EQ2330 Image and Video Processing

EQ2330 Image and Video Processing, Project 2

Yuqi Zheng
yuqizh@kth.se

Ernan Wang
ernan@kth.se

Shuyi Chen
shuyic@kth.se

December 22, 2021

Summary

This project investigates three video coding algorithms: intra-frame video coder, conditional replenishment video coder and video coder with motion compensation. They were implemented in order and the comparison of their performance of them was made. The Peak Signal to Noise Ratio (PSNR) is used as the performance measure. The videos used for this investigation is 176 x 144 pixels, with a frame rate of 30 fps. They are represented by 8 bits per pixel, so the video bitrate is about 6080 Kbps.

The intra-frame coding exploits the redundancy of the single frame, while the conditional replenishment coding exploits the correlation between two successive frames. And the inter-frame encoder with motion compensation is expanded from the conditional replenishment video coder. It predicts a frame in a video based on successive blocks. The rate-PSNR curves are plotted to evaluate their performance and to compare their efficiency. In general, the video coder with motion compensation has the best performance.

1 Introduction

Uncompressed video requires a very high data rate, so video compression is important for video production and transmission, In the field of video compression, A video is compressed utilizing different algorithms centred mainly around the amount of data compression.

The first method, intra-frame coding, is used within a video frame taking advantage of the similarity of neighbouring pixels within an image. It divides the frame image into blocks and encodes minor differences of neighbouring pixels with fewer bits, which finally leads to smaller file sizes and lower bitrates.

The conditional replenishment video coder is expanded for the first one via adding block-based conditional replenishment. It makes use of the property that frames usually do not change a lot from one to the next. It transmits only the changing part of an image and retains the static part. Through this, it varies the effective sampling rates at differences within an image and realizes low data rates.

Motion compensation is applied to reduce the redundancy during video compression. It is based on the observation that the motion in images is made of the displacement of an element and then moves block to the place that fits best. It usually finds that the block match is not perfect thus encodes changes and transmits them with the motion vector. This minimizes the amount of information that must be transmitted and has the potential to obtain an overall low data rate for images with small amounts of motion.

2 System Description

In order to investigates three kinds of coding modes– intra-frame encoder, conditional replenishment encoder and inter-frame encoder with motion compensation, we use the PSNR and the bit rate as key performance indicators. And we vary the quantizer step-size over the range 8, 16, 32, 64 to measure a rate-PSNR curve.

We carry the analysis according to the following steps. First, each frame was divided to 16×16 blocks. Then, according to Lagragian cost function Equation 1, the best mode that minimizes the function for each block can be chosen. Finally, we encode the frame with the chosen modes and estimate the bit rate of the chosen mode.

$$L = D_n + \lambda R_n. \tag{1}$$

Intra-frame encoder encodes each frame independently from previous frames. For the intra-frame video coder, each frame was subdivided into 16×16 blocks and for per 16×16 block, we apply 4 two-dimensional 8×8 DCTs. And then each block was quantized in DCT domain with 8, 16, 32 and 64 quantization steps. After quantization, the inverse DCT was applied for each block separately. The distortion was computed as the mean square error of the encoded block and the original block. The bit rate was estimated using the ideal code word length of a variable length code (VLC) on the block in DCT domain. The bit rate is the entropy coding of pixels probability in the DCT domain.

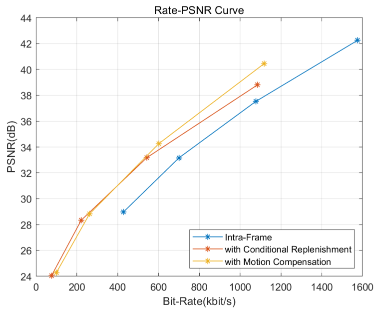
In the conditional replenishment video coder, we choose to use intra mode or copy mode based on the Lagrangian cost function. In the case of copy mode, the block is simply copied from the same location in the previous frame, therefore less bits are used for coding. The D in the Lagrangian cost function is the Mean Square Error between copied block from the previous frame and the block from the same location in the original video frame.

Motion compensation mode assumes that some blocks might just change the position from one frame to the other. The motion compensation mode allow a displacement range of [-10, . . . ,10]×[-10, . . . ,10] pel for motion-compensated prediction. There are 21 possible shifts in x direction and 21 possible shifts in y direction. So the bit rate needed to encode the best motion vector is $2 \times log_2 21 = 2 \times 5$ bits. In addition, 2 bits are needed for the mode selection (possible modes are: intra-frame coding, copy mode or motion compensation mode). Then the motion compensation uses residual coding. Subtracting the motion-compensated block from the original block determines the residual signal. The residual block is encoded using the intra-frame encoder and added to the block copied from the best motion direction. The final bit rate for motion compensation is then 10 bits that are necessary to code the motion vector and the residual transform coefficients plus 2 bits to indicate the inter mode. The distortion error is the mean square error of the reconstructed block (sum of the motion-compensated block and the residual block).

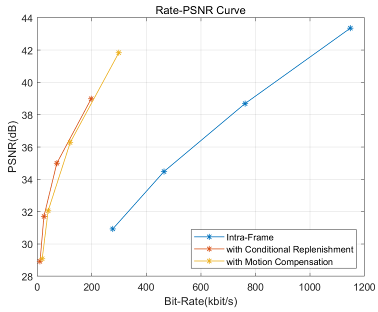
3 Results

3.1 Intra-Frame Video Coder

An intra-frame coder encodes each frame of a video sequence independently from other pictures. In this problem, we utilize the DCT transform as investigated in **Project 2**. Firstly, we subdivide the frame into 16×16 blocks and apply four two-dimensional 8×8 DCTs in each block. Then, we quantize the DCT transform coefficients by a scalar uniform quantizer with the same quantizer step-size Q for all coefficients. Next, we encode the DCT transform coefficients, assuming the ideal code word length of VLC, which is the entropy of the coefficients, as studied in **Project 2**. To be specific, we use a different VLC for the transform coefficients within each block, but use the same VLC for the i-th transform coefficient in each block. We encode *Foreman* and *MotherandDaughter* with the intra-frame coder with the quantizer step-size varying over the range 2^3, 2^4, 2^5, 2^6. The rate-PSNR results are shown as the blue curves in Fig 1(a) and Fig 1(b).



(a) Rate-PSNR Curve of *Foreman*



(b) Rate-PSNR Curve of *Mother&Daughter*

Figure 1: Rate-PSNR Curve

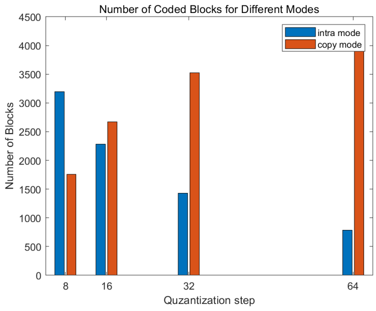
We can see that as the bit-rate is significantly reduced compared to the original video. The performance of the compression is good. Also, as the quantizer step-size decreases, the bit-rate and the PSNR increase. At high bit-rates, the gain in performance is about 6 dB per added bit/pixel for each video, which is consistent with the result we obtained in **Project 2**.

3.2 Conditional Replenishment Video Coder

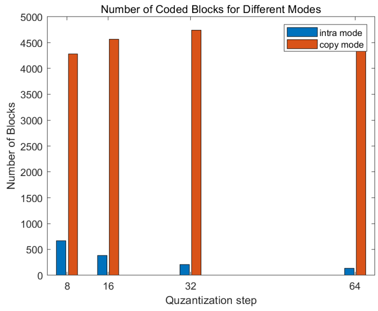
In this problem, we expand the intra-frame video coder by adding block-based conditional replenishment. The encoder decides whether to use the *intra mode* as investigated in **3.1** or the *copy mode* which copies the co-located 16×16 block in the previous frame as the current 16×16 block. In this case, we assume that the difference between one frame and another is not significant so that the *copy mode* can still guarantee the video quality. The choice of the mode is based on the *Lagrangian* cost function $J_n = D_n + \lambda R_n$ for each 16×16 block as described in **System Description**. Especially, we assume that λ is proportional to the squared quantizer step-size Q^2, i.e., $\lambda = 0.001Q^2$ for *intra mode* and $\lambda = 0.002Q^2$ for *copy mode*. We choose the mode that leads to the smallest cost to encode the frames as we did in **3.1**.

Fig 2(a) and Fig 2(b) show the number of coded blocks of these two modes as the quantizer step-size varies over the range 2^3, 2^4, 2^5, 2^6. We can see that as the quantizer step-size increases, the number of blocks encoded by the *copy mode* becomes larger while the number of blocks encoded by the *intra mode* becomes smaller. Intuitively, this is because the difference between frames becomes less obvious as the quantizer step-size increases. The rate-PSNR results are shown as the red curves in Fig 1(a) and Fig 1(b). Similar to the *intra mode* (the blue curves), the PSNR of the conditional replenishment video coder increases with bit-rate as the quantizer step-size decreases. This is also consistent with the results of the number of encoded blocks of these two modes we just obtained.

In addition, the performance of the conditional replenishment video coder is better than that of the intra-frame video coder. This is because copying blocks saves bits without compromising video quality. Especially, for the same bit-rate, the PSNR of the conditional replenishment coder of the *Foreman* is on average 1-2 dB greater than that of the intra-frame coder. While as for the *Mother and Daughter*, the conditional replenishment video coder performs much better since this video is more static than the *Foreman*. Thus, the error caused by copying is smaller.



(a) Number of Coded Blocks of *Foreman*



(b) Number of Coded Blocks of *Mother&Daughter*

Figure 2: Number of Coded Blocks for Two Modes

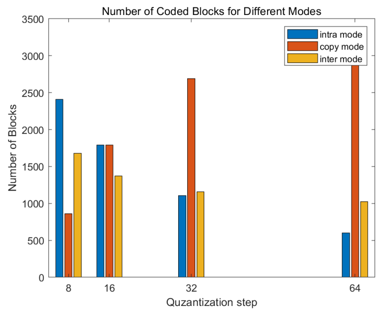
3.3 Video Coder with Motion Compensation

In this section, we look into video coder with motion compensation. If the position of each block does not shift much from frame to frame, in addition to the *intra mode* and the *copy mode*, we could implement a 16×16 block-based *inter mode* that allows block-based integer-pel motion-compensated prediction. That is, if the block moves less than ten pixels from frame to frame, we can predict the block in the next frame based on the motion vector, which is determined by the block motion estimation. Next, we need to use the intra-frame coder to encode the residual of each block. Similar to **3.2**, we choose the mode that leads to the smallest cost to encode the frames. Fig 3(a) and Fig 3(b) show the number of coded blocks of the *intra mode*, the *copy mode* and the *inter mode* as the quantizer step-size varies over the range 2^3, 2^4, 2^5, 2^6. We can see that as the quantizer step-size increases, the number of blocks encoded by the *copy mode* becomes larger while the number of blocks encoded by the *intra mode* and *inter mode* becomes smaller since larger quantizer step-size means fewer details and less information. Hence, just copying blocks from the previous frames is enough.

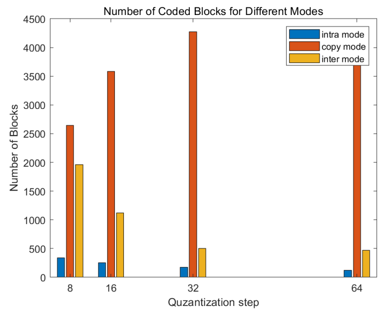
The rate-PSNR results are shown as the yellow curves in Fig 1(a) and Fig 1(b). Similar to **3.1** and **3.2**, the PSNR of the video coder with motion compensation increases with bit-rate as the quantizer step-size decreases, which is also consistent with the results we obtained previously of the number of encoded blocks of these three modes.

Also, we can see that the performance of the video coder with motion compensation is better than that of the intra-frame video coder. This is because motion compensation takes the block movement and the residual into account. Especially, for the same bit-rate, the PSNR of the video coder with motion compensation of the *Foreman* is on average 2 dB greater than that of the intra-frame coder. While as for the *Mother and Daughter*, the video coder with motion compensation performs much better. For each video, at high bit-rates, the gain in performance of the video coder with motion compensation is about 6 dB per added bit/pixel as well.

Furthermore, we can see that for larger quantizer step-size, the conditional replenishment video coder and the video coder with motion compensation perform close to each other. However, for smaller quantizer step-size, the video coder with motion compensation performs better than the conditional replenishment video coder.



(a) Number of Coded Blocks of *Foreman*



(b) Number of Coded Blocks of *Mother&Daughter*

Figure 3: Number of Coded Blocks for Three Modes

4 Conclusions

In this project, three different video compression algorithms were implemented and PSNR-rate curves were plotted to compare their performances. In conclusion, the motion compensation outperformed the others.

As shown in Fig 1(a) and Fig 1(b), the conditional replacement coding and the motion compensation coding improved video PSNR. The performance of the conditional replenishment coding is generally better than the intra-frame coding, as blocks for larger quantization steps can save more bits and maintain the PSNR.

The conditional replenishment mode performs similarly to the motion compensation mode when the quantization steps are large. However, when the quantization steps are smaller, the motion compensation mode outperforms the conditional replenishment in terms of PSNR. This is because motion compensation mode takes into account blocks movements and residual coding. Also, it can be noticed that the gain in performance of motion compensation mode and intra-frame mode is roughly 6 dB per added bit/pixel for large bit-rates.

Appendix

Who Did What

Report

Summary and Introduction: Shuyi Chen
System description: Ernan Wang
Results: Yuqi Zheng
Conclusion: Shuyi Chen, Ernan Wang

Code

Yuqi Zheng

MatLab code

The MATLAB code is attached as matlab.rar.

References

[1] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Prentice Hall, 2nd ed., 2002