Transmission of Audio and Video Signals Over the Internet

**Homework 2**

**Alexander Fok, 308669944**

**Igal Lankri, 316225622**

# Question 1

1. קראו את המאמר של Kumar וממשו את שיטת ה Intra Prediction של סטנדרט ה H.264 עבור בלוקים בגודל 4x4. המימוש יכלול את כל מודי האינטרה לבלוקים בגודל 4x4, התחשבות במצבים ההתחלתיים כאשר לא כל הפיקסלים של המסגרת קיימים, התמרת DCT, קוונטיזציה וכן התמרת DCT הפוכה.

יש לתעד היטב את תוכנית המחשב המבצעת את הדחיסה.

The program with implementation of all intra prediction modes is attached. It also can be found here: <https://github.com/alexfok/AVSignalsTransmission/blob/main/HW2/ex2_main.py>. The running instructions and results can be found here: <https://github.com/alexfok/AVSignalsTransmission/blob/main/HW2/README.md>

The selected images are frame#1 from akiyo\_cif.yuv and frame#1 from waterfall\_cif.yuv.

To run the program:

python ./HW2/ex2\_q2\_main.py ./HW1/stv/akiyo\_cif.yuv 288 352 4

python ./HW2/ex2\_q2\_main.py ./HW1/stv/waterfall\_cif.yuv 288 352 4

# Question 2

## A

עבור כל אחת מהתמונות שבחרתם, ועבור ערכי QP של 6, 12, 18 ו 30 הציגו את התמונה המקורית, תמונת שערוכי האינטרה, תמונת השארית וכן את התמונה המשוחזרת.

### Original image - Akio\_cif



### Predicted images

#### Akio\_cif\_predicted QP=6



#### Akio\_cif\_predicted QP=12



#### Akio\_cif\_predicted QP=18



#### Akio\_cif\_predicted QP=30



### Residual Images

#### Akio\_cif\_residual QP=6



#### Akio\_cif\_residual QP=12



#### Akio\_cif\_residual QP=18



#### Akio\_cif\_residual QP=30



### Reconstructed Images

#### Akio\_cif\_reconstructed QP=6



#### Akio\_cif\_reconstructed QP=12



#### Akio\_cif\_reconstructed QP=18

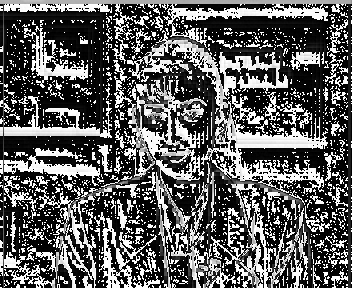


#### Akio\_cif\_reconstructed QP=30



### Selected residual frames for most different intra-prediction modes

#### Akio\_cif\_residual QP=6, mode=0



#### Akio\_cif\_residual QP=6, mode=1



#### Akio\_cif\_residual QP=6, mode=5



### Original Image Waterfall\_cif



### Predicted images

#### Waterfall\_cif\_predicted QP=6



#### Waterfall\_cif\_predicted QP=12



#### Waterfall\_cif\_predicted QP=18

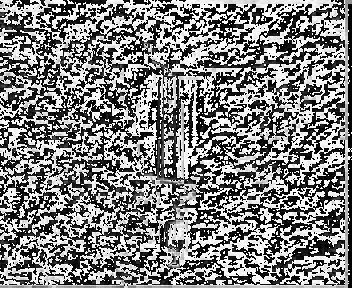


#### Waterfall\_cif\_predicted QP=30

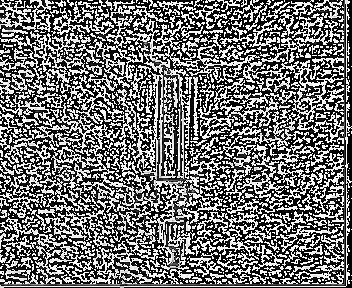


### Residual Images

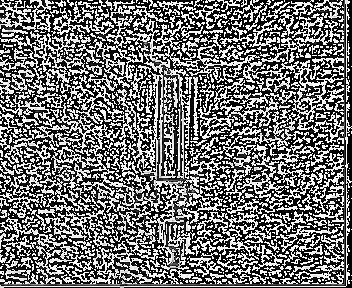
#### Waterfall\_cif\_residual QP=6



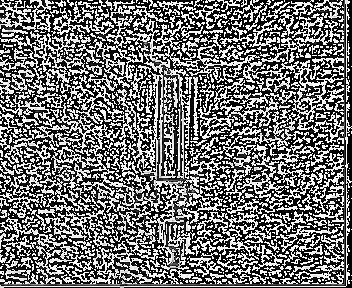
#### Waterfall\_cif\_residual QP=12



#### Waterfall\_cif\_residual QP=18



#### Waterfall\_cif\_residual QP=30



### Reconstructed Images

#### Waterfall\_cif\_reconstructed QP=6



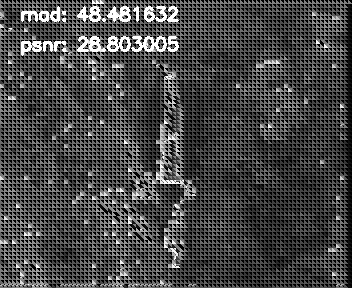
#### Waterfall\_cif\_reconstructed QP=12



##### Waterfall\_cif\_reconstructed QP=18

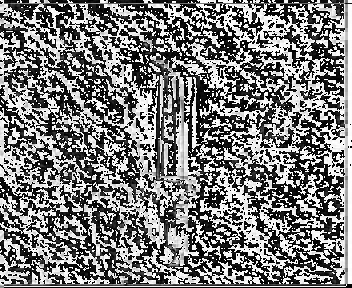


##### Waterfall\_cif\_reconstructed QP=30

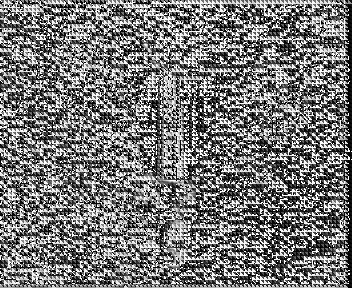


### Selected residual frames for most different intra-prediction modes

##### Waterfall\_cif\_residual QP=6, mode=1



##### Waterfall\_cif\_residual QP=6, mode=7



## B

חשבו את ה PSNR בין התמונות המקוריות ובין התמונות המשוחזרות. את ה PSNR המתאים רשמו בכותרת התמונה המשוחזרת.

## C

חשבו את ה MAD אותו ניתן להוציא מתהליך חיפוש המודים הטובים ביותר ורשמו אותו לצד תמונת השארית.

### Akiyo image

test\_intra\_pred\_modes: test\_mode 0

psnr: 32.912112

mad: 10.123490

test\_intra\_pred\_modes: test\_mode 1

psnr: 32.901050

test\_intra\_pred\_modes: test\_mode 2

psnr: 34.946536

mad: 6.403015

test\_intra\_pred\_modes: test\_mode 3

psnr: 35.390134

mad: 6.118006

test\_intra\_pred\_modes: test\_mode 4

psnr: 33.001738

mad: 10.156650

test\_intra\_pred\_modes: test\_mode 5

psnr: 33.001758

mad: 10.151450

test\_intra\_pred\_modes: test\_mode 6

psnr: 33.000738

mad: 10.152650

test\_intra\_pred\_modes: test\_mode 7

psnr: 30.155924

mad: 42.782997

test\_intra\_pred\_modes: test\_mode 8

psnr: 33.003728

mad: 10.153640

### Waterfall image

test\_intra\_pred\_modes: test\_mode 0

psnr: 29.350348

mad: 16.383947

test\_intra\_pred\_modes: test\_mode 1

psnr: 30.870920

mad: 10.175397

test\_intra\_pred\_modes: test\_mode 2

psnr: 30.681791

mad: 10.193645

test\_intra\_pred\_modes: test\_mode 3

psnr: 29.377085

mad: 16.106022

test\_intra\_pred\_modes: test\_mode 4

psnr: 29.376075

mad: 16.106312

test\_intra\_pred\_modes: test\_mode 5

psnr: 29.375265

mad: 16.107021

test\_intra\_pred\_modes: test\_mode 6

psnr: 29.382065

mad: 16.118012

test\_intra\_pred\_modes: test\_mode 7

psnr: 28.803005

mad: 48.481632

test\_intra\_pred\_modes: test\_mode 8

psnr: 29.377085

mad: 16.106022

# Question 3

## A

עבור שתי התמונות ו- QP=12, ערכו בטבלה את שכיחות המודים השונים.

For two images and QP=12, put in the table the frequency of different modes

mode\_freq\_list: [mode0=1431, mode1=943, mode2=3392, mode3=560, mode4=0, mode5=0, mode6=0, mode7=4, mode8=6]

From the modes frequency table we can see that the mode 2 (DC: Average) is the most frequent mode chosen due to minimal SAD.

## B

מצאו עבור כל תמונה את מספר הרצפים שיש מכל מוד. כיצד נשפר את ה bitrate תוך שימוש במידע שמוד הבלוק הנוכחי זהה למוד של הבלוק הקודם?

For every image, find number of mode repetition sequences.

mode\_pred\_list: [mode0=459, mode1=155, mode2=2088, mode3=58, mode4=0, mode5=0, mode6=0, mode7=0, mode8=6]

From the modes repetition table we can see that the mode 2 (DC: Average) is the most repetitive mode, while modes 4,5,6 and 7 are not used at all.

By exploiting the fact that adjacent blocks are often visually similar within a frame, video encoders can reduce redundancy by encoding the intra prediction mode used for a block once and then applying it to neighboring blocks. This avoids the need to transmit or store redundant mode information for each individual block. Using the same intra mode for adjacent blocks can help improve compression efficiency by capturing and representing the spatial correlation within a frame. It allows for better reconstruction of the original image by exploiting similarities and reducing the amount of information needed to encode each block independently.

For image **akiyo**, the number of adjacent blocks encoded with the same intra mode is 2760. It is about half of total encoded blocks – 6336. It means that significant bit rate reduction can be achieved or every encoded video frame since no need to transmit the intra mode encoding for almost half of the blocks.

For image **waterfall**, the number of adjacent blocks encoded with the same intra mode is 3485. It is more than half of total encoded blocks – 6336. It means that bit rate reduction for this image can be even higher.

# Question 4

## A

MSE measures the average squared difference between the predicted and original pixel values. It provides a measure of the overall distortion or error between the two images. However, MSE can be sensitive to outliers, which means that extreme pixel values, such as those introduced by salt and pepper noise, can have a significant impact on the computed value. This can lead to misleading results and potentially prioritize smoothness over accurately capturing the original image details.

On the other hand, SAD measures the sum of the absolute differences between the predicted and original pixel values. It provides a measure of the total intensity difference between the two images, regardless of the sign or magnitude of the differences. SAD is less sensitive to outliers and can be more robust in the presence of salt and pepper noise. It focuses on capturing the magnitude of the differences, which can be important in scenarios where preserving sharp edges or fine details is crucial.

Considering the characteristics of salt and pepper noise, where extreme pixel values are present as isolated outliers, SAD might be a more suitable metric for evaluating the performance of intra mode prediction in the presence of such noise. It can effectively capture the overall intensity differences, including the impact of the noisy pixels, while still providing a meaningful measure of the prediction quality.