SUBBAND IMAGE CODING WITH JOINTLY OPTIMIZED QUANTIZERS

Faouzi Kossentini, Wilson C. Chung, and Mark J. T. Smith

School of Electrical & Computer Engineering,
Georgia Institute of Technology, Atlanta GA 30332

ABSTRACT

An iterative design algorithm for the joint design of complexity- and entropy-constrained subband quantizers and associated entropy coders is proposed. Unlike conventional subband design algorithms, the proposed algorithm does not require the use of various bit allocation algorithms. Multistage residual quantizers are employed here because they provide greater control of the complexity-performance tradeoffs, and also because they allow efficient and effective high-order statistical modeling. The resulting subband coder exploits statistical dependencies within subbands, across subbands, and across stages, mainly through complexity-constrained high-order entropy coding. Experimental results demonstrate that the complexity-rate-distortion performance of the new subband coder is exceptional.

1. Introduction

The conventional approach to subband image coding has been to design separate optimal or near-optimal quantizers and associated entropy coders for each of the subband images. A bit allocation algorithm is then used to distribute bits among the subbands [1].

The subband image coder proposed here is different in that the subband quantizers and associated entropy coders are optimized jointly within and across the subbands in a complexity- and entropy-constrained framework. The algorithm used to design the coder employs multistage residual vector quantizers [2] which feature very low complexity and memory, and provide much greater control over both design and encoding complexity. An important aspect of the algorithm is that no explicit bit allocation is needed. The bits are indirectly (but optimally) allocated among the subbands during the design process. The new coder exploits both statistical intra-band and inter-band dependencies simultaneously, mainly through complexity-constrained high-order conditional entropy coding, as discussed in the remainder of this paper.
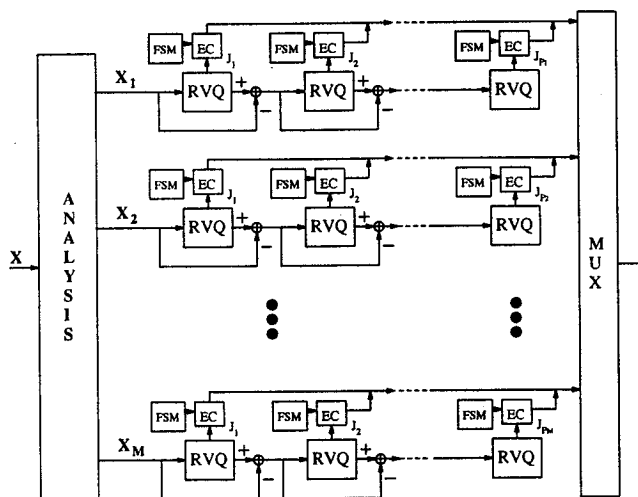
Figure 1: Basic block diagram of the subband encoder

2. System Description

As shown in Figure 1, the input signal is first decomposed into M subbands using an analysis transformation. Each subband is then encoded using a sequence of P_m ($1 \leq m \leq M$) residual vector quantization (RVQ) fixed length encoders. Any vector quantization (VQ) encoder can be used, although we find RVQs to achieve excellent rate-distortion-complexity performance [3]. The output symbol for each stage vector quantizer is fed into an entropy coder driven by a high-order stage statistical model that is controlled by a finite state machine (FSM). The FSM allows the statistical model to switch among several zero-order conditional models (represented by first order probabilities) based on the state transitions. In this work, a nonlinear mapping F given by $u = F(s_1, s_2, \ldots, s_n)$, where $s_1, s_2, \ldots, s_n$ are n previously coded symbols or outputs of some stage quantizers, will be used to determine the conditioning state u. How to construct the best mapping F will be described in Section 4. Finally, the output bits of the entropy coders are combined together and sent to the channel. Since only previously coded symbols are used by the FSM, no side information is necessary and the decoder can track the state of the encoder.

3. The Design Algorithm

Given fixed analysis/synthesis transformations, the proposed design algorithm minimizes the expected distortion $E\left[d(\boldsymbol{X}, \hat{\boldsymbol{X}})\right]$, where $\boldsymbol{X}$ is the input and $\hat{\boldsymbol{X}}$ is the output, subject to a constraint on the overall entropy of the product of the M subband VQs. This design algorithm is an iterative descent algorithm based on a Lagrangian minimization, and is a generalization of the entropy-constrained algorithms described in [4, 2, 5]. Given a fixed Lagrangian parameter λ, the algorithm attempts to satisfy simultaneously optimality conditions, requiring the subband encoders, decoders, and entropy coders be designed jointly. Details of the optimality conditions and convergence of the algorithm can be found in [3].

The parameter λ is chosen based on the overall rate and distortion of the subbands, and is used in the entropy-constrained design of all subband quantizers. Therefore, explicit bit allocation is not needed in the design process. In fact, it can be shown [3, 6] that maintaining the same slope λ for all subband operational R(D) curves results in a locally optimal allocation of bits.

4. Complexity Issues

The complexity and memory associated with the design algorithm grows *exponentially* as a function of the quantization and entropy coding parameters. To reduce them substantially, constrained VQs must generally be employed. In this work, we choose to use multistage residual vector quantizers, mainly because they require relatively low encoding complexity and memory, and because they simplify the design process by providing greater control over the complexity-performance tradeoffs.

As shown in [3], optimal encoding requires that the synthesis transformation be embedded in the design loop, which can result in much larger computational and memory requirements. However, experimental results show that the subbands can be encoded accurately by minimizing the distortion between the input and the output of the subband quantizers (independently) instead of minimizing the overall distortion of the analysis-quantizer-synthesis system.

Exhaustive searching, which is generally necessary for optimal encoding, can be circumvented effectively by exploiting the multistage residual structure and using tree-structured searching techniques such as the (M, L) algorithm. Such algorithms can be used to search the stage codebooks in all subbands with reasonable computational requirements. In particular, we choose to use the dynamic M-search algorithm [7], which employs a thresholding technique to decide the best number of paths that should be saved at each RVQ stage. Dynamic M-search provides a flexible way of
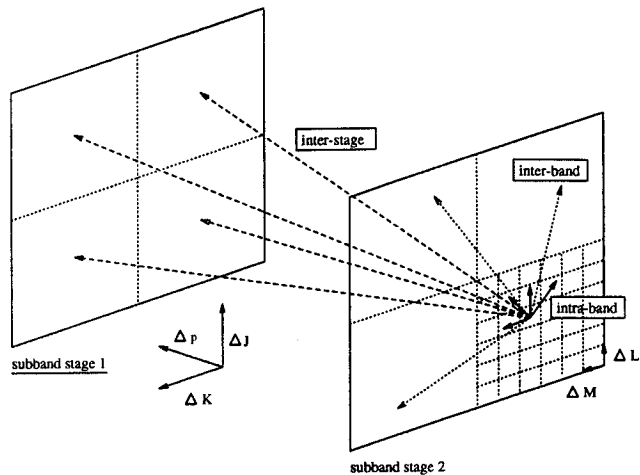

Figure 2: Inter-stage, inter-band, and intra-band conditioning scheme within an image.

trading complexity for performance, and it can achieve rate-distortion performance that is very close to that of exhaustive searching while requiring only 20%-50% more computations than sequential searching.

Optimal decoding requires a 2-dimensional optimization procedure (see Figure 1) which consists of using the iterative Gauss-Seidel algorithm [5] to minimize the average distortion between the input and the synthesized reproduction of all stage codebooks in all subbands. Although the joint decoding optimization is potentially more demanding than that associated with non-constrained quantizers, its complexity can be drastically reduced by, for example, grouping neighboring stage codebooks in neighboring subbands and jointly optimizing each group independently. This typically results in less than a 0.10 dB loss in PSNR.

The most important advantage of the multistage residual structure is that it can substantially reduce the complexity and memory by making the output alphabet of the stage quantizers small (e.g., 2, 3, or 4). For example, a 4-stage RVQ with 4 code vectors per stage codebook using a second-order conditional entropy coder for each stage, generally requires that 64 probabilities per stage be computed and stored. For a single-stage conventional VQ with 256 code vectors, as many as 256^3 probabilities may need to be computed and stored. Multistage RVQs provide another dimension upon which to capitalize. To illustrate this point, Figure 2 shows the inter-stage, inter-band, and intra-band conditioning scheme used in the system. Each image shown is a multistage approximation of the input image, and statistical dependencies among these images generally exist. For each stage (m, p) in each subband m, a 5-dimensional initial region of support $\mathcal{R}_{m,p}$ containing a sufficiently large number $R_{m,p}$ of conditioning symbols, or previous outputs of

fixed-length stage encoders, is first chosen. Since using the region of support $\mathcal{R}_{m,p}$ in the conditioning process generally results in unbearable complexity, we locate the $n_{m,p}$, $n_{m,p} \ll R_{m,p}$, conditioning symbols $s^1, \ldots, s^{n_{m,p}}$ such that the $n_{m,p}$th order conditional entropy $H(J_{m,p}|s^1, \ldots, s^{n_{m,p}})$ is minimized. To do this we build a tree where the levels represent the orders (or number of conditioning symbols) and the branches represent the possible combinations of conditioning symbols at each level. As is described in [8], this tree is symmetric, which can simplify the search process. Our experimental results show that dynamic M-search provides an excellent balance between conditional entropy and search complexity.

Since the objective is to minimize the average entropy of all stage statistical models in all subbands given a fixed level of complexity and memory of the joint entropy coder, our approach is to first build a tree with $\sum_{m=1}^{M} P_m$ branches, where P_m is the number of stage codebooks in the mth subband. Each branch is a unary tree of length $L_{m,p}$, where $L_{m,p}$ is the number of complexity-entropy pairs. The dynamic M-search algorithm is used to find the best $n_{m,p}$ ($1 \leq n_{m,p} \leq L_{m,p}$) conditioning symbols given $R_{p,m}$ conditioning symbols. For each complexity-entropy pair, complexity is given by $\mathcal{N}_{m,p} = S_{m,p} N_{m,p}$, where $S_{m,p}$ is the number of all combinations of realizations of the conditioning symbols and $N_{m,p}$ is the output alphabet size of stage p in band m. Once all complexity-entropy pairs are obtained, we then use the generalized BFOS algorithm [9] to minimize the overall output entropy subject to a constraint $\mathcal{N}^{max}$ on the total number of conditional probabilities, which is used here as a measure of complexity and memory.

The FSM statistical model for each stage (m,p) employs a mapping F to determine the state given $n_{m,p}$ available symbols. The mapping F is one-to-one and is actually given by a table that contains the numbers $0, 1, \ldots, S_{m,p} - 1$, representing each of the possible combinations. As discussed in [10], a large number of the tables representing the $S_{m,p}$ states are usually either not populated or scarcely populated. In addition to this inefficiency, some of the empty tables may be visited during actual encoding even though they were never visited during the design process. This is the so-called empty state problem that often arises in FSM design. To address this, we use the PNN algorithm [11], as described in [12], to drastically reduce the number of states while still bounding the loss in entropy performance to 1%. The PNN algorithm used here first merges all of the empty states with the least probable state into *one* conditioning state, thereby completely removing empty states. Then, the two conditioning states resulting in the lowest increase in entropy (when merged) are combined into one conditioning state, and so on until only one state, which represents one table

of first-order probabilities, is obtained. As described above, another tree with $\sum_{m=1}^{M} P_m$ branches is built, where the nodes now represent complexity-entropy pairs obtained by the PNN algorithm. The BFOS algorithm is again used to minimize the overall output entropy subject to the use of a much smaller number of conditional probabilities.

Quantizing the conditioning states has the additional advantage that the stage statistical model orders can be allowed to grow to relatively large numbers, which tends to lower the overall entropy with only a small increase in encoding complexity. Moreover, the merging process improves the robustness of the subband coder because only global statistics are carried through, and the possibility of a strong mismatch between the test sequence and the coder is less likely.

5. Experimental Results

Several images of size 512×512 (excluding the test images) were used to design jointly optimized complexity- and entropy-constrained subband residual scalar quantization (RSQ) codebooks (i.e. the vector size is 1×1). In the proposed framework, the use of larger vector sizes resulted in a large increase in complexity but no increase in rate-distortion performance.

For analysis/synthesis, we employ a 3-level balanced tree-structured IIR allpass polyphase filter bank as described in [13], resulting in 64 uniform subbands. For this implementation, all stage codebooks in all subbands contain 3 scalars. The initial maximum allowed number of conditional probabilities N^{max} is set to 2048. After using the BFOS algorithm, we employed the PNN algorithm to populate another tree as described before. The BFOS algorithm is used again to locate the best numbers of conditioning states for each stage in each subband subject to using a maximum number of 256 probabilities. The output of each of the stage fixed-length RSQ encoders is encoded using an adaptive arithmetic coder driven by the corresponding FSM statistical model just generated.

For each rate-distortion point, the total memory required to store the subband RSQ codebooks and associated mapping tables as well as tables of conditional probabilities is approximately 1.5 kilobytes. For encoding (including analysis) using dynamic M-search, approximately 8 multiplies and 11 adds per pixel are required. Only 3 multiplies and 9 adds are required for decoding (including synthesis). In this example, the design time is approximately 4 CPU hours on a Sun Sparc 10 workstation. Not only are the complexity and memory relatively small, but the performance is exceptional. For instance, the PSNRs obtained for the image LENA at 0.50 and 0.25 bpp are approximately 37 and 34.1 dB, respectively. Moreover, the subjective quality is also very good. Figure 3 shows the coding result of the image LENA at a bit rate of 0.25 bpp using

(a) ORIGINAL

(b) IIR/RSQ SUBBAND CODER

Figure 3: The LENA image coded using the IIR/RSQ SUBBAND CODER. The bit rate is 0.25 bpp and the PSNR is 34.07 dB.

the subband coder. This is almost 4 dB better than the JPEG standard at the same rate.

These experimental results show that the proposed subband coder can achieve very good compression results while maintaining relatively low complexity and memory. Moreover, it also compares favorably with standard JPEG in terms of complexity. The proposed coder does require more encoding/decoding computations, but multiplication-free implementations of it [14] can achieve complexity very close to that of JPEG with only a small loss (less than 0.50 dB) in quality.

6. References

[1] J. W. Woods, ed., *Subband Image Coding.* Norwell, MA: Kluwer Academic Publishers, 1991.

[2] F. Kossentini, M. Smith, and C. Barnes, "Entropy-constrained residual vector quantization," in *Proc. IEEE ICASSP*, vol. V, (Minneapolis, MN, USA), pp. 598–601, Apr. 1993.

[3] F. Kossentini, W. Chung, and M. Smith, "A jointly optimized subband coder," *Submitted to Trans. on Image Processing*, July 1994.

[4] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. on ASSP*, vol. ASSP-37(1), pp. 31–42, Jan. 1989.

[5] F. Kossentini, M. Smith, and C. Barnes, "Necessary conditions for the optimality of variable rate residual vector quantizers," *Submitted to Trans. on Information Theory in June 1993. Revised in May 1994.*

[6] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. on Image Processing*, vol. 2, pp. 160–174, April 1993.

[7] F. Kossentini and M. Smith, "A fast searching technique for residual vector quantizers," *Signal Processing Letters*, vol. 1, pp. 114–116, July 1994.

[8] F. Kossentini, W. Chung, and M. Smith, "Conditional entropy-constrained residual VQ with application to image coding," *Submitted to Trans. on Image Processing*, July 1994.

[9] E. A. Riskin, "Optimal bit allocation via the generated BFOS algorithm," *IEEE Trans. on Information Theory*, vol. 37, pp. 400–402, Mar. 1991.

[10] S. M. Lei, T. C. Chen, and K. H. Tzou, "Subband HDTV coding using high-order conditional statistics," *IEEE Journal on Selected Areas in Communications*, vol. 11, pp. 65–76, Jan. 1993.

[11] W. H. Equitz, "New vector quantization clustering algorithm," *IEEE Trans. on ASSP*, vol. 37, pp. 1568–1575, Oct. 1989.

[12] F. Kossentini, W. Chung and M. Smith, "Subband coding of color images with multiplierless encoders and decoders," submitted to IEEE Int. Sym. on Circuits and Systems, (Seattle, WA), May 1995.

[13] M. Smith and S. Eddins, "Analysis/synthesis techniques for subband image coding," *IEEE Trans. on ASSP*, vol. 38, pp. 1446–1456, Aug. 1991.

[14] F. Kossentini, W. Chung, and M. Smith, "Subband coding of color images with multiplierless encoders and decoders," in *Abstract Submitted to IEEE ISCAS*, (Seattle, WA, USA), May 1995.