Joint block-based video source/ channel coding for packet-switched networks

Raynard Hinds, Thrasyvoulos Pappas, Jae Lim

Raynard O. Hinds, Thrasyvoulos N. Pappas, Jae S. Lim, "Joint block-based video source/channel coding for packet-switched networks," Proc. SPIE 3309, Visual Communications and Image Processing '98, (9 January 1998); doi: 10.1117/12.298322

Event: Photonics West '98 Electronic Imaging, 1998, San Jose, CA, United States

Joint block-based video source/channel coding for packet-switched networks

R. O. Hinds[a], T. N. Pappas[b], and J. S. Lim[a]

[a]Advanced Television and Signal Processing Research Group, Rm 36-647
Massachusetts Institute of Technology Cambridge, MA 02139 USA

[b]Bell Laboratories Innovations for Lucent Technologies, Rm 2D-336
600 Mountain Avenue Murray Hill, NJ 07974-0636 USA

ABSTRACT

Block-based video coders rely on motion compensated block prediction for more data compression. With the introduction of video coding over packet-switched networks such as the Internet and the resulting packet loss that occurs on congested networks, coding mode selection for each macro-block is significant in determining the overall distortion on the decoded video sequence. In this work, we examine the problem of mode selection for macro-blocks in the presence of loss and a channel rate constraint. We present and evaluate several methods for mode selection that attempt to minimize perceptual distortion from packet loss. We formulate a simplified problem which is useful for gaining some insight, and present an efficient algorithm to finding an optimal solution.

Keywords: Video coding, Packet loss, Internet, Joint source/channel coding

1. INTRODUCTION

Block-based coding is used in several current video coding standards. When coding video with these coders, the quantization step-sizes and the coding mode for each macro-block must be chosen to result in the most efficient coding scheme meeting all constraints. There is a trade-off between the coding rate and the resulting distortion in the decoded video sequence. For constrained rate-distortion optimization, the video sequence should be coded at the lowest rate that meets the distortion constraint or the lowest distortion that meets the rate constraint. This is a difficult problem to solve completely because of the dependencies that exist between the quantization step-sizes, coding modes, and resulting bit-rates for macro-blocks. In past work, simpler problems have been investigated 1–4.

With the introduction of video coding over packet-switched networks such as the Internet, the coding problem just described has become even more complex. Packet loss on congested networks is a major problem for real-time applications. The user is usually forced to send at a very low bit-rate in order not to exceed the available bandwidth. It is safer to have all information loss occur at the source coder where the user can control what is lost with compression. With network transmission, optimal coding methods now have to consider the distortion from packet loss along with the distortion from source coding. The coding rate must be increased to allow the decoder to recover from packet loss or to effectively conceal errors. This paper will look at how the coding problem has been changed with the introduction of potential loss as well as a solution to a simplified problem in network video coding. We specifically address the problem of making coding mode decisions which has a major impact on the perceptual distortion in the presence of loss.

2. SOURCE CODING

When coding video with a block-based coder, each frame is partitioned into S macro-blocks. Let X_n^i denote the original macro-block to be source coded in the n^{th} frame at spatial location i, and $\boldsymbol{X} = \{X_n^i \; \forall\, i, n\}$. Let Y_n^i be the result of decoding the compressed X_n^i at the source, and $\boldsymbol{Y} = \{Y_n^i \; \forall\, i, n\}$. The coding parameters must be chosen at the source for optimal rate-distortion performance. Solving the joint problem of quantization step-size and mode selection for optimal source coding has proven to be very difficult. However, the simpler problem of finding one set of parameters while holding the others fixed has been addressed. For example when fixing the quality of each frame, the coding mode for each macro-block must be found to code the sequence at the lowest bit-rate for most efficient coding. Here we assume that the quantization step-sizes are either constant or determined through some method defined to result in consistent frame quality. In both cases, the rate-distortion function is just a horizontal line at the fixed distortion. Since the frame quality is fixed, the distortion remains constant as other coding parameters are varied. Changes in mode selection will only affect the bit-rate of the coded sequence.

Without dependencies among macro-block coding parameters, this problem is easily solved by selecting the coding mode for each macro-block independently. The macro-block is coded in both modes at the same quality, and the mode with the lowest bit-rate is chosen. However, dependencies do exist among coded macro-blocks from differential coding of macro-block parameters. Kwok offers a solution to this problem of mode selection with fixed quality when these inter-macro-block dependencies are taken into consideration.[2]

Ramchandran[3] addresses the dual problem of finding the optimal quantization step-sizes given a predetermined MPEG Group of Picture (GOP) structure. In this case, the frame type is specified and optimal quantization is determined. The GOP consists of 3 types of picture frames I, B, and P. Every macro-block in the I frame is coded in intra mode. The macro-blocks in P frames can be coded in either intra or inter mode. It is often assumed they will be coded in inter mode. When coded in inter mode the prediction is formed from macro-blocks in previous I or P frames. The prediction for every macro-block in B frames is interpolated from macro-blocks in nearest I and P frames. The macro-blocks in B frames can not be used for prediction.

When coding in inter mode, the quantization step-size for macro-blocks used for prediction determines the rate-distortion function for the macro-blocks to be coded in P and B frames. Ramchandran acknowledges dependent quantization when solving the problem and takes advantage of the monotonicity property of the rate-distortion function of dependent frames to avoid exhaustive search of all quantization parameter combinations. Lee[4] extends the solution to solve the problem of finding the number and position of the P frames along with the quantization values in the GOP. In both problems, it is still unclear exactly how the macro-block coding mode for each macro-block is chosen in the P frames and no argument is ever made that the particular mode choices are optimal.

3. NETWORK VIDEO CODING

With the onset of video coding over networks and the resulting loss that may occur, the first problem described above of solving for optimal mode selection when the quantization step-sizes are fixed becomes more difficult. We address the problem of selecting the modes for each macro-block to minimize distortion after concealment when coded at constant frame quality at the source. In this case, the total distortion at the receiver consists of 2 separate components. Let $\hat{Y}_n^i$ represent the macro-block in the n^{th} frame at spatial location i at the receiver after decoding and concealment, and $\hat{\boldsymbol{Y}} = \{\hat{Y}_n^i \; \forall \, i, n\}$. Then the total distortion becomes

$$D(\boldsymbol{X}, \hat{\boldsymbol{Y}}) = D^s(\boldsymbol{X}, \boldsymbol{Y}) + D^c(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) \tag{1}$$

The first component (D^s) is distortion introduced at the source from compression, and the second (D^c) is distortion from packet loss that can occur on the network channel. The resulting quality at the receiver is not constant. The potential for loss makes mode selection more critical for each macro-block, and has different effects when macro-blocks are coded in intra mode and inter mode. When coding in inter mode, error persistence from lost macro-blocks is a problem, while errors are short-lived when macro-blocks are coded in intra mode. However, coding every macro-block in intra mode may lead to inefficient use of the channel.

Source coding at constant quality is equivalent to compressing at a fixed distortion. Since by problem specification the coding distortion for each frame is constant at the source regardless of mode selection for the macro-blocks ($D^s(X_n^i, Y_n^i) = K \; \forall \, i, n$), we can ignore that component of the distortion at the receiver and look at choosing optimal coding modes in the rate-channel-distortion sense. This leads to a joint source/channel coding problem.

3.1. Problem formulation

We first impose 3 simplifications to the problem for tractability. To solve this problem, the coded data must be divided into packets, the network channel modeled, and the distortion metric to be used chosen. Data should be packed to minimize the effects of a single lost packet. Macro-blocks are the smallest coding units in block-based video coders. On a single priority network, coded data should be divided into packets at macro-block boundaries. In addition, coding state information must be put into the header of every transmitted packet so that each packet can be independently decoded. In our investigation this strategy is taken to its extreme, and our first simplification is to restrict a packet to contain a single macro-block.

When motion compensation is used, motion vectors are transmitted to allow a macro-block size region from any location in the previous frame to be used to predict the current macro-block. When non-zero motion vectors are allowed, optimal mode selection is more difficult because of the resulting dependency among coded macro-blocks at different spatial locations throughout the sequence. Our second simplification is to restrict the coder to zero-vector displacement which decouples the macro-block coding mode selection at different spatial locations. When a macro-block is coded in inter mode, the decoded macro-block at the same spatial location in the previous frame is used for prediction. We independently solve the mode selection problem for the sequence of macro-blocks at each spatial location i.

Neglecting isolated bit errors which are rare on the Internet as well as detectable from error detection codes in the packet header, the major cause of distortion on packet-switched networks is from loss of contiguous bits of data with each lost packet. With the addition of sequence numbering in the packet, lost data packets can be detected at the receiver, and packet loss becomes channel erasure. Attempts to accurately characterize this erasure on the networks have been very problematic for researchers. In line with other attempts to analyze this problem, we have decided to use probabilistic models to describe the packet loss that occurs in the network. It is generally known that loss on packet-switched networks is bursty, and it has been shown that certain probabilistic models can describe the loss observed on a network. Our third simplification is to model the channel loss as an independent Bernoulli process where each macro-block is lost with probability p.

By packing at the macro-block level, each arriving macro-block can be decoded at the receiver regardless of past losses that may have occurred on the network. However when decoding macro-blocks coded in inter mode, the macro-block used for prediction at the transmitter may not be available at the receiver. The arriving residual even though it can be decoded is useless. Therefore in our problem, we consider any arriving residual with lost prediction as being lost.

The problem to be solved is to find the mode selection that will result in minimal channel distortion at the receiver after decoding and concealment. Let $\boldsymbol{M^i} = M_0^i, M_1^i, \ldots, M_{N-1}^i$ denote the coding modes for the sequence of transmitted macro-blocks at spatial location i. We impose a total rate constraint to ensure efficient use of the channel. Given a distortion metric $D(Y_n^i, \hat{Y}_n^i, \boldsymbol{M^i})$ where we explicitly include the dependence upon the mode selection, we want to select the modes for each macro-block to minimize the total distortion for a total bit-rate constraint. We formulate the problem as

$$\min_{\boldsymbol{M}} D(\boldsymbol{Y}, \hat{\boldsymbol{Y}}, \boldsymbol{M}) = \sum_{i=0}^{S} \sum_{n=0}^{N-1} D(Y_n^i, \hat{Y}_n^i, \boldsymbol{M^i}), \qquad \textit{subject to } R(\boldsymbol{M}) \leq R_c \tag{2}$$

We use a Lagrange multiplier to solve this constrained optimization problem. An equivalent problem is to minimize for an appropriate choice of λ the Lagrangian cost function:

$$J(\boldsymbol{Y}, \hat{\boldsymbol{Y}}, \boldsymbol{M}) = \lambda D(\boldsymbol{Y}, \hat{\boldsymbol{Y}}, \boldsymbol{M}) + R(\boldsymbol{M}) \tag{3}$$

3.2. Channel distortion metrics

An appropriate distortion metric which can characterize the perceptual effects of packet loss on the coded video data must be determined. In this paper, we examine three reasonable distortion metrics which are used in our problem formulation and evaluate the results to optimal mode selection in each case in the presence of loss.

3.2.1. Maximize the number of intra mode macro-blocks

From our knowledge of the relationship between distortion from packet loss and mode selection, we know we want to encourage the use of intra coded macro-blocks. A reasonable method for mode selection would be to code as many macro-blocks as possible in intra-mode while meeting the given rate constraint. This is equivalent to minimizing the number of inter coded macro-blocks at that rate. The first distortion metric we consider $D(Y_n^i, \hat{Y}_n^i, \boldsymbol{M^i}) = D_I(M_n^i)$ penalizes the use of inter coded macro-blocks.

$$D_I(M_n^i) = \begin{cases} 0 & \text{if } M_n^i = \text{intra mode} \\ 1 & \text{if } M_n^i = \text{inter mode} \end{cases} \tag{4}$$

This distortion metric is a function of only the parameters chosen at the coder. No knowledge of the channel loss probability or concealment at the receiver is used to optimize mode selection.

3.2.2. Minimize probability of error

To measure distortion at the receiver due only to the loss that occurs in the channel, we next consider the Hamming distortion metric $D_H(Y_n^i, \hat{Y}_n^i)$, where

$$D_H(Y_n^i, \hat{Y}_n^i) = \begin{cases} 0 & \text{if } \hat{Y}_n^i = Y_n^i \\ 1 & \text{if } \hat{Y}_n^i \neq Y_n^i \end{cases} \tag{5}$$

This metric is attractive to use since there are an unlimited number of methods to conceal the distortion at the receiver which work to varying degrees. To decouple the choice of concealment method from the problem of optimal coding mode selection, we make the assumption that lost macro-blocks will not be able to be adequately concealed, and it is equally bad to lose any macro-block. This assumption is not far from correct in many examples. We look at $D(Y_n^i, \hat{Y}_n^i, M^i)$ equal to the expected distortion $E[D_H(Y_n^i, \hat{Y}_n^i)]$ between the macro-block at the transmitter and receiver. When trying to code to minimize the expected distortion, this formulation results in coding to minimize the probability of error (MPE) which is consistent with traditional attempts to specify channel coding criteria. The probability of error for Y_n^i is equal to $1-(1-p)^{k+1}$ where k is the number of consecutive inter coded macro-blocks up to and including Y_n^i. It is easy to show the total expected distortion for a sequence of N transmitted macro-blocks at spatial location i with loss probability p is bounded above and below by:

$$Np \leq \sum_{n=0}^{N-1} E[D_H(Y_n^i, \hat{Y}_n^i)] \leq N - \frac{(1-p)}{p}[1-(1-p)^N] \tag{6}$$

The new rate-expected-distortion curve at a fixed source coding quality is not a horizontal line. Figure 1 shows an example form of the rate-expected-distortion curve. The most robust coding method will code all macro-blocks in intra mode to minimize the expected distortion at the lower bound (point C). This will result in a very high coding rate. The least robust coding strategy will code all macro-blocks beyond the first frame in inter mode with the expected distortion reaching the upper bound at point A. The expected distortion for the rate optimal selection in which the coding modes are selected to minimize the bit-rate will lie at point B with expected distortion somewhere in-between the distortion at points A and C. When fairly accurate predictions can be made from the decoded macro-block in the previous frame, the current macro-block is coded in inter mode to minimize the rate. Otherwise, the macro-block is coded in intra mode. A trade-off exists between coding to minimize rate and coding to minimize expected channel distortion. Coding more macro-blocks in inter mode than what is used to code at the minimum bit-rate (point B) will only lead to higher distortion. There is no reason to consider mode selections that lead to both a higher rate and distortion than the mode selection for minimum bit-rate selection.

3.2.3. Minimize mean-square-error

For a given concealment method used at the decoder, we can select modes at the encoder to minimize the mean-square-error $(MMSE)$ distortion between the original and decoded sequence after concealment. The third distortion metric we consider is $D_{MSE}(X_n^i, \hat{Y}_n^i)$. In this case, we actually measure the total error between the original and reconstructed macro-block at the receiver.

$$D_{MSE}(X_n^i, \hat{Y}_n^i) = ||X_n^i - \hat{Y}_n^i||^2 \tag{7}$$

Again, we select modes to minimize the expected distortion for a given rate constraint.

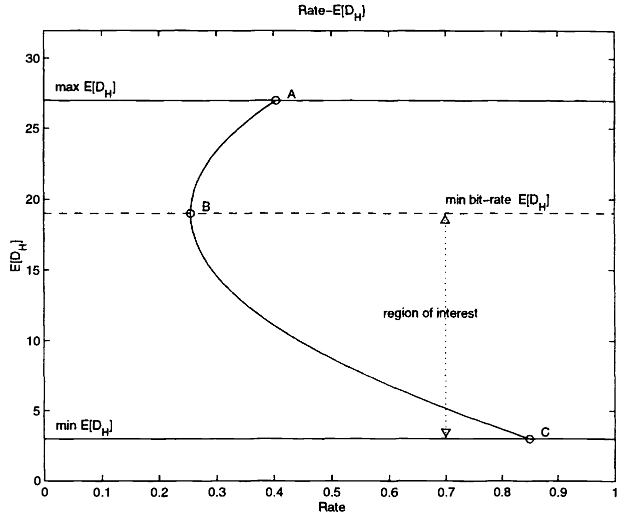


Figure 1. Example rate-expected-distortion function.

4. JOINT SOURCE/CHANNEL CODING

For efficient channel coding, the minimum number of additional bits is used to reduce the probability of decoding error after transmission. If the total rate is constrained, bits must be allocated between source and channel coding. In previous work, Lagrange multipliers have been used to allocate bits optimally for chosen source and channel coders.[5] Our problem can be viewed to be similar to these former problem formulations. Let $R_Q(n)$ equal the number of bits needed to code a macro-block at time n in intra mode and $R'_Q(n; \boldsymbol{M})$ equal the number of bits needed to code it in inter mode all at a given quality Q. We momentarily ignore the dependency of $R'_Q(n; \boldsymbol{M})$ on mode selection $\boldsymbol{M}$. We define the function $B_Q(n) = R_Q(n) - R'_Q(n)$ to be the difference in bits used to code a macro-block in intra and inter mode. When $B_Q(n) > 0$, those additional bits to code in intra mode can be viewed as channel coding bits when used, since they add redundancy to the compressed bit stream to reduce the probability of error. The function $B_Q(n)$ is determined by the video sequence and the coded video quality. In the following section, we show a method to select the modes to minimize the Lagrangian cost function $J(\boldsymbol{Y}, \hat{\boldsymbol{Y}}, \boldsymbol{M})$ for any of the previous distortion metrics described.

4.1. Minimization

We form a trellis for each macro-block location i in the picture frame. Each state at time n accounts for a different possible number of consecutive inter-coded macro-blocks up to time n. Figure 2 shows an example of a trellis for mode selection. Since all macro-blocks in the first frame are coded in intra mode, there is only one state at time 0. The trellis diverges linearly with time. By constructing the trellis in this manner and expanding out all possible states, there is only a single Lagrangian cost value associated with each state transition between time steps. We use the Viterbi algorithm to efficiently optimize mode selection for the sequence of macro-blocks at spatial location i. The computational burden is in calculating the cost function associated with each link.

Since the number of states in the trellis is not constant at every time step but increases by 1 at each stage, the number of link costs to calculate for N stages is equal to $N(N-1)$. Each link on the trellis has a cost equal to the λ times the distortion plus the number of bits needed to code the macro-block in a particular mode given the coding modes for the past macro-blocks. $R'_Q(n; \boldsymbol{M})$ depends on the previous mode selections, and a total of $(N-1)(N+2)/2$ different rates have to be calculated to fill an N stage trellis.

We can reduce the number of calculations needed for minimization by dividing the original N stage problem into several shorter stage problems. Using $R_Q(n)$ and $R'_Q(n; \boldsymbol{M})$, we can determine *a priori* macro-blocks that will always be coded in intra mode regardless of mode selections for the macro-blocks at the same spatial location in other frames. When the $\max B_Q(n; \boldsymbol{M}) \leq 0$, it is always better to code the macro-block at time n in intra mode

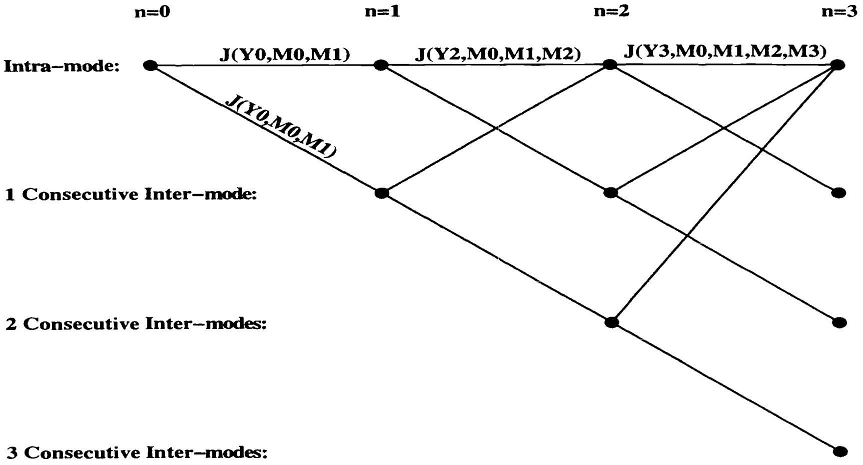

n=0
n=1
n=2
n=3
Intra–mode:
J(Y0,M0,M1)
J(Y2,M0,M1,M2)
J(Y3,M0,M1,M2,M3)
J(Y0,M0,M1)
1 Consecutive Inter–mode:
2 Consecutive Inter–modes:
3 Consecutive Inter–modes:

Figure 2. Trellis

since there is no trade-off between reducing the distortion and rate. In addition, if the cost function for coding a macro-block in intra mode at a stage is less than the minimum possible value for coding it in inter mode, then you want to code it in intra mode. Intra mode coded macro-blocks decouple the problem, and mode selection can be done independently for the stages between these macro-blocks.

5. CODING RESULTS

To evaluate the different distortion metrics chosen, we code 100 QCIF resolution frames of the Carphone sequence at a constant quality using an $H.261$ encoder and simulate packet loss in an erasure channel. Each macro-block is lost independently with probability $p = .10$. To code at constant quality, we code all intra-coded macro-blocks with fixed quantization step-size $Q = 3$. To ensure the quality is consistent when a macro-block is coded in inter mode, it is first coded in intra mode with $Q = 3$ so that the mean-squared-error between the original and decoded macro-block can be measured. The parameter Q is then adjusted so that the mean-square-error of the decoded inter-coded block is as close as possible to the intra-coded value. We use the algorithm described above to compute the optimal coding mode selection for a chosen distortion metric to meet a rate constraint. This is equivalent to minimizing the Lagrangian cost function $J(\boldsymbol{Y}, \hat{\boldsymbol{Y}}, \boldsymbol{M})$ for a particular value of λ which can be found by using a bisection algorithm. We use temporal extrapolation to conceal lost macro-blocks, and a lost macro-block at spatial location i in frame n is replaced by the most recent correctly received macro-block at spatial location i.

To justify the computational effort used to select modes in the rate-distortion optimal methods, we compare our results to mode selection using the MPEG TM5 mode selection with forced intra updates. The algorithm for MPEG TM5 is summarized in figure 3. When the variance of the prediction error is less than 64 or less than the variance of the original macro-block, the macro-block is coded in inter mode, otherwise it is coded in intra mode. This simple method for mode selection is a cursory attempt to choose the modes to code the sequence at the lowest possible rate R_{min} for the given frame quality. Since the rate constraint R_c is set much higher then R_{min}, a threshold T is chosen, and the macro-block at each location i in the image frame is forced to be coded in intra mode at least once every T times it is transmitted for robustness to loss. The smallest integer T is chosen such that the coded sequence bit-rate is less than the rate constraint. We code the sequence using MPEG TM5 with $T = 10$ for mode selection at $950Kbps^*$. Figure 4 shows the mode selection results for each of the 99 macro-block locations in the frame using this algorithm. A small rectangular mark is printed when the macro-block is coded in intra mode otherwise the space is left blank. All macro-blocks in the first frame are coded in intra mode. This plot shows a quasi-periodic structure resulting from the forced intra mode updates.

*This high bit-rate is a result of the fine quantization step-size used as well as the zero-vector displacement restriction.

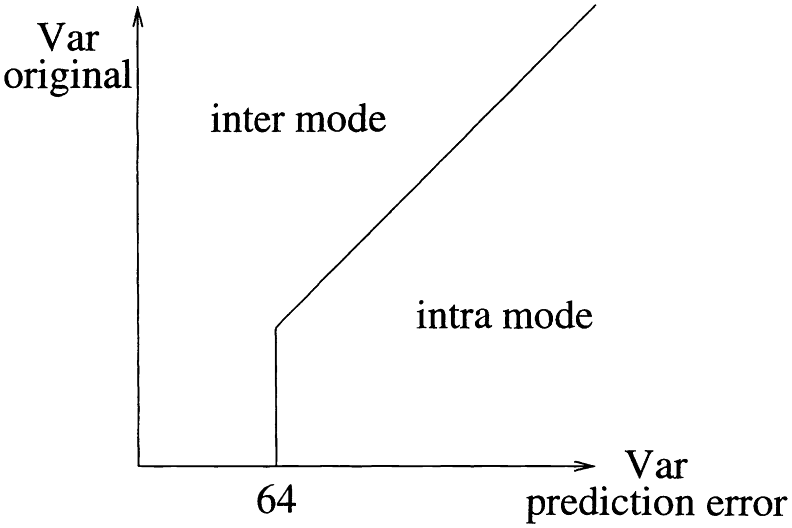


Figure 3. MPEG TM5 mode selection

Figure 4 also shows the optimal mode selection when the sequence is coded to minimize the distortion metric $D_I(M_n^i)$ at the same bit-rate. This results in many more intra coded macro-blocks in the sequence and improved perceptual video quality when viewing the decoded sequence. One drawback with this distortion measure is the localization in space of intra coded macro-blocks. Macro-blocks at certain spatial locations are always coded in intra, while there are other regions whose macro-blocks are rarely coded in intra mode. Perceptually this results in regions in the video sequence with no noticeable distortion while there are other locations with very distracting artifacts.

Figure 5 shows the optimal mode selection to minimize $E[D_H(Y_n^i, \hat{Y_n^i})]$ which is equal to the probability of error. Since the probability of error is an increasing function with the distance between intra coded macro-blocks at each spatial location i, mode selection to minimize the average probability of error results in a more even distribution of intra coded macro-blocks. This results in overall improved perceptual quality over the previous methods for the examples we tried. The decoded video sequence after concealment typically has fewer annoying artifacts. However, the method that maximizes the number of intra coded macro-blocks sometimes outperforms this method in local regions in the video sequence frame, particularly in the macro-block locations that are always coded in intra mode.

Even though MPE mode selection gives improved results over previous methods, it does not use an accurate measure of error in its distortion metric. There are some macro-blocks which may be effectively concealed when lost. An extreme example is in the case of stationary regions in the sequence when temporal extrapolation is used for concealment. Thus in the MPE mode selection method, too many intra coded macro-blocks may be allocated to regions of the video sequence that are simple to conceal. The optimal mode selection to minimize $E[D_{MSE}(X_n^i, \hat{Y_n^i})]$ uses the known concealment method to get an accurate measure of the expected error. We would expect this result to perceptually outperform the previous method. The question is how much. The mode selection for the $MMSE$ mode selection is shown in figure 5. Qualitatively the mode selection is similar to MPE mode selection in that it evens out the distribution of intra coded macro-blocks, but it actually codes a smaller number of macro-blocks in intra mode. In the examples we tried, $MMSE$ mode selection offers a slight perceptual improvement over MPE mode selection. It typically avoids even more distracting artifacts than MPE mode selection at the same rate. Even though it results in fewer intra coded macro-blocks, $MMSE$ chooses to code macro-blocks in intra mode that will not be concealed well with high probability.

6. CONCLUSIONS

In this paper, we investigated several methods for macro-block coding mode selection in the presence of random loss. We defined a simplified problem and presented an efficient algorithm to a solution with different distortion metrics.

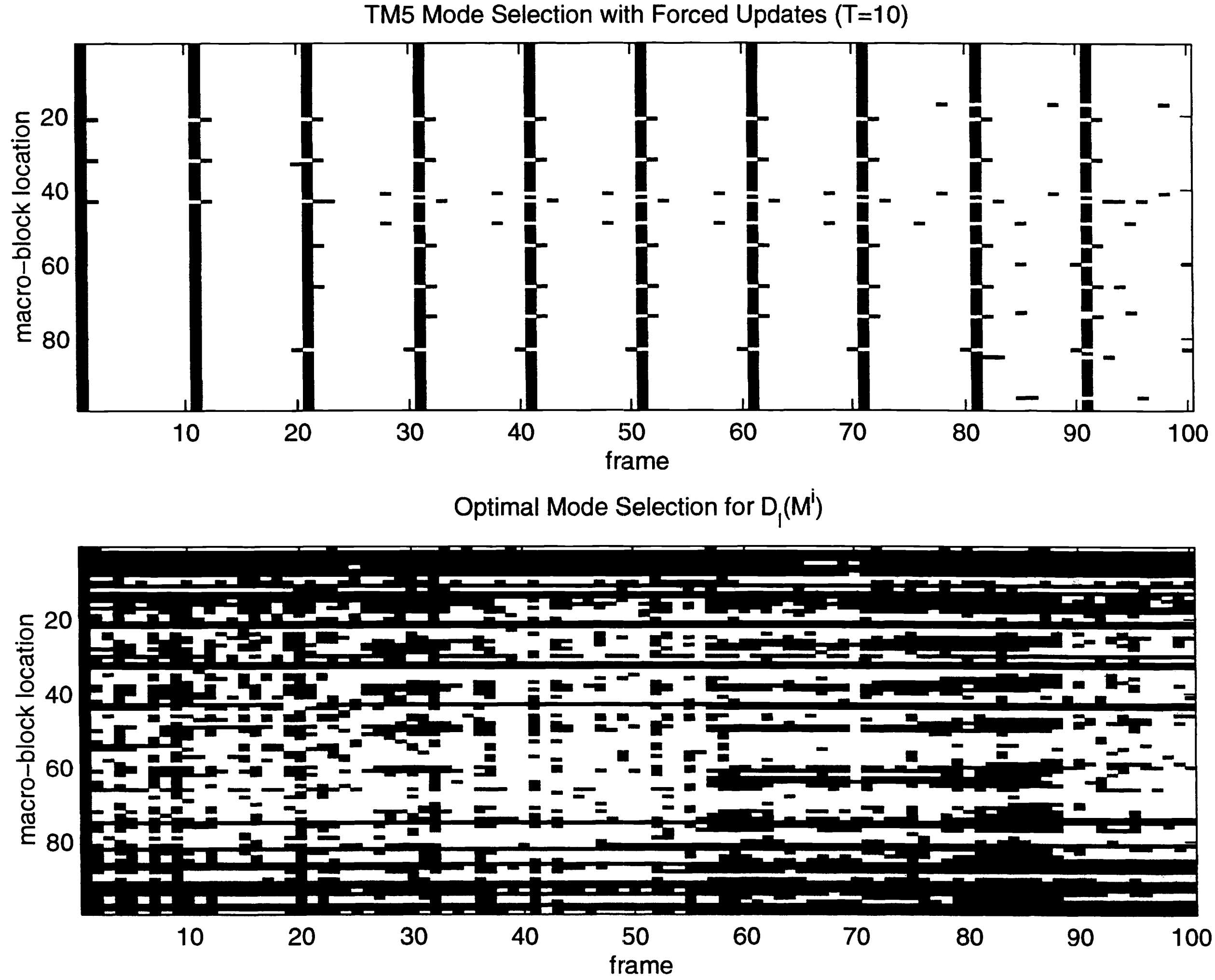
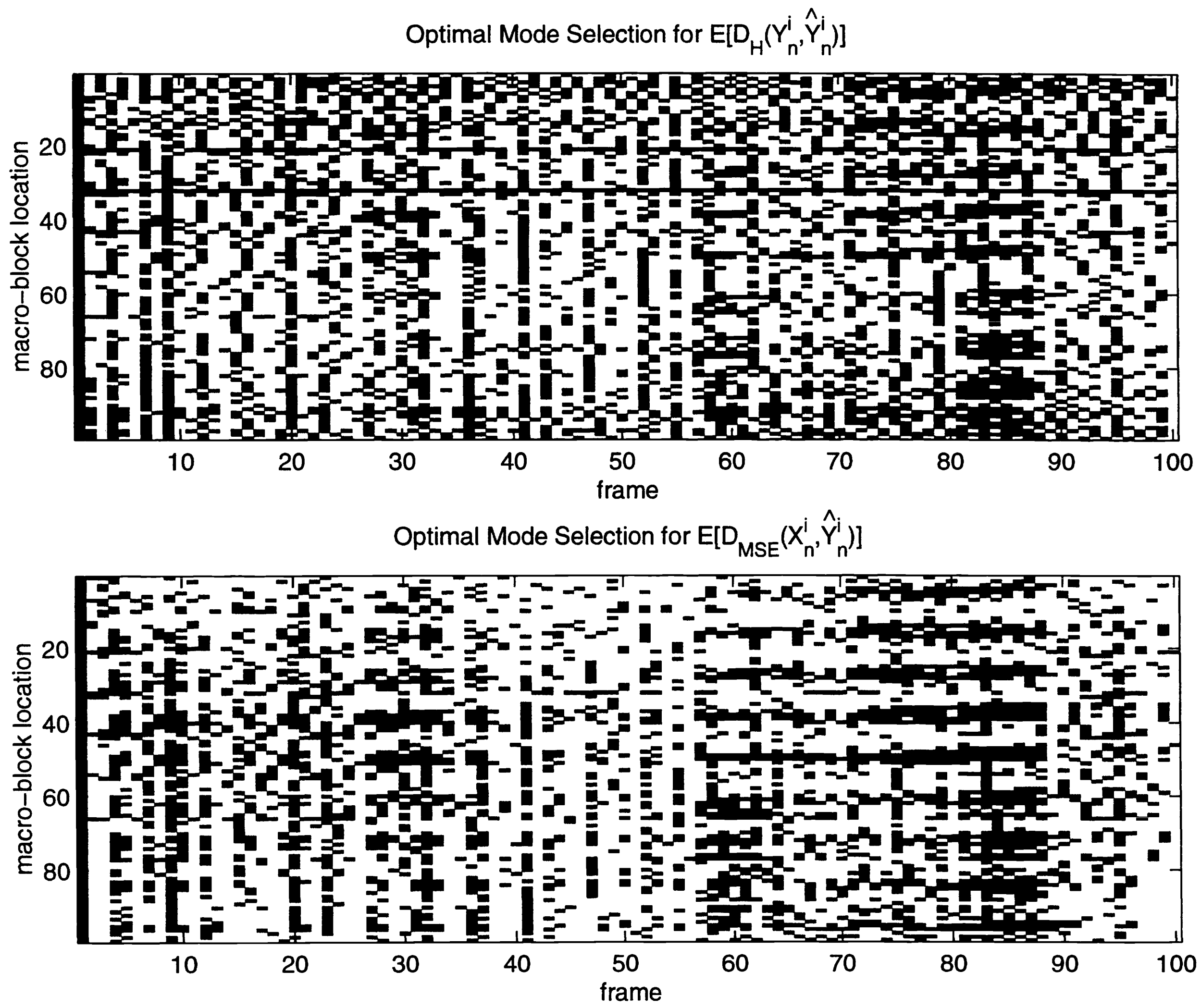


Figure 4. mode selection

We evaluated the results and discussed some of the advantages and drawbacks using certain reasonable distortion metrics. Further research must be done to determine a method for mode selection which is best for many possible concealment methods. We need to determine how much can be done by optimal mode selection at the encoder and concealment at the decoder to improve overall video quality when macro-block loss occurs. In addition, further analysis must be done to determine optimal mode selection when some of the simplifying restrictions are relaxed, namely the allowance of non-zero motion vectors.

ACKNOWLEDGEMENTS

This research was supported in part by the Cooperative Research Fellowship Program from Lucent Technologies, by summer employment at Bell Laboratories in Murray Hill, New Jersey, and the Advanced Television Research Program at M.I.T. The results and viewpoints presented in this paper do no represent those of the research sponsors.


Optimal Mode Selection for $E[D_H(Y^i_n,\hat{Y}^i_n)]$

Optimal Mode Selection for $E[D_{MSE}(X^i_n,\hat{Y}^i_n)]$

Figure 5. mode selection

REFERENCES

1. T. Wiegand, M. Lightstone, *et al.*, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging h.263 standard," *IEEE Transactions on Circuits and Systems for Video Technology* **6**, pp. 182–190, April 1996.
2. W. Kwok, H. Sun, and J. Ju, "Obtaining an upper bound in mpeg coding performance from jointly optimizing coding mode decisions and rate control," in *SPIE*, vol. 2501, pp. 2–10, 1995.
3. K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and mpeg video coders," *IEEE Transactions on Image Processing* **3**, pp. 533–545, September 1994.
4. J. Lee and B. W. Dickinson, "Joint optimizations of frame type selection and bit allocation for mpeg video encoders," in *Proceedings ICIP*, pp. 962–966, 1994.
5. G. Davis and J. Danskin, "Joint source and channel coding for image transmission over lossy packet networks," in *SPIE*, vol. 2847, pp. 376–386, 1996.