR variations.
- The better performance of the Daala solution with the subjectively weighted metrics is surprising, especially as it has not been confirmed by the subjective test (except for "Woman").
- The sometimes inferior results of the reference in the subjective test lead to some doubts about the selected method.

Also some recommendations for the future were made:

- If coding performance with respect to fidelity is to be judged then a different method (e.g. comparison) should be used, because otherwise it will be difficult to distinguish between coding efficiency and post-processing.
- The subjective test has to be further optimized. Goal should not be "which image do you like best" but "which image comes closest to the original". Hence, for future work, we need a much more sophisticated test procedure.
- There should be more emphasis on the features offered by each codec and characteristics such as complexity, power consumption, etc. beyond compression efficiency.
- Many of the codecs claim to be able to cope with more than 8-bit images. They should be compared to each other also for such content.
- Too few images were used in objective and subjective comparisons. A much larger set of content, with mean RD performance and standard deviation across content for at least the objective metrics would be useful. Also some gray scale only images should be considered.
- Presentation of interesting cases (for instance for which type of content a given codec performs best or fails) will be useful to show in future.
- More than one encoder per codec should be explored as many of the codecs have different encoder implementations; even the old JPEG has several encoders such as two-pass versus one-pass, etc.
- Better evaluation of chrominance (both subjective and objective) and a comparison of the distribution of bit budget between luminance and chrominance in codecs will be useful information to have.

It was mentioned that the assessment work performed by QUALINET was remarkable.

Finally, it was recommended to retain this or a similar format session for the next Picture Coding Symposium. However, if similar events are to be organized at future Picture Coding Symposia, it could be useful and even desirable if the attendees can have a chance to actually participate in subjective evaluations of the content prior to the panel and while at the Picture Coding Symposium. This will give them a chance to see for themselves the type of distortions that codecs exhibit and have a better feeling on the adequacy of the subjective

evaluation used. Research groups should be encouraged to bring new ideas compression ideas to the Picture Coding Symposium without worrying about how they do relative to the latest standard.