BIT ALLOCATION FOR DEPENDENT QUANTIZATION WITH APPLICATIONS TO MPEG VIDEO CODERS

*Kannan Ramchandran** *Antonio Ortega*† *Martin Vetterli*‡

Department of Electrical Engineering
and Center for Telecommunications Research,
Columbia University, New York.

ABSTRACT

We address the problem of efficient bit allocation in a dependent coding environment. While optimal bit allocation for independently coded signal blocks has been studied in the literature, we extend these techniques to the more general dependent coding scenarios. Of particular interest is the topical MPEG [1] video coder. We show how a certain monotonicity property of dependent operational rate-distortion curves, verified through MPEG simulations, can be exploited in formulating fast ways to obtain optimal and near-optimal solutions (in the R-D sense) for the MPEG bit allocation problem.

1. INTRODUCTION

The topic of optimal bit allocation for independently coded signal sets has been studied in the literature [2]. In this work, we generalize the allocation problem to include dependent coding blocks. By dependent coding environments, we mean scenarios where the operational rate-distortion (R-D) curves of some coding blocks depend on the particular operating R-D point of other blocks (see Fig. 1). Typical examples include DPCM, the Laplacian spatial pyramid with quantization feedback (which, along with bit allocation for the spatio-temporal pyramid video coder has been analyzed in [3]), and MPEG [1]. In this paper, we address the theory of bit allocation in temporally dependent coding environments. In particular, the MPEG coder is covered in detail. We formulate the optimal strategy for a special case of the MPEG coder, and show how the Viterbi algorithm provides the optimal solution to this case. We then point out the importance of the monotonicity property in the R-D curves of dependent coding blocks, and show how it can be exploited to reduce computational complexity. Finally, we show how the intuition gained with the special case can be used to solve the general MPEG bit allocation problem.

2. DEPENDENT BIT ALLOCATION

We now address the general temporal dependency quantization problem of which MPEG [1] is an example. Fig. 2 shows the MPEG temporal dependency framework. Let us first consider a 2-layer dependency as in Fig. 1. Shown are the R-D characteristics for the first independent frame and the second dependent frame. Our constrained optimization problem (COP) is: what quantization choice do we use for each frame such that the total distortion is minimized subject to a maximum total bit budget constraint?

*Work supported in part by the New York State Science and Technology Foundation's CAT at Columbia University

†Work supported in part by the Fulbright Commission and the Ministry of Education of Spain

‡Work supported in part by the National Science Foundation under grants ECD-88-11111 and MIP-90-14189.

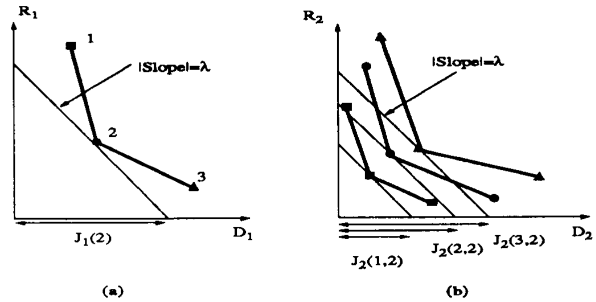

Figure 1: Operational R-D characteristics of 2 frames in a dependent coding framework, where frame 2 depends on frame 1. (a) Independent frame's R-D curve. (b) Dependent frame's R-D curves. Note how each quantizer choice for frame 1 leads to a different $R_2 - D_2$ curve.

$$\min_{q_1,q_2} [D_1(q_1)+D_2(q_1,q_2)] \text{ s.t. } R_1(q_1)+R_2(q_1,q_2) \leq R_{budget}. \quad (1)$$

This problem can be solved by introducing the Lagrangian cost $J = D + \lambda R$ corresponding to the Lagrange multiplier $\lambda \geq 0$ as in [2] as follows:

$$J_1(q_1) = D_1(q_1) + \lambda\ R_1(q_1), \quad (2)$$
$$J_2(q_1,q_2) = D_2(q_1,q_2) + \lambda\ R_2(q_1,q_2), \quad (3)$$

and considering the following unconstrained minimization problem:

$$\min_{q_1,q_2}[J_1(q_1) + J_2(q_1,q_2)]. \quad (4)$$

Then, by a direct extension of Theorem 1 of Shoham and Gersho in [2], the following result follows:

Theorem 1 *If (q_1^*, q_2^*) solves the unconstrained problem of Eq. (4), then it also solves the constrained problem of Eq. (1) for the particular case of $R_{budget} = [R_1(q_1^*) + R_2(q_1^*, q_2^*)]$.*

The above result implies that if we solve the unconstrained problem of Eq. (4) for a fixed value of λ, and if the total bit rate happens to be R_{budget}, then we have also optimally solved the constrained optimization problem of Eq. (1). Further, as λ is swept from 0 to ∞, one traces out the convex hull of the composite R-D curve of the dependent allocation problem. The monotonic relationship between λ and the expended bit rate [2] makes it easy to search for the "correct" value of λ for a desired R_{budget}.

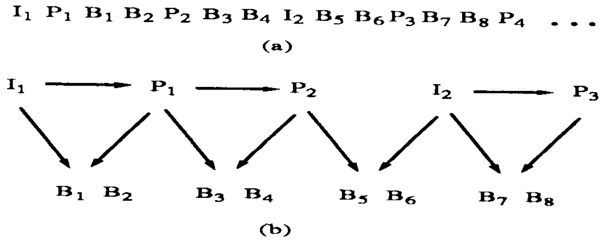

Figure 2: Typical MPEG coding framework. (a) The MPEG frames: the I frames are independently coded, the P frames are predicted from previous I or P frames, and the B frames are interpolated from adjacent I and/or P frame pairs. (b) Temporal dependency in the MPEG framework. Note that the B frames are leaves in the dependency tree.

Note how for the independent case ($J_2(q_1, q_2) = J_2(q_2)$), Eq. (4) becomes the familiar result of [2], where each frame is minimized independently. Thus, the 2-frame problem becomes the quest for q_1^*, q_2^* that solve:

$$J_1(q_1^*) + J_2(q_1^*, q_2^*) = \min_{q_1, q_2}[J_1(q_1) + J_2(q_1, q_2)] \quad (5)$$
$$= \min_{q_1}[J_1(q_1) + J_2(q_1, q_2^*(q_1))], (6)$$

where $J_2(q_1, q_2^*(q_1)) = \min_{q_2}[D_2(q_1, q_2) + \lambda R_2(q_1, q_2)]$ is the minimum Lagrangian cost (for quality condition λ) associated with the dependent layer when the independent layer is quantized with q_1. See Fig. 1. Thus, for the desired operating quality λ, we find the optimal solution by finding, for all choices of q_1 for the independent layer, the optimal $(q_2^*(q_1))$ which "lives" at absolute slope λ on the (dependent) R_2–D_2 curve associated with q_1.

By a simple extension of this result, it follows that the optimal solution to our general N-frame dependency problem consists in introducing $J_i(q_1, q_2, \ldots, q_i) = D_i(q_1, q_2, \ldots, q_i) + \lambda R_i(q_1, q_2, \ldots, q_i)$ for $i = 1, 2, \ldots, N$ and solving the following unconstrained problem for the "correct" value of λ which meets the given R_{budget}:

$$\min_{q_1, q_2, \ldots, q_N} [J_1(q_1) + J_2(q_1, q_2) + \ldots J_N(q_1, q_2, \ldots, q_N)]. \quad (7)$$

We will show how to use the above results to solve the general MPEG (with I,P,B frames as in Fig. 2) allocation problem. As a first step towards this end, we begin with a simpler special case of MPEG that is easier to analyze and which provides the intuition for the more complex general problem.

2.1. A particular case of MPEG: I-B-I

We consider a special case of MPEG having only I and B frames (see Fig. 3), i.e. the predicted P frames of the more general MPEG format are missing. The dependency tree is shown in the form of a more compact trellis. The "states" of the trellis represent the quantization choices for the *independently coded* I frames (ordered from top to bottom in the direction of finest to coarsest), while the "branches" denote the quantizer choices associated with the two B frames.

The trellis is populated with Lagrangian costs (for a fixed λ) associated with the quantizers for each frame. Let us focus on the $I_1 - B_1 - B_2 - I_2$ stage of the trellis. The state nodes are populated with the costs of the respective I frame quantizers $J(q) = (D(q) + \lambda R(q))$. Each (i, j) branch connecting quantizer state i of I_1 to quantizer state j of

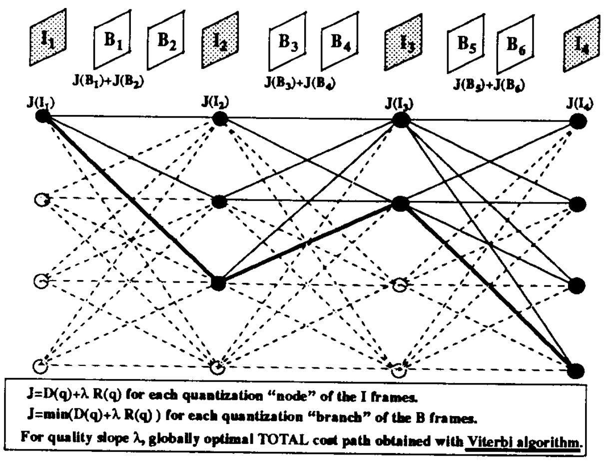

Figure 3: The I-B-I special case of MPEG. Finding an R-D convex hull point corresponding to a λ is equivalent to finding the smallest cost path through the trellis. Each trellis node corresponds to a quantizer choice for the I frames, monotonically ordered from finest to coarsest, and is populated with the associated Lagrangian cost. The branches correspond to the B frame pairs, and are populated with their minimum Lagrangian costs for the particular I frame quantizer choices given by each branch's end nodes. The "dark line" path joins the smallest cost I frame nodes. Monotonicity implies that all dashed line paths can be pruned out.

I_2 is populated with the sum of the minimum Lagrangian costs of the B_1 and B_2 frames, i.e. with $J_{B_1} + J_{B_2}$, with:

$$J_{B_l} = \min_{q_{B_l}}[D(q_{B_l}) + \lambda R(q_{B_l})] \text{ for } l = 1, 2 \quad (8)$$

where the R-D curves for B_1, B_2, are generated from the i, j quantizer choices for I_1, I_2 respectively. From Eq. (7), it is clear that the optimal path is that which has the minimum total cost across all trellis paths. Since the independent I frames "decouple" the B frame pairs from one another, it is obvious that the popular Viterbi algorithm (VA) [4] will provide the minimum cost path through the trellis!

2.2. Complexity

The VA which provides the optimal solution is obviously computationally intensive. An important point to be made is that the computational complexity is dominated by the *data generation* phase, i.e. in the trellis population phase. In order to ease the computational burden, we are therefore interested, not so much in fast methods which approximate the VA given the entire trellis (like the stack algorithm [4]), but rather in methods which ***will eliminate the very need to populate the entire trellis***! We now examine an important property which enables us to do exactly this.

3. MONOTONICITY

The key to obtaining a fast solution to the complex dependent allocation problem of Eq. (1) is the monotonicity property of the R-D curves of the dependent components (frames). We now explain what this means. Consider the example of 2 frames, with the operational R-D curve of the second frame depending on that of the first (see Fig. 1).

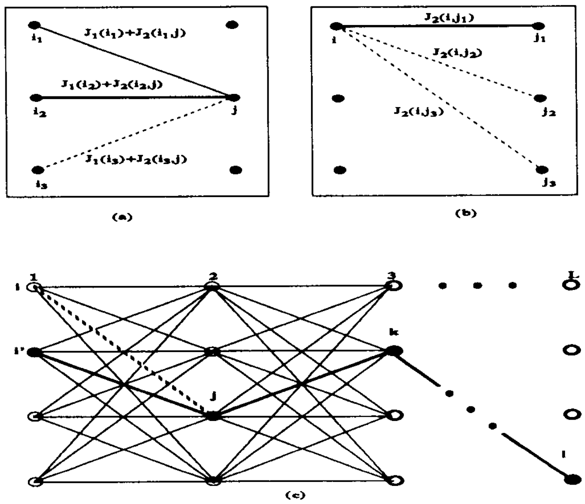

Figure 4: Pruning conditions obtained from monotonicity. (a) $J_1(i_2) + J_2(i_2, j)$ is the minimum Lagrangian cost of all branches terminating in node j. Therefore (see Lemma 1), the (i_3, j) branch can be pruned. (b) $J_2(i, j_1))$ is the minimum Lagrangian cost of all branches originating from node i. Therefore (see Lemma 2), the (i, j_2) and the (i, j_3) branches can be pruned. (c) Diagram used for the proof of Lemma 1.

The ordering of the quantizer grades is, as before, monotonic from finest to coarsest. Then, the monotonicity property holds if, for q_1 and q_2 as the quantizer choices for the I and P frames respectively:

$$J_2(q_1, q_2) \leq J_2(q_1', q_2), \quad \text{for } q_1 \leq q_1', \tag{9}$$

Stated in words, the monotonicity condition simply implies that a "better" (i.e. finer quantized) predictor will lead to more efficient coding, in the rate-distortion sense, of the residue (whose energy increases as the predictor quality gets worse). That is, the dependent frame's family of R-D curves will be monotonic in the fineness of the quantizer choice associated with the parent frame from which they are derived. See Fig. 1. Experimental results involving MPEG *verify this monotonicity property* for all the cases that we studied. Thus, monotonicity appears to be a realistic property, which has favorable theoretical implications as well, as we now describe.

3.1. Pruning conditions implied by monotonicity

The monotonicity condition stated earlier implies the following two pruning conditions in the quest for the optimal path for the dependent coding problem. The first lemma is associated with Fig. 4(a). As a reminder, the quantizer states are ordered monotonically from finest to coarsest.

Lemma 1 *If*

$$J_1(i) + J_2(i, j) < J_1(i') + J_2(i', j) \quad \textit{for any } i' > i, \tag{10}$$

then the (i', j) branch cannot be part of the optimal path and can be pruned out.

Proof: We prove the lemma by contradiction. Assume that (i', j) for any $i' > i$ is part of the optimal path (see Fig. 4(c)). Let the optimal quantizer sequence path be $(i', j, k, \ldots, l)$. But, by monotonicity, we have:

$$J_3(i, j, k) \leq J_3(i', j, k) \tag{11}$$

$$\cdots$$

$$J_L(i, j, k, \ldots, l) \leq J_L(i', j, k, \ldots, l) \tag{12}$$

Summing up Eqs. (10), (11), ..., (12), we get the contradiction that the total Lagrangian cost of the path $(i, j, k, \ldots, l)$ is smaller than that of the optimal path $(i', j, k, \ldots, l)$. □

The above lemma is associated with pruning branches that merge into a common destination state. A dual result holds for the pruning of branches that originate from a common source state, as stated below. The proof will be omitted as it is similar to that of Lemma 1. See Fig. 4(b) for the diagram associated with the following lemma.

Lemma 2 *If $J_2(i, j) < J_2(i, j')$ for any $j' > j$, then the (i, j') branch cannot be part of the optimal path and can be pruned out.*

The two pruning conditions of Lemmas 1 and 2 can be used to lower the complexity of the VA. In the special case of MPEG of Section 2.1. (refer to Fig. 3), Lemmas 1 and 2 eliminate the need to consider the full trellis on which to run the VA making it unnecessary to consider any paths lying below the (dark line) path connecting the minimum cost state nodes of the I frames. This is because any path with excursions below the path connecting the minimum cost state nodes (corresponding to the I frames only) can be replaced by one which lies *above* this boundary, by monotonicity. The reduction in complexity is due to the lack of need to grow branches out of nodes that have been eliminated so that it is not necessary to populate the fully connected trellis.

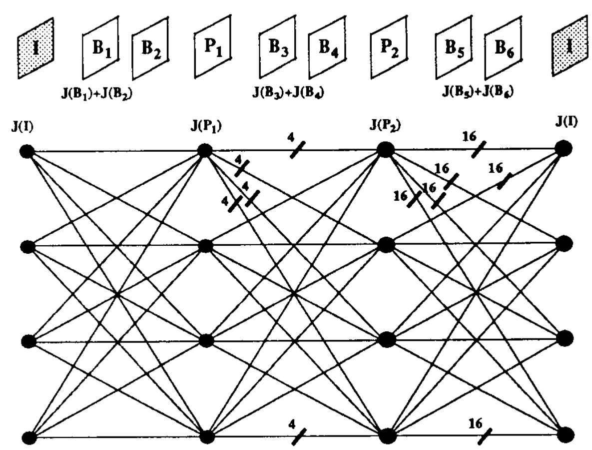

Figure 5: General MPEG "trellis" diagram extension of Fig. 3. Here, the inclusion of the P frames prevents the decoupling of the B frame pairs, and the entire tree has to be grown. Note that each stage of the trellis is represented by "vector" branches whose dimension grows exponentially with the dependency tree depth.

4. GENERAL MPEG BIT ALLOCATION

Having established the intuition behind dependent allocation and the power of monotonicity, we now evolve to the more complex (general) MPEG format of Fig. 5. The presence of the P frames extends the dependency tree depth, and the decoupling between successive stages of the trellis is lost. We can thus no longer resort to the Viterbi algorithm, but must instead retain the entire tree, which grows exponentially with the number of dependent levels. The good news, however, is that the monotonicity conditions still apply, and the pruning conditions of Lemmas 1 and 2 can aid in reducing complexity dramatically. As an example, see Fig. 6 where we consider an I-B-P-B-P sequence of MPEG frames and a choice of 3 quantizer grades for each frame. While the exhaustive search would have us grow as many as 363 Lagrangian costs, in our example, only 36 costs need to be grown, an order of magnitude reduction in complexity with no loss of optimality if the monotonicity conditions apply (as verified in our examples)! The complexity reduction due to monotonicity is dependent on the desired quality slope λ, with higher quality targets achieving better reduction. In the limit, as λ goes to 0, the minimum cost path is always the one corresponding to the finest quantizers and thus only a single "highest quality" path has to be grown. Conversely, if λ goes to ∞ the monotonicity property provides no gain. See [5] for details.

4.1. Suboptimal heuristics

As pointed out, the amount to which the monotonicity property can be exploited is λ dependent, and may not suffice for some applications. To this end, it is advisable to come up with fast heuristics, which used in combination with monotonicity, can approach the optimal performance at a fraction of the complexity. In trying to formulate a fast MPEG heuristic, it is necessary to consider some important points: (i) the "anchor" I-frame is the most important of the group of pictures and must not be compromised, (ii) most signal sequences enjoy a finite memory property, where the influence of a parent frame diminishes with the level of its dependency. The folllowing heuristic seems to work well: (i) retain all paths that *originate from* each of the I frame quantization states; (ii) use a "greedy" pruning condition (in combination with the monotonicity property) to keep only the lowest cost branch (so far) at all other nodes in the trellis. This heuristic, as shown in Fig. 7, leads to near optimal performance at a fraction of the computational cost. See [5] for details.

Acknowledgments: We thank Prof. D. Anastassiou of Columbia University and Dr. F.-M. Wang for providing the MPEG-1 software used in our simulations.

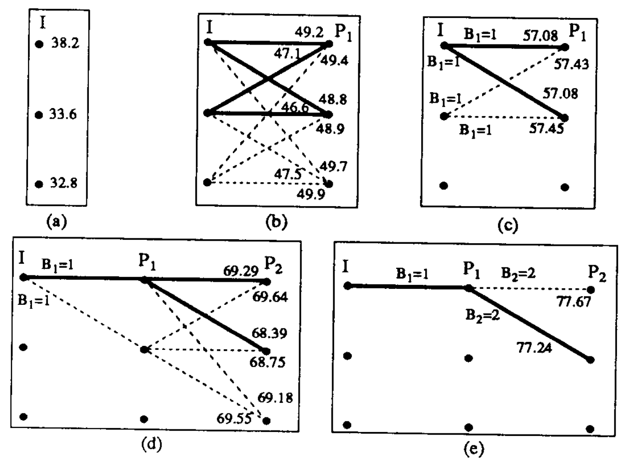

Figure 6: Tree pruning using the monotonicity property (Lemmas 1,2). The numbers are the cumulative Lagrangian costs for a typical example for $\lambda = 10$. Branches pruned at each stage are shown with dashed lines. In this example, the number of R–D points generated is cut down from 363 (exhaustive) to only 36 with no loss of optimality.

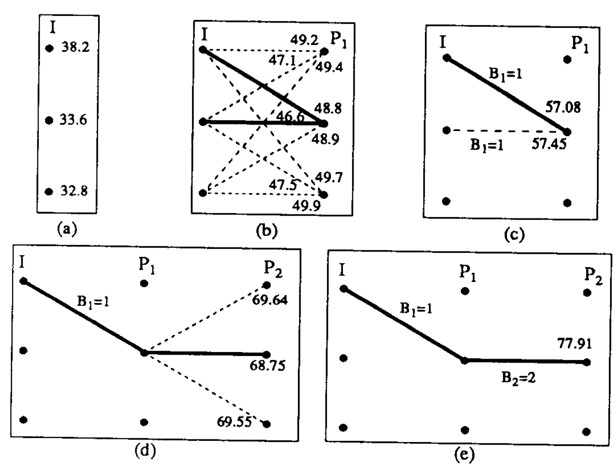

Figure 7: Tree pruning using monotonicity as well as a "greedy" heuristic for the same conditions as those of Fig. 6. The number of R-D points generated is now 24, at a slight loss of optimality (total Lagrangian cost is 77.91 versus optimal cost of 77.24).

REFERENCES

[1] "MPEG video simulation model three, ISO, coded representation of picture and audio information," 1990.

[2] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoust., Speech, Signal Proc.*, vol. 36, pp. 1445–1453, Sept. 1988.

[3] K. Ramchandran, A. Ortega, and M. Vetterli, "Efficient quantization for a multiresolution HDTV source coder," in *Proc. Fifth Intl. Workshop on HDTV, Tokyo*, vol. II, pp. 87.1–87.8, Nov. 1992.

[4] R. E. Blahut, *Theory and practice of error control codes.* Addison-Wesley, 1983.

[5] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *To be submitted to IEEE Trans. on Image Proc.*, 1993.