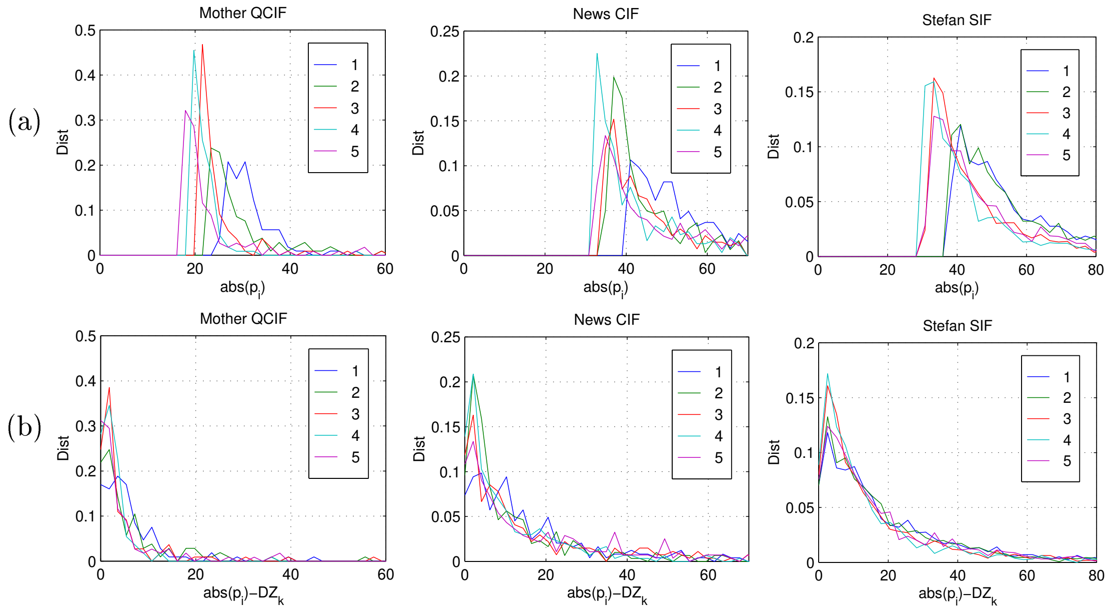
Figure 4.6: Absolute modulus distributions for single matching pursuit residual frames coded sequentially without quantization. (a) Sequences coded without DZ-subtraction. (b) Same sequences with framewise DZ-subtraction. Note the number of atoms per frame match the standard MPEG-4 sequence runs of Mother at 24 kbit/s, News at 112 kbit/s, and Stefan at 1 Mbit/s.

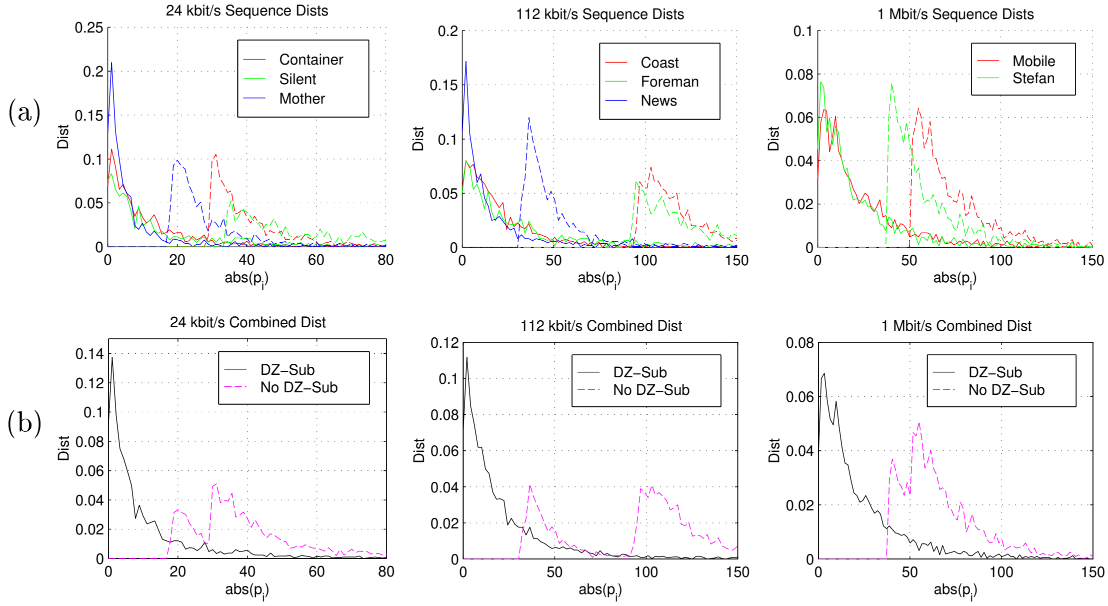


such as Stefan, where a large number of atoms per frame allows more reliable computation of the distribution curves.

Because the distribution shifts between frames, we propose a simple transformation to the original modulus source. Define DZ_k as the minimum absolute modulus value found in encoding frame k. A "DZ-Subtraction" operation is then performed on the modulus source, that is DZ_k is subtracted from all absolute modulus values within the frame. The distributions of the resulting DZ-Subtracted source are more consistent from frame to frame, as illustrated in Figure 4.6b.

The effect is even more drastic when modulus distributions are combined over multiple sequences or across multiple bit rates. Figure 4.7 shows the combination of multiple sequences at the same rate. Each plot is generated by encoding all of the MPEG-4 test

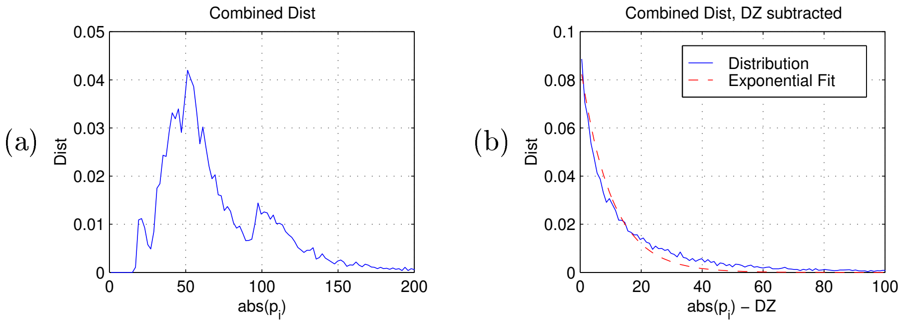
Figure 4.7: Combined modulus distributions for each of three MPEG-4 test rate categories. (a) Separate distribution curves for each sequence in the given bit rate category. Solid lines show runs with DZ-subtraction. Dotted lines show runs without DZ-subtraction. (b) Combined distribution curve for all sequences in the given rate category.



a simple combination of the absolute moduli, while Figure 4.8b shows the DZ-subtracted result. Without the DZ-subtraction, the combined distribution is somewhat random and matches neither the individual frame distributions from Figure 4.6a nor the combined distributions for the various rate categories shown in Figure 4.7. With DZ-subtraction, the final combined distribution is again a sharp exponential, which well matches the corresponding DZ-subtracted distributions in Figures 4.6b and 4.7. To verify the exponential form of the distribution, a least squares best fit exponential curve is also shown in Figure 4.8b.

We compare the information content of the two distributions of Figure 4.8 by

Figure 4.8: Modulus distributions combined across all MPEG-4 sequences and rates. Enough coded frames were taken to provide approximately 3000 atoms from each of the five bit rate categories (10, 24, 48, 112, and 1024 Kbit/s). (a) Without dead-zone subtraction. (b) With frame-wise dead zone subtraction. Dotted line shows best exponential fit.



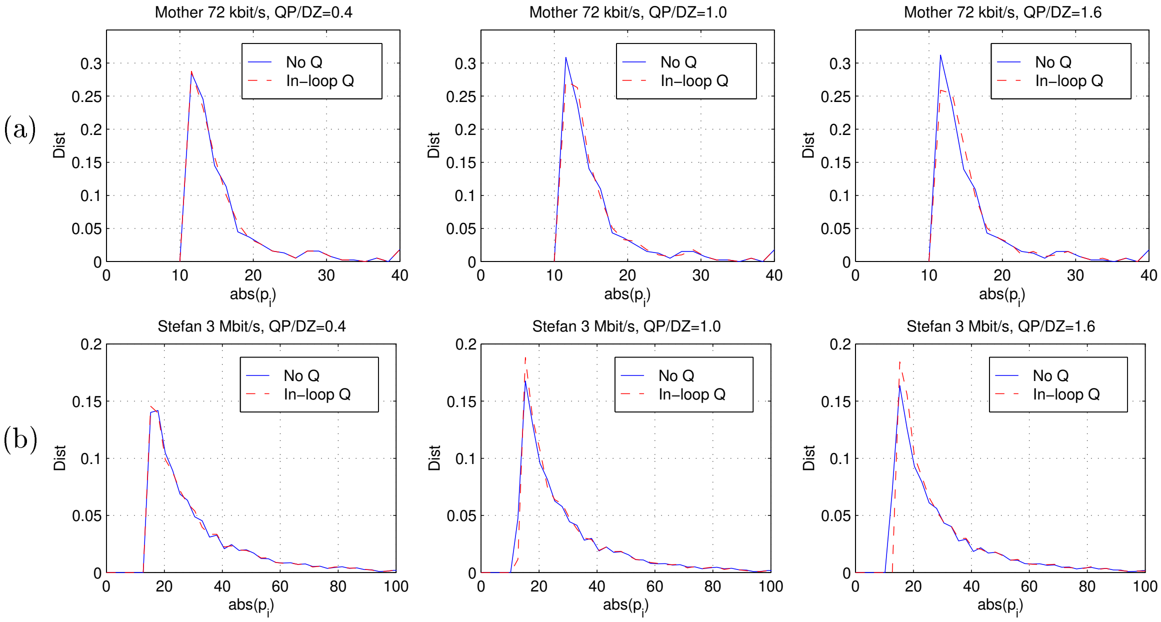
computing the Shannon differential entropy, $h(X)$, for each distribution.[2] Simple numerical integration [4] gives a value $h(X) = 6.914$ for the combined distribution in Figure 4.8a, and a value $h(X) = 5.682$ for the DZ-subtracted distribution in Figure 4.8b. Suppose the modulus source is to be quantized using the optimal entropy constrained scalar quantizer as defined in [23]. Dead-zone subtraction will then improve the rate-distortion performance $r(d)$ of the quantizer by the difference of 1.279 bits per symbol, a savings of about 18 percent. The dead zone parameter DZ_k must be transmitted once per frame, but this is a negligible cost even at low bit rates.

The above experiments were done in the absence of quantization. We now show that the addition of in-loop quantization does not significantly change the observed distribution. Figure 4.9 shows single-frame modulus distributions for two sample frames coded with and without in-loop quantization. The top row of plots, (a), shows a frame from

[2]Shannon differential entropy reflects the information content of a continuous source, similar to entropy for discrete valued sources [31]. It is defined in terms of the probability density function $p_x(x)$ as:

$$h(X) = \int_{-\infty}^{\infty} p_x(x) \log_2 p_x(x) dx$$

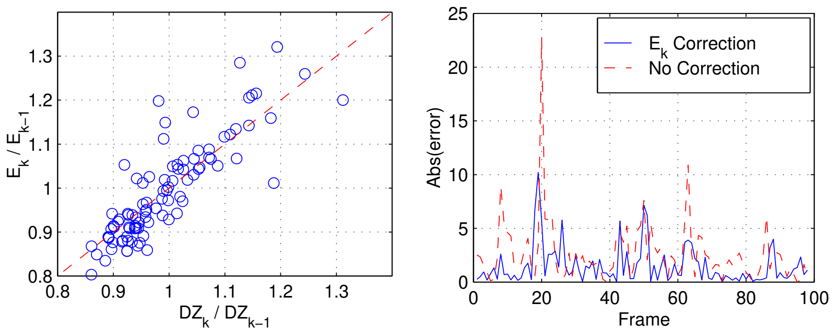
Figure 4.9: The effect of in-loop quantization on the absolute modulus distribution for single coded frames. (a) Mother QCIF at 72 kbit/s. (b) Stefan SIF at 3 Mbit/s. Three different quantizers with progressively larger stepsize to deadzone ratios are used for each frame.



Mother QCIF at 24 kbit/s, while the bottom row, (b), shows a frame of Stefan SIF at 3 Mbit/s. For each plot, the solid line depicts matching pursuit without quantization, and the dotted line shows in-loop quantization. Clearly the amount of error introduced by the quantizer will determine the amount of change seen in the resulting distribution. For this reason, three quantizers are used with varying stepsizes. The quantizer is uniform with an expanded dead zone, which is equivalent to DZ-subtraction followed by UTQ. The dead zone size DZ is set to be the minimum absolute modulus value seen in the non-quantized encoding. The step size QP is then set so that the ratio QP/DZ takes on values of 0.4, 1.0 and 1.6. Note that this stepsize to dead-zone ratio provides a convenient description of the quantizer, and so will be used in the remainder of this chapter.

Scanning the plots of Figure 4.9 from left to right, it is clear that for ratio values of 1.0 and below, in-loop quantization produces no noticeable change to the absolute modulus

Figure 4.17: (a) Justification of the energy correction for DZ prediction. Horizontal axis shows ratio of current to previous frame DZ values as found by the full 2-pass adaptive algorithm. Vertical axis shows ratio of current to previous frame residual energy before atom coding. (b) Absolute error in predicting DZ_k for each frame of Mother at 24 kbit/s. Dotted line shows error for simple prediction. Solid line shows reduced error when energy corrected prediction is used.



(a) (b)

the horizontal axis and the corresponding ratio of initial frame energies on the vertical. A correlation between the two ratios can be seen, and the dotted line $y = x$ suggests the following energy-corrected prediction for the current frame dead zone:

$$DZ_k = DZ_{k-1} \cdot \left(\frac{E_k}{E_{k-1}} \right) \tag{4.17}$$

The effectiveness of these prediction methods is evaluated in Figure 4.17b. Data is from the 2-pass encoding of Mother at 24 kbit/s, and the value of DZ_k is predicted with energy correction according to Equation (4.17) and without correction as in Equation (4.16). The energy corrected prediction is seen to be better on average, however it is not guaranteed to be better for individual frames. If we compute mean square prediction error in DZ_k, then energy correction reduces the error from 15.03 to 5.45.

The improved DZ prediction method is thus used to define a one-pass adaptive quantization algorithm. This algorithm is similar to the two-pass algorithm of Figure 4.14,

Figure 4.21: Luminance PSNR vs. frame for four sample sequences at standard MPEG4 test rates.

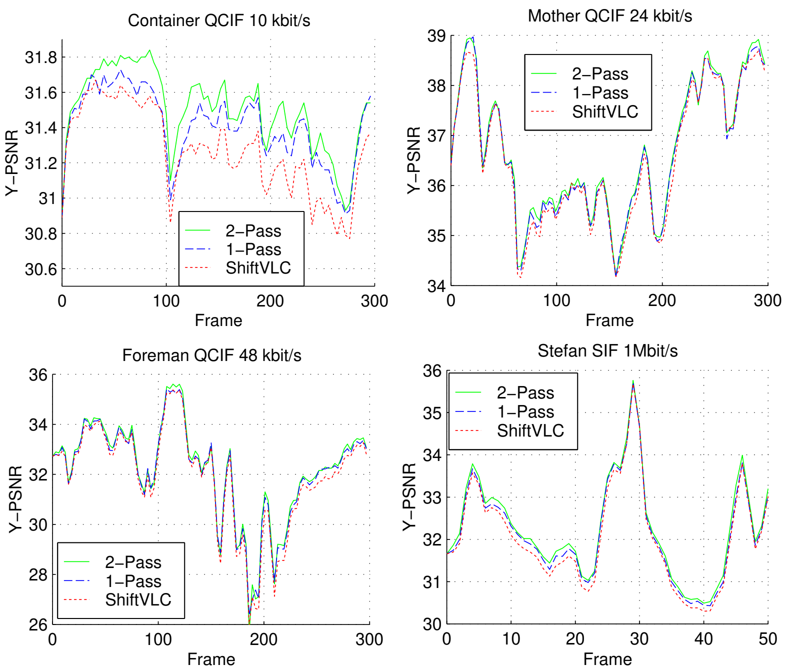
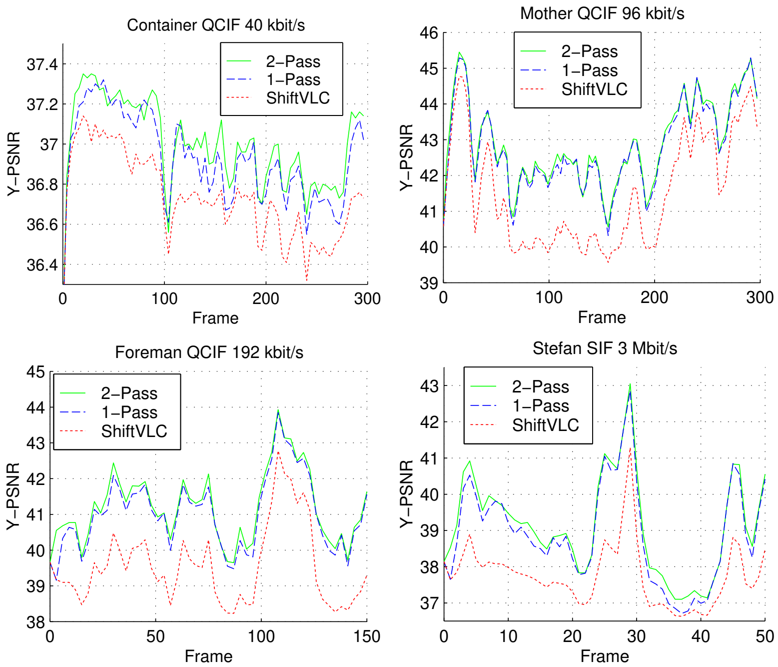


Figure 4.22. It is clear from these plots that the gain for the adaptive schemes is much higher at these rates, and that again the 1-Pass and 2-Pass schemes are nearly equivalent.

4.5 Conclusions

Frame adaptive quantizer designs have been shown to outperform the fixed quantizer used in earlier matching pursuit video coding work. Our design uses a combination of adaptive dead zone subtraction and a UTQ with adaptive stepsize. A 2-pass algorithm was shown to choose a nearly optimal balance between the quantizer precision and the number of coded atoms, and hence comes close to minimizing overall coding distortion D for a given target rate R. A 1-pass algorithm was shown to have nearly the same performance, but with complexity similar to that of the original fixed quantizer design. Small improvements

Figure 4.22: Luminance PSNR vs. frame for four sample sequences at higher bit rates.



of .1 to .4 dB are shown for the MPEG sequences at standard rates, and larger gains of up to 1.7 dB can be shown for the same sequences at higher bit rates.

Figure 5.10: Illustration of a simple distortion model with p^2 distortion contribution ignored. (a) Real and predicted change in p^2 contribution $\Delta\rho_i$. (b) Real and predicted change in d^2 contribution $\Delta\delta_i$. (c) Pass 1 distortion D_i compared to real and predicted values of Pass 2 distortion D'_i. (d) Absolute prediction error. (e) Percent differential relative prediction error, $e_{|D'-D|}$. (e) Percent Relative prediction error, $e_{|D'|}$.

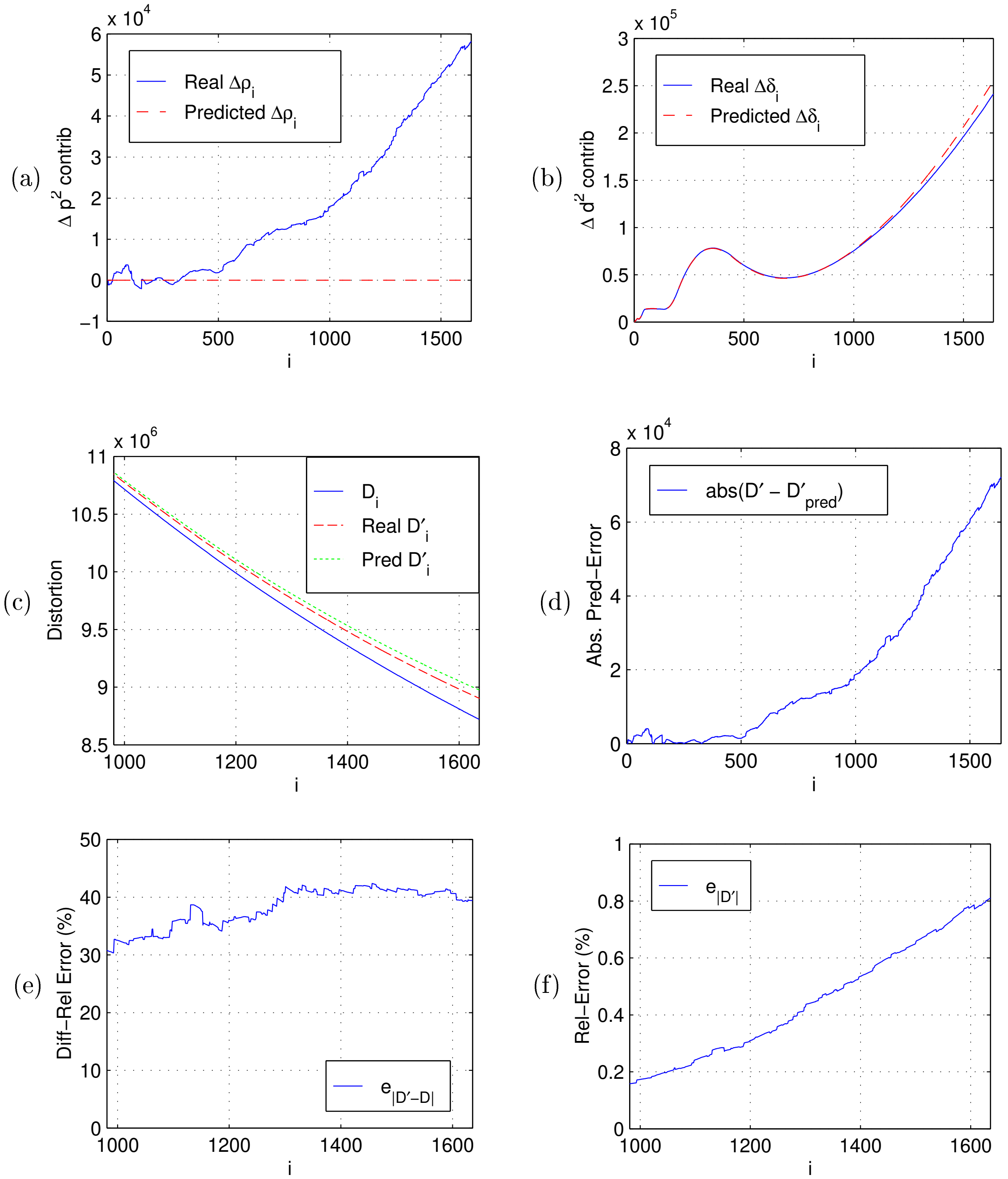


Figure 5.12: Illustration of the full distortion model with both p^2 and d^2 distortion contributions. (a) Real and predicted change in p^2 contribution $\Delta\rho_i$. (b) Real and predicted change in d^2 contribution $\Delta\delta_i$. (c) Pass 1 distortion D_i compared to real and predicted values of Pass 2 distortion D'_i. (d) Absolute prediction error. (e) Percent differential relative prediction error, $e_{|D'-D|}$. (e) Percent relative prediction error, $e_{|D'|}$.

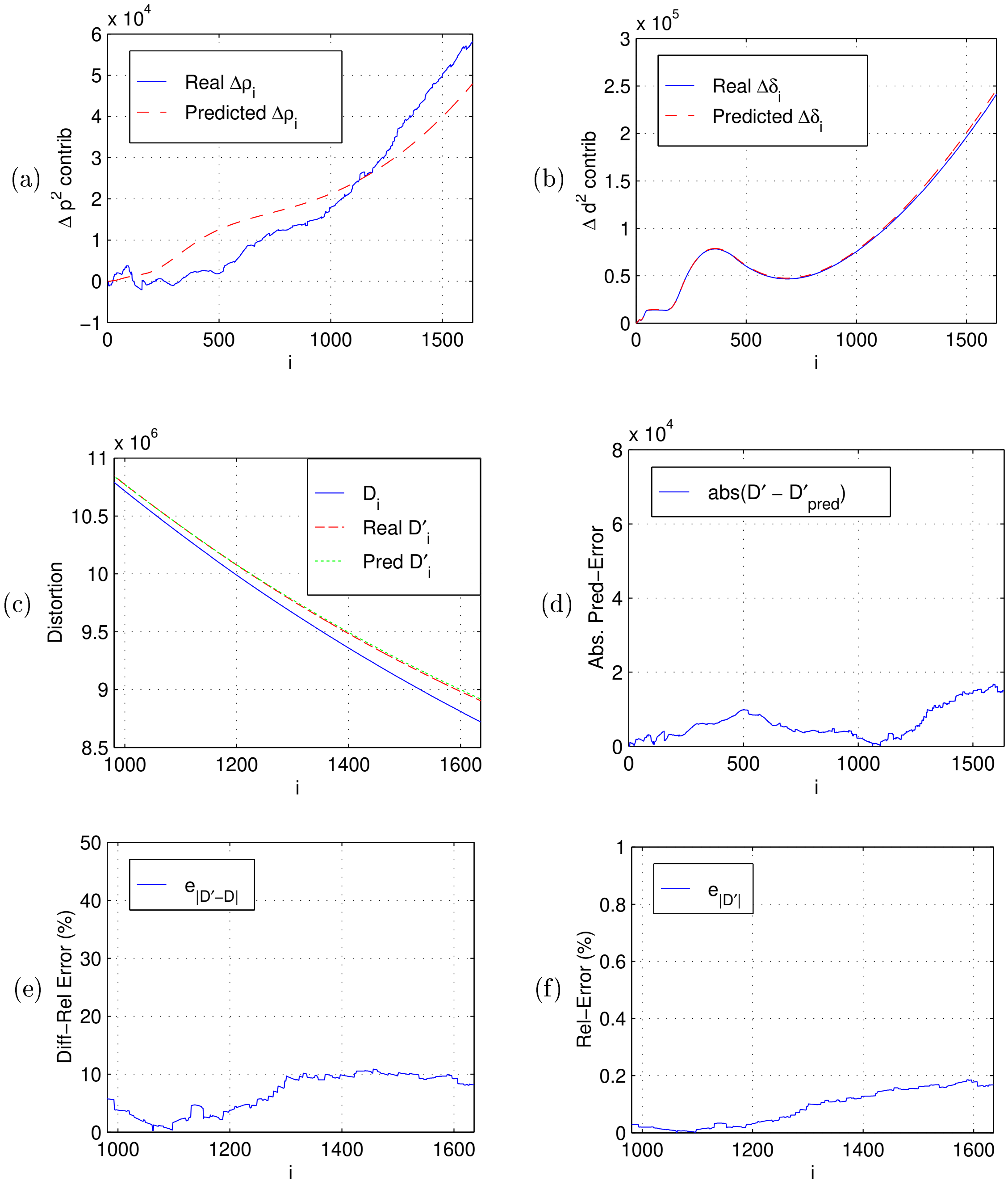
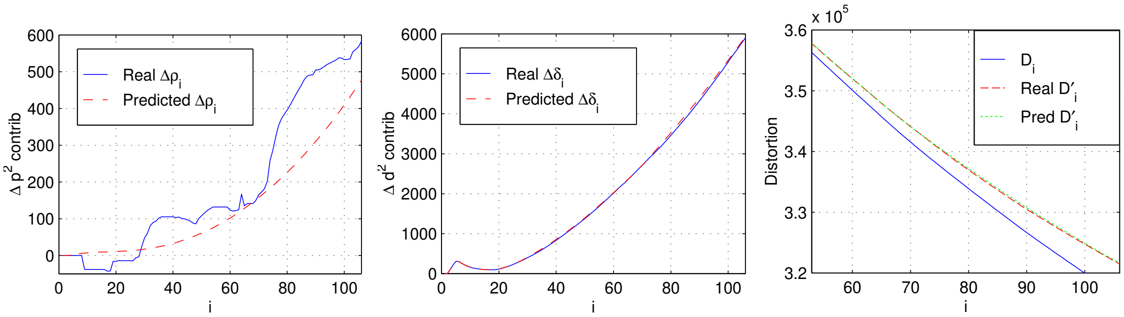
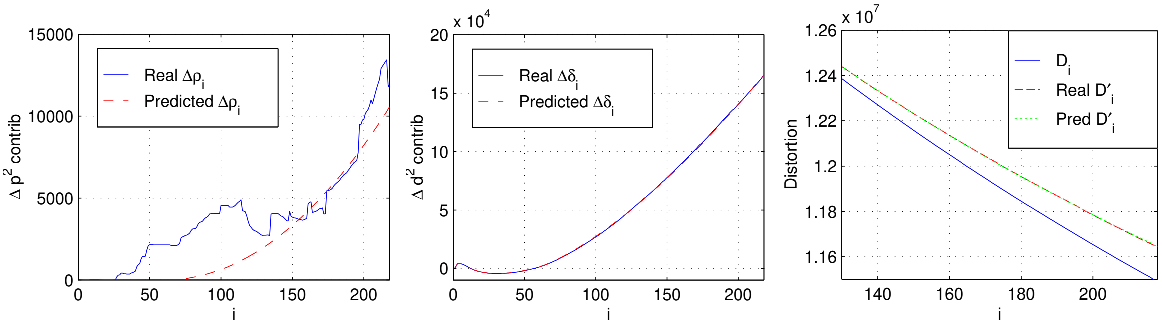


Figure 5.14: Distortion prediction results for sample frames from three sequences with quantizer ratio change $0.6 \rightarrow 1.0$.

Mother at 24 kbit/s:



Coast Guard at 112 kbit/s:



Stefan at 1 Mbit/s:

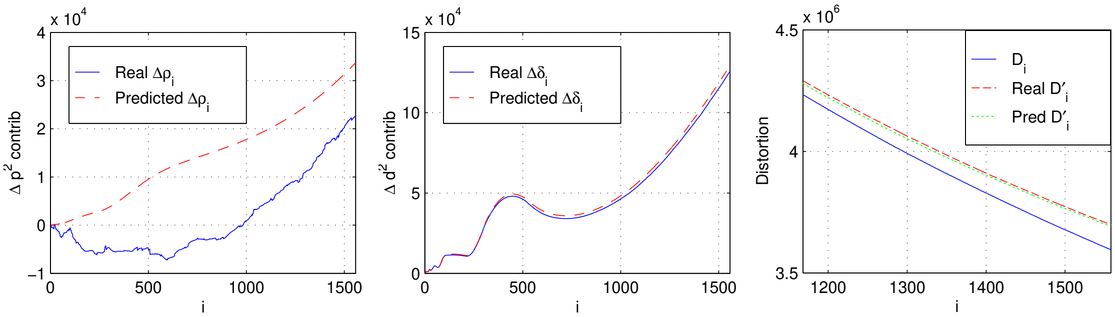
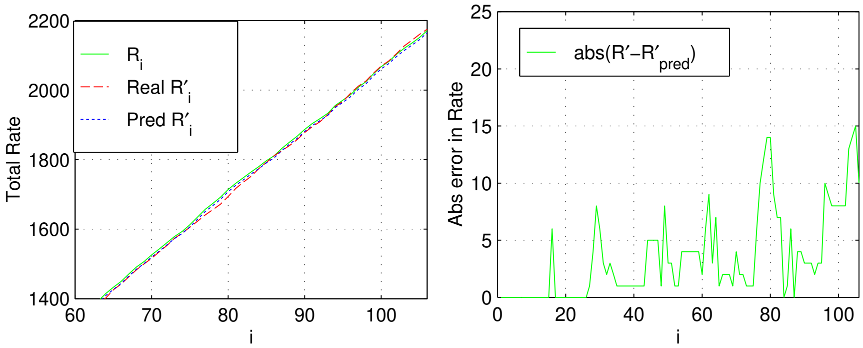
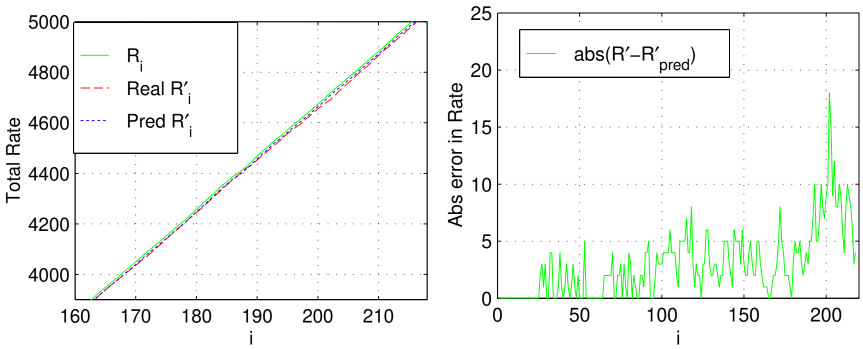


Figure 5.16: Rate prediction results for sample frames from three sequences with quantizer ratio change 0.6 → 1.0. Rates shown are for encoding the first i atoms, with motion and header bit rates ignored.

Mother at 24 kbit/s:



Coast Guard at 112 kbit/s:



Stefan at 1 Mbit/s:

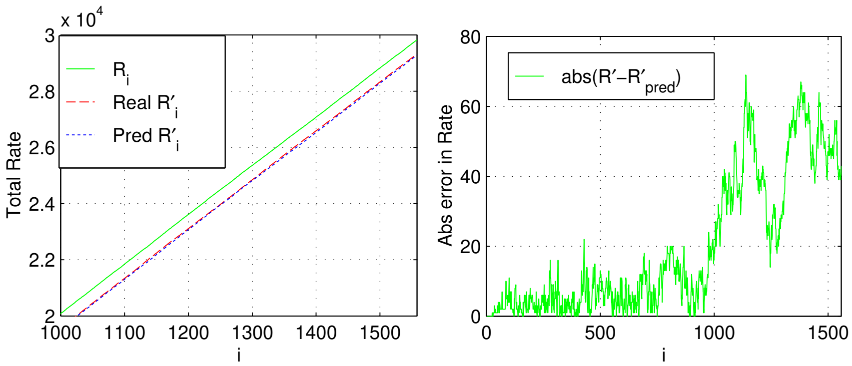


Figure 5.25: Improvement in accuracy due to the hybrid prediction model. Original *Red*3 prediction is compared to the hybrid model with anchor point set at $a1$ and $a2$. Full Pre-Code prediction is also shown. Results are shown for 9 frames encoded using the 2-pass quantizer encoder.

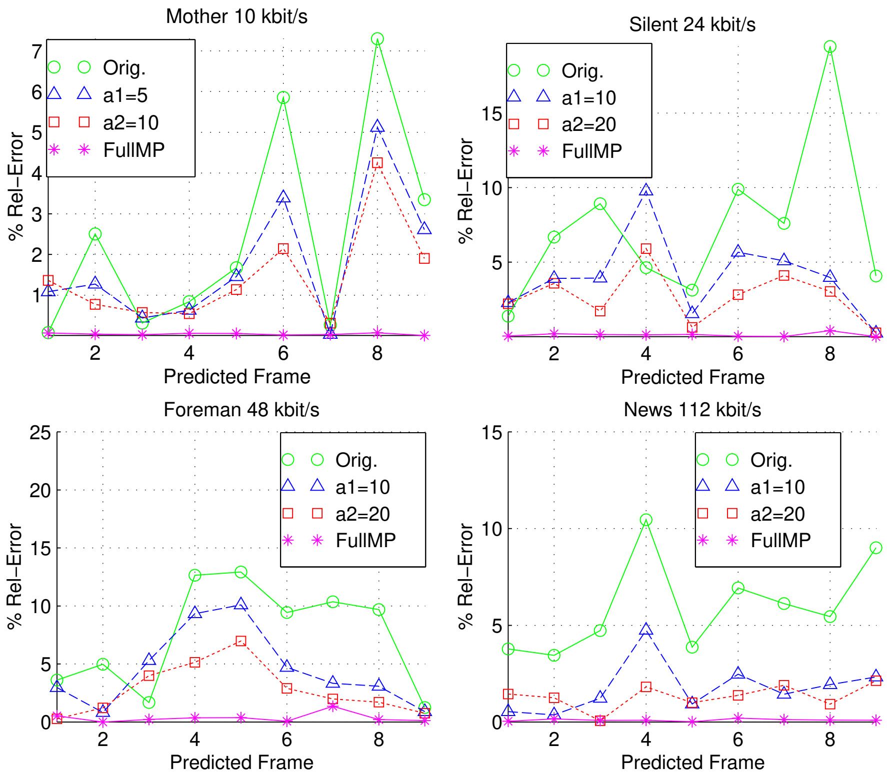


Figure 5.26: Complexity vs. Accuracy tradeoffs, averaged across all sequences in a bit rate category. Full Pre-Code method is (*). Other symbols show Reduced Pre-Code models. Connecting lines show the Hybrid Model tradeoffs achieved as additional full dictionary atoms are added.

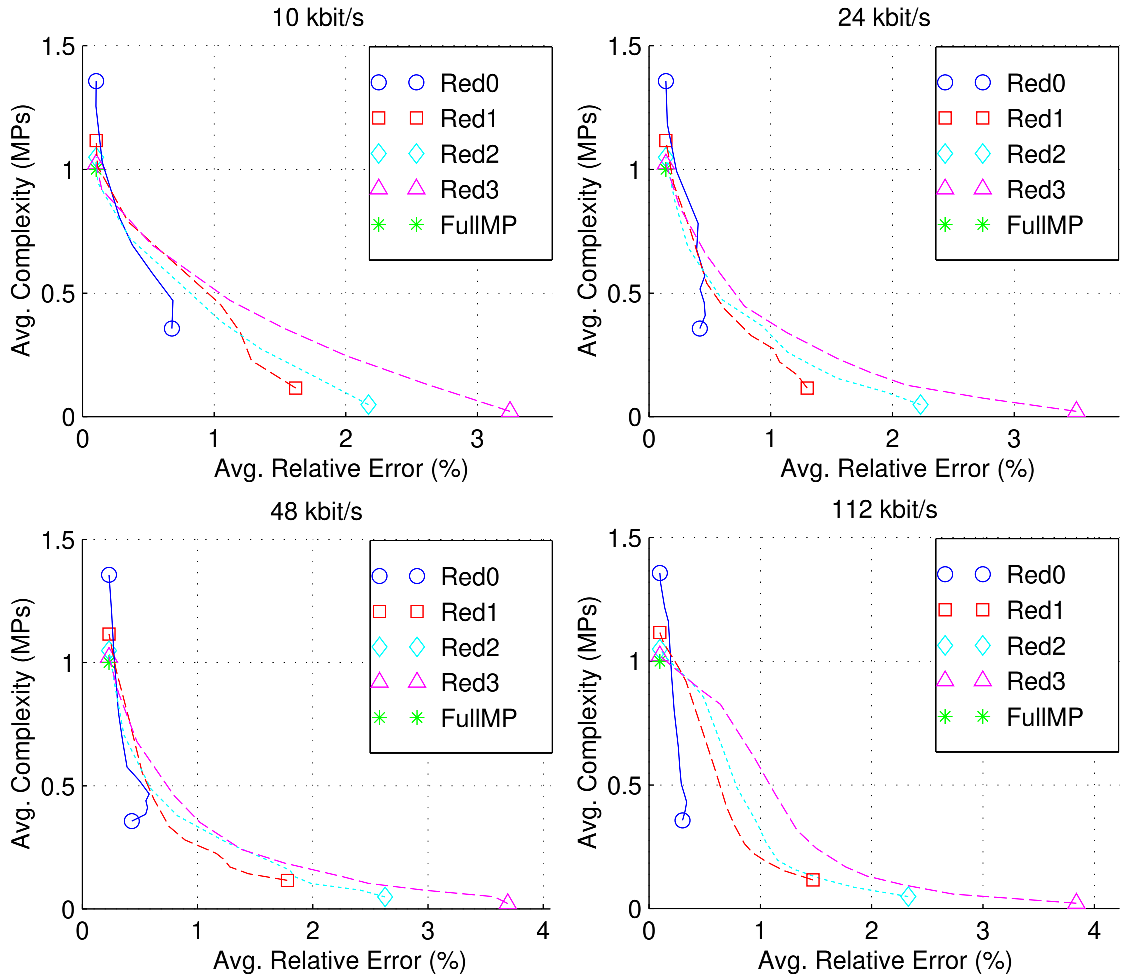
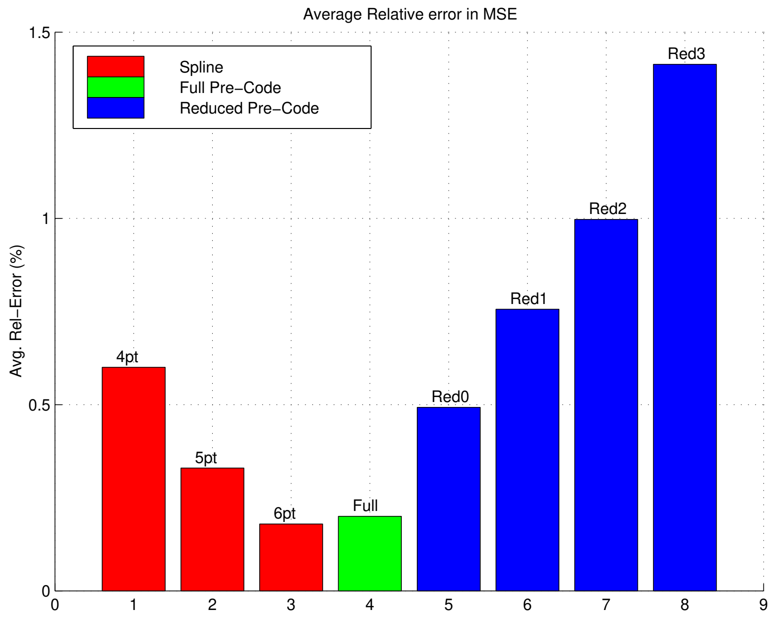


Figure 5.28: Average Relative-Error Comparison.



Complexity of Spline Model

At this point, it is logical to ask the following question: "If the accuracy of the cubic spline is equal to or greater than that of the matching pursuit models developed earlier, then why not replace these models with the spline?" The answer stems from complexity considerations, which must be considered separately for each matching pursuit quantizer design. To simplify the discussion, we consider predicted (R, D) for a single frame over some known set of σ choices.

Suppose either the Fixed or 1-Pass quantizer designs are to be used. For these cases, the quantizer is known at the start of coding, and so a single encoding pass to N_{max} atoms is required in order to define the exact (R, D) control points to be used with the spline model. However, this same single encoding pass will trace out all the exact (R, D)