Video Compression and Coding using Transforms, Subband Filters and Motion Compensation

Philip Gertzell
Rickard Larsson
Group 5

2018-05-02

A: INTRA-Frame Image Compression

Task 1.1: DCT-based Image Compression

When applying the DCT method an impressive compression can be achieved. Even though 90 % of the DCT-coefficients have been removed, the image can be recovered without losing the information in the image, however the quality is impaired. Since the human eye is a subjective perception of quality, a mathematical measurement is needed. This can be done by measuring the mean SSIM (structural similarity) of the image and the PSNR (Peak Signal to Noise Ratio).

The index $th = 0.9 \cdot k_1 \cdot k_2 = 69624$ correspond to a threshold value of 0.1154. The resulting PSNR due to compression was 29.8 dB and the mean SSIM resulted in 78 %. These values verify that the compressed image has an acceptable quality (considering the compression ratio). The original and compressed image are shown together with the error and DCT domain image in figure 1. As seen in the figure most of the information is located in the low frequency domain with many strong DC-components. All the high frequency components have been removed by thresholding since they do not carry the most important data to represent the image.

Task 1.2: Block-based DCT-domain Image Compression

In the previous task the most valued information was gathered from the entire frame. By dividing the image in small blocks (8x8 pixels), the most valuable information from each block can be extracted. As seen in figure 2 each block's most valuable frequency and DC-components have been extracted. The index $th = 0.9 \cdot k_1 \cdot k_2 = 69624$ correspond to a threshold value of 0.0758. The resulting PSNR due to compression was 34.33 dB and the mean SSIM resulted in 88.6 % which is an improvement from last method. Even though the mathematical measurements claim a better image quality, the subjective mind might dislike the fact that the blocks are noticeable.

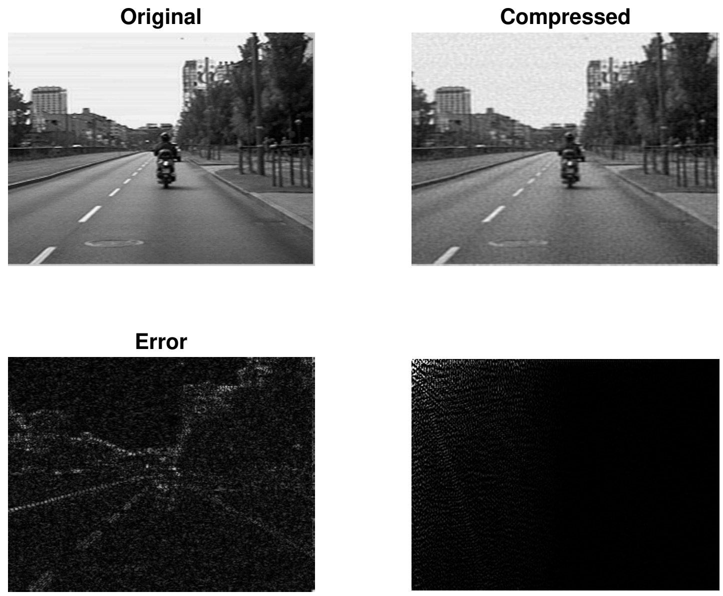

Figure 1: DCT-based image compression with 90 % compression rate. The original image is shown in top left, compressed in top right, 30 times the error is shown in bottom left and the DCT-domain in the bottom right.

Task 1.3: Subband Image Compression using Wavelets

The qualitative metrics PSNR and MSSIM speak a clear language when comparing block-based DCT and wavelet compression, as can be seen in table 1. Block-based DCT yielded roughly a +5dB improvement in Peak-SNR and 4% improvement in Mean-SSIM. However, visually comparing the two compressed images side by side, as in figure 3, the qualitative differences are not as obvious at first glance.

The two images have its own unique characteristics and the perceived quality is highly subjective. The wavelet compressed image strikes us as more "pixelated", especially in the sky and the dashed line in the middle of the street. In the DCT compressed image, the blocks are more visible and relatively sharp edges can be seen in the lower part of the image. Therefore, in our opinion, the DCT compressed image is of higher quality because the transitions between different areas of the image appear smoother (in most regions of the image).

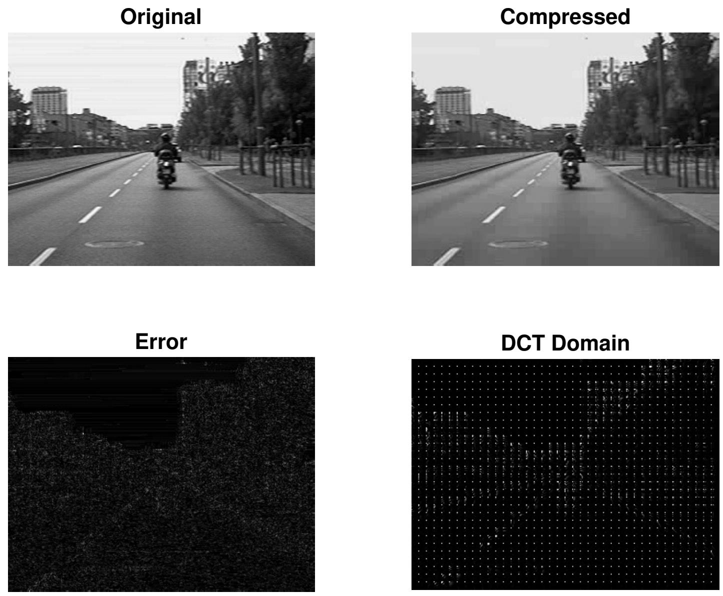

Figure 2: Block-based DCT image compression with 90 % compression rate. The original image is shown in top left, compressed in top right, 30 times the error is shown in bottom left and the DCT-domain in the bottom right.

Table 1: Resulting qualitative metrics MSSIM and PSNR

	Mean SSIM [%]	PSNR [dB]
Block DCT	88.6	34.3
Wavelets	84.42	29.4

B: Inter-Frame Compensation and Compression

Task 2.1 to 2.3

In order to compensate for motion between consecutive image frames a binary image map, I_{motion}, had to be computed. This was done by computing the differences between two frames by a simple matrix subtraction. The resulting image, I_{diff}, was then thresholded in order to create the binary image map that only indicated areas of the images where motion was clearly visual. This motion map can be seen in figure 4, where white indicates an area of motion. The threshold value chosen to generate a fair motion map was 45/255.

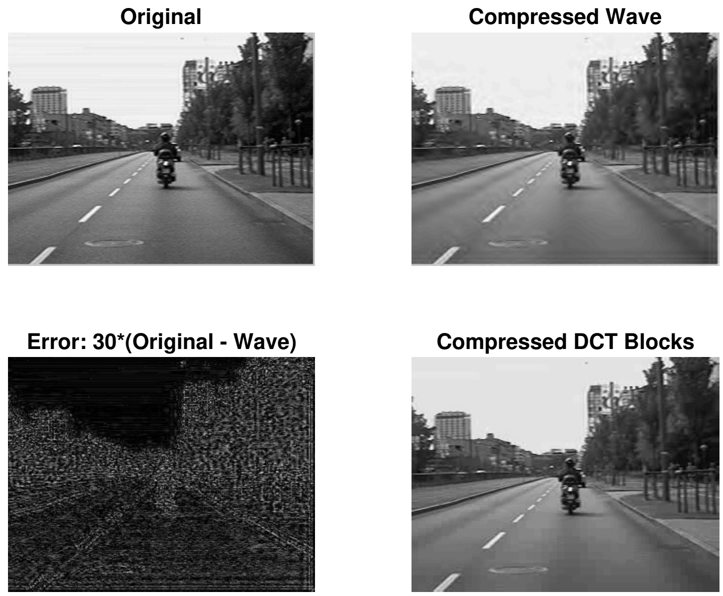

Figure 3: Wavelet-based image compression with 90 % compression rate. The original image is shown in top left, compressed in top right, 30 times the error is shown in bottom left and the DCT-domain in the bottom right.

Task 2.4 and 2.5

In order to reconstruct the next consecutive frame, the motion image was used to determine which pixels in the old frame needed to be changed. Since most information can be used in the old frame a full global scan was performed in order to compute the motion vectors under the MAE criterion. Once the scan for each block was done, the motion vectors are transmitted to the receiver. The receiver can then reconstruct the consecutive frame by saving the old pixels that doesn't need to change and apply the motion vectors to the pixels that was mapped in I_{motion}. The results are seen in figure 5. The MSSIM = 65.1% and PSNR = $21.9dB$, which indicates a rather poor quality, but with a great compression rate.

C: Answers to question 1-6

1)

Image compression can be done by excluding redundant information, in our case we exclude high frequency components. Video compression can be done

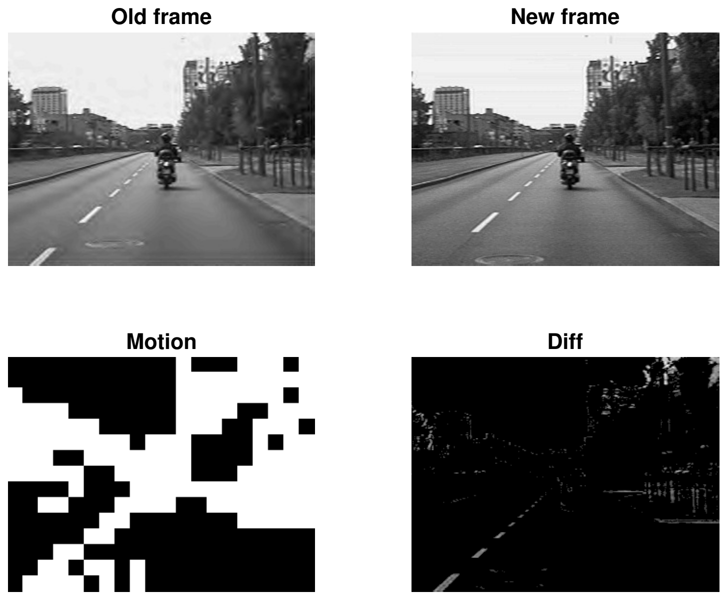

Figure 4: The top left picture shows the original compressed first frame, top right is the original next consecutive frame. The motion image is shown in bottom left and the difference between the old and new frame in bottom right.

by first transmitting the first compressed frame (Intra) then only transmit how e.g. frame 2 has changed from frame 1 (Inter).

2)

The threshold of which information should remain and which should be excluded will determine the rate of compression.

3)

Intuitively, we believed wavelet would outperform DCT since it seemed to be a more complex technique and because it is used in JPEG2000. However, when comparing them side by side in figure 3, wavelet did not seem to be superior despite the PSNR and MSSIM were higher.

The main issue with DCT compression is that blocks starts to be visible when the compression rate is pushed too high. Depending on the application this can be experienced as a decrease in quality. Wavelet compression counters this issue since it does not use blocks nor rely on the similarity of neighbouring pixels.

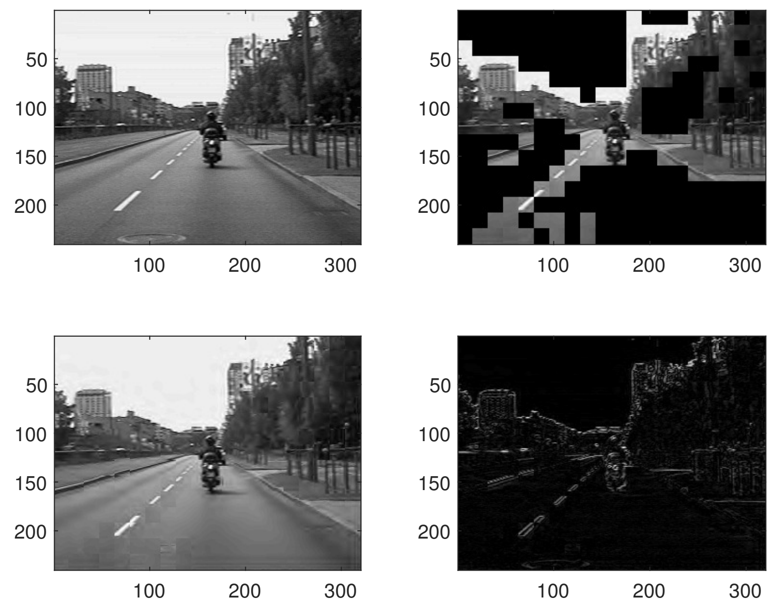

Figure 5: The top left image shows the original compressed new frame, the top right shows which of the old pixels that will need to be changed from previous frame (the black area will remain the same). In the bottom left the reconstructed image from both INTER and INTRA modes with the corresponding error in the bottom right

However, wavelet compression is not necessarily superior since it comes with a price. It is considerably more computationally heavy and a big amount of data usage can be saved by using DCT if the application allows it.

In summary, none of the compression techniques are inherently superior since it depends on the application and it all boils down to a design problem where the trade-offs (image quality vs. complexity for instance) need to be considered. Furthermore, as mentioned earlier, image quality is highly subjective and therefore it is a difficult task to objectively compare compression techniques.

4)

The fact that most information in the previous frame can be used to reconstruct the next frame.

5

Local searching means that each block is only searched for in a limited search area in close proximity of the original block. By knowing the application of the motion compensation one could compute an average block length motion and use this as an initial radius for the search area. Thereafter, calculating the success of the motion compensation, the search area can be adjusted accordingly.

6)

Since a global search is more carefully and profound it will result in a better quality at the cost of higher computation than the local search.