BLOCKWISE MIXED FIXED-LENGTH/VARIABLE-LENGTH CODING IN JPEG2000

Till Halbach, Tor A. Ramstad and Andrew Perkis

Department of Telecommunications,
Norwegian University of Science and Technology (NTNU)
7491 Trondheim, Norway
Email: {halbach,tor,andrew}@tele.ntnu.no

ABSTRACT

A still image subband codec based on mixed fixed-length coding and variable-length coding is proposed. The proposed system combines a trellis coded quantizer with subsequent entropy encoder - which generates variable-length code words - with a nonuniform scalar quantizer - producing fixed-length code words - in order to make the resulting bit stream robust to channel errors during transmission. After subband decomposition, the codec organizes the wavelet samples blockwise, hereby exploiting local energy variations. Finally, the parameter specifying the ratio between fixed-length coding and variable-length coding and its influences are investigated. The Verification Model 2.1 of the standardization project JPEG2000 is used as a reference.

1. INTRODUCTION

Entropy coding is known to be prone to channel errors since synchronization is usually lost when an error is encountered in the variable-length (VL) encoded bit stream. Appropriate error resilience methods therefore have to be introduced for channels with bit or burst errors. In the following, subband coding using a Mallat subband decomposition is considered.

The first approach made in the JPEG2000 core experiments on error resilience inserts resynchronization markers in the bit stream and hereby improves the image quality of the decompressed image greatly [1, 2, 3]. The approach in these works is done subbandwise, i.e., the markers are inserted between two subbands. However, if the bit stream is affected by errors the bit stream between two resynchronization markers, i.e., one subband, has to be discarded since the error cannot be located exactly.

A further gain can be obtained by introducing reversible VL codes (RVLC) [4], by which the error can be located more exactly. Samples of the forefront and rear region of the subband can usually be recovered, and not all are lost. But still the performance of variable-length coding (VLC) is limited when mobile communication channels are considered, implying large bit and burst error probabilities. Searching for better error resilience strategies, the research on the aforementioned fields is thus still continuing.

Fixed-length (FL) coding as an alternative solution to VLC/RVLC removes the problem of resynchronization during decoding since every channel error will affect only one wavelet sample. On the other hand, the rate-distortion performance of fixed-length coding (FLC) is poorer than the performance of VLC. Several approaches with FLC have been studied so far.

In the first method, the entire subband information is coded using FL codes [5, 6]. When compared to entropy codec performances, this leads to improvements under channel conditions as in mobile and wireless communicationa, whereas in the error-free case a small degradation of image quality has to be accepted. It was also considered to encode only the most important subband information with FLC, i.e., the lowest-frequency subband [7]. By doing so, the degradation in the error-free case could be minimized, maintaining a better codec performance under error conditions when compared to the performance of a codec which uses VLC only.

Furthermore, [8] approaches the problem by partly encoding subband information with FL codes and the remaining wavelet samples with VL codes. However, this is done subbandwise, assuming the important information, i.e., sample sets with a large variance, to be found only in the lower-frequency subbands. Local energy variations are thus not exploited. Also, the ratio of FL encoded samples to VL encoded samples is fixed. The quality improvement of the decompressed image under error conditions therefore depends on the image to be encoded and its content.

Hence, a codec based on blockwise encoding with an adaptive coding ratio (CR) specifying the ratio between FL and VL code words is proposed in this paper. For this, the subband information is further split into blocks. The CR has to be kept adaptive in order to allow for coding of images of various size, type and quality under differing channel conditions.

2. SYSTEM DESCRIPTIONS

In this section, the reference system based on the Verification Model VM2.1 of the JPEG2000 project [9] and the proposed system are described in detail.

2.1. Reference Codec: The VM2.1

The block diagram describing the basic principle of the encoder side in the VM2.1 is shown in Fig.1.

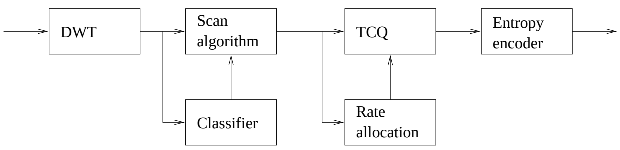

Figure 1: VM2.1 encoder block diagram.

First, a discrete wavelet transform (DWT) with a Mallat decomposition tree is applied to the image array. This results in a two-dimensional decomposition into 4 subbands for one decomposition level.

After DWT, the VM classifier complements the decomposition process by splitting one subband into 2 sequences of similar statistics [9]. The motivation is that subsequent phases of the encoder will take benefit of the common properties of each sequence.

Then, the wavelet samples are passed to the trellis coded quantizer. Each sample is quantized using one of 4 separate (scalar) quantizers with adaptive step sizes. The allowable sequence of choices among these quantizers is specified in the form of a trellis. The rate allocation is a Lagrangian method which minimizes the MSE under a rate constraint.

Finally, a binary arithmetic bit plane encoder with causal contexts is employed to VL encode the quantizer output indices.

2.2. Proposed Codec

The proposed system is illustrated in Fig.2. It can be seen that the structure of the VM is basically maintained but extended in certain areas.

DWT and the VM classification are maintained. Between the subband decomposition and quantization two further blocks are inserted which are described in the following. Also, an FLC block parallel to the VLC section is added. At the end, all data streams are merged to the final bit stream. The streams containing side information are not depicted for simplicity reasons.

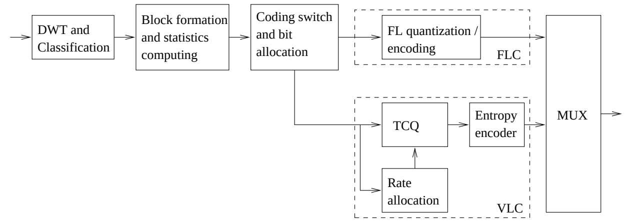

Figure 2: Block diagram of the encoder of the proposed system.

The block formation of the sequences is done with regard to the decomposition tree. Beginning with a block size of 4 at the highest level, i.e., the lowest-frequency sequence, the respective sizes for blocks of lower decomposition levels are 16, 64, 256, etc. The correspondence of samples is shown in Fig.3. It can also be seen that one decomposition level contains 3 subbands, except the highest level. After block formation, the statistical properties of each block are computed, e.g., mean and variance.

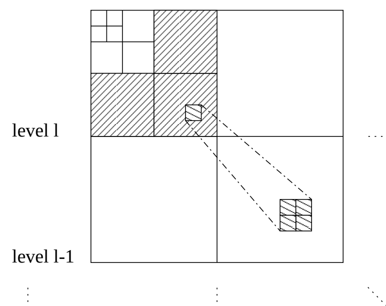

Figure 3: Mallat decomposition tree and sample correspondence among different levels.

Then, a coding switch passes the blocks to either the FL or the VL encoder, where the latter is identical to the one from the VM2.1. This is done with respect to a certain coding ratio (CR), the image size and the desired (target) bit rate. The CR is further discussed in Paragraph 3.1. It has to be selected such that the most important information is FL encoded, whereas VL encoding is used for less important information. It hereby takes channel signal-to-noise ratio (CSNR), quality of the received image and error robustness of the bit stream to be transmitted into account.

After that, The samples of one block are quantized, where the functionality of the VLC part is as mentioned in Paragraph 2.1. The FL quantization consists of 6 different scalar mid-tread quantizers with nonuniform decision and reconstruction levels optimized for a Laplacian distribution of samples in one block, using the mean-squared error (MSE) criterion.

An internal (local) bit allocation procedure preceeds the quantization and allocates every sample 0 to 7 bits according to the variance of the respective block. The mid-tread characteristic is chosen because it prevents coding artifacts like granular noise.

Finally, the multiplexer (MUX) merges the different data streams to the one that is transmitted. Its syntax is depicted in Fig.4. The header contains important side information and must be protected completely by forward error correction codes. The FLC part of the header consists of bit allocation table and quantized block variances. The VLC part contains necessary decoder information like step size of the quantizers, etc. It should be noted that the bit stream syntax and the aforegoing block formation make it impossible to progressively decode the bit stream since the blocks with important information are typically distributed over all decomposition levels and therefore over all subbands or sequences, respectively.

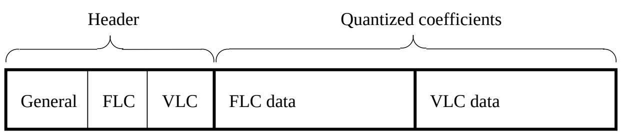

Figure 4: Bit stream syntax for the Core Experiment.

A drawback of the new codec is that it adds approxemately 53% to the header size without FLC. For the image used later, the header size takes a reasonable 7% of the compressed file size when coded with a bit rate of 0.25 bpp. The relative header size grows further the smaller the image is. Since it achieves even 15% for a bit rate of 0.125 bpp, it should therefore be attempted to reduce the amount of side information and keep it small near 5% in future investigations. However, the implementation of coding ratio and its corresponding bit allocation mechanism does not require further side information. All necessary information can be taken out of the FLC header.

Another disadvantage is increased complexitiy of both encoder and decoder. The computation time is not affected significantly. Unfortunately, it is not possible to determine *the* optimal CR prior to transmission since its choice is mainly influenced by the PSNR of the reconstructed image.

3. TESTING

Here, the newly developed system has to compete with the reference codec. The testing is done according to the Common Testing Conditions of the JPEG2000 community [10]. Only the image *Goldhill* is used.

3.1. Objective Evaluation

For objective evaluation, i.e., the determination of peak-signal-to-noise ratio (PSNR) of the reconstructed image, 3 cases are investigated: The error-free channel (EFC), the binary symmetric channel (BSC) defined by the bit error rate (BER), and a Raleigh fading channel (RFC) or burst channel which parameters are total BER and burst length (BL). To cover a wide spectrum of error pattern possibilities, 20 runs with different seeds are done. The PSNR is averaged arithmetically over the different runs. The results are discussed in the following.

Fig.5 shows the PSNR as a function of CR for some BSCs. The more CR tends to 0, the larger the VLC part in the encoded bit stream, and the smaller the FLC part. On the other way round, the FLC part grows the more CR tends to 1, wheras the VLC part at the same time shrinks. It can be seen that for the choosen BERs, FLC is superior to VLC, i.e., $\lim_{CR \to 0}(PSNR) < \lim_{CR \to 1}(PSNR)$, and that the curves are almost monotone. It is obvious that PSNR is the larger the smaller BER is. The curve of EFC is additionally depicted, for a dominating FLC part (CR $\to$ 0) showing a degradation in PSNR when compared to the case where the bit stream consists mostly of VL encoded coefficients (CR $\to$ 1).

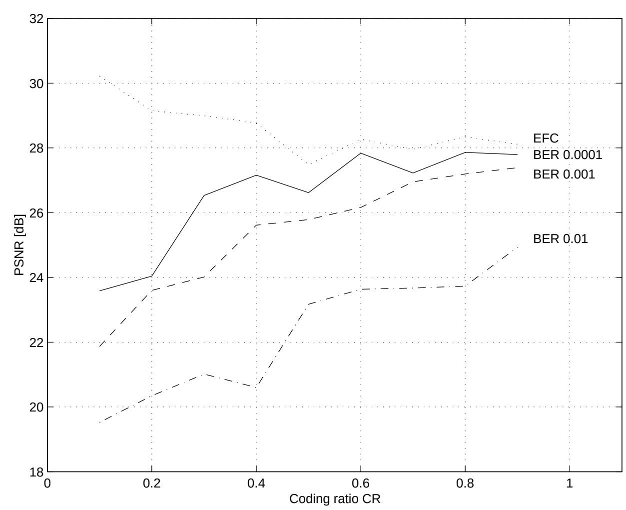

Figure 5: Image *Goldhill* coded at 0.25 bpp and transmitted over both EFC and several BSCs with the BER as a parameter.

Fig.6 is analogous to Fig.5, that means the codec shows the same behaviour. The PSNR yielded for burst lengths of 10 ms is generally larger than when the bursts are only 1 ms long. The reason for this is that the bit errors are more scattered over the code words for a shorter burst length and the same (total) BER. For 10 ms, the channel errors act more concentrated and result therefore in a smaller MSE of the reconstructed image, corresponding to a higher PSNR.

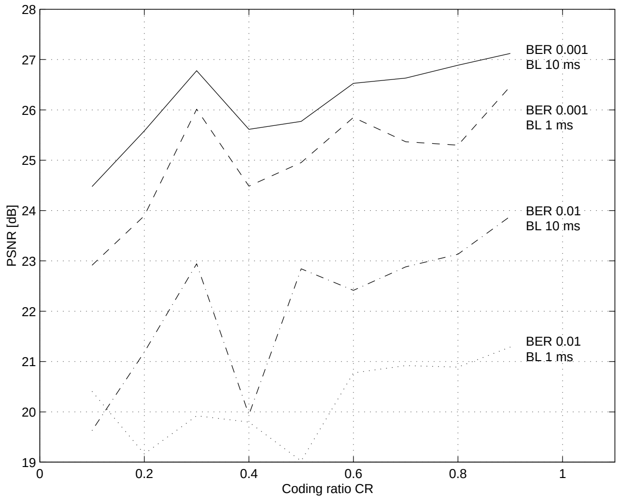

Figure 6: Image *Goldhill* coded at 0.25 bpp and transmitted over several RFCs with BER and BL as parameters.

In Fig.7 finally, three codecs are compared to each other. First, the codec generating VL code words only, i.e., the VM2.1, shows the best performance in the EFC since it can operate closest to the information theoretical bound. Above a certain threshold (here: BER $> 10^{-6}$), the codec's performance collapses rapidly until it converges to a particular value. This is due to the error resilience method of the codec that sets one or more sequences to zero if an error is encountered. Sequences are separated by resynchronization markers which are detected by the decoder. Thereby, not the whole bit stream is lost under error conditions but the sequences nearest to the error, typically one. The PSNR threshold is reached when all sequences are set to zero except the lowest frequency subband which coefficients are set to the image mean. Such an image hereby appears completely gray without further contexts.

Furthermore, the performance of a codec that only FL encodes the subband samples is depicted. In the EFC, it typically achieves a somewhat lower PSNR than the VLC codec, whereas it outperforms other codecs clearly for medium BERs. This is because a bit stream consisting of FL code words is most robust to bit or burst errors since only one or a few coefficients are altered by the channel. All other samples are maintained. However for (in real conditions unusually) high BERs, eventually its curve degrades quickly as well.

In this work it can be shown that a mixed FLC/-VLC codec combines some advantages of both aforementioned codecs and fills the gap between them. In the following, this codec is called Core Experiment (CE) according to JPEG2000 conventions. As mentioned before, the CE migrates into the FLC codec for $CR \to 1$ and turns into the VLC codec for $CR \to 0$. This can be seen for CR=0.9, CR=0.6 and CR=0.2 in Fig.7 as well. Here at approxemately $4 \cdot 10^{-6}$, the curves lie closest together.

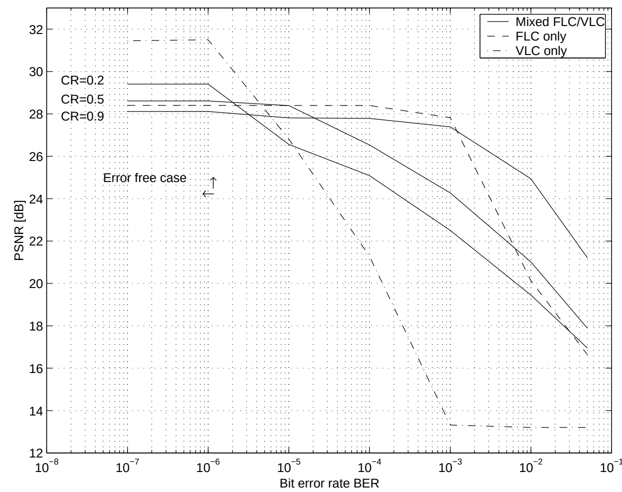

Figure 7: Image *Goldhill* coded at 0.25 bpp and transmitted over a BSC.

By modifying CR, it is possible to adjust the robustness of the bit stream to the channel conditions. If the conditions are known prior to transmission - which is often the case for two-direction channels - the quality of the reconstructed image can adaptively be optimized. Further in the interval $\left[4 \cdot 10^{-6}, 1 \cdot 10^{-2}\right]$, it can be observed that the curves of the CE do not degrade neither as fast as the FLC nor as fast as the VLC codec. The reasons for this behaviour is, first, that the CE bit allocates to blocks scattered over all subbands according to visual importance and not to whole subbands or sequences as the VLC codec. Second, the bit stream of the CE additionally holds VL encoded coefficients grouped together in blocks as well which are significantly smaller than the sequence they were taken from. As with the sequences of the VLC codec, after every sample block there is a resynchronization marker. The small block sizes lead to the result that - in case of errors - only very few coefficients are lost. This explains why the CE is much more stable in the interval mentioned above.

A further positive result is that the CE outperforms the FLC codec for large bit error rates and a high CR (e.g. 0.9). This is due to the fact that VL encoded blocks which are not affected by channel errors add significantly (partly more than 4 dB) to the received PSNR.

3.2. Subjective Evaluation

Since PSNR is not a reliable subjective measure, the images are also evaluated subjectively. Because of space limitations, here only the two most important cases are investigated: The nearly error-free case (neglectible BER) and a BSC with a (high) BER of 0.01. For reasons of comparison the original image is given as well (Fig.8).

Figure 8: Original image *Goldhill*, 720x576 pixel, gray, 8 bpp representation.

Comparing Fig.9 and Fig.8, the image qualities are nearly comparable. The CE generates a high-contrast image which is partly slightly blurred (e.g., in the dark region on the right). Also, some high-frequency information is missing (street, roofs). No artifacts are observed. In conclusion, the image quality can be rated as satisfactory.

Fig.10 shows the effects of channel errors on the reconstructed image. The conspicuous artifacts are due to bit alterations in FL code words when one or more bits of a higher bit plane are affected. Because of the inverse wavelet transform on the decoder side, modified samples affect small regions. The higher the sample in the decomposition tree, the larger the region (e.g. sky, left image edge on the ground).

The image has also lost sharpness (dark region on the right). This is due to the error resilience of VL encoded blocks that sets all block samples to zero if contaminated. However, the image content is simply recognized, and still there are a lot of details (e.g., windows), despite of the very high BER! The CE codec is therefore most appropriate for such error-prone channels. As already mentioned before, the VM produces a completely gray image under the same conditions.

Figure 9: *Goldhill* coded at 0.25 bpp (compression ratio 32) with CE codec and transmitted over an error-free channel. Further values: PSNR=29.404, CR=0.2, relative header size 0.077.

Figure 10: *Goldhill* coded at 0.25 bpp with CE codec and transmitted over a BSC with BER=0.01. Further values: PSNR=25.129, CR=0.9, relative header size 0.068.

4. CONCLUSIONS

A new codec was proposed that consists of a mixture of existing codecs, one using VLC and one using FLC only. This codec, called CE, can adaptively react on changing channel conditions by choosing an appropriate ratio between FLC and VLC. For good conditions the CR is low and hereby VLC dominates, whereas under bad conditions a high CR is recommended which makes the bit stream more robust to errors.

It is shown that for large BERs, the CE outperforms the VM2.1 significantly. Furthermore for even poorer channel conditions, it achieves better image qualities than the FLC codec. By choosing a low CR, the performance converges to the one of the VLC codec which operates closest to the information theoretical bound.

Concluding, the newly developed codec including its bit stream syntax is proposed as a possible candidate for flexible image compression and decompression. It is also recommended for implementation in the current JPEG2000 Verification Model. Finally, it is proposed to group the samples of each subband in blocks before subsequent quantization and encoding (both FLC and VLC) in order to exploit local energy variations. By this, also the error resilience capability of a VL encoded bit stream is increased.

5. REFERENCES

[1] Jie Liang. *Data partitioning and resynchronization for robust image coding and transmission.* ISO/IEC JTC 1/SC 29/WG 1 N764, March 1998.

[2] Mark Banham and Salma Soudagar. *Results for Core Experiment E2: Error resilient resynchronization and data partitioning.* ISO/IEC JTC 1/SC 29/WG 1 N747, March 1998.

[3] Iole Moccagatta, Osama Alshaykh, Homer Chen, Te-Chun I. Yang and C.-C. J. Kuo. *Proposal for JPEG2000 error resilience baseline and test conditions.* ISO/IEC JTC 1/SC 29/WG 1 N753, March 1998.

[4] Òscar S. Jimenez. *A family of reversible variable-length codes for robust still image compression systems.* Master's thesis. The Norwegian University of Science and Technology (NTNU) Trondheim, May 1999.

[5] Andrew Perkis and Tor A. Ramstad. *Impact of fixed-length coding versus variable-length coding in JPEG2000.* ISO/IEC JTC 1/SC 29/WG 1 N738, March 1998.

[6] Till Halbach. *Robust still image compression for mobile and wireless communications.* Master's thesis. The Norwegian University of Science and Technology (NTNU) Trondheim, December 1998.

[7] Hong Man, Mark J.T. Smith, Patrick Rault and Faouzi Kossentini. *Mixed FLC/VLF coding for higher error resilience.* ISO/IEC JTC 1/SC 29/WG 1 N1196, March 1999.

[8] Daniel G. Cardelo. *Error resilience in still image compression for mobile communications.* Master's thesis. The Norwegian University of Science and Technology (NTNU) Trondheim, March 1999.

[9] Charilaos Christopoulos. *The JPEG2000 Verification Model 2.0/2.1.* ISO/IEC JTC 1/SC 29/WG 1 N988, October 1998.

[10] Error Resilience Ad-Hoc Group. *Description of the error resilience core experiments and the common testing conditions.* ISO/IEC JTC 1/SC 29/WG 1 N737, February 1998.