A THESIS PROPOSAL ON

RATE ALLOCATION FOR SCALABLE HIGH EFFICIENCY VIDEO CODING

UNDER THE GUIDANCE OF

DR. K.R RAO, DEPARTMENT OF ELECTRICAL ENGINEERING

SUBMISSION DATE: 26th JANUARY 2015

SUBMITTED BY

UMA SAGAR MADHUGIRI DAYANANDA

UT ARLINGTON ID: 1001002964

EMAIL ID: uma.sagar@mavs.uta.edu

DEPARTMENT OF ELECTRICAL ENGINEERING

UNIVERSITY OF TEXAS, ARLINGTON
Contents
1. Motivation ............................................................................................................................ 4
2. Overview .................................................................................................................................. 4
3. Key Technologies .................................................................................................................... 4
   3.1 High Efficiency Video Coding .......................................................................................... 4
      3.1.1 Picture Partitioning ...................................................................................................... 6
      3.1.2 Prediction ..................................................................................................................... 6
      3.1.3 Entropy Coding ............................................................................................................ 7
      3.1.4 In-loop Filtering .......................................................................................................... 7
   3.2 Scalable High Efficiency Video Coding ............................................................................. 7
      3.2.1 Features ....................................................................................................................... 7
      3.2.2 Architecture ................................................................................................................ 8
      3.2.3 Interlayer Prediction .................................................................................................. 8
   3.3 Dynamic Adaptive Streaming over HTTP (MPEG-DASH) .................................................. 9
4. Need for SHVC Rate Allocation Algorithm ........................................................................... 10
5. SHVC Rate Allocation Algorithm ......................................................................................... 11
   5.1 Inputs and Outputs ............................................................................................................ 11
   5.2 Techniques to solve rate allocation problem ..................................................................... 11
6. Tentative Work Plan ............................................................................................................... 12
7. References .............................................................................................................................. 12
List of Acronyms and Abbreviations

AVC – Advanced Video Coding
BL – Base Layer
CABAC – Context Adaptive Binary Arithmetic Coding
CTB – Coding Tree Block
CTU – Coding Tree Unit
CU – Coding Unit
DASH – Dynamic Adaptive Streaming over HTTP
EL – Enhancement Layer
FPS – Frames per second
HD – High Definition
HEVC – High Efficiency Video Coding
HLS – High Level Syntax
HTTP – Hyper Text Transfer Protocol
ILR – Inter Layer Reference
JCTVC – Joint Collaborative Team on Video Coding
Mbps – Megabits per second
MPD – Media Presentation Description
MPEG – Moving Picture Experts Group
MV – Motion Vector
OTT – Over the Top
PSNR – Peak Signal to Noise Ratio
PU – Prediction Unit
SAO – Sample Adaptive Offset
SHVC – Scalable High Efficiency Video Coding
SNR – Signal to Noise Ratio
SPIE – Society of Photo-Optical Instrumentation Engineers
TU – Transform Unit
UHD – Ultra High Definition
URL – Uniform Resource Locator
1. Motivation

The multimedia environment is growing with:

- Increasing HD -1920x1080 progressive (p) and UHD -3840x2160p video content
- State of the art video coding technologies such as High Efficiency Video Coding (HEVC) [1] and Scalable HEVC (SHVC) [3]
- Adaptive HTTP Streaming technologies such as MPEG-DASH [10]
- Heterogeneous clients such as UHD televisions and mobile devices, supporting different codecs (AVC or HEVC), resolutions (HD or UHD) and varying bandwidth characteristics

In order to efficiently deliver this high resolution video content to heterogeneous clients having varying characteristics, these video coding and streaming technologies can be harnessed.

2. Overview

Traditionally, multiple versions of the same video are stored on the servers to satisfy varying client characteristics and are delivered using simulcast coding. This leads to increased video bitrate and hence increases storage costs. However, using scalable video coding - where multiple versions of the video are embedded into different layers of the bit stream, results in bitrate savings. This bitrate savings come at a cost of reduced coding efficiency due to addition of layers, known as scalability overhead. Thus, while using scalable video coding both bitrate savings and the scalability overhead needs to be considered.

A scalable bit stream can be delivered to clients using adaptive HTTP streaming technology. Dynamic Adaptive Streaming over HTTP (DASH), also known as MPEG-DASH, is one such streaming technique that enables high quality streaming of multimedia content over Internet to be delivered through conventional HTTP web servers.

A brief review of HEVC, SHVC and MPEG-DASH technologies is presented in the next section, followed by discussion on need for rate allocation in scalable video coding with SHVC in the context of HTTP streaming. The inputs needed for rate allocation algorithm are discussed and also the techniques that can be used to solve this allocation problem are given briefly.

3. Key Technologies

3.1 High Efficiency Video Coding

High Efficiency Video Coding (HEVC) [1] is an international standard for video compression developed by a working group of ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG). The main goal of HEVC standard is to significantly improve compression performance compared to existing standards (such as H.264/AVC [29]) in the range of 50% bit rate reduction at similar visual quality [1].

HEVC uses block-based hybrid video coding techniques. Redundancies in video sequences can be categorized into spatial, temporal, statistical and perpetual redundancies. Various video coding techniques are used to remove these redundancies in HEVC video codec such as: spatial redundancy removal using intra prediction and block transforms, temporal redundancy removal using inter prediction, statistical redundancy removal using entropy coding and perpetual redundancy using quantization [30].

For encoding source video consisting of video frames using HEVC, each frame is partitioned into blocks. Each of the blocks is predicted using intra-prediction or inter-prediction. The prediction
difference is obtained by subtracting the predicted block from the original block. The resulting difference is transformed, followed by quantization of transformed coefficients. The encoded bit stream is then obtained using entropy coding. HEVC decoder inverses all these operations in encoding and decodes the bit stream.

Fig. 1. represents the block diagram of HEVC Encoder [1]:

Fig. 1. Block Diagram of HEVC Encoder (with decoder modelling in shaded light grey) [1]

Fig. 2. Partitioning of Picture in HEVC [31]
3.1.1 Picture Partitioning

In HEVC, each picture is partitioned into square-shaped Coding Tree Blocks (CTBs), where its size varies from 16x16 to 64x64 pixels. One Luma CTB and two chroma CTBs with syntax elements form the Coding Tree Unit (CTU), which is the basic processing unit in HEVC. This is similar to macro-blocks (16x16) in previous H.264/AVC standard. CTU is sub-divided into square regions called Coding Units (CUs) using quad-tree structure. CU size ranges from 8x8 to 64x64 pixels. Each CU is partitioned into Prediction Units (PUs) which is predicted using intra or inter prediction. The difference of original and prediction in each CU is transformed using one or more block transforms of size varying from 32x32 to 4x4 pixels [31].

Fig. 2. represents the picture partitioning into CTUs, CU, PU and TU [31] and Fig. 3. represents the partitioning of a video frame of KristenAndSara test sequence using Elecard HEVC Analyser [38]:

![Partitioning of KristenAndSara Test Sequence](image)

3.1.2 Prediction

Prediction in HEVC can be intra prediction or inter prediction. Intra prediction is where the blocks are predicted using the neighboring pixels reconstructed from the same frame, exploring the spacial redundancy. Inter-prediction uses temporal redundancy between the adjacent frames in order to predict the current block of frame.

Fig. 4. shows the various predictions used in HEVC [38].
3.1.3 Entropy Coding

The HEVC encoded stream consists of quantized transform coefficients, prediction information such as prediction modes and motion vectors, partitioning information and other header data. All of these data is encoded using Context Adaptive Binary Arithmetic Coding (CABAC).

3.1.4 In-loop Filtering

HEVC uses two in-loop filters such as de-blocking and Sample Adaptive Offset (SAO) filters. They are used in encoding and decoding loops after inverse quantization and before saving the picture in decoded picture buffer. De-blocking filter is applied first and it reduces the discontinuities at the prediction and transform block boundaries. SAO is applied for the output of de-blocking filter and it reduces the ringing artifacts.

3.2 Scalable High Efficiency Video Coding

Scalable High Efficiency Video Coding (SHVC) standard is the scalable extension of High Efficiency Video Coding (HEVC/H.265). With SHVC, basic version of the video is coded as base layer and improved versions of video in terms of spatial resolution, temporal frame rate or quality is coded as enhancement layers. The design philosophy of SHVC standard is to achieve high scalable coding efficiency using a system architecture that requires only high level syntax (HLS) changes to slice header level and above [26].

3.2.1 Features

SHVC supports resolution up to 4K [40] and 8K [41]. It reuses the coding tools present in HEVC along with interlayer prediction to provide gain in coding efficiency.

SHVC provides the following scalability features:

- Temporal scalability
- Spatial scalability
- Coarse grain SNR scalability
- Hybrid codec scalability
- Bit depth scalability
- Interlaced-to-progressive scalability
- Colour gamut scalability
- Combination of these scalabilities

Fig. 5. represents various scalability options in SHVC:
Fig. 5. Scalability Options in SHVC

The HEVC design provides temporal scalability when hierarchical temporal prediction is used. The other scalability features in SHVC are enabled using the layered approach.

UHD video content is gaining importance, but it will not completely replace HD contents because of increased data rates and backward compatibility with legacy devices [11]. In particular, backward compatibility with legacy devices can be supported using SHVC. UHD contents can be encoded as the enhancement layer (EL) of HD contents, so legacy devices capable of decoding HD contents can be used continuously, while new devices can decode both UHD and HD contents. This is supported by hybrid codec feature of SHVC where the base layer can be:

- HEVC codec or
- Non-HEVC codec such as H.264/AVC codec [29]

3.2.2 Architecture

SHVC encodes the original input video into L layers. The first layer represents the base quality of video, and decoding more layers allows to further enhance the video. It uses a multi-layered decoding structure. To decode Lth layer, all intermediate layers from l = 1 to L-1 have to be fully decoded to perform inter-layer predictions.

For spatial coding with L-layers, the SHVC encoder consists of ‘L’ HEVC encoders, one for each layer. The Base Layer (BL) HEVC or H.264/AVC encoder (l=1) encodes the down-sampled version of the original video and feeds the HEVC encoder corresponding to next Enhancement Layer (EL, l=2) with the decoded picture and its motion vectors (MVs). The Lth HEVC encoder encodes the original video using the up-sampled picture from the lower layer HEVC encoder (l=L-1) and its up-scaled MVs as an additional reference picture for interlayer predictions. The output of the L encoders are multiplexed to form a conforming SHVC bit-stream. Fig. 6. represents the architecture of SHVC encoder and decoder in the context of two layer spatial scalability.

3.2.3 Interlayer Prediction

The inter-layer prediction is a powerful tool whereby the enhancement layer is predicted from a base layer. It uses redundancies between layers to improve coding efficiency. In the context of spatial scalability, the base layer picture is up-sampled and used as interlayer reference (ILR) picture along with temporal references for predicting enhancement layer.

Fig. 7. represents the interlayer prediction used in SHVC [28].
3.3 Dynamic Adaptive Streaming over HTTP (MPEG-DASH)

In HTTP streaming, multimedia file is partitioned into segments and delivered to client over HTTP. Different resolutions or bit rates for each representation will be available. MPEG-DASH [10] is an international standard developed by MPEG for adaptive bit-rate streaming. It enables high quality streaming of media content over the Internet delivered from conventional HTTP web servers. The content exists on the server in two parts: Media Presentation Description (MPD), which describes a manifest of the available content, its various alternatives, their URL addresses, and other characteristics; and segments, which contain the actual multimedia bit streams in the form of chunks, in single or multiple files. To play the content, the DASH client first obtains the MPD.
By parsing the MPD, the DASH client gets information about the program timing, media-content availability, media types, resolutions, minimum and maximum bandwidths, and the existence of various encoded alternatives of multimedia components, media-component locations on the network, and other content characteristics. Using this information, the DASH client selects the appropriate encoded alternative and starts streaming the content by fetching the segments [10]. Fig. 8 represents a sample streaming scenario between HTTP server and DASH client [10].

Fig. 8. Streaming Scenario between HTTP Server and DASH Client [10]

4. Need for SHVC Rate Allocation Algorithm

In order to effectively satisfy varying client bandwidth characteristics with HTTP video streaming, an optimal bitrate allocation algorithm is necessary to allocate bitrate to layers of SHVC. Heuristic bitrate allocation is a tedious process and also error prone. Hence, there is a need for an optimal rate allocation algorithm that can decide the number of layers in SHVC and also bitrate of each layer depending on the client distribution.

Fig. 9 represents an end to end SHVC streaming system. Here, the original video is encoded into layers having different versions of the same video. The encoded bit stream can be stored on the server and delivered to client as chunks of video using adaptive bitrate HTTP streaming. The client having a bitrate requirement of 1 Mbps will be satisfied as well the clients with 5 Mbps and 15 Mbps, using a single bit stream. As a result of this scalable video coding, bitrate savings can be obtained compared to simulcast coding, but at the cost of scalability overhead due to addition of layers. Hence, to have an effective bitrate allocation algorithm, the scalability overhead should also be considered.

Fig. 9. End to End System representing HTTP Streaming of SHVC Encoded Video
The SHVC rate allocation algorithm is needed particularly in the case of UHD video delivery via HTTP, where the UHD video content can be encoded as enhancement layer of the HD video content. In this thesis, rate allocation for SHVC with spacial scalability can be considered.

5. SHVC Rate Allocation Algorithm

5.1 Inputs and Outputs

The inputs and outputs of the SHVC rate allocation algorithm [21] are represented in Fig. 10.

![Fig. 10. Inputs and Outputs of SHVC Rate Allocation Algorithm [21]](image)

Inputs:

- **Scalability Overhead Function** – Takes into account the overhead due to reduced coding efficiency and additional codec protocols.
- **Client Distribution** – Represents distribution of clients in the real time HTTP video streaming scenario having varying bandwidth characteristics.
- **Client Utility Function** – Represents the utility function of the given client distribution that best satisfies the client according to given utility such as PSNR or received bandwidth.

Outputs:

- **Number of Layers** – Represents optimum number of layers in SHVC.
- **Bitrate of each layer** – Represents optimal bitrate of each layer in SHVC to satisfy a given client distribution.

5.2 Techniques to solve rate allocation problem

Once the rate allocation problem is defined, various techniques can be used to solve this problem. A survey of available literature [17-25] is carried out and various approaches are identified, that can be used to solve this problem such as:

- Dynamic programming: The literature in [21] defines a rate allocation problem for multi-version and multi-layer bit streams including scalable video coding. A dynamic programming based approach is used to solve this problem.
- Game theory: The literature in [22-25] describes game theory approach to solve rate control and allocation algorithms. These approaches can be adapted to solve SHVC rate allocation problem.

Further exploration of additional techniques needs to be done to solve the rate allocation problem.

In this thesis, a SHVC rate allocation problem can be formulated and an approach to solve this problem has to be selected. An optimal rate allocation algorithm should be designed and evaluated.
6. Tentative Work Plan

The thesis can be carried out in various steps and a tentative plan is given below:

- Formulation of SHVC rate allocation problem – 20th February, 2015
- Selection of approach to solve the problem – 6th March, 2015
- Design an algorithm to solve the problem – 10th May, 2015
- Evaluation and results – 14th June 2015

7. References


[34] SHVC Reference software – SHM6.1 available online: https://hevc.hhi.fraunhofer.de/svn/svn_SHVCSoftware/

[35] HEVC Reference Software – HM 15.0 available online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/


