Communications&VideoCoding,Oct. 11--13, 1995, Brooklyn, NY, U.S.A.

A RATE-CONSTRAINED ENCODING STRATEGY FOR H.263 VIDEO COMPRESSION

Thomas Wiegand[1], Michael Lightstone[2], T. George Campbell[3]
and Sanjit K. Mitra[1]

[1]Center for Information Processing Research
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106

[2]Chromatic Research, Inc.
800A East Middlefield Road
Mountain View, CA 94043-4030

[3]Compression Labs, Inc.
2860 Junction Avenue
San Jose, CA 95134-1900

INTRODUCTION

In recent years numerous standards such as H.261 [1], MPEG-1 [2], and MPEG-2 [3] have been introduced to address the compression of video data for digital storage and communication services. Together, the applications for these standards span the gamut from low bit-rate video telephony to high quality HDTV with a new emerging standard, H.263 [4], targeting the low bit-rate end. More specifically, the primary mission for H.263 is traditionally regarded as the coding of digital video at rates suitable for transmission over public switched telephone network (PSTN) lines. Fast modems suited for this application typically run at 28.8 Kbits per second (Kb/s) within which video, audio, data, and overhead must be transmitted. This places a demanding rate constraint on the video coder which in most cases must operate at less than 24 Kb/s. In terms of wireless mobile networks whose capacities are often less than 19.2 Kb/s [5], this range of operation is also very conducive. Not surprisingly then, in addition to traditional telephony, there has been a significant and growing interest in the extension of the H.263 standard to mobile and wireless applications [6].

With regards to efficient realization, a key problem for H.263 as well as the other standards is the operational control of the encoder. Whereas most video standards uniquely stipulate the bit-stream syntax and, in effect, the decoder operation, the exact nature of the encoder is generally left open to user specification. Ideally, the encoder should balance the quality of the decoded images with channel capacity. This problem is compounded by the fact that typical video sequences contain widely varying content and motion that is often more effectively quantized if different strategies are permitted to code different regions. Currently, most video coding standards, including H.263, block-by-block basis. The advantage of the multi-mode approach is that its inherent adaptability lays the foundation for better coding results. In this paper, the goal is to provide a practical, yet efficient, algorithm for exploiting this potential in a real H.263 video encoder. It is important to note that while H.263 shares many common elements with past standards, several unique attributes such as hierarchical motion estimation and overlapped motion compensation distinguish it from the others.

To improve overall encoder performance, past papers have successfully applied rate-distortion theory to optimize the frame type and/or the quantizer selection in an MPEG system [7], [8]. A potential drawback for these approaches, however, is that the problem of selecting the best encoding strategy for a frame is not considered at the macroblock level. Rather, the optimization is accomplished by assuming a fixed number of quantization choices for each frame. For a given number of frames, a diverging trellis is generated whose paths correspond to all possible combinations of quantization choices. The job of the encoder is to find the path with the lowest total cost in the rate-distortion sense. Unfortunately, the size of the tree grows exponentially with the tree depth, and only if the number of quantization choices is relatively small can the optimal solution be feasibly found. For systems like H.263 [4], and even MPEG [2], [3], this scenario is not very likely unless the multi-mode flexibility is restricted so as to lessen the number of possible quantization choices for each frame.

In this paper, we employ a new technique [9] that formalizes the problem of encoder optimization on a macroblock-by-macroblock basis using a rate-constrained product code framework [10]. An associated Lagrangian formulation leads to an unconstrained cost function and, in the special case of mode selection, a non-diverging trellis whose associated paths correspond to all possible operational rate-distortion points for the specified image region. The best path in the trellis can be efficiently located using a dynamic programming solution based on the Viterbi algorithm [11]. The final result is an H.263 encoder algorithm that, when applied to a macroblock slice, selects the optimum combination of macroblock modes and associated mode parameters so as to minimize the overall distortion for a given bit-rate budget. To this end, the paper is organized as follows. First, we formulate the mode selection problem as it pertains to a general block-based multi-mode video coding system, and then examine a solution for obtaining the best achievable performance in the rate-distortion sense. Next, the application of this technique to H.263 is described, and results are presented for various video phone sequences at data rates from 8 to 20 Kb/s.

EFFICIENT MODE SWITCHING

Consider an image region which is partitioned into a group of blocks (GOB) given by $\mathcal{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$. For a multi-mode video coder, each macroblock in $\mathcal{X}$ can be coded using only one of K possible modes given by the set $\mathcal{I} = \{I_1, \ldots, I_K\}$. Let $M_i \in \mathcal{I}$ be the mode selected to code block $\mathbf{X}_i$. Then for a given GOB, the modes assigned to the elements in $\mathcal{X}$ are given by the N-tuple, $\mathcal{M} = (M_1, \ldots, M_N) \in \mathcal{I}^N$. The problem of finding the combination of modes that minimizes the distortion for a given GOB and a given rate constraint R_c can be formulated as

$$\min_{\mathcal{M}} D(\mathcal{X}, \mathcal{M}) \;\; \text{subject to} \;\; R(\mathcal{X}, \mathcal{M}) \leq R_c. \tag{1}$$

Here, $D(\mathcal{X}, \mathcal{M})$ and $R(\mathcal{X}, \mathcal{M})$ represent the total distortion and rate, respectively, resulting from the quantization of the GOB $\mathcal{X}$ with a particular mode combination $\mathcal{M}$.

product code framework [10]. Assuming an additive distortion measure, the cost function and rate constraint can be simultaneously decomposed into a sum of terms over the elements in $\mathcal{X}$ and rewritten using an unconstrained Lagrangian formulation so that the objective function becomes

$$\min_{\mathcal{M}} \sum_{i=1}^{N} J(\mathbf{X}_i, \mathcal{M}), \tag{2}$$

where $J(\mathbf{X}_i, \mathcal{M})$ is the Lagrangian cost function for block $\mathbf{X}_i$ and is given by

$$J(\mathbf{X}_i, \mathcal{M}) = D(\mathbf{X}_i, \mathcal{M}) + \lambda \cdot R(\mathbf{X}_i, \mathcal{M}). \tag{3}$$

It is not difficult to show that each solution to (2) for a given value of the Lagrange multiplier λ corresponds to an optimal solution to (1) for a particular value of R_c [12], [13]. Unfortunately, even with the simplified Lagrangian formulation, the solution to (2) remains rather unwieldy due to the rate and distortion dependencies manifested in the $D(\mathbf{X}_i, \mathcal{M})$ and $R(\mathbf{X}_i, \mathcal{M})$ terms. Without further assumptions, the resulting distortion and rate associated with a particular block in the GOB is inextricably coupled to the chosen modes for every other block in $\mathcal{X}$. On the other hand, for many video coding systems, the bit-stream syntax imposes additional constraints that can further simplify the optimization problem.

For example, we can restrict the codec so that both the rate and distortion for a given image macroblock are impacted by only (i) the content of the current block and its respective operational mode, (ii) the content of the current and previous block, and (iii) the content of the current, previous, and ensuing block. These three cases correspond to simplified Lagrangians given by

$$\text{(i)} \quad J(\mathbf{X}_i, \mathcal{M}) = J(\mathbf{X}_i, M_i), \tag{4}$$

$$\text{(ii)} \quad J(\mathbf{X}_i, \mathcal{M}) = J(\mathbf{X}_i, M_{i-1}, M_i), \quad \text{and} \tag{5}$$

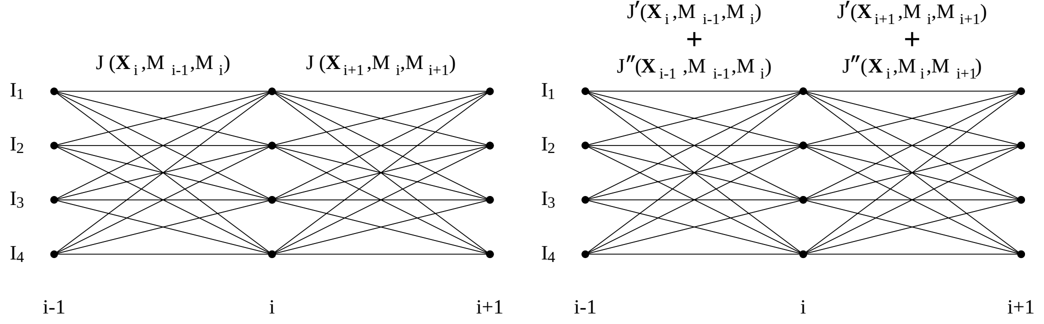
$$\begin{aligned} \text{(iii)} \quad J(\mathbf{X}_i, \mathcal{M}) &= J(\mathbf{X}_i, M_{i-1}, M_i, M_{i+1}) \\ &= J'(\mathbf{X}_i, M_{i-1}, M_i) + J''(\mathbf{X}_i, M_i, M_{i+1}), \end{aligned} \tag{6}$$

respectively. For scenario (i), the optimization problem of (2) can be easily minimized by independently selecting the best mode for each macroblock in the GOB. For this particular case, the problem formulation is equivalent to the bit allocation problem for an arbitrary set of quantizers, proposed earlier by Shoham and Gersho in [13], and specifically for video coding by Wu and Gersho in [14]. The drawback is that this structural constraint is rather restrictive and does not correspond to the way macroblocks are coded in most video coding standards such as H.261, MPEG-1, MPEG-2, and especially H.263.

For instance, a block-to-block dependency typically exists due to the differential encoding of motion vectors such that the rate term for a given macroblock is dependent not only on the current mode but on the modes of previous blocks as well, leading to case (ii) and the Lagrangian of (5). For overlapped motion compensation (as found in H.263), the dependency also manifests itself in the distortion terms, introducing dependencies on past and future blocks. This behavior corresponds to case (iii) and the Lagrangian of (6)[1].

---

[1] Here we assume that the influence of the previous block can be separated from the influence of



(a) (b)

Figure 1. Resulting multi-mode trellis for the cases when the rate and distortion dependencies are (a) on past macroblocks and (b) on past and future macroblocks.

For both (5) and (6), we can obtain the optimal solution by viewing the search for the best combination of N modes in the GOB as an equivalent search for the best path in a trellis of length N. For case (ii), the nodes in the trellis for $i = 1, \ldots, N$, are given by the elements in $\mathcal{I}$, and the transitional costs from node M_{i-1} to node M_i are given by Lagrangian cost terms specified in (5). This trellis, shown in Fig. 1(a) for $K = 4$, can be efficiently searched using the Viterbi algorithm to obtain the optimal solution to (2). Similarly, the Viterbi algorithm can be implemented to obtain an optimal solution for case (iii) by searching the trellis described in Fig. 1(b). Further details on this mode selection technique, including the joint optimization of the mode selection with the selection of macroblock parameters such as the quantization step size, can be found in [9].

A final critical consideration with regards to mode selection is the determination of the Lagrange multiplier λ. Recall that while the solution to the unconstrained Lagrangian cost function for any value of λ results in minimum distortion for some rate, the final rate cannot be specified a priori. Often it is desirable to find a particular value for λ so that upon optimization of (2), the resulting rate closely matches a given rate constraint R_c. Because of the monotonic relationship between λ and rate, a possible solution is the bisection search algorithm [15], [16]. However, the computation associated with the re-optimization of (2) for numerous values of λ may preclude such a search in a practical encoder. As an alternative, we have considered a variety of successful approaches including a frame-to-frame update of λ using least-mean-squares (LMS) adaptation [17]. In our experiments (provided at the end of the paper), we employ a method for determining the LMS step-size dynamically for each frame or GOB (indexed by k) of the video sequence [18]. The strategy effectively reduces the bursty behavior of adaptation and results in an update procedure for the Lagrange multiplier given by

$$\begin{aligned} \lambda_{k+1} &= \lambda_k + \frac{1}{R_k^2}(R_c - R_k)R_k \\ &= \lambda_k + (\frac{R_c}{R_k} - 1). \end{aligned} \tag{7}$$

In summary, it is important to note that whether the bisection or LMS algorithm is uti-

---

the subsequent block—which is the case in H.263. For this situation, the transitional cost from node M_{i-1} to node M_i is given by the sum of two terms, $J'(\mathbf{X}_i, M_{i-1}, M_i)$ and $J''(\mathbf{X}_{i-1}, M_{i-1}, M_i)$ which constitute the contribution from the preceding and ensuing macroblocks, respectively.

outcome that—no matter what bit rate results—the distortion of the GOB will be minimum for that rate. This is in striking contrast to other encoder strategies that typically scale a single parameter such as the quantizer step size to control the instantaneous rate, but cannot guarantee any type of optimal rate-distortion performance.

APPLICATION TO H.263

We now consider the application of the rate-constrained mode switching algorithm to H.263, the International Telecommunication Union's (ITU) draft recommendation for video coding over narrow telecommunications channels [4]. As is the case with the other standards, in H.263 each frame of the image sequence is first subdivided into unit regions called macroblocks which relate to 16 pixels by 16 lines of the luminance component (Y) and the spatially corresponding 8 pixels by 8 lines of both chrominance components ($\mathrm{C_B}$ and $\mathrm{C_R}$). As part of H.263, each macroblock can also be coded using any one of several possible modes, the allowable set of which is determined by the picture coding type.

The recommendation for the standard contains two picture coding types, INTRA and INTER which specify the possible macroblock modes that may be used for the current frame. The INTRA picture type is more limiting in that it only allows intra coding for macroblocks. It is typically used only for special purposes, e.g., coding the first frame of a video sequence. In this paper, we concern ourselves with the INTER picture type because within this picture type, individual macroblocks can be coded using a large variety of macroblocks modes, including intra and inter. Specific to H.263 is an additional capability called Advanced Prediction which enforces overlapped motion compensation and permits the use of four motion vectors per macroblock. This function can be set by a single bit and impacts the macroblock modes for an entire frame. For our simulations we include the following standard and optional macroblocks modes: intra (I-mode), inter with one motion vector (P-mode), inter with four motion vectors ($P4$-mode), and uncoded (U-mode) which we now briefly describe.

In the I-mode, the luminance and chrominance components are quantized using a "JPEG-like" coding scheme. The components are initially segmented into 8×8 blocks which are subsequently transformed by the DCT. All AC transform coefficients are then identically scalar quantized with an even step-size value ranging from 2 to 62. Next, the coefficients are "zig-zag" scanned and losslessly encoded using a look-up table that exploits long runs of zeros. Special attention is paid to the quantization of the DC transform coefficient as it is uniformly scalar quantized using an 8 bit codeword. Typically, the quantizer step size is fixed for all macroblocks in a GOB. However, as part of the H.263 standard, the encoder can set a two-bit option in the macroblock header which permits a change in the quantizer step-size of ± 1 or ± 2 for all succeeding macroblocks. For our experiments, this type of macroblock-by-macroblock parameter adjustment is not considered for now due to the associated complexity required for its optimization, though in principle it is not a fundamental obstacle.

In the P-mode, the current macroblock is first predicted using a single, half-pixel accurate motion vector. Each motion vector points to a 16×16 luminance region and two 8×8 chrominance regions in the previously decoded frame within a horizontal and vertical range of -16 to $+15.5$ pixels. Once determined, the motion vectors are differentially encoded after each vector is first predicted using the median of three candidate vectors. The candidate vectors correspond to the three surrounding motion vectors located directly above, above and to the right, and directly left of the current

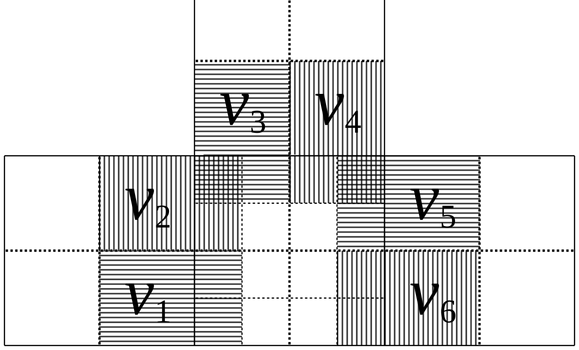


Figure 2. Illustration of H.263 overlapped motion compensation.

motion vector, respectively. Each motion-error term is encoded without loss using a single variable-length codeword from a fixed look-up table. Next, the resulting motion-compensated prediction error is transformed and quantized in the same manner as the I-mode, with the exception that the DC coefficient is not treated separately. The incremental modification of the quantizer step size for individual macroblocks, while allowed by the H.263 standard, is not considered in this paper.

When Advanced Prediction is turned off, both the I and P-modes act very similarly as in past standards such as H.261 and MPEG. In contrast, when the Advanced Prediction bit is set, the P-mode is modified to include overlapped motion compensation [19]. Moreover, by flipping this bit, an additional macroblock mode can be utilized that not only includes overlapped motion compensation, but also specifies four motion vectors per macroblock. In this mode, which we refer to as the $P4$-mode, the macroblock is segmented into four smaller 8×8 blocks, each compensated by one of the four specified motion vectors in the same manner that the larger 16×16 blocks are compensated in the P-mode. An important point is that the $P4$-mode must be used in conjunction with another special functionality of H.263, called the unrestricted motion vector mode, in order to allow the lapping of pixels located outside the frame boundaries. This function is similarly set by a single bit for an entire frame and is defined such that the pixels from the border of the picture are copied to the regions outside. The lapping from the outside into the current macroblock is depicted in Fig. 2. The vectors $\{\boldsymbol{v}_1 ... \boldsymbol{v}_6\}$ are the motion vectors from the neighboring macroblocks, and the lapping is performed using fixed weighting windows. Within a macroblock, each of the four smaller luminance blocks is similarly predicted by internally applying overlapped motion compensation between the blocks. The exact procedure for differentially encoding the four motion vectors is detailed in the recommendation. Otherwise, the same prediction loop as previously described is applied, and the quantization is performed as explained for the P-mode. Finally, the uncoded mode (U-mode) (which is indicated by just a single bit for a given macroblock) specifies that the current macroblock is to be represented by simply duplicating the contents of the corresponding macroblock in the previous frame.

According to the standard [4], "the criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the coding control strategy." In what follows, we consider the application of the mode selection strategy described previously as an encoder control solution for the H.263 standard. Our goal is to determine the optimum mode selection for a given GOB. For all simulations, the GOB is defined as a single, horizontal macroblock stripe across a given frame. For example, a 176×144 QCIF-image consists of 9 macroblock stripes, each containing 11 macroblocks. We restrict ourselves to this scenario so that the Viterbi algorithm. This approach also lends itself to wireless scenarios in that the generation of GOB's on a regular interval facilitates the recovery from bit errors which are more likely in the wireless environment.

We note that whereas, in general, the coding of a given macroblock in H.263 is influenced by the selected mode of neighboring blocks, there are two notable exceptions for this type of dependency: the I-mode and the U-mode in which the mode selection can be carried out independently of the surrounding macroblocks. Because there is no transitional cost between modes, the costs for these nodes can be assigned using (4). For the P-mode, the rate term is dependent on three neighboring macroblocks due to the differential encoding of the motion vectors. By restricting the GOB to a horizontal macroblock stripe, we can eliminate the impact on the trellis from above and need only consider those dependencies resulting from the immediately preceding macroblock. Consequently, we can assign a transitional cost from the previous node to the current node using (5).

In the case of Advanced Prediction, for both the P and $P4$-mode, rate and distortion are dependent on the previous choice for the macroblock mode, while the distortion is dependent on the succeeding macroblock mode as well. Using (6), we can compute the cost for the incoming and outgoing transitions of the current node assigned for the P and $P4$-modes as follows. As illustrated in Fig. 2, the distortion of the left half of the macroblock is only influenced by the motion vectors of the macroblock to the left and from the above. The macroblocks modes from above are fixed because they are determined in the previous GOB, and thus, we need only consider the influence from the left when computing the distortion component of $J'(\cdot)$ in (6). Analogously, all distortion influences except those from the right can be eliminated when computing the distortion component in $J''(\cdot)$. Likewise, the distortion for both chrominance components are equally distributed to the in and outcoming transitions. In terms of rate, the cost assignment to the trellis branches is slightly more complicated because the motion vectors on the right half of the $P4$-mode are predicted from the motion vectors to the left. Consequently, a dynamic update for $J'(\cdot)$ and $J''(\cdot)$ based on the decisions for the incoming transitions is required. Finally, the quantizer step size parameter, QUANT, is optimized using the strategy outlined in [9] for each GOB.

CODING RESULTS

Simulation results for the proposed mode switching strategy as applied to H.263 are provided in Figs. 3–4. In the first experiments, the frame rate is held constant at 8.33 frames per second and the Lagrange multiplier λ is varied to generate coded sequences with an overall average rate from 2.9 Kbits per second (Kb/s) to 400 Kb/s. Though fixing λ for the video sequence does not represent a practical implementation since the maximum rate is not constrained, it does provide a means for assessing the relative importance of each mode at different bit rates. For example, Fig. 3 demonstrates the probability of selecting the I, P, $P4$, and U modes after running the algorithm on the well-known color video sequence, "Grandmother." Similar plots for "Mother-Daughter" and "Car Phone" can be found in [9]. Upon close examination, several intuitively appealing results are confirmed by the plots. For instance, the U-mode, as expected, drops to zero at high rates (when $\lambda = 0$) for all of the test sequences. For the "Grandmother" and "Mother-Daughter" sequences, where the motion model is fairly accurate, the I-mode is almost never chosen except at the highest rates. In contrast, for the "Car Phone" sequence which has more complex motion, the algorithm erratic behavior in the mode probabilities at the highest rates can be attributed to the logarithmic scale of the plots, and the fact that only a few data points were computed at these rates since they are outside the recommended usage for H.263.

Next, the overall rate-distortion performance for the three sequences is shown for the mean absolute error (MAE) in Fig. 4. The plots are generated by varying λ from 0 to 50. It is interesting to note that the relationship between λ and distortion (for rates above 10 Kb/s) is rather consistent between the three different sequences, i.e., the same value of λ corresponds roughly to the same value of MAE in all cases. If the primary objective is a constant-distortion coder, then this is good news in that the Lagrange multiplier need not be substantially modified from one frame to the next. Unfortunately, the same desirable relationship does not manifest itself for rate and λ. In fact, depending on the sequence, the same value of λ may correspond to widely varying bit rates. Thus, if the goal is coding for a specified rate, which is more often the case, a method for for controlling the Lagrange multiplier is required. For this reason, we consider a more sophisticated encoder control strategy than before in which the frame-skip is adaptive, and the LMS algorithm is used to update λ on a frame-by-frame basis. These steps are undertaken to generate a more constant rate, and thus, a more practical system. The coding results of the scheme are very encouraging, producing usable video with rates as low as 6.9 Kb/s, 11.0 Kb/s, and 18.3 Kb/s for the "Grandmother", "Mother-Daughter", and "Car Phone" sequences, respectively. The maximum rates for these scenarios correspond to 8 kb/s, 12 Kb/s, and 20 Kb/s. Note that the average rates are very close to the maximum allowable rates for all three of the sequences, confirming the applicability of the LMS algorithm with regards to rate control. Interested readers can also refer to [9] in order to view sample still images taken from these coded sequences.

ACKNOWLEDGMENTS

This work was supported in part by the Ditze Foundation, a National Science Foundation Graduate Fellowship and in part by a University of California MICRO grant with matching supports from Hughes Aircraft, Signal Technology Inc., and Xerox Corporation. The authors are greatful for the helpful comments of Jong Dae Kim.

REFERENCES

[1] ITU-T Recommendation H.261, "Video codec for audiovisual services at $p \times 64$ kbit/s", Dec. 1990, Mar. 1993 (revised).

[2] ISO/IEC 11172-2, "Information technology–coding of moving picture and associated audio for digital storage media at up to about 1.5 mbit/s: Part 2 video", Aug. 1993.

[3] ITU-T Recommendation H.262—ISO/IEC 13818-2, "Information technology–generic coding of moving picture and associated audio for digital storage media at up to about 1.5 mbit/s: Video", (Draft), Mar. 1994.

[4] ITU-T Recommendation H.263, "Video coding for narrow telecommunication channels at less than 64 kbit/s", (Draft), April 1995.

[5] ITU-T, SG15, WP15/1, Expert's group on Very Low Bitrate Video Telephony, LBC-95-193, Delta Information Systems, "Description of mobile networks", June 1995.

of recommendation H.263 towards mobile applications", June 1995.

[7] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders", *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 533–545, Sept. 1994.

[8] J. Lee and B.W. Dickinson, "Joint optimization of frame type selection and bit allocation for MPEG video coders", in *Proc. of the Int. Conf. on Image Proc.*, 1994, vol. II, pp. 962–966.

[9] T. Wiegand, M. Lightstone, T.G. Campbell, and S.K. Mitra, "Efficient mode selection for block-based motion compensated video coding", in *Proceedings of the 1995 IEEE International Conference on Image Processing (ICIP '95)*, Washington, D.C., Oct. 1995, (To appear).

[10] M. Lightstone, D. Miller, and S.K. Mitra, "Entropy-constrained product code vector quantization with application to image coding", in *Proceedings of the First IEEE International Conference on Image Processing*, Austin, Texas, Nov. 1994, vol. I, pp. 623–627.

[11] G.D. Forney, "The Viterbi algorithm", *Proceedings of the IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[12] H. Everett III, "Generalized lagrange multiplier method for solving problems of optimum allocation of resources", *Operations Research*, vol. 11, pp. 399–417, 1963.

[13] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers", *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, Sept. 1988.

[14] S.W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 1, no. 1, pp. 100–112, March 1991.

[15] J.E. Dennis and R.B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall,, Englewood Cliffs, NJ, 1983.

[16] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense", *IEEE Trans. on Image Processing*, vol. 2, no. 2, pp. 160–175, Apr. 1993.

[17] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, 1991.

[18] M. Rupp, "Bursting in the LMS algorithm", 1995, Submitted for publication.

[19] H. Watanabe and S. Singhal, "Windowed motion compensation", in *Proc. of the SPIE Symposium on Visual Comm. and Image Proc.*, 1991, vol. 1605, pp. 582–589.

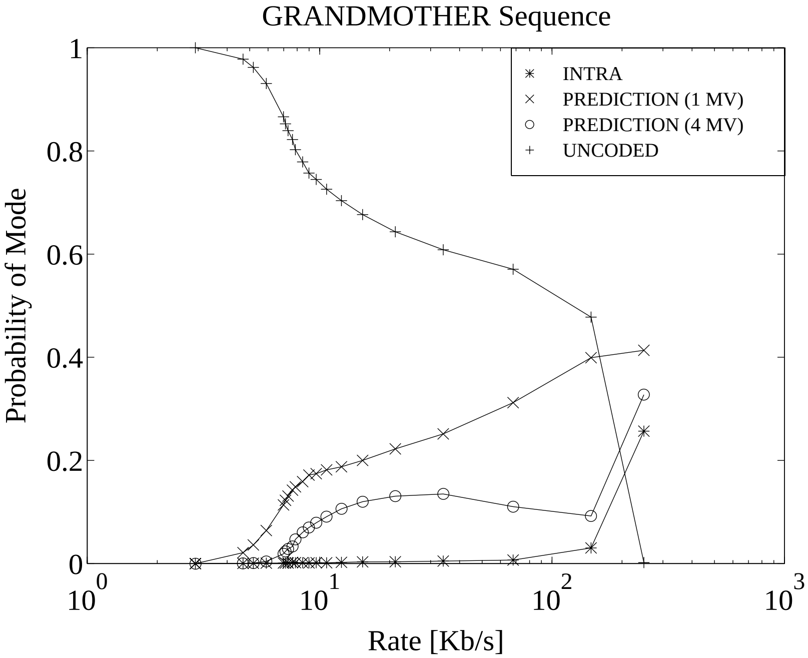


Figure 3. Probability of mode versus rate for the "Grandmother" sequence.

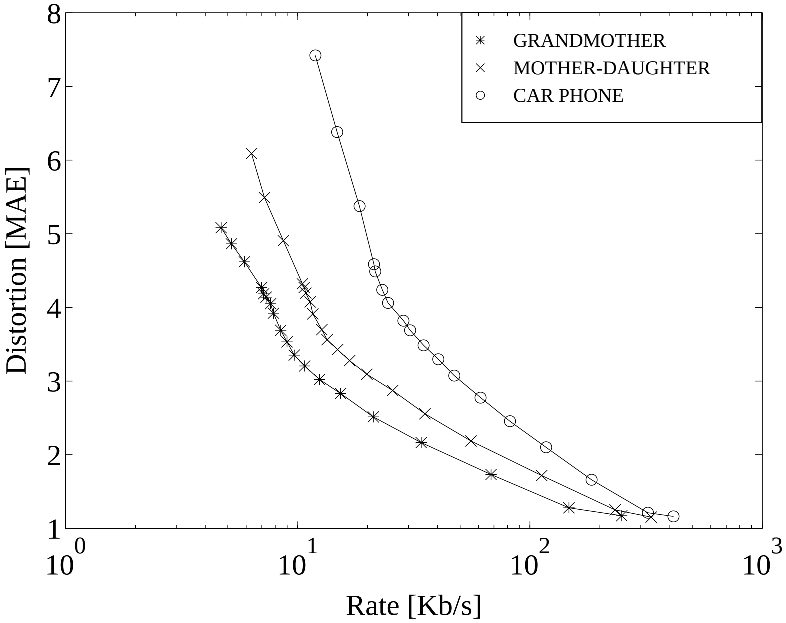


Figure 4. Plot of rate versus distortion for the "Grandmother", "Mother-Daughter", and "Car Phone" video sequences. The Lagrange multiplier λ is varied from 50 to 10 with an equidistant step size of 10, from 10 to 5 with step 1 and from 5 to 0 with step 0.5. Note: the frame skip is held constant at 2 for a frame rate of 8.33 frames per second.