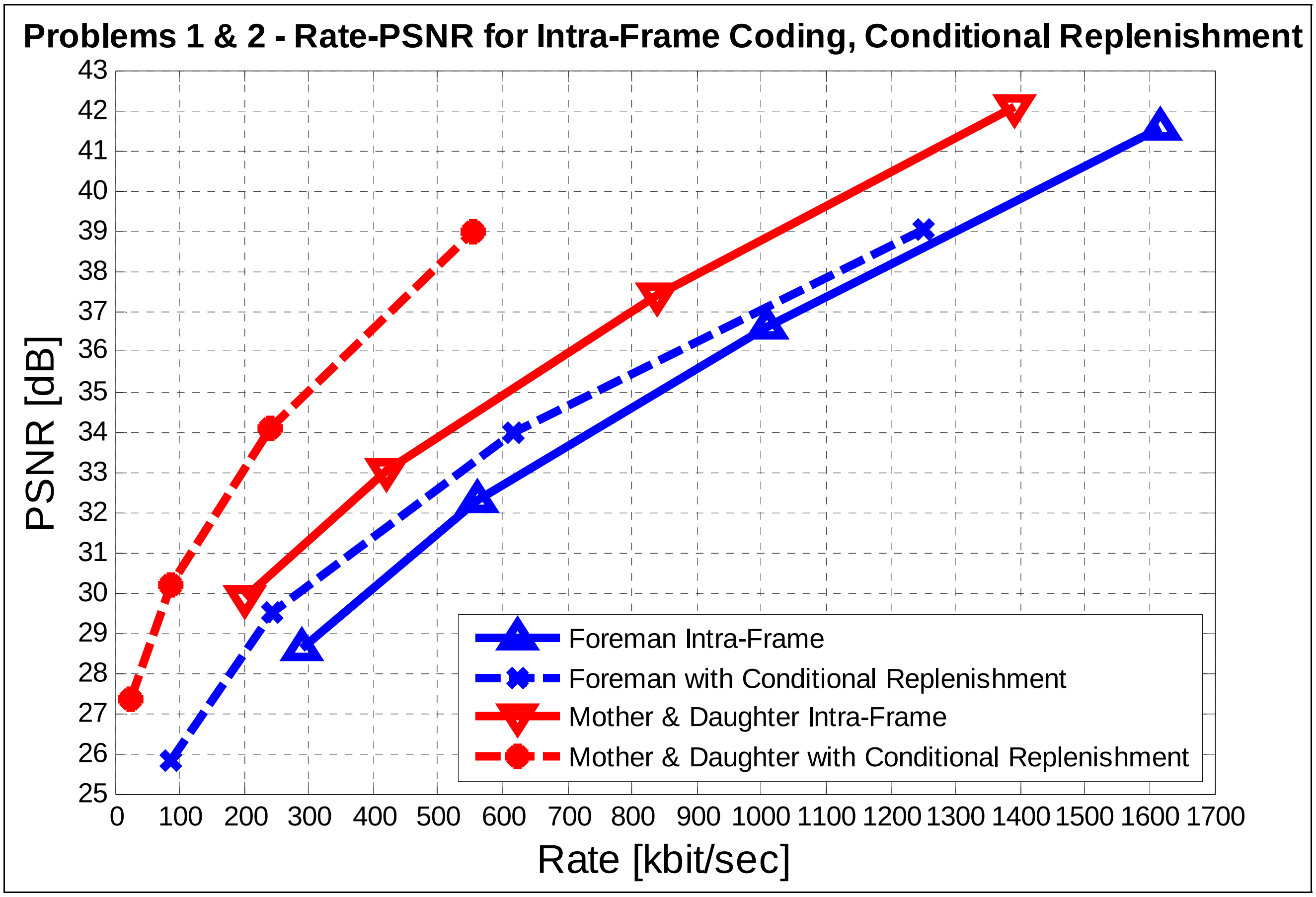
Problem Set VII – Video Sequence Coding

Problems 1 & 2 – Intra-Frame Coding with Conditional Replenishment

For both the "Foreman" and "Mother & Daughter" video sequences, we encode the luminance signal of the first fifty frames (divided into 16 × 16 blocks for the Discrete Cosine Transform) with the same variable-length codebook across all frames. To trace the Rate-PSNR curve for intra-frame coding, we implement a variety of quantization step sizes: 8, 16, 32, and 64 length quantization intervals. If we assume that the frame rate of the sequence is 30 frames per second and measure the bit rate in kilobits per second, we can plot approximate rate versus the average PSNR of the fifty frames, tapping Parseval's Theorem to determine the block distortion. After stacking our fifty consecutive frames in an image stack, we transform each block independently of the others with the DCT and quantize the coefficients, as in JPEG compression. Finally, we invert the quantized transform coefficients and measure the resulting difference with mean-square error criterion.

Furthermore, we can apply block-based conditional replenishment into our coder by calculating the Lagrangian cost function $J_n = D_n + \lambda R_n$ for each 16 × 16 block to determine whether or not we use the *intra mode* or the *copy mode*, which simply repeats a block. We again sweep a quartet of quantization step sizes, now setting $\lambda = 0.2\ Q^2$, where Q represents our various quantization step sizes. Thus, we employ a different Lagrange multiplier for each quantization step size, and encode the video sequence in the mode yielding the minimal Lagrangian cost function. This additional computation might seem cumbersome but yields higher PSNR for lower rates because the copy mode significantly reduces the overall complexity of the video sequence. We compare its effect on the "Foreman" sequence to its gain in the "Mother & Daughter" sequence:



We notice that, while conditional replenishment with Lagrangian mode decision lowers the average bit rate for both sequences across all quantization schemes, the PSNR has increased. The "Foreman" sequence benefits less from the copy mode than the "Mother & Daughter" sequence, likely because the "Mother & Daughter" sequence contains more still frames and hence benefits more from copy coding. The intense motion throughout the foreman's face also foils copy coding, in many cases increasing computational cost through video coding. Overall, conditional replenishment with Lagrangian mode decision lowers the bit rates but boasts higher PSNR at the same bit rate as the purely intra-mode counterpart. Whether or not this technique offers substantial improvement depends intimately on the proportion of still frames to relative pixel motion, as tranquility increases the frequency and effectiveness of the newly augmented copy mode.

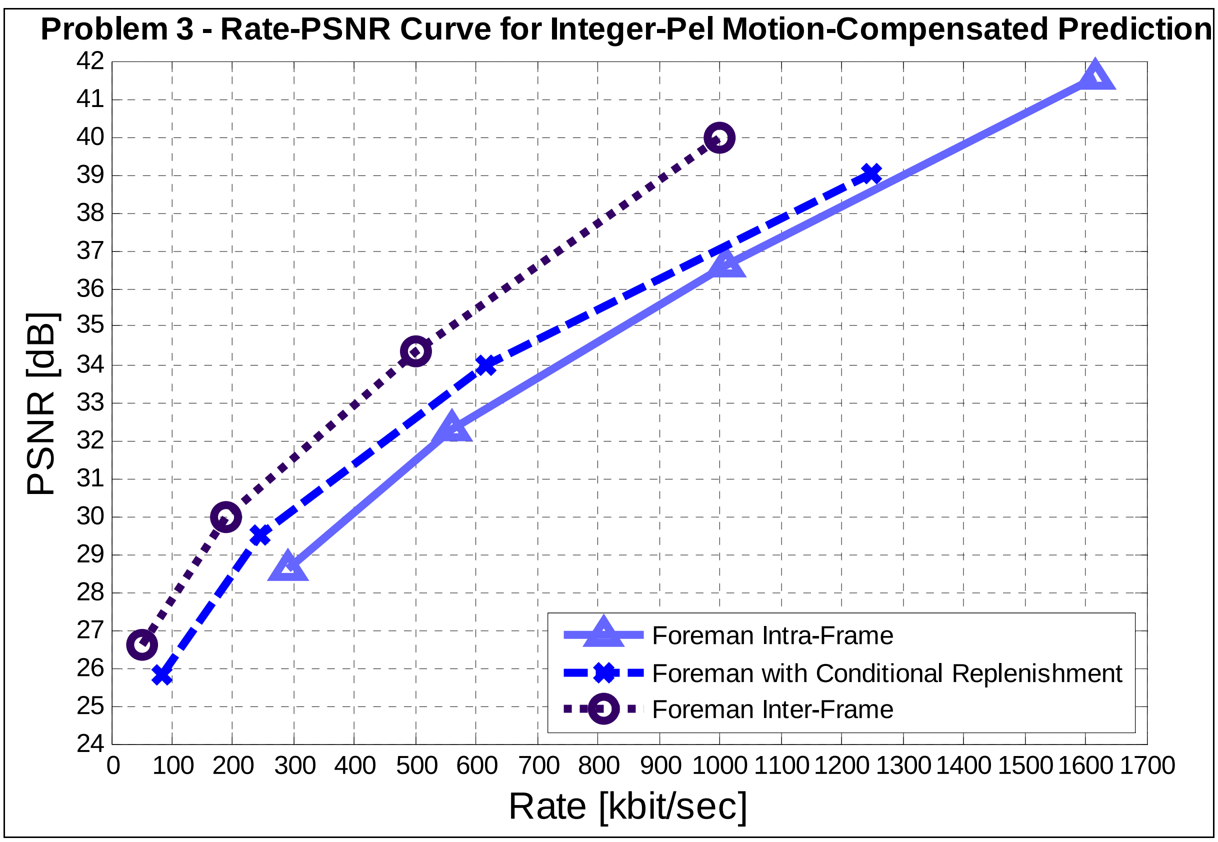
Problem 3 – Integer-Pel Motion Compensated Prediction

We now expand the intra and copy mode video sequence coder to accommodate block-based integer-pel motion-compensated prediction in an *inter mode*, which comprises two new components: motion estimation for displacements within 10 pixels, and residual coding on the difference between motion-compensated blocks. We assume a Laplacian distribution for the motion vectors and select the variance such that a displacement of 10 can be coded with approximately 16 bits:

```
variance = 0.95;
b = sqrt(variance/2);
```

```
pX = 1/(2*b) * exp(-abs(i-1)/b);
pY = 1/(2*b) * exp(-abs(j-1)/b);
```

where the values of *i* and *j* vary across the search range of 10-pixel displacements. Juxtaposing our rates and PSNR values with those obtained without the *inter mode*, we witness noticeable improvement:



On the "Foreman" video alone, the motion estimation boosts PSNR by approximately 2 dB for lower rates and up to 3 dB for higher rates! Thus, motion prediction captures non-trivial information in a sequence as animated as the Foreman's constantly changing face, as the improvement ascertains.