Comparative and performance analysis of HEVC and H.264 Intra frame coding and JPEG2000

EE5359 Multimedia Processing
Interim Report
Spring 2013

The University of Texas at Arlington
Department of Electrical Engineering

Submitted by:
Harshdeep Brahmasury Jain
ID: 1000845657
Email: harshdeep.brahmasuryjain@mavs.uta.edu

Under the Guidance of:
Dr. K.R.Rao
LIST OF ACRONYMS

CABAC: Context adaptive binary arithmetic coding

CU: Coding unit.

CTU: Coding tree unit.

FDCT: Fast discrete cosine transform.

HD: High definition

HEVC: High Efficiency Video Coding.

IDCT: Inverse discrete cosine transform.

JPEG2000: Joint photographic experts group.

JCT-VC: Joint collaborative team on video coding.

JM software: Joint model software.

MSE: Mean square error.

MPEG: Moving picture experts group.

PU: Prediction unit

PSNR: Peak signal to noise ratio.

QCIF: Quarter common intermediate format.

QP: Quantization parameter

SSIM: Structural similarity index.

TU: Transform units.
HEVC: HIGH EFFICIENCY VIDEO CODING

Objective:
The objective of this project is to study, implement and compare video coding standards like HEVC, H.264 and JPEG 2000 [1][5][7]. The analysis will be carried out on the intra frame coding and different image performance metrics like MSE, PSNR, bitrate, SSIM [23] and video quality will be evaluated for high resolution videos.

Motivation:
With the ever increasing demand for high definition video, several different video coding standards have been developed to address the needs of HD video coding. HEVC is the newest video coding standard of the ITU-T video coding experts group and the ISO/IEC moving picture experts group [1]. The main goal of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standards such as H.264 [5].

Introduction:
THE high efficiency video coding (HEVC) standard is the most recent joint video project of the ITU-T video coding experts group (VCEG) and the ISO/IEC moving picture experts group (MPEG) standardization organizations, working together in a partnership known as the joint collaborative team on video coding (JCT-VC) [1]. While it is widely accepted that video coding algorithms, e.g. H.264/AVC [5], exploit both the temporal redundancy between adjacent frames and spatial redundancy within a frame can achieve higher compression efficiency than those algorithms designed only for images. There are some particular scenarios in which intra coding or image coding is the only
choice, such as the editing within post production where easy access to each individual frame is of great importance [19]. HEVC encoder block diagram is shown in the figure 1.

Figure: 1 HEVC video encoder block diagram [1].
Further, the lossless intra coding is especially required for applications where images acquired are for extracting specific information in the future. HEVC aims at doubling the compression ratio of H.264/AVC High profile [5] with comparable image quality, at the expense of increased computational complexity [1]. HEVC decoder block diagram is shown in the figure.2.

The main structure of HEVC is very similar to that of H.264. Both of them utilize the hybrid video coding schemes with spatial and temporal predictions, transform quantization followed by intra-picture estimation and motion estimation and checks for the prediction errors and entropy coding is performed. The significant compression ratio improvement of HEVC compared with H.264 is achieved by the refinement of various coding tools [1].
**Need for compression:**
Consider an image of resolution $640 \times 480$. To calculate the size of the picture in RAW format, each of the RGB color is represented by 8 bits. Then for each pixel it needs 24 bits. Total number of pixels in the image is $640 \times 480 = 307200$ pixels. Therefore, the size of the image turns to $307200 \times 3$ bytes $= 921600$ bytes. But, an image in compressed format with the same resolution can be reduced to 100 KB.

Hence, compression is very useful for storage and transfer of images. Compression also removes redundant bit pixels of the image, thereby reducing the size. However, compression comes with a price affecting the quality of image.

Therefore, various standard image compression methods [7] that make a best tradeoff between these properties and compression are studied and implemented in this project. However, there can be lossy and lossless compression which also affects these properties. Lossy being permanent loss of some image data while lossless means complete retrieval of data after decoding.

**HEVC intra frame coding:**
Like H.264/AVC, HEVC uses block-based intra prediction to take advantage of spatial correlation within a picture. HEVC follows the basic idea of H.264/AVC intra prediction but makes it far more flexible. HEVC intra angular prediction modes are shown in figure.3 and angular directions of different modes is shown in figure.5. Furthermore, intra prediction can be done at different block sizes, ranging from 4 X 4 to 64 X 64 (Fig.4). HEVC also includes a planar intra prediction mode, which is useful for predicting smooth picture regions. In planar mode, the prediction is generated from the average of two linear interpolations (horizontal and vertical) [18].
Figure 3: Modes and directional orientations for intra picture prediction for HEVC [1] [18].

Figure 4: H.264 intra prediction modes [3].

Figure 5: Angular intra-prediction modes for HEVC [1].
**H.264 Encoder and Decoder Block Diagrams:**

H.264/MPEG-4 AVC is a block-oriented motion-compensation-based codec standard developed by the ITU-T video coding experts group (VCEG) together with the ISO/IEC moving picture experts group (MPEG) [1]. H.264 encoder block diagram is shown in figure 6.

![H.264 Encoder Block Diagram](image)

Figure 6: H.264 encoder block diagram [4]:

An H.264 encoder converts the raw video into a compressed version and the decoder converts the compressed video back to its original format. The encoder performs transform, quantization, prediction and encoding to produce compressed video. It also aims at having significant improvements in coding efficiency using CABAC entropy coder, error robustness and network friendliness. Parameter set concept, arbitrary slice ordering, flexible macro block structure, redundant pictures, switched predictive and switched intra pictures have contributed to error resilience / robustness of this standard [23].
Figure 7: H.264 decoder block diagram [15]:

The decoder on the other hand does the inverse operations to obtain the uncompressed video. H.264 decoder block diagram is shown in figure.7.

**JPEG2000 [16]:**

JPEG 2000 is an image compression standard and coding system. It was developed by the joint photographic experts group in 2000 with the intention of superseding the original discrete cosine transform-based JPEG standard [15] (developed in 1992) with a newly designed, wavelet-based method.

The JPEG2000 image coding algorithm consists of 5 parts, pre-processing, discrete wavelet transform, uniform quantizer with dead zone, adaptive binary arithmetic coder and bit-stream organization.

It provides functionalities that current standards can neither address efficiently nor address at all i.e., both lossless and lossy compression, encoding of very large images, progressive transmission by pixel accuracy and by resolution, robustness to channel noise, region of interest coding and random code stream access. It is also designed to
address the requirements of very different kinds of applications, e.g. internet, color facsimile, printing, scanning, digital photography, remote sensing, mobile applications, medical imagery, digital libraries and e-commerce.

The preprocessing part includes three functions, image tiling, DC level shifting and component transformations. The term tiling refers to the partition of the original image into rectangular non-overlapping blocks, which are compressed independently as if they are totally distinct images.

![Diagram of JPEG 2000 codec](image)

Figure 8: Structure of JPEG 2000 codec. The structure of the (a) encoder (b) decoder [23]

The main advantage offered by JPEG 2000 is the significant flexibility of the code stream. JPEG2000 encoder and decoder block diagram is shown in figure.8. The code stream obtained after compression of an image with JPEG 2000 is scalable in nature. Higher-resolution images tend to benefit more, where JPEG-2000's spatial-redundancy prediction can contribute more to the compression process. In very low-bitrate applications, studies have shown JPEG 2000 to be outperformed by the intra-frame coding mode of H.264 [5]. Good applications for JPEG 2000 are large images, images with low-contrast edges — e.g., medical images [5].
**IMAGE QUALITY MEASURES:**

There are two types of image quality measures.

1) **Objective quality measures**- MSE, PSNR.

2) **Subjective quality measures**- SSIM

**MSE and PSNR:**

MSE and PSNR for a N x M pixel image are defined as where x is the original image and y is the reconstructed image. M and N are the width and height of an image and ‘L’ is the maximum pixel value in the NxM pixel image.

\[
MSE = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} [x(m,n) - y(m,n)]^2 
\]

(1)

\[
PSNR = 10 \log_{10} \frac{L^2}{MSE}, \text{dB}
\]

(2)

**Structural Similarity Index:**

The structural similarity (SSIM) [22] index is a method for measuring the similarity between two images. SSIM emphasizes that the human visual system is highly adapted to extract structural information from visual scenes. Therefore, structural similarity measurement should provide a good approximation to perceptual image quality.

SSIM is designed to improve on methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proved to be inconsistent with human eye perception.

SSIM considers image degradation as perceived change in structural information. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close.

\[
SSIM(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) (\sigma_x^2 + \sigma_y^2 + C_2)}
\]

(3)

where x and y correspond to two different signals that need to be compared for similarity, i.e. two different blocks in two separate images;
\( \mu_x \) the average of \( x \);
\( \mu_y \) the average of \( y \);
\( \sigma_x^2 \) the variance of \( x \);
\( \sigma_y^2 \) the variance of \( y \);
\( \sigma_{xy} \) the covariance of \( x \) and \( y \);
\( c_1 = (k_1 L)^2 \), \( c_2 = (k_2 L)^2 \) two variables to stabilize the division with weak denominator;
\( L \) the dynamic range of the pixel-values (typically this is 2\#bits per pixel – 1);
\( k_1 = 0.01 \) and \( k_2 = 0.03 \) by default.

Figure 8: Absolute error map, SSIM index map, original image and JPEG compressed image [23]
Evaluation Methodology

Different video formats are used to do the analysis and comparative tests, from low quality video to high definition quality video. HEVC encoder is used to encode the test sequences, which are present in the .yuv format. Reference software (HM 10.0) [10] is used for HEVC encoder and each frame of the test sequences is coded in the I-frame mode. The test sequences of different resolutions are used in the analysis of H.264 codec. Reference software (JM 18.2) [9] is used for H.264 encoder and each frame of the test sequences is coded in the I-frame mode.

The configuration of the H.264/AVC JM18.2 encoder [9] is chosen as follows:

- ProfileIDC = 100  # Profile IDC (80 = main, FREXT Profiles: 100=High)
- LevelIDC = 40  # Level IDC   (e.g. 20 = level 2.0)
- IntraProfile = 1  # Activate Intra Profile for FRExt (0: false, 1: true)
- Deblocking filter = off
- QPISlice = 32  # Quant. parameter for I Slices (0-51)

The configuration of the H.265 HM10 encoder [10] is chosen as follows:

- IntraPeriod : 1    # Period of I-Frame ( -1 = only first)
- DecodingRefreshType : 0  # Random Accesss 0:none, 1:CDR, 2:IDR
- GOPSize : 1    # GOP Size (number of B slice = GOPSize-1)
- QP : 32       #Quant. Parameter (I- 51)

The command line arguments for HM10 software are:

\texttt{TAppEncoder \[-h\] \[-c \textit{config.cfg}\] \[-\textit{parameter=value}\].}

The command line arguments for JM18.2 software are:

\texttt{lencod \[-h\] \[-d \textit{defenc.cfg}\] \{-\textit{f curenc1.cfg}...\{-\textit{f curencN.cfg}\}\}
\{-\textit{p EncParam1=EncValue1}...\{-\textit{p EncParamM=EncValueM}\}\}

Reference software jasper [8] is used in the JPEG2000. The video sequence is opened in YUV player deluxe. In YUV file is saved by saving any frame in .bmp format. The .bmp file is used to compress by using JPEG2000 encoder.

\texttt{jasper \[-f \textit{Input sequence}\] \[-f \textit{Output sequence}\] \[-O rate=0.01\].}
The following test sequences are used for simulation:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ParkScience</td>
<td>1920X1080</td>
</tr>
<tr>
<td>Race Horses</td>
<td>416X240</td>
</tr>
<tr>
<td>KristenandSara</td>
<td>1280X720</td>
</tr>
<tr>
<td>PeopleonStreet</td>
<td>2560X1600</td>
</tr>
</tbody>
</table>
**Tabulation:**

**HEVC encoder:**

Test sequence: Race horses.
QP: 32
Total number of frames: 1
GOP size: 1

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>416X240</td>
<td>1929.12</td>
<td>34.89</td>
<td>2.2</td>
</tr>
</tbody>
</table>

Test sequence: ParkScence.
QP: 32
Total number of frames: 1
GOP size: 1

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1280X720</td>
<td>13761.216</td>
<td>36.77</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Test sequence: PeopleonStreet
QP: 32
Total number of frames: 1
GOP size: 1

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2560X1600</td>
<td>32818.32</td>
<td>37.9</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Test Sequence: KristenandSara
QP: 32
Total number of frames: 1
GOP size: 1

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920X1080</td>
<td>7958.88</td>
<td>35.655</td>
<td>2.98</td>
</tr>
</tbody>
</table>
**H.264 Encoder:**

Test sequence: Race horses  
QP: 32  
Total number of frames encoded: 1  

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>416X240</td>
<td>2323.68</td>
<td>34.56</td>
<td>1.82</td>
</tr>
</tbody>
</table>

Test sequence: ParkScence  
QP: 32  
Total number of frames encoded: 1  

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1280X720</td>
<td>16221.89</td>
<td>36.56</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Test sequence: PeopleonStreet  
QP: 32  
Total number of frames encoded: 1  

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2560X1600</td>
<td>41045.28</td>
<td>37.71</td>
<td>2.15</td>
</tr>
</tbody>
</table>

Test sequence: KristenandSara  
QP: 32  
Total number of frames encoded: 1  

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Bitrate (kbit/s)</th>
<th>PSNR (dB)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920X1080</td>
<td>10620</td>
<td>35.17</td>
<td>1.92</td>
</tr>
</tbody>
</table>

**JPEG2000**  
Rate = 0.01.

<table>
<thead>
<tr>
<th>Test Sequences</th>
<th>Resolution</th>
<th>PSNR in dB</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Racehorses</td>
<td>416X240</td>
<td>32.8</td>
<td>2.03</td>
</tr>
<tr>
<td>PeopleonStreet</td>
<td>2560X1600</td>
<td>39.4</td>
<td>2.21</td>
</tr>
<tr>
<td>ParkScence</td>
<td>1920X1080</td>
<td>37.8</td>
<td>1.98</td>
</tr>
<tr>
<td>KristenandSara</td>
<td>1280X720</td>
<td>35.12</td>
<td>2.79</td>
</tr>
</tbody>
</table>
Figure: 9

Comparison between HEVC & H.264 for one frame and JPEG2000

<table>
<thead>
<tr>
<th></th>
<th>Race Horses-416X240</th>
<th>KristenAndSara-1280X720</th>
<th>ParkScence-1920X1080</th>
<th>PeopleonStreet-2560X1600</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H.265</strong></td>
<td>34.89</td>
<td>35.655</td>
<td>36.77</td>
<td>37.9</td>
</tr>
<tr>
<td><strong>H.264</strong></td>
<td>34.56</td>
<td>35.17</td>
<td>36.56</td>
<td>37.71</td>
</tr>
<tr>
<td><strong>JPEG2000</strong></td>
<td>32.8</td>
<td>35.12</td>
<td>37.8</td>
<td>39.4</td>
</tr>
</tbody>
</table>

Figure: 10

Comparison between HEVC & H.264 for one frame with JPEG2000

HEVC has a higher compression ratio compared to H.264 and JPEG2000 for one frame, which can be interpreted that the encoder sequence using HEVC has less file size.
Figure: 11

Comparison between HEVC & H.264 for one frame

Conclusions:

HEVC encoder takes more time to encode one frame of the video sequence compared to encoding time by H.264 encoder for one frame. This is due to the additional blocks (figure.1) added in the encoder of HEVC, when compared with H.264 encoder block diagram (Figure.5). The bit rate of HEVC is less when compared with the bit rate of the H.264 encoder. From the figure 11 it can be seen that there is gain of 20% in the bit rate of HEVC encoder compared to H.264 encoder. The compression ratio of the encoded frame stored on file by HEVC encoder is compared with H.264 encoded file and JPEG2000 encoded file is shown in figure 10. From figure 9 it can be derived that the PSNR value is more for HEVC when compared to H.264 and as JPEG2000 has low PSNR value for low-quality video and greater than HEVC and H.264 in high quality video.
References:


