The source code and the report are mostly wrote by myself

Mini Project Report
Department of Electronic Engineering
The Chinese University of Hong Kong

Course: ELEG5502

Name: Su, Shang-Yen

SID: 1155032058

Objective

- To familiarize with the hybrid (predictive + transform) video coding framework;
- To investigate the functions of intra and inter prediction in hybrid video coding;

Introduction

The hybrid video coding includes prediction coding and transformation coding. There are two kinds of prediction coding, one is intra prediction and another one is inter prediction. In Section I, we will do the intra prediction coding for frame 0, then do the DCT transformation and quantization for the intra prediction results, reconstructed frame 0. In Section II, we use the previous reconstructed frame and the next frame to do the motion estimation. Here, we use 3-step search for the motion estimation. After the motion estimation, we use motion compensation to get the completed image. By DCT transformation and quantization of the image, the reconstructed frame is generated. The whole process of Section II is called intra prediction coding. In this project, we will do the intra prediction coding for frame 1 till frame 9.

Methodology

Section I

The first step of intra prediction is finding the appropriate intra coding modes. There are five intra coding modes in 4*4 size (see Figure 1): mode 0 (vertical), mode 1 (horizontal), mode 2 (DC), mode 3 (diagonal down-left), and mode 4 (diagonal down-right). We make frame 0 divided by 4*4 blocks, then check every block in frame 0 by every mode. The mode with smallest SAD would be selected for the corresponding block.

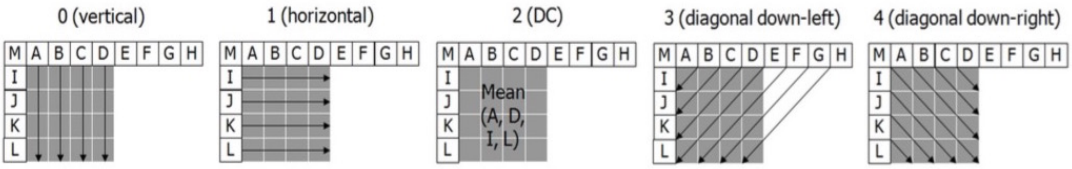

Figure 1. Five intra coding modes

After the intra prediction coding, we do the DCT transformation. First, we find the difference of original image, frame 0, and the intra predicted frame 0. Then, we input every 4*4 block, X, into the formula shown below:

$$Y' = (C_f X C_f^T) \otimes E_f$$

$$= \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} X \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \right) \otimes \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix}$$

Here, a = 0.5 and b = 0.632.

After getting the DCT result, Y', we do the inverse DCT transformation with following formula by input Y':

$$Y = C_i^T (Y' \otimes E_i) C_i$$

$$= \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} \left(X \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{bmatrix}$$

Y is the final result of inverse DCT.

The final step is to do the quantization. The transform coefficient block $W_{ij} = C_f X_{ij} C_f^T$. Here X_{ij} is the block at position (i, j) in the final result of inverse DCT. Then, we multiply W_{ij} with PF/Q_{step} to get the integer result, Z_{ij}. In this case, because that we set the quantize parameter to 1, the value of Q_{step} is equal to 0.6875. The value of PF is shown below:

Position	PF
(0,0), (2,0), (0,2) or (2,2)	a^2
(1,1), (1,3), (3,1) or (3,3)	$b^2/4$
others	$ab/2$

Here, a = 0.4 and b = 0.632.

To do the inverse quantization, we multiply the quantized result with Q_{step}, PF, and 64. The Value of PF is shown below:

Position	PF
(0,0), (2,0), (0,2) or (2,2)	a^2
(1,1), (1,3), (3,1) or (3,3)	b^2
others	ab

Here, a = 0.4 and b = 0.632.

The final result is the reconstructed frame 0.

Section II

In Section II, we use 3-step search method to find the moving vector of each block. We use the reconstructed frame 0 with frame 1. The search range is form -16 to 16, and we search from the left-top area. At first, the search blocks have size 8*8. Once the targets are found (the smallest SAD block), the sizes of blocks become a half. With the motion vector outputted by motion estimation step, we can get the motion compensation result. In motion estimation and motion compensation parts, I referenced tutor's source code "motion_estimation.m" and "warp_image.m".

With the motion compensation result, we do the DCT transformation and quantization, But here, the quantize parameter is 8 so that the Q_{step} is equal to 1.625. With repeating these steps introduced in Section II, we can get the reconstructed frame from frame 1 to frame 9.

Simulation Results

Section I

The PSNR of the reconstructed frame 0 is 62.66. For totally 6336 blocks, mode 0 is selected for 2251 times, mode 1 is selected for 2028 times, mode 2 is selected for 143 times, mode 3 is selected for 670 times, and mode 4 is selected for 1244 times. The intra-predicted image is shown in Figure 2, and the intra-predicted image with DCT transformation and quantization is shown in Figure 3.

Figure 2. Intra-predicted image

Figure 3. Intra-predicted image with DCT transformation and quantization

Section II

Each reconstructed frame and PSNR are shown below:

Figure 4. Reconstructed frame 1 (PSNR = 55.26)

Figure 5. Reconstructed frame 2 (PSNR = 55.13)

Figure 6. Reconstructed frame 3 (PSNR = 55.08)

Figure 7. Reconstructed frame 4 (PSNR = 55.04)

Figure 8. Reconstructed frame 5 (PSNR = 55.00)

Figure 9. Reconstructed frame 6 (PSNR = 54.99)

Figure 10. Reconstructed frame 7 (PSNR = 54.98)

Figure 11. Reconstructed frame 8 (PSNR = 54.91)

Figure 12. Reconstructed frame 9 (PSNR = 54.90)

Discussion

Compare the results of intra prediction and inter prediction, we can found that the PSNR of intra predicted image is higher, which means the performance is better. Besides, for inter prediction, the PSNR is smaller and smaller frames by frames. It is because that we used the reconstructed image (with inter prediction) to reconstruct another image.

Besides, for quantization, we can found that if the quantize parameter is smaller, the PSNR value of reconstructed image is smaller. And for 3-step search method, if we use the larger search blocks, the result cannot present in very detailed.

Conclusion

With prediction coding and transformation coding, we can present the video specification as H.264 and so on. The intra prediction coding and inter prediction coding also improve the compression efficiency.