Perceptually optimized low bit-rate image encoding

Eli Ben-David, Sharon Carmel, Boris Filippov, Dror Gill, Alexey Martemyanov, Tamar Shoham, Nikolay Terterov, Pavel Tiktov, Tom Vaughan, Alexander Zheludkov

elibd@beamr.com, sharon@beamr.com, bfilippov@beamr.com, dror@beamr.com, martin@beamr.com, tamar@beamr.com, nick@beamr.com, titkov@beamr.com, tv@beamr.com, azheludkov@beamr.com

Abstract

In this paper we describe a system for high quality encoding of a given image set to a pre-determined, target average Bit-Per-Pixel (BPP). The proposed system uses our proprietary, patent protected, perceptual quality measure to determine the optimal allocation of bits among the images in the image set, and encodes each image using the HEVC/H.265 video encoder with a per image optimal encoding configuration and optional pre- and post-process. We employ learning methodologies both within the quality measure, and to ascertain optimal per image encoding configurations.

1. Introduction

The challenge of compressing images to low Bit-Per-Pixel (BPP) values, while maintaining high quality reconstructed images, has been a holy grail of digital image processing since its dawn in the 60s. Undoubtedly, the biggest advance in this field was achieved by the JPEG compression standard [1] when it was introduced in 1992. This is still the most common image format used today, over 25 years later. Many excellent image processing approaches have been presented since, but so-far, none of these have succeeded in becoming the de-facto cross-application standard the way JPEG has. When Beamr developed the technology behind JPEGmini, the goal was to offer a solution combining perceptual optimization, thus providing the ability to compress each image to an optimal degree, while remaining within the realm of the standard JPEG format, used in many applications, environments and work-flows. For this challenge we take it a step further, and propose to use HEVC Intra frames, i.e. a format, similar to the HEIC or HEIF files [2], [3] which are receiving wide industry attention [4], [5]. To this we add perceptual optimization for better resource allocation, and encoder configurations, thus providing high quality low BPP images.

The remainder of this paper is organized as follows: We present the structure of the proposed system, and provide some details on each component, followed by a high-level overview of the perceptual quality measure used. Then we will present some results and conclude. Note that throughout this paper basic knowledge of HEVC concepts and terms is assumed. Readers can refer to [6] and [7] for further details on this video coding standard.

2. Proposed image compression system

2.1. System architecture

The block diagram of the proposed system is provided in Fig 1. The system is designed to guarantee convergence to a selected average BPP over an input image set, while optimizing perceptual quality per image. In the first stage of preliminary processing, we determine a set of per image properties, such as relative visual quality obtained for the image when using a common compression parameter for all images, or the presence of film grain or high frequency textures in the image. We use these to configure the per image encoder and proceed to encode each image in an optimal manner in the main processing unit. In the third part, we iteratively allocate any “leftover” bits to the lowest quality images. In addition, since an HEVC codec is used in the system, the input RGB PNG images are converted to even-dimensioned YUV images prior to encode, and back to RGB after decoding and prior to output.

2.2. Preliminary processing

The goal of the preliminary processing block is to identify the inherent properties of each image, so as to best configure the encoding process for each image in the set. This consists of two parts. First, we analyze each image to determine if it would benefit from pre-, and/or, post-process operations such as grain removal and reinsertion. If so, we perform the appropriate pre-process, and prepare header data to be added to the image bitstream, which contains all the information needed by the decoder to perform corresponding post process. Details re approaches

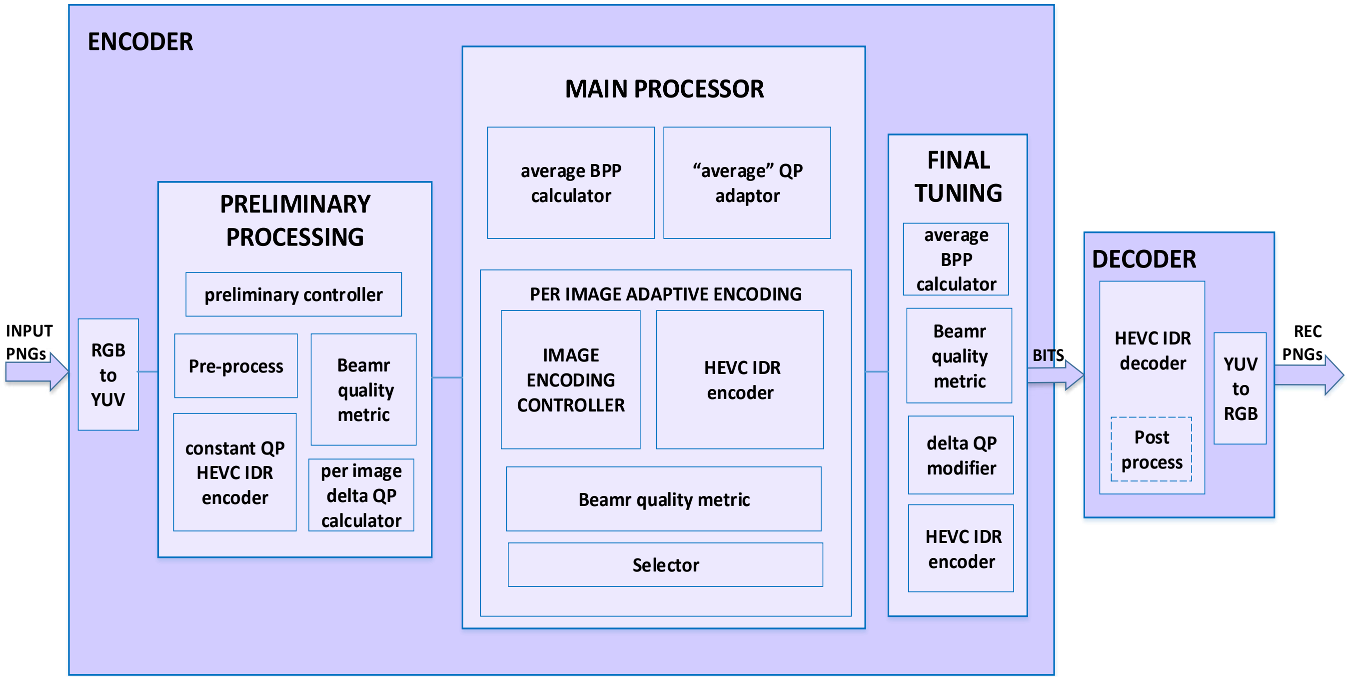

Figure 1: Block diagram of system for optimized image compression

we use for grain modeling, filtering and recreation can be found for instance in our corresponding patent application [8]. Then, we encode all images in the set using a constant picture QP value, and determine the relative perceptual degradation incurred by compression of each image in the set. This quality is measured using the Beamr proprietary quality measure, described briefly below. Based on the obtained per image perceptual quality we calculate the appropriate per image QP modulation, to be used in the main processing stage. We may also determine various per image properties, used for optimal encoder configuration.

2.3. Main Processor

Most of the encoding effort is concentrated in the main processor block, which has two main sub-components or logical units. The first operates at the level of the entire image batch and seeks the "average" QP which yields a BPP value below and close to the target. For each image this "average" QP value acts as the baseline for the per image QP modulation, according to the results of the preliminary processing.

The second part operates at a per image level. Its goal is to find the optimal encoder configuration for each image encode, according to various image properties. Some of the configuration decisions are based on offline learning, which sought the best encoder configurations or parameter sets for still image encoding using a training image set. The final selection between candidate configurations, is done as part of the encoding process, using the selector module which seeks a result that provides the best rate-distortion performance for a specific image or image class.

2.4. Final tuning

The goal of the last block in the encoding process is to improve convergence to target BPP. To obtain the best possible overall perceptual quality on the image set, we iteratively improve the worst image of the set until we converge to the target BPP. In each iteration we seek the currently most degraded image, for example the one with lowest Beamr quality measure score. Then we re-encode this image with less aggressive quantization. This process is repeated until we are as close as we wish to the target BPP value.

2.5. Beamr quality metric

The Beamr proprietary quality measure is an objective, low complexity, full-reference metric, which can reliably indicate the level of perceptual degradation introduced when compressing an image or video frame using a block based (hybrid) encoder. The score calculation consists of splitting the frame into regions, calculating multiple score components per region, and applying sophisticated perceptually weighted pooling to the per-component and per-region results, to obtain a video frame, or image, quality score. Score components include pixel wise differences, changes in texture, added artifactual edges and edge preservation. For video there is also a temporal score component which is not in use when measuring still image quality. In addition, some pre-analysis of the image or video is performed to guide and configure the quality calculation process, for example by identifying cases where a certain degree of degradation will have a more severe perceptual impact. Some of this pre-process is based on prior learning of training data. Further details on for the Beamr perceptual metric can be found for instance in [10], [11].

2.6. Decoder

The decoder utilizes the Beamr HEVC Intra frame decoder. The Beamr5 decoder has undergone extensive performance optimization, thus decoding of compressed images in our proposed system is very fast, even when using some additional post processing such as grain re-insertion. As mentioned above, the decoder block also converts the reconstructed, even-dimensioned, YUV image to the final PNG RGB image, applying cropping of single row or column where needed, as indicated in the image header.

3. Results

Applying the proposed system on the CLIC challenge validation image set resulted in an average bit-rate of 0.1493 with total PSNR of 30.24. For the test image set we obtained an average bit-rate of 0.1499 with total PSNR of 29.02.

Encode and decode times for the 268 images of the test set are about 42.5 minutes for encoding and 2.5 minutes for decoding when running on a c4.8xlarge Amazon Web Service (AWS) EC2 instance.

Visual quality of the results for both sets is superior to the quality obtained when using only HEVC encoding, without any per image adaptation of configuration and QP value.

4. Conclusion

As we have shown, the HEVC Intra frame encoder is an excellent candidate for next generation low bit rate image compression. It offers sophisticated encoding tools, a large and constantly expanding install base, and the possibility of fast decoding. Combined with our secret sauce of perceptually oriented encoding, to provide maximal viewing quality while maintaining the standard format, we believe this system is well posed to be the first image compression technology able to compete with JPEG's wide adoption and install base.

References

[2] ISO/IEC IS 10918-1 | ITU-T T.81, "Information technology - Digital compression and coding of continuous-tone still images - Requirements and guidelines"

[3] ISO/IEC 23008-12 MPEG-H Part 12, "High Efficiency Image File Format (HEIF)"

[4] https://en.wikipedia.org/wiki/High_Efficiency_Image_File_Format

[5] S. Horaczek, "Your photos just had a massive change with iOS 11. Here's what happened". Popular Science, September 2017.

[6] D. Gill, "Move over JPEG, Here Comes HEIF!", JPEGmini Blog, 2016, http://blog.jpegmini.com/move-over-jpeg-here-comes-heif/

[7] ISO/IEC 23008-2 | ITU-T H.265, " High Efficiency Video Coding (HEVC)".

[8] G.J. Sullivan, J. Ohm, Woo-Jin Han, T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.

[9] M. Terterov, A. Martemyanov, B. Filippov, T. Shoham, PCT/IL2017/050995, "Method of pre-processing of video information for optimized video encoding".

[10] T. Shoham, D. Gill, S, Carmel, "A novel perceptual image quality measure for block based image compression", Proceedings of the SPIE, Volume 7867, id. 786709 (2011).

[11] S. Carmel, D. Gill, T. Shoham, "Apparatus and Methods for Re-Compression of digital images", U.S. Patent 8,908,984, issued 2014-12-09.