Compression of Infrared images

Claire Mantel, Søren Forchhammer, Departments of Photonics Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

Abstract

This paper investigates the compression of infrared images with three codecs: JPEG2000, JPEG-XT and HEVC. Results are evaluated in terms of SNR, Mean Relative Squared Error (MRSE) and the HDR-VDP2 quality metric. JPEG2000 and HEVC perform fairy similar and better than JPEG-XT. JPEG2000 performs best for bits-per-pixel rates below 1.4 bpp, while HEVC obtains best performance in the range 1.4 to 6.5 bpp. The compression performance is also evaluated based on maximum errors. These results also show that HEVC can achieve a precision of $1°C$ with an average of 1.3 bpp.

Introduction

Thermal cameras have mainly been used in military or medical applications, however, decrease in the price of infrared sensors makes it affordable for a much wider range of applications: inspection of houses or energy systems, search and rescue operations, wildlife monitoring, firefighting, etc.

Infrared (IR) images differ from those in the visible spectrum commonly used, because they do not relate to consumer electronics, i.e. the end user will be a professional or a computer vision algorithm that aims at performing a task and not an observer watching a movie for example. Most common tasks are inspection/detection, navigation and tracking.

Infrared sensors can be split in 3 categories: Short wave (SWIR), medium wave (MWIR) and Long wave infra red (LWIR). This paper focuses on the latter (LWIR), also called the thermal band. It is the most commonly used type of sensor. The reason being that in the band acquired by LWIR sensors (7.5-15μm), there is no need for solar or infrared illumination for objects to emit radiations.

Even though thermal sensors usually have lower resolution than the average visible camera, there may still be a need to compress IR images, sequences and videos. One scenario is the case of an IR camera mounted on a small unmanned aerial vehicle (UAV). The images shall be stored on a light and small system and/or transmitted via a wireless network, and therefore compression can improve the overall system performance.

Thermal images differ from visible spectrum images in various ways. Most thermal sensors nowadays offer 16 bits bitdepth, which is different from the classical 8 bit chain for visible images. In order to process them, two options then exist: either dynamic range reduction (tone mapping) to 8 bits before a standard compression or compression with higher bitdepth. HDR displays are still rare and in most cases nowadays the final user will watch the images on a regular 8 bit display. However, HDR displays are becoming more common and in case of automatic computer vision processing, it can be a great advantage to have the original 14 bits data available. Moreover, among the developed tone mapping operators (designed for visible images) some are local, which may

not fit in case of infrared images as it would modify the relative temperature values within the image.

Depending on the applications, the requirements in terms of precision can vary, for most sensors the advertised temperature resolution is around 50mK.

Although many studies investigate the presence of noise in IR images: in order to reduce it [1, 2, 3] or to assess its effect on quality [4], very few investigate the effect of compression on IR images.

The aim of this paper is to investigate compression performance and the effect of compression of high bitdepth infrared images on the objective quality.

The remainder of the paper is organized as follow: firstly some previously published studies performed on compression of infrared images or compression of high bitdepth visible images are introduced. Secondly the performed experiment is presented and finally the results obtained are detailed.

Previous Work

Lossless compression of infrared images was studied in the context of medical images by Schaefer et al. in [5]. They compare five image compression schemes (JPEG, JPEG-LS, JPEG2000, PNG and CALIC) and conclude that the best performance is achieved by JPEG-LS for lossless coding in terms of compression ratio obtained and compression speed. However, their images are 8 bit bitdepth.

The Natural Scene Statistics of infrared images is studied by Goodall et al. in [4] where they study the impact of noise and blur on the perceived quality of infrared images. However, their images are either to start with available with 8bits bitdepth or they linearly scale them to that range.

The Motion Imagery Standards Board, released a report on the compression of infrared images and videos [6] in 2014. In this, they compare h.264 FRExt compression, JPEG2000 (and MPEG-2 with tone mapping) compressions and approve both h.264 FRExt and JPEG2000 as viable compression standards. No quality evaluation (objective nor subjective) is however performed.

For JPEG-XT compression, HDR-VDP2 was found to perform best with regard to correlation with Mean Opinion Scores [7] in comparison with the other tested metrics (SNR, MRSE and SSIM) on a dataset of 20 HDR images.

A study comparing HEVC and h.264 standards for the compression of 12bits visible videos was done by Banitalebi-Dehkordi et al. in [8]. They found that HEVC performs better both in terms of average obtained MOS (10%) and bit rate (25%).

Experiment
Image Datasets

In this study we have focused on IR images/video databases with a bitdepth higher than 8 bits. The images and videos used in the experiment come from three datasets:

- The LTIR dataset from Linkping University [9], which contains 13 videos of 16 bit bitdepth (plus additional 8 bits bit-depth videos). These were acquired with various thermal cameras of resolutions ranging from 320x256 to 640x480
- The ASL dataset from ETH in Zurich [10], which contains 9 videos of 16bits bitdepth. The videos were acquired using a FLiR Tau 320 camera and have a resolution of 324x256.
- An image dataset from DTU in Copenhagen, which contains 90 images of 16 bitdepth of photovoltaic panels acquired by a UAV.

It can be noted that for all these datasets, these images are provided in 16 bits formats, however to the best of our knowledge they have been acquired with 14 bits sensors.

Codecs and settings

The datasets contain both images and videos, however this preliminary study was performed on individual images only. We chose to compress only the first frame of each video. This yields a total of 112 images in the experiment. Three different codecs were evaluated: JPEG-XT [11], JPEG 2000 [12] and HEVC [13]. The settings for each of them is detailed in this section.

For the JPEG-XT compression, two of the three profiles were used: profile b and c. For a detailed description of the codec, the reader is referred to [7]. The JPEG-XT compression is controlled by two quantization parameters: q for the base layer (the LDR image) and Q for the residual layer. In this experiment, a set of 100x100 quantizer pairs covering uniformly the possible range (1-100) were used. In order to decompose the image into base and residual layers, the tone mapping operator from Reinhard et al. [14] was used.

For the HEVC compression, we used the Format Range Extension (RExt) of the HEVC test Model (HM) for compression. It can deal with videos up to 16 bitdepth [15]. The full possible range of quantization parameters was explored, i.e. from -48 to 51 (for 16 bitdepths data).

The JPEG2000 compression was done using the OpenJPEG implementation of the standard (http://www.openjpeg.org/). The targeted bitrate is defined in terms of compression ratio. We have aimed at compression ratios ranging from 1 to 50.

Results and Quality Metrics

The image values are mapped into luminance values for all the metrics using a Gamma of 2.2 and a Peak White of 350 cd/m^2.

Objective quality metrics

In order to evaluate the performance of the various codecs, the following objective metrics were applied: Signal-to-Noise Ratio (SNR [16]), the Mean Relative Squared Error (MRSE [17]) and the HDR Visual Difference Predictor (HDR-VDP2 [18]). The SNR is defined according to JPEG-XT committee recommendation:

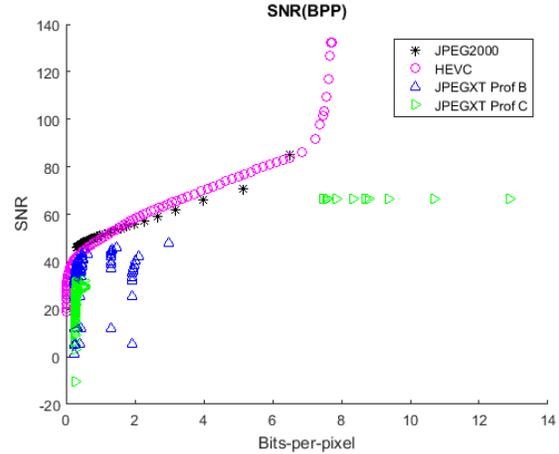$$SNR = 10 \cdot log\left(\frac{max_{i \in N}(x_i)}{\sum_{i \in N}(x_i - y_i)^2}\right), \qquad (1)$$

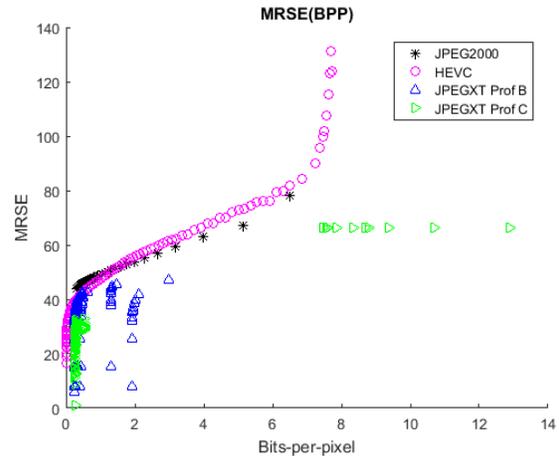**Figure 1.** *SNR as a function of bits per pixel for all encodings*

Figure 2. *MRSE as a function of bits per pixel for all encodings*

where x_i and y_i are the targeted and actual luminance of pixel i, respectively, and N the total number of pixels. It represents the PSNR for HDR content: indeed, as there exists no pre-defined range for the images, each evaluation is normalized according to its own maximum.

The MRSE evaluator is computed as:

$$MRSE = -10 \cdot log\left(\frac{1}{N}\sum_{i \in N}\frac{(x_i - y_i)^2}{x_i^2 + y_i^2}\right), \qquad (2)$$

where the notations are similar.

The direct comparison of codecs can be seen in Figure 1 for SNR, in Figure 2 for MRSE and in Figure 3 for HDR-VDP2. In Figures 1, 2 and 3 the data are averaged over all images of all datasets for each parameter setting of a codec.

It should be noted that for JPEG2000 compression, the points with lossless compression have been removed as their RMSE or SNR values would be infinite. That explains why there is no point for JPEG2000 compression above 6.5 bpp. Indeed the lossless compression is achieved with JPEG2000 at 7.5 bpp on average.

At mid-range and high bitrates, it is seen that HEVC is performing best. Below 1.4 bpp for SNR and MRSE (and 1.8 for
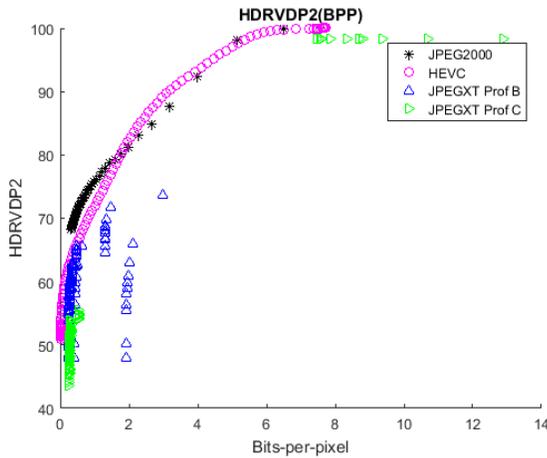
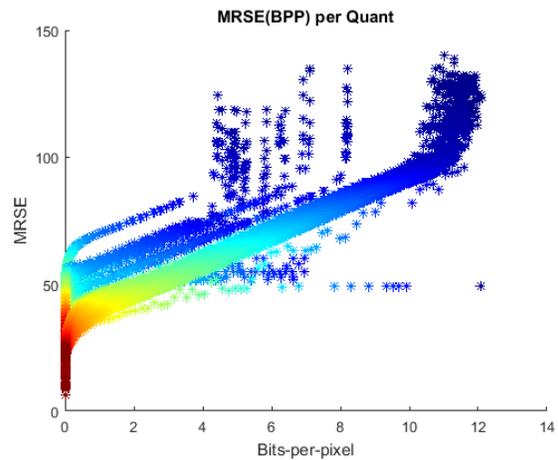**Figure 3.** HDR-VDP2 as a function of bits per pixel for all encodings

Figure 5. MRSE as a function of bits per pixel for the HEVC encoding
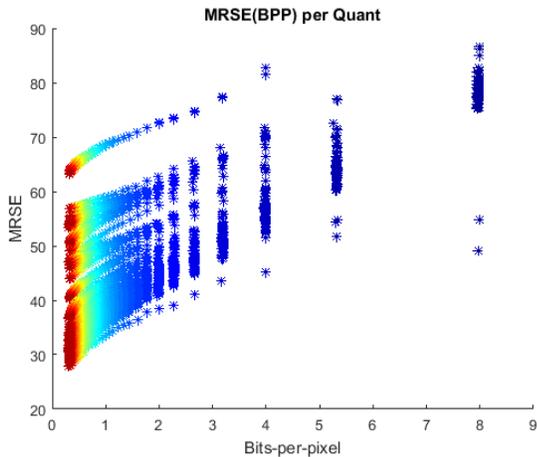

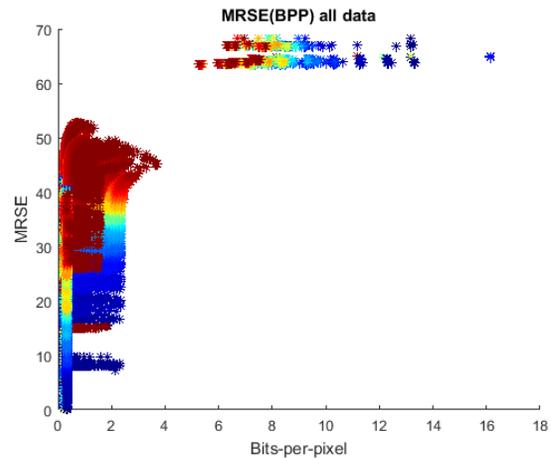**Figure 4.** MRSE as a function of bits per pixel for the JPEG2000 encoding

Figure 6. MRSE as a function of bits per pixel for the JPEG-XT encoding

HDRVDP2), JPEG2000 gives better performance. The JPEG-XT codec is inferior in performance globally and seems to saturate for the MRSE and SNR evaluators. It should be noted that what appears as outliers for JPEG-XT can be due to unrealistic/extreme sets of quantizer pairs (e.g. $q = 1$ and $Q = 100$). It should also be noted that the tone mapping to the LDR images has not been optimized for any of the images.

The MRSE values for each of the different codecs are depicted in Figures 4, 5 and 6 respectively, as a function of the bits-per-pixel used to code each image. The corresponding HDR-VDP2 values obtained are shown in Figures 7, 8, and 9. In those figures (Fig. 4-9), the color corresponds to the compression settings: compression ratio for JPEG2000, quantization parameter for HEVC and LDR quantization parameter q for JPEG-XT.

Compared to MRSE, the HDR-VDP2 plots have a larger number of points as they also include the points with lossless quality that were removed from MRSE as the measure quality is then infinite.

Finally, despite uniformly covering the range of possible quantizer values, the JPEG-XT results are not regular in terms of obtained bitrate or objective quality. It may be noted that for
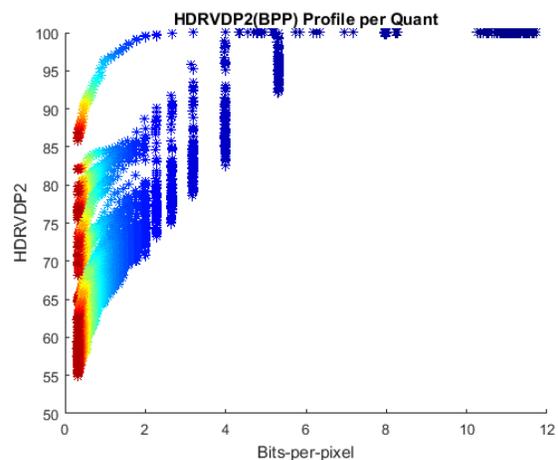

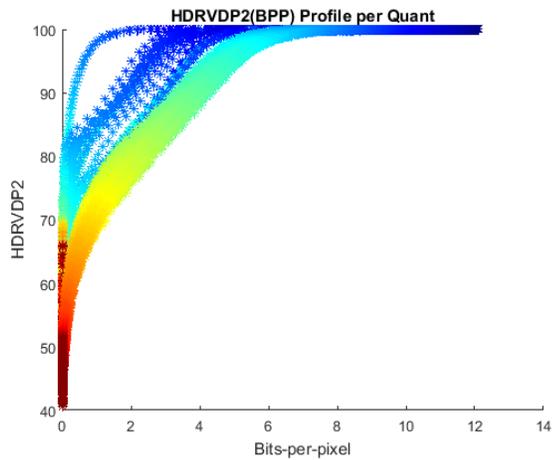**Figure 7.** HDRVDP2 as a function of bits per pixel for the JPEG2000 encoding

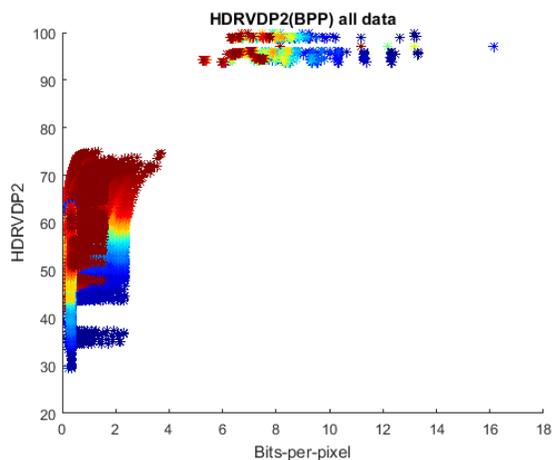**Figure 8.** *HDRVDP2 as a function of bits per pixel for the HEVC encoding*

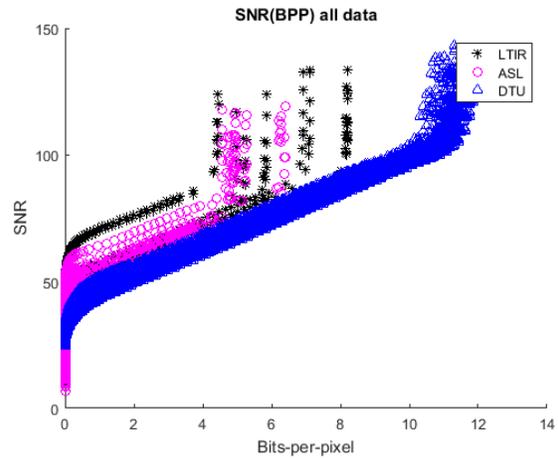**Figure 9.** *HDRVDP2 as a function of bits per pixel for the JPEG-XT encoding*

Figure 10. *SNR as a function of bits per pixel for HEVC encoding grouped by dataset*
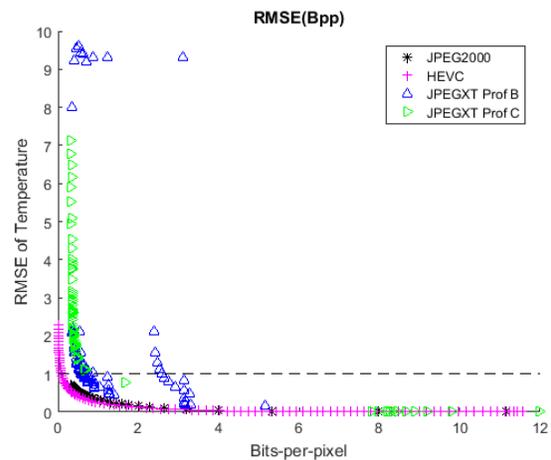

Figure 11. *RMSE as a function of bits per pixel for the DTU dataset*

JPEG-XT the results are split between the two clusters.

One interesting point is also the effect of the content on the compression performance. As can be seen in Figure 10, depending on the content, the performance for quasi lossless encoding can vary from 5 bpp to around 11 bpp. (What we refer to as quasi-lossless coding are the almost vertical parts of the curve with very high SNR values above 100 dB).

Temperature errors

As stated in the introduction, IR images will be used for a specific application, whether by a computer vision algorithm or by a professional looking for a target.

For the images from the DTU database, the conversion formula from digital numbers to temperatures in °C was available to the authors.

The results are shown in Figure 11 for the Root Mean Squared Error (RMSE) of the error and in Figure 12 for the maximum temperature error in each image averaged over the set of images. The horizontal dashed line indicates a specification of accuracy at 1°C.

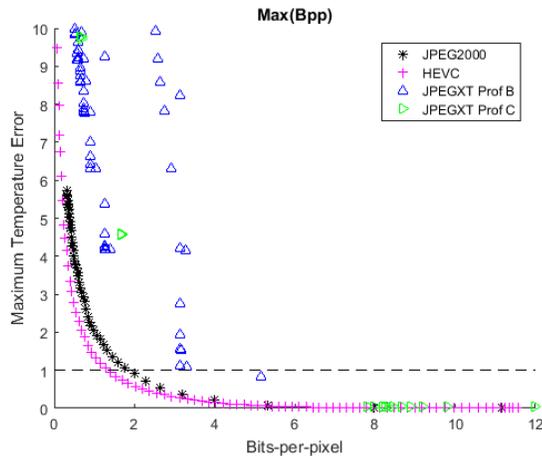As can be seen from the figure, results are coherent with the

Figure 12. *Maximum temperature error as a function of bits per pixel for the DTU dataset*

objective evaluations depicted in the previous section. It can however be noted that even at low bpp HEVC outperforms the other compression schemes. HEVC achieves a maximal error of $1°C$ for 1.3 bpp (the maximum error within each image is averaged over the 90 images), a maximal error of $1/2°C$ for 2.2 bpp and a maximal error of $1/4°C$ for 3.3 bpp.

Conclusion

In this paper the compression of infrared images is presented through three codecs: JPEG2000, JPEG-XT and HEVC. Evaluations by objective metrics (SNR, RMSE and HDR-VDP2) show that HEVC performs best for average to high bits-per-pixel (on average from 1.4 to 6.5 bpp), and that for lower bitrates JPEG2000 outperforms HEVC. JPEG2000 also achieves lossless compression using 7.5 bpp on average. According to the two signal-based evaluators (SNR and MRSE), JPEG2000 and HEVC perform fairly similar and clearly better than JPEG-XT. HDR-VDP2 evaluates that the three performances are relatively more comparable but still with JPEG2000 and HEVC as the better performing codecs. Evaluating results by maximum temperature error also show that an accuracy better than $1°C$ (in terms of average maximum error) can be achieved with 1.3 bpp on average with HEVC.

Acknowledgments

References

[1] J. M. Mooney and F. D. Shepherd, "Characterizing ir fpa nonuniformity and ir camera spatial noise," *Infrared Physics & Technology*, vol. 37, no. 5, pp. 595 – 606, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/1350449595001336

[2] F. Prez, J. E. Pezoa, M. Figueroa, and S. N. Torres, "Empirical frequency domain model for fixed-pattern noise in infrared focal plane arrays," *Infrared Physics & Technology*, vol. 67, pp. 413 – 426, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1350449514001893

[3] Y. Cao, M. Yang, and C. L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1d guided filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[4] T. R. Goodall, A. C. Bovik, and N. G. Paulter, "Tasking on natural statistics of infrared images," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 65–79, Jan 2016.

[5] G. Schaefer, R. Starosolski, and S. Y. Zhu, "An evaluation of lossless compression algorithms for medical infrared images," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, Jan 2005, pp. 1673–1676.

[6] M. I. S. B. (MISB), "Standard st0404.1: Compression for infrared motion imagery," MISB, Tech. Rep., 2014.

[7] A. Artusi, R. K. Mantiuk, T. Richter, P. Hanhart, P. Korshunov, M. Agostinelli, A. Ten, and T. Ebrahimi, "Overview and evaluation of the jpeg xt hdr image compression standard," *Journal of Real-Time Image Processing*, pp. 1–16, 2015. [Online]. Available: http://dx.doi.org/10.1007/s11554-015-0547-x

[8] A. Banitalebi-Dehkordi, M. Azimi, M. T. Pourazad, and P. Nasiopoulos, "Compression of high dynamic range video using the hevc and h.264/avc standards," in *Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine), 2014 10th International Conference on*, Aug 2014, pp. 8–12.

[9] A. Berg, J. Ahlberg, and M. Felsberg, "A thermal object tracking benchmark," in *12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2015.

[10] J. Portmann, S. Lynen, M. Chli, and R. Siegwart, "People detection and tracking from aerial thermal views," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, May 2014, pp. 1794–1800.

[11] I. J. S. 29, *ISO/IEC 18477 JPEG-XT*, 2016.

[12] ——, *ISO/IEC 15444 JPEG2000*, 2000.

[13] ITU-T, *h.265 High efficiency video coding*, 2015.

[14] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, Jul. 2002. [Online]. Available: http://doi.acm.org/10.1145/566654.566575

[15] D. Flynn, D. Marpe, M. Naccari, T. Nguyen, C. Rosewarne, K. Sharman, J. Sole, and J. Xu, "Overview of the range extensions for the hevc standard: Tools, profiles, and performance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 4–19, Jan 2016.

[16] K. Fliegel, L. Krasula, and O. behalf of COST Action IC1003 QUA-LINET', *"JPEG XT verification tests by CTU in Prague", JPEG document WG1N6584-3*, 2014.

[17] T. Richter, "Evaluation of floating point image compression," in *Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on*, July 2009, pp. 222–227.

[18] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, 2011.

[19] V. Hulusic, G. Valenzise, E. Provenzi, K. Debattista, and F. Dufaux, "Perceived dynamic range of hdr images," in *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, June 2016, pp. 1–6.

[20] M. Narwaria, M. P. D. Silva, and P. L. Callet, "Hdr-vqm: An objective quality measure for high dynamic range video," *Signal Processing: Image Communication*, vol. 35, pp. 46 – 60, 2015. [Online]. Available:

http://www.sciencedirect.com/science/article/pii/S0923596515000703

[21] G. Bjontegaard, *VCEG-M33 Calculation of average PSNR differences between RD-curves*, ITU - Telecommunications Standardization, STUDY GROUP 16 Question 6, Video Coding Experts Group (VCEG), 2001.

Author Biography

Claire Mantel received the M.S. and Ph.D. degrees in signal processing from Grenoble Polytechnic Institute, France, in 2007 and 2011, respectively. She is currently working as a Researcher at the Department of Photonics Engineering of the Technical University of Denmark. Her research interests include image and video coding and visual quality assessment.

Søren Forchhammer received the M.S. degree in engineering and the Ph.D. degree from the Technical University of Denmark, Lyngby, in 1984 and 1988, respectively. Currently, he is a Professor with DTU Fotonik, Technical University of Denmark, where he has been since 1988. He is Head of the Coding and Visual Communication Group at DTU Fotonik. His interests include source coding, image and video coding, distributed source coding, processing for image displays, two-dimensional information theory, and visual communications.