NTNU

Norwegian University of Science And Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

TTT4135 MULTIMEDIA SIGNAL PROCESSING
Assignment 2

Note

The assignment consists of two parts. First part deals with theoretical problems while the second one deals with practical issues. Both parts should be done in groups of two students.
Please write the solutions in English or Norwegian, and submit it electronically on It's Learning in "Mandatory Assignments → Assignment 2". DEADLINE FOR HANDING IN THE SOLUTIONS IS February 28, 2022. Solutions submitted after the deadline WILL NOT be accepted.

Preliminaries

The following materials should be studied before starting this assignment. The related papers can be found in the folder 'References' on its learning.

- Chapters 11.1, 11.2 and 11.3 of the textbook (Digital compression for multimedia) for details of MPEG video coding standard
- Video Compression From Concepts to the H.264/AVC Standard
- Overview of the H.264/AVC Video Coding Standard
- New features and applications of the H-264 video coding standard

Part I (50 points)

1 Temporal prediction

The inherent temporal correlation that exists in video data can be utilized for compression purposes.

a) Explain the principles of *motion compensation* and how this generally is done. (5 points)

b) What do we mean by I, P and B-frames, and what types of motion compensation is done in each of these? What does this imply for the compression gains and visual quality for each of these frame types? (10 points)

c) Why is there a difference between the *transmit order* and the *display order* of frames?(10 points)

d) Give a short summary (no details) of how H.264 separates itself from other standards with respect to how motion compensation is done. (10 points)

2 Decomposition/Transform

As is usual in still image coders, video coders do a frequency decomposition prior to quantization and entropy coding. It is the residual image (what remains after prediction) that is transformed

a) Which transform has usually been used in classic video coders (such as MPEG-1)? What is used in H.264? (5 points)

b) Block-transform-based video/still image coders usually exhibit so-called blocking artifacts. How is this effect minimized in H.264? (10 points)

Part II (50 points)

3 Video coding

A video signal is modelled as a stationary time-discrete gaussian process $Z[n]$ with average 0 and autocorrelation function $r_Z[k] = E\{Z[n]Z[n+k] = 0.9^{|k|}\}$. First, every second pixel is coded predicatively using the predictor:

$$Z'[n] = a\, x\, Z''[n+1] \tag{1}$$

Then, the remaining pixels are coded using the predictor

$$Z'[n] = b\, x\, Z''[n-1] + c\, x\, Z''[n+1] \tag{2}$$

(It is ok to use both backward and forward prediction, since both the preceding and the following pixel are known to the decoder.) The prediction errors are quantized finely and the quantized errors are entropy coded. Determine the optimal performance (distortion as a function of rate) for this coding method. (20 points)

4 Motion estimation

In the following task three consecutive frames from a video are given, the frames are found in the assignment folder.

a) Using optimal displacement, calculate and plot all the motions vectors for each macro-block of size (16x16) for a chosen search range V of size (32x32) for the whole frame. Plot the vectors for both frame N and N+1 as illustrated in Figure 1. (20 points)

b) Do the same for a macro-block of size (8x8) and compare the complexity for the 2 searches (e.g. number of operations). (10 points)

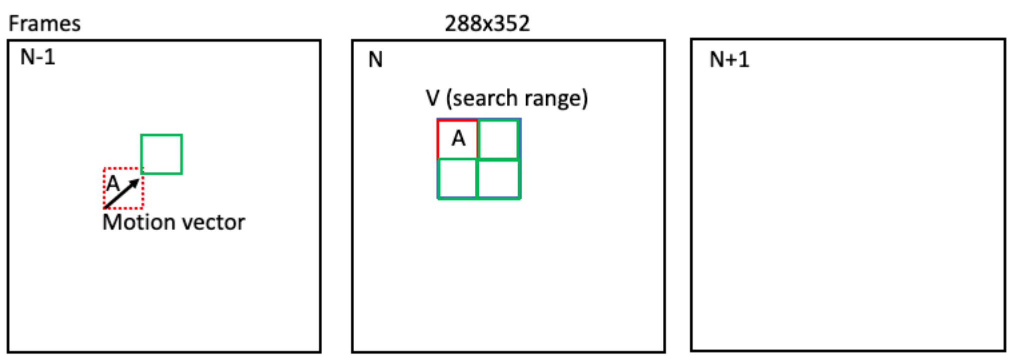

Figure 1: Sketch of the task

Optimal displacement is given by:

$$v(i)^{*} = \min \sum_{\underline{X} \in M} D(I(\underline{X}, t) - I(\underline{X} - \underline{v}, t - \tau)), \underline{v} \in V$$

$\underline{X}$	Co-ordinate of picture element in Frame M
$\underline{v}$	Motion vector relevant to I($\underline{\text{X}}$,t)
V	Search range
$D(\cdot)$	Distortion measure (Cost function)

Figure 2: Optimal displacement