Bandwidth Optimization For Real Time Video Streaming

Submitted By

Sarthak S. Trivedi

14MCEN28

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2016

Bandwidth Optimization For Real Time Video Streaming

Major Project

Submitted in partial fulfillment of the requirements

For the degree of

Master of Technology in Computer Science and Engineering(Networking Technologies)

Submitted By

Sarthak S. Trivedi

(14mcen28)

Guided By

Dr. Priyanka Sharma

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2016

Certificate

This is to certify that the Major Project Report entitled "**Bandwidth Optimization For Real Time Video Streaming**" submitted by **Sarthak S Trivedi (Roll No: 14MCEN28)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Networking Technologies(CSE) of Institute of Technology, Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this major project, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Dr Priyanka Sharma
Guide & Professor,
Coordinator M.Tech - CSE,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Gaurang Raval
Associate Professor,
Coordinator M.Tech-CSE(NT),
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad

Dr. Sanjay Garg
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr P N Tekwani
Director,
Institute of Technology,
Nirma University, Ahmedabad

Statement of Originality

I, **Sarthak S Trivedi**, Roll. No. **14MCEN28**, give undertaking that the Major Project entitled "**Bandwidth Optimization For Real Time Video Streaming**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering (Networking Technologies)** of Nirma University, Ahmedabad, is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student
Date:
Place:

Endorsed by
Dr Priyanka Sharma
(Signature of Guide)

Acknowledgements

It gives me immense pleasure in expressing thanks and profound gratitude to **Dr Priyanka Sharma**, Professor, CSE Department, Institute of Technology, Nirma University, Ahmedabad for her valuable guidance and continual encouragement throughout this work. The appreciation and continual support she has imparted has been a great motivation to me in reaching a higher goal. Her guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

It gives me an immense pleasure to thank **Dr. Sanjay Garg**, Hon'ble Head of Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr P N Tekwani**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout course of this work.

I would also thank the Institution, all faculty members of Computer Engineering Department, Nirma University, and Ahmedabad for their special attention and suggestions towards the project work.

The blessings of God and family members make the way for completion of Project. I am very much grateful to them.

- Sarthak S Trivedi

14MCEN28

Abstract

There is an emerged demand to increase efficiency for various video codecs for development in network infrastructure, storage capacity, Vod, 3D-TV, Quadcopter, smartphones, HD-TV broadcasting, video conference etc. Besides modern ways for transmission using the internet, there are wide range connection qualities which are the result of resource sharing mechanism. Due to rise in video streaming in smartphone and tablets along with the difference in their computational capability has given rise for a scalable extension. HEVC is recently introduced Video Coding standard; it is an advance version of H.264/Advance Video Coding which is currently used in various applications to reduce the bandwidth by half with same video quality. This paper describes the existing features, techniques, related work done on the various modes of prediction and few approaches are discussed.

Keywords : HEVC,inter-mode,intra-mode,prediction,video,scalability,H.265/AVC.

Abbreviations

AVC : Advanced Video Coding

ABR : Adaptive Bit Rate

AMVP : Advanced motion vector prediction

B: : Bi-directionally Predicted Frame

BD-PSNR : Bjontegaard metric calculation

BL : Base Layer

CABAC : Context Adaptive Binary Arithmetic Coding

CTB : Coding Tree Block

CTU : Coding Tree Unit

CU : Coding Unit

CAVLC : Context Adaptive Variable Length Coding

CIF : Common Intermediate Format

CUDA : Compute Unified Device Architecture

DCT : Discrete Cosine Transforms

DASH : Dynamic Adaptive Streaming over HTTP

EL : Enhancement Layer

FPGA : Field Programmable Gate Array

FPS : Frames per second

GPU: Graphics Processing Unit

HM : HEVC Model

HEVC : High Efficiency Video Coding

HD : High Definition

I : Intra Frame

IEC : International Electrotechnical Commission

ISO : International Organization for standardization

ITU : International Telecommunication Union

Abbreviations

JCT-VC	:	Joint Collaborative Team on Video Coding
MC	:	Motion Compensation
ME	:	Motion Estimation
MPEG	:	Moving Picture Experts Group
MV:	:	Motion Vector
PSNR	:	Peak Signal to Noise Ratio
P:	:	Predicted Frame
QP:	:	Quantization Parameter
QCIF:	:	Quarter Common Intermediate Format
PSNR:	:	Peak Signal To Noise Ratio
PU:	:	Prediction Unit
RD:	:	Rate Distortion
SAO:	:	Sample Adaptive Offset
SAD:	:	Sum of Absolute Differences
SATD:	:	Sum of Absolute Transformed Differences (SATD)
SDK	:	Software Development Kit
SHVC	:	Scalable HEVC
SIMD	:	Single Instruction Multiple Data
SSIM	:	Structural Similarity
SVC	:	Scalable Video Coding
SNR	:	Signal to Noise Ratio
UHD	:	Ultra High Definition
TU	:	Transform Unit
URQ	:	Uniform Reconstruction Quantization
VCEG	:	Video Coding Experts Group
VOD	:	Video On Demand

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

There is the growing demand of videos in this technological world,as number of devices that works on internet becomes easy to use.This need and demand gives rise to high demand of bandwidth.Various video codecs are been developed to optimize the codecs over the various previous versions.There are various factor that are been coming onto factor but that does not affect the video quality.Better algorithms are been developed for the less bandwidth required to transmit the video.

Internet is been conventionally been used for various different purposes and many data applications like chatting,web surfing, multimedia technology. This technology of multimedia was from quite along time but it came into demand after some time , and nowadays it has become an most important part of the web . There are various application ranging from field of video lectures to catalogues and gaming videos to video . In today's world the main source of extraction is internet and multimedia technology gaining the attention from developers researchers and other variety of users.[1]Transmission of various video in various format with different display screens have been a necessity for every internet user. This transmission of video requires much bandwidth.

Compression is necessary as it has to transmitted over various media wireless or wired networks.Due to high demand of devices and their capability the capturing quality has increased and there quality has also increased . There are various traditional protocols

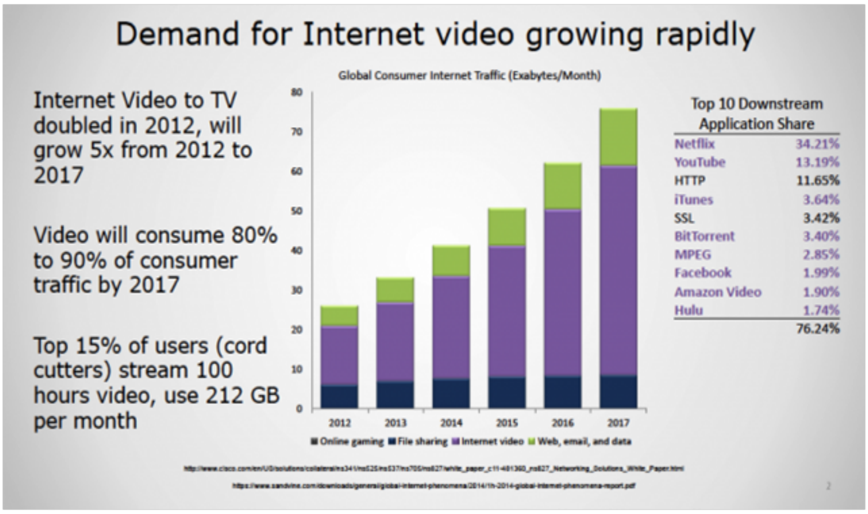


Figure 1.1: Demand of video

used for such high volume multimedia.hence we need good protocols .There are many researches which are still developing various protocols working on real network. Video with high element range more exile shading are more noteworthy features and ability for HEVC. Besides modern ways for transmission using internet there are wide range connection qualities which are result of resource sharing mechanism.[1]

Some factors that are taken into consideration while designing a video codec are:

- Encoding Time.
- Video Quality (Measured by using objective measurement metrics such as PSNR, SSIM, BDRATE etc).
- File size of the encoded video (More the file size, better will be the video quality.)
- Bandwidth requirement over the network.
- Quality of video watched by the user.
- Storage capacity of any server that stores and transmits the encoded video.
- Storage capacity of device that records and stores the compressed video.

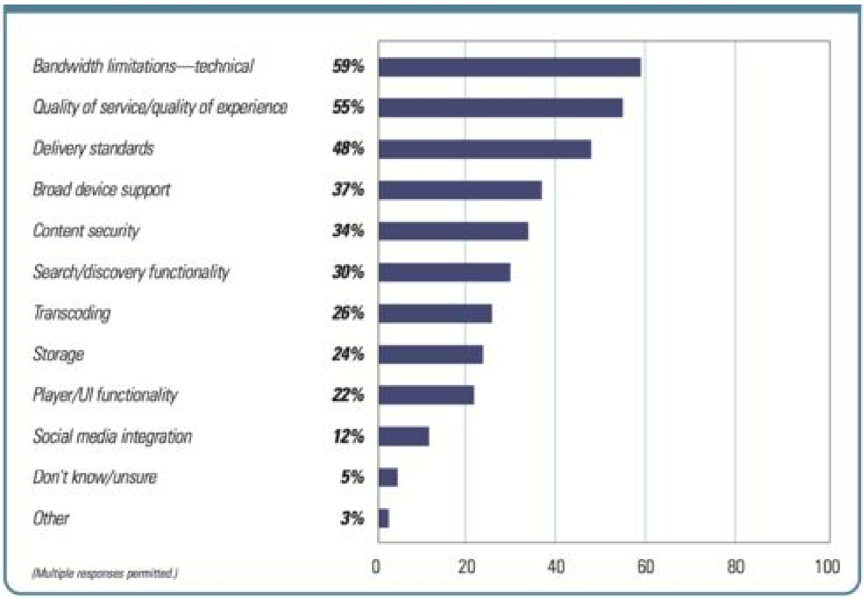


Figure 1.2: Technical challenges for video

1.2 Demand of internet video going rapidly

Video is transportation of any moving image,which is usually accompanied with sound. Due to increases n demand of various mobile devices like smart-phones, tablets , computers , palmtops , notebooks with different screen sizes and computational capability.This technology of multimedia was from quite along time but it came into demand after some time , and nowadays it has become an most important part of the web . So there arise a need of scalable extension. Total cost saving, easy of complexity, scalability, improved productivity,are the key drivers for this technology.

Internet is been conventionally been used for various different purposes and many data applications like chatting,web surfing, multimedia technology. This technology of multimedia was from quite along time but it came into demand after some time , and nowadays it has become an most important part of the web . There are various application ranging from field of video lectures to catalogues and gaming videos to video . In today's world the main source of extraction is internet and multimedia technology gaining the attention from developers researchers and other variety of users.[1]

Transmission of various video in various format with different display screens have been a necessity for every internet user. This transmission of video requires much bandwidth. Compression is necessary as it has to transmitted over various media wireless or wired networks.Due to high demand of devices and their capability the capturing quality has increased and there quality has also increased . There are various traditional protocols used for such high volume multimedia.hence we need gud protocols .There are many researches which are still developing various protocols working on real network. Video with high element range more exible shading are more noteworthy features and ability for HEVC. Besides modern ways for transmission using internet there are wide range connection qualites which are result of resource sharing mechanism.[1]

1.3 Application of Video

There are wide ranging applications of Video. It can vary significantly in application requirements.

- Broadcast TV: Cable TV over wireless network, satellite, cables, terrestrial, HDTV etc
- Live event video streaming: Corporate webcasts, university events on TV
- Surveillance video
- Interactive video conferencing
- On-demand video
- First person view quadcopter
- Multimedia messaging services (MMS) over ISDN, DSL, ethernet, LAN, wireless and mobile networks, etc.

1.4 Objective of the Work

- Optimization bandwidth utilization.[1]
- Optimization of parameters for live video streaming.
- Coding efficiency.
- Data loss resilience.
- Bitrates reduction: same subjective visual quality.

Chapter 2

Literature Survey

2.1 CODECS

- JPEG2000
- MPEG-2
- MJPEG
- MPEG-4
- H.264/AVC
- HEVC

2.2 HEVC

Robustness,flexibility and other various frameworks are been used by HEVC for getting associated information of video content and can be used for various different applications.Network abstraction layer(NAL) is a sequence of data units in HEVC . High level information belongs to the one or more slices or pictures of bitstreams[2].

2.2.1 Simplified block diagram of HEVC

Video is divided into blocks which are called coding tree units (CTUs),which provides macroblocks.The coding unit (CU) characterizes a district having the same prediction mode (intra, inter or skip) and it is spoken to by the leaf hub of a quadtree structure.

The prediction unit (PU) characterizes an area having the same expectation data. The transform unit (TU), specified by another quadtree, defines a region sharing the same transformation and quantization.The best intra mode among a total of 35 modes (Planar, DC and 33 angular directions) is selected and coded. Mode dependent context sample smoothing is applied to increase prediction efficiency and the three most probable modes (MPM) are used to increase symbol coding efficiency.[3, 4, 5]

2.3 Picture partitioning

2.3.1 CTU partitioning

Group of coding tree units(CTUs) consist of the partitioned images.The CTU idea is extensively closely resembling that of the macroblock in past models, for example AVC.For a image that has exhibits three examples, a CTU comprises of a NxN square of luma tests together with two relating pieces of chroma tests. The maximum allowed size of the luma hinder in a CTU is determined to be 64x64 in Main profile.[6]

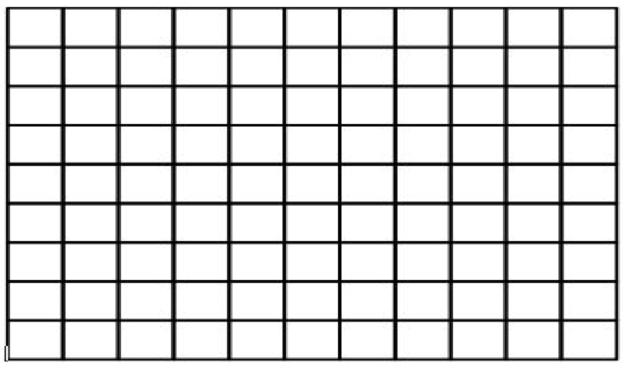


Figure 2.1: Example of a picture divided into CTUs

2.3.2 Slice and tile structures

Entire picture or region of picture is called slice.Data structure is sliced which can be decoded independently from other slices of the same picture,consisting of sequence of slice segments starting with all subsequent dependent slice segments that precede the next independent slice segment within same access unit.Sequence of CTUs make slice segment,this slice segments are independent as they are not inferred from the values for

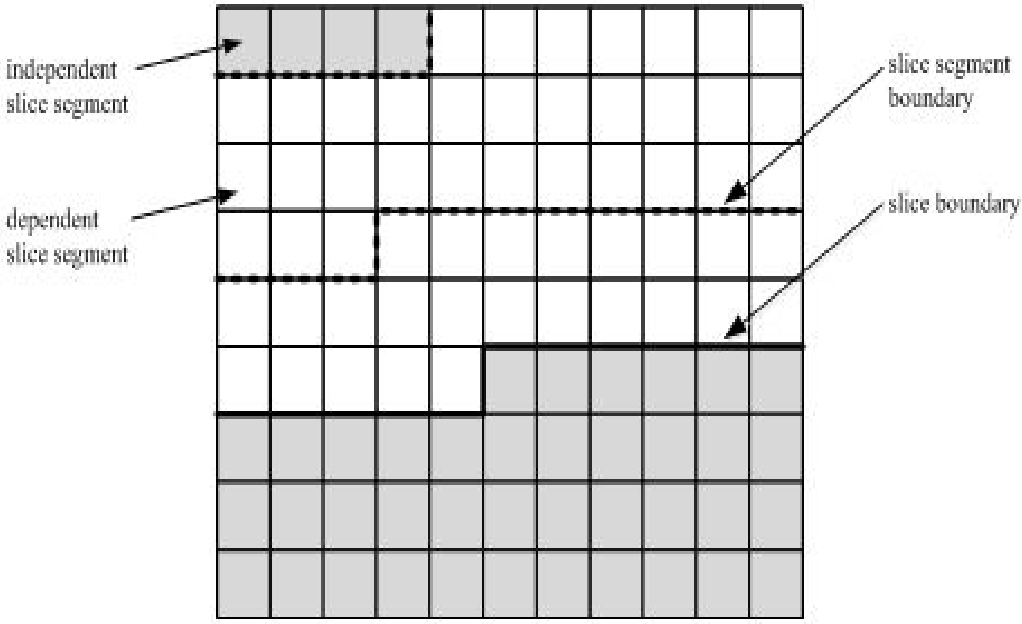


Figure 2.2: Example of slices and slice segments

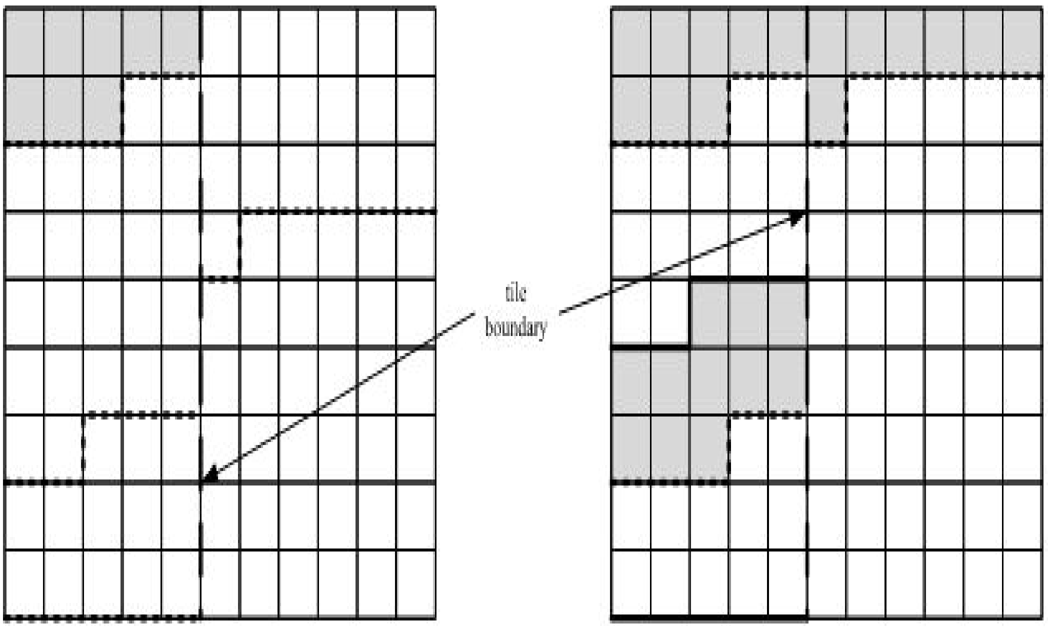


Figure 2.3: Examples of tiles and slices

Raster scan of coding tree a tile is a rectangular region containing an integer number.A specific sequential ordering of coding tree blocks partitioning a picture in which the coding tree blocks are ordered consecutively in coding tree block raster scan in a tile, whereas tiles in a picture are ordered consecutively in a raster scan of the tiles of the picture.[7, 4, 8]

2.3.3 Prediction unit (PU) structure

PU sizes from 64x64 down to 4x4 samples. However, in order to reduce the memory bandwidth of motion compensation, the 4x4 block size is not allowed for an inter-coded PU.[9]

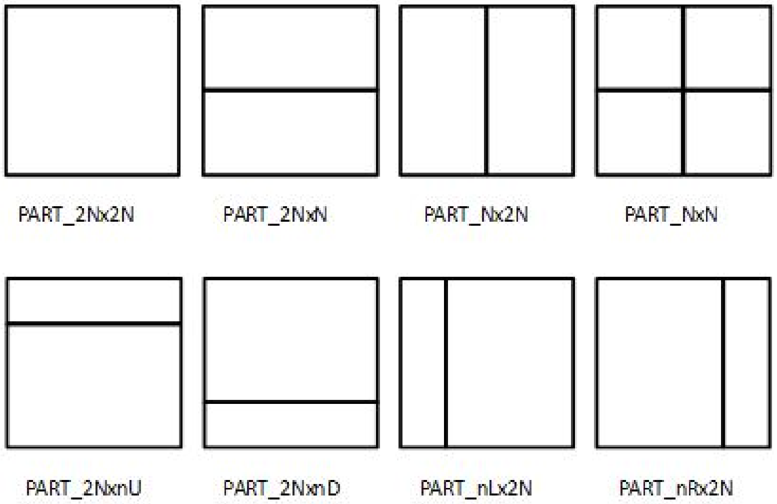


Figure 2.4: 8 partition modes for inter PU

2.3.4 Transform unit (TU)

The transform unit (TU) is a square region, defined by quadtree partitioning of the CU, which shares the same transform and quantization processes. The quadtree structure of multiple TUs within a CU.[1]

The TU shape is always square and it may take a size from 32x32 down to 4x4 samples. The maximum quadtree depth is adjustable and is specified in the slice header syntax.For an inter CU, the TU can be larger than PU, i.e. the TU may contain PU boundaries. However, the TU cannot cross PU boundaries for an intra CU.

2.4 Intra Prediction

2.4.1 Prediction modes

The decoded boundary samples of adjacent blocks are used as reference data for spatial prediction in regions where inter picture prediction is not performed. All TUs within

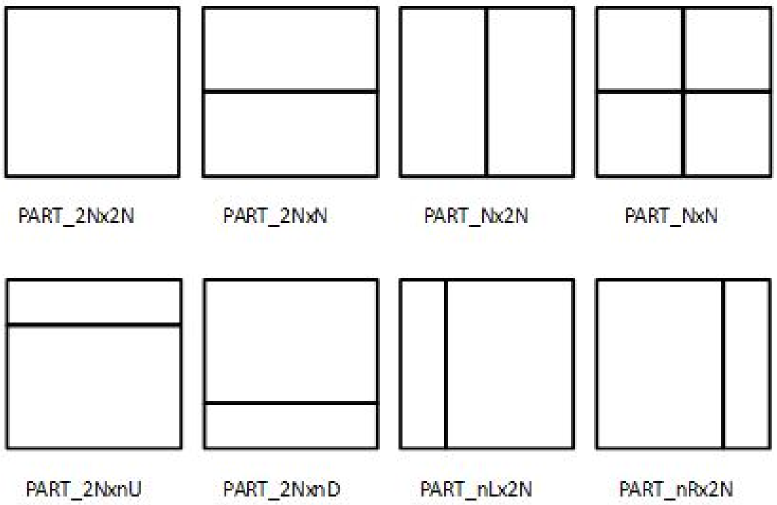


Figure 2.5: Example of transform tree structure within CU

a PU use the same associated intra prediction mode for the luma component and the chroma components. The encoder selects the best luma intra prediction mode of each PU from 35 options: 33 directional prediction modes, a DC mode and a Planar mode. The 33 possible intra prediction directions are illustrated in Figure below.[10, 11]

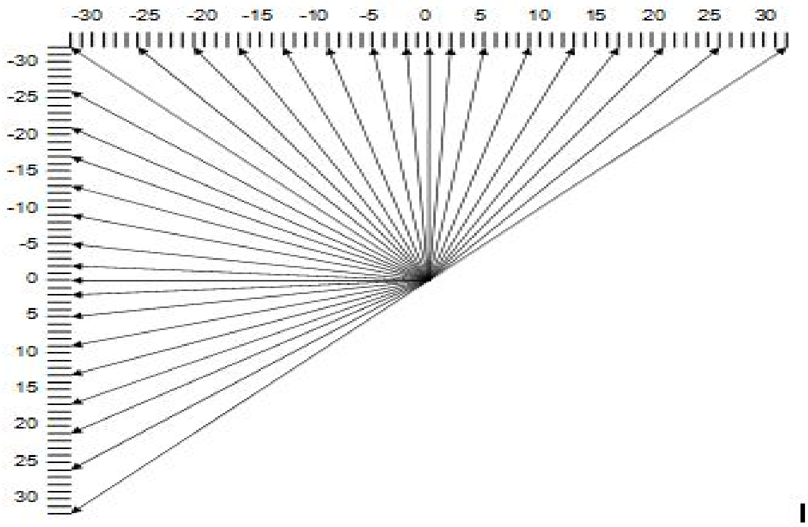


Figure 2.6: The 33 intra prediction directions

The mapping between the intra prediction direction and the intra prediction mode number is specified in Figure below.

Mapping between intra prediction direction and intra prediction mode for chroma.When the intra prediction mode number for the chroma component is 4, the intra prediction direction for the luma component is used for the intra prediction sample generation for the

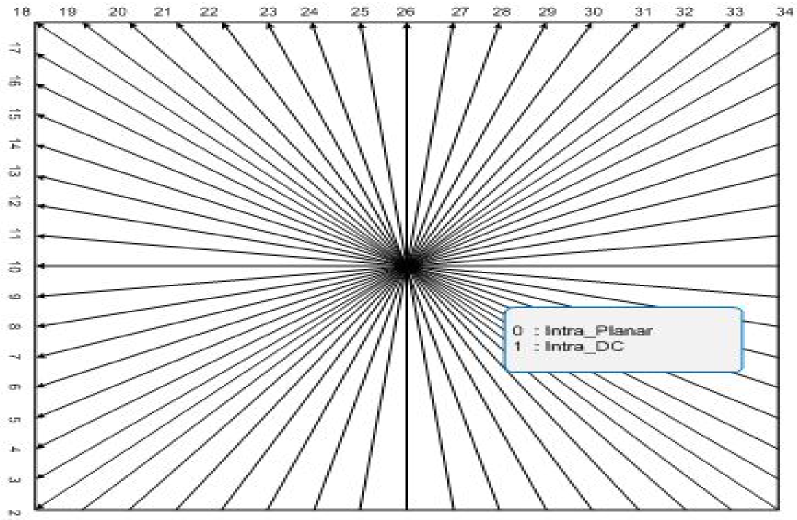


Figure 2.7: Mapping between intra prediction direction and intra prediction mode

chroma component. When the intra prediction mode number for the chroma component is not 4 and it is identical to the intra prediction mode number for the luma component, the intra prediction direction of 34 is used for the intra prediction sample generation for the chroma component.[12]

2.5 Inter Prediction

2.5.1 Prediction modes

Each inter coded PU has a set of motion parameters consisting of motion vector, reference picture index, reference picture list usage flag to be used for inter prediction sample generation, signalled in an explicit or implicit manner. When a CU is coded with skip mode, the CU is represented as one PU that has no significant transform coefficients and motion vectors, reference picture index and reference picture list usage flag obtained by merge mode. The merge mode is to find the neighbouring inter coded PU such that its motion parameters (motion vector, reference picture index, and reference picture list usage flag) can be inferred as the ones for the current PU.[13] The encoder can select the best inferred motion parameters from multiple candidates formed by spatially neighbouring PUs and temporally neighbouring PUs, and transmit the corresponding index indicating the chosen candidate. The merge mode can be applied to any inter coded PU, not only for skip mode. In any inter coded PUs, the encoder can use merge mode or explicit transmission of motion parameters, where motion vector, corresponding reference picture index for each reference picture list and reference picture list usage flag are signalled explicitly per each PU. For inter coded PU, significant transform coefficients are sent to decoder.[14, 15]

Two types of merge candidates are considered in merge mode: spatial merge candidate and temporal merge candidate. For spatial merge candidate derivation, a maximum of four merge candidates are selected among candidates that are located in five different positions. In the process of candidate selection, duplicated candidates having the same motion parameters as the previous candidate in the processing order are removed from the candidate list. Also, candidates inside the same merge estimation region (MER) are not considered, in order to help parallel merge processing. Redundant partition shape is avoided in order to not emulate a virtual 2Nx2N partition.[16, 17] For temporal merge candidate derivation, a maximum of one merge candidate is selected among two candidates. Since constant number of candidates for each PU is assumed at decoder, additional candidates are generated when the number of candidates does not reach to maximum number of merge candidate (MaxNumMergeCand) which is signalled in slice header. For B slices, combined bi-predictive candidates are generated utilizing the candidates from the list of spatio-temporal candidates. For both P- and B-slices, zero merge candidates are added at the end of the list. Between each generation step, the derivation process is stopped if the number of candidates reaches to MaxNumMergeCand. In the current common test condition, MaxNumMergeCand is set equal to five. Since the number of candidates is constant, index of best merge candidate is encoded using truncated unary binarization (TU).

2.5.2 Spatial merge candidates

In the derivation of spatial merge candidates, a maximum of four merge candidates are selected among candidates that are located in positions as depicted in Figure. The order of derivation is A1 B1 B0 A0 (B2). Position B2 is considered only when any PU of position A1, B1, B0, A0 is not available or is intra coded.[6]

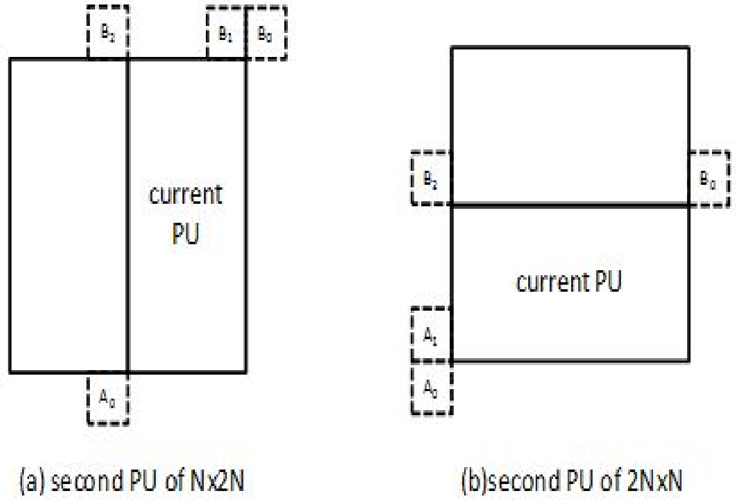
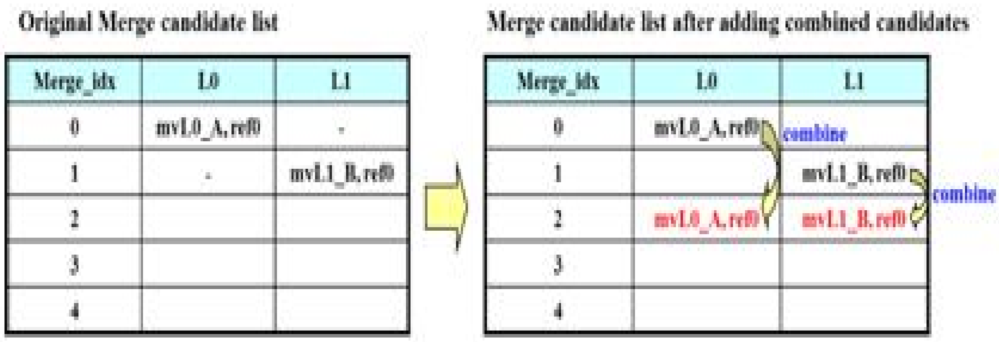


Figure 2.8: Positions for the second PU of Nx2N and 2NxN partitions

2.5.3 Generated merge candidates

Besides spatio-temporal merge candidates, there are two additional types of merge candidates: combined bi-predictive merge candidate and zero merge candidate. Combined bi-predictive merge candidates are generated by utilizing spatio-temporal merge candidates. Combined bi-predictive merge candidate is used for B-Slice only. For example, two candidates in the original merge candidate list, which have mvL0 and refIdxL0 or mvL1 and refIdxL1, are used to create a combined bi-predictive merge candidate as illustrated in Figure.



Original Merge candidate list

Merge_idx	L0	L1
0	mvL0_A, ref0	-
1	-	mvL1_B, ref0
2		
3		
4		

Merge candidate list after adding combined candidates

Merge_idx	L0	L1
0	mvL0_A, ref0	
1		mvL1_B, ref0
2	mvL0_A, ref0	mvL1_B, ref0
3		
4		

Figure 2.9: Example of combined bi-predictive merge candidate [9]

2.5.4 Motion vector prediction

Motion vector prediction exploits spatio-temporal correlation of motion vector with neighbouring PUs, which is used for explicit transmission of motion parameters. It constructs a motion vector candidate list by firstly checking availability of left, above temporally neighbouring PU positions, removing redundant candidates and adding zero vector to make the candidate list to be constant length. Then, the encoder can select the best predictor from the candidate list and transmit the corresponding index indicating the chosen candidate. Similarly with merge index signalling, the index of the best motion vector candidate is encoded using truncated unary, as maximum number is equal to 2. In the following sections, details about derivation process of motion vector prediction candidate are provided.[18]

2.5.5 Spatial motion vector candidates

In the derivation of spatial motion vector candidates, a maximum of two candidates are considered among five potential candidates, which are derived from PUs located in positions as depicted in Figure. The candidate positions of motion vector prediction are the same as those of motion merge. The order of derivation for left side of the current PU is set as A0 A1 scaled A0 scaled A1. The order of derivation for above side of the current PU is set as B0 B1 B2 scaled B0 scaled B1 scaled B2. For each side, there are four cases which can be used as motion vector candidate. Although two cases are not required to do spatial scaling, the other two cases are. The four different cases are summarized as follows. No spatial scaling (1) Same reference picture list, and same reference picture index (same POC) (2) Different reference picture list, but same reference picture (same POC) Spatial scaling (3) Same reference picture list, but different reference picture (different POC) (4) Different reference picture list, and different reference picture (different POC) No spatial scaling cases are checked first and spatial scaling cases are checked sequentially. Spatial scaling is considered when POC is different between the reference picture of the neighbouring PU and that of the current PU regardless of reference picture list. If all PUs of left candidates are not available or are intra coded, scaling for the above motion vector is allowed to help parallel derivation of left and above MV candidates. Otherwise, spatial scaling is not allowed for the above motion vector.

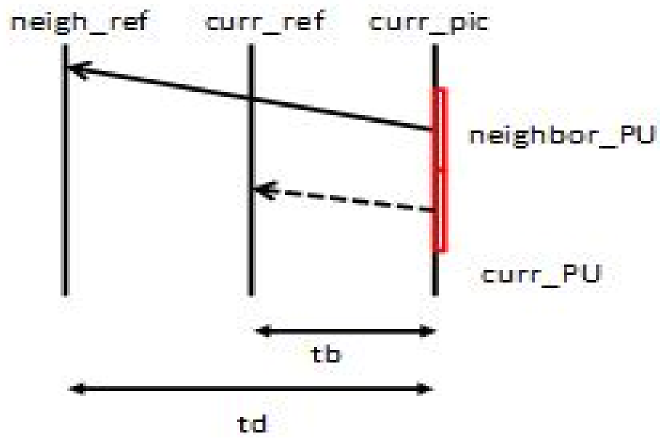


Figure 2.10: Illustration of motion vector scaling for spatial motion vector candidate

In a spatial scaling process, the motion vector of the neighbouring PU is scaled in a similar manner as for temporal scaling, as depicted as Figure. The main difference is that the reference picture list and index of current PU is given as input; the actual scaling process is the same as that of temporal scaling.

2.5.6 Temporal motion vector candidates

Apart for the reference picture index derivation, all processes for the derivation of temporal merge candidates are the same as for the derivation of spatial motion vector candidates. The reference picture index is signalled to the decoder.

2.6 Weighted Prediction

A weighted prediction (WP) tool is included in the HM encoding software. This WP tool corresponds to the equivalent tool for AVC and it is intended to improve the performance of inter prediction when the source material is subject to illumination variations, e.g. when using fading or cross-fading. It should be noted that WP is not enabled in the HM common test conditions. The principle of WP is to replace the inter prediction signal P by a linear weighted prediction signal P (with weight w and offset o):

The applicable weights and offsets are calculated by the encoder, using some mechanism that is not defined by the HEVC specification, and are conveyed within the bitstream. L0 and L1 suffixes define List0 and List1 of the reference pictures list, respectively. Bit depth is maintained to 14 bit accuracy before averaging the prediction signals, as for interpolation filters. In the case of biprediction with at least one reference picture available in each list L0 and L1, the following formula applies to the explicit signalling of weighted prediction parameters relating to the a corresponding formula applies to the chroma channel and to the case of uniprediction.

2.7 Loop Filtering

2.7.1 Overview of Loop filtering

HEVC includes two processing stages in the in-loop filter: a deblocking filter and then a Sample Adaptive Offset (SAO) filter. The deblocking filter aims to reduce the visibility of blocking artefacts and is applied only to samples located at block boundaries. The SAO filter aims to improve the accuracy of the reconstruction of the original signal amplitudes and is applied adaptively to all samples, by conditionally adding an offset value to each sample based on values in look-up tables defined by the encoder.[19]

2.7.2 Deblocking filter

A deblocking filter process is performed for each CU in the same order as the decoding process. First vertical edges are filtered (horizontal filtering) then horizontal edges are filtered (vertical filtering). Filtering is applied to 8x8 block boundaries which are determined to be filtered, both for luma and chroma components. 4x4 block boundaries are not processed in order to reduce the complexity (unlike the earlier AVC standard). Figure illustrates the overall flow of deblocking filter processes. A boundary can have three filtering status values: no filtering, weak filtering and strong filtering. Each filtering decision is based on boundary strength, Bs, and threshold values, and tC.

2.7.3 Boundary decision

Two kinds of boundaries are involved in the deblocking filter process: TU boundaries and PU boundaries. CU boundaries are also considered, since CU boundaries are necessarily also TU and PU boundaries. When PU shape is 2NxN (N ¿ 4) and RQT depth is equal

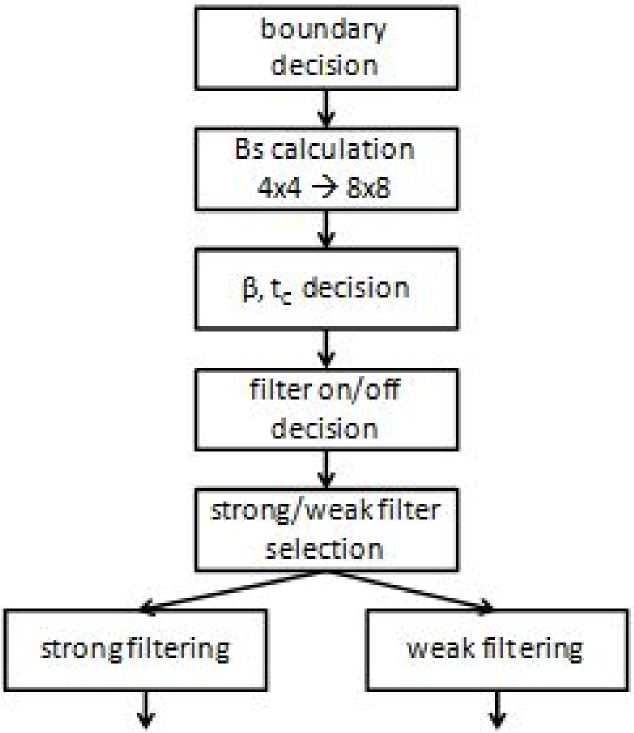


Figure 2.11: Overall processing flow of deblocking filter process

to 1, TU boundaries at 8x8 block grid and PU boundaries between each PU inside the CU are also involved in the filtering.

2.8 Entropy Coding

A single entropy coding scheme is used in all configurations of HEVC: Context Adaptive Binary Arithmetic Coding (CABAC). There is no equivalent to the alternative second entropy coding scheme that was included in the earlier AVC standard.[10, 11] The core of the CABAC algorithm for HEVC is essentially the same as the CABAC algorithm for used AVC, but with some changes in the details. As a result, substantially fewer contexts are used in HEVC compared to the AVC standard, despite slightly improving the compression performance and throughput speed.

2.9 Challenges

- Challenges for video streaming

- Latency
- Image quality
- Storage requirement
- Bandwidth consumption
- Distributed denial of Service
- Poising the network
- Fairness in Sharing
- Blocking of peer to peer traffic

Chapter 3

EXISTING SEARCH PATTERNS

3.1 BLOCK-BASED MOTION ESTIMATION

All CODEC standards uses the block based motion estimation and compensation techniques,where each frame is divided into equal-size.In reference frame are associated with the search region of each source block.The main objective is to find a candidate block in the search region which best matches to the source block.this relative distance between the source and candidate blocks are motion vectors.

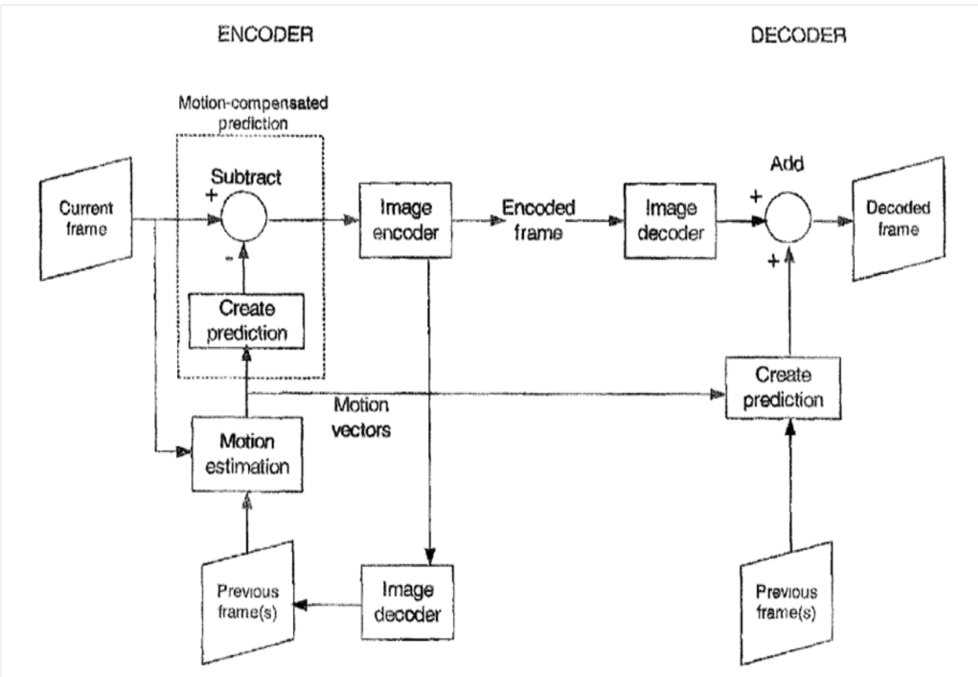


Figure 3.1: Block diagram of motion estimation and compensation

3.2 Motion Estimation Algorithm

Each current frame is divided into equal-size blocks. The reference frame of these source blocks is associated with search region. A candidate block in the search region best matched to the source is the objective of the block-matching. Motion vectors are the relative distance between a source block and its candidate.[12] Various types of algorithms are Hexagonal Search Algorithm, Diamond search, full Search Algorithm.

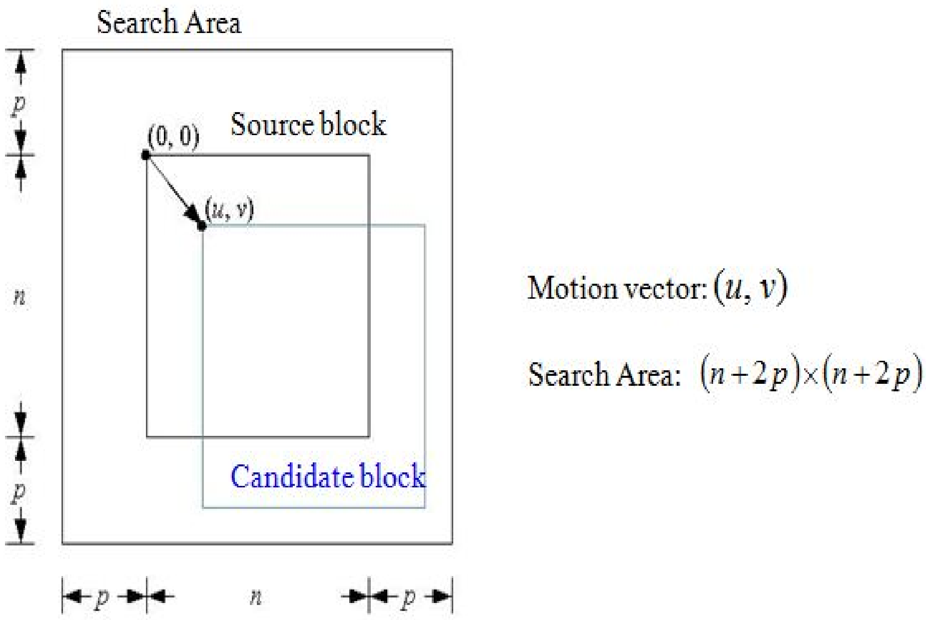


Figure 3.2: Block matching scenario

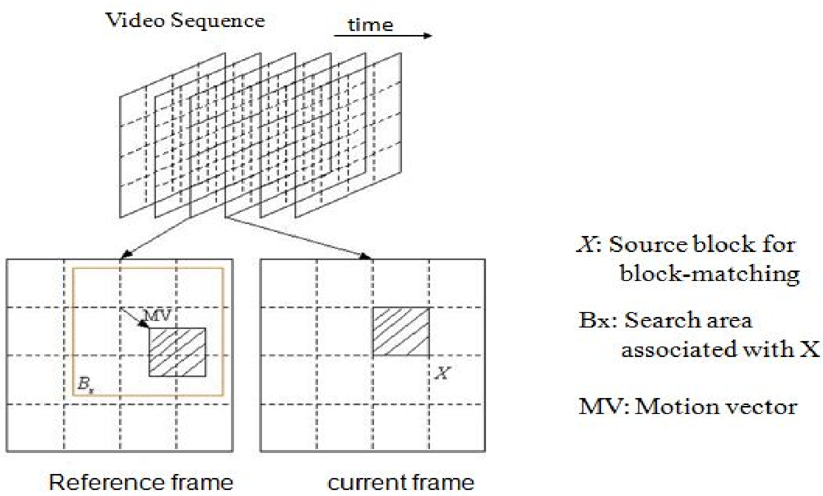


Figure 3.3: Block matching scenario

3.2.1 Full Search Algorithm

The current block is compared with the reference frame of candidate block as to get the best match block in the reference frame.[5] The sum of absolute difference values at each possible location in the search window is calculated by full search motion estimation. All the candidate blocks for large search windows are calculated by full search. Computational complexity is of order n2.The sum of absolute difference values at each possible location in the search window is calculated by full search motion estimation. All the candidate blocks for large search windows are calculated by full search. Computational complexity is of order n2.[20]

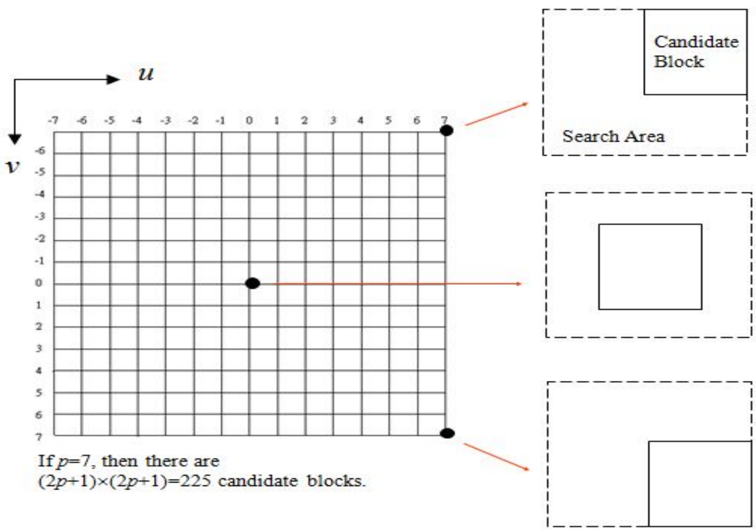


Figure 3.4: Full Search Algorithm

3.2.2 Diamond Search Algorithm

There are two search patterns in Diamond search algorithms. Small diamond search pattern: A small diamond shape comprising of 5 check points. Large Diamond Search Pattern: Comprises 9 check points from which eight points surround the center one to compose a diamond shape.[21]

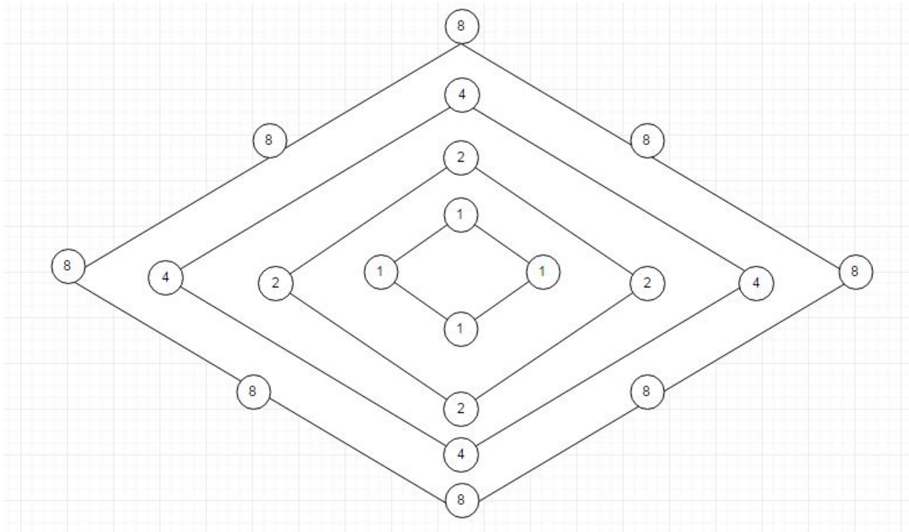


Figure 3.5: Diamond Search Scenario for ME

3.2.3 Hexagonal Search Algorithm

HEXBS has the minimum cost compared to the DS algorithm with same motion vectors and fewer search points.[4] The more search points can be saved when there are larger motion vectors. The last block distortion occurs at the center point if LDSP is repeatedly used.

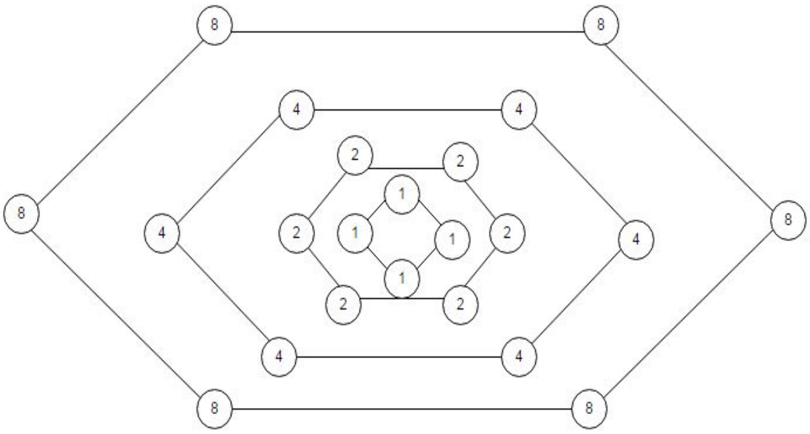


Figure 3.6: Hexagonal Search Algorithm

Chapter 4

Proposed System

4.1 Proposals of Modified Cross Hexagon Diamond Search Algorithm(Fast search)

Hexagon helps in decreasing the computational complexity to the greater extend.To reduce the complexity of time and bandwidth new hybrid search pattern is designed.[22] Three different patterns to achieve the minimum number of search points at each step.ME complexity with comparable performance are basic features of these algorithm.

4.1.1 Algorithm Development

The algorithm is designed using Symmetric Cross pattern with step size 2, modified asymmetric Hexagon with step size 1 and small Diamond pattern with step size 1 from center point. The steps for proposed algorithm are given below:

- Step 1: The search starts with the Symmetric Cross search pattern with pixel distance 2. The minimum cost point is obtained using Mean Absolute Difference computation. If the minimum cost is found to be at center point, then go to Step 3.

- Step 2: The Modified Asymmetric Hexagon search is performed with pixel distance 1. second search with minimum cost point from Step 1 as center point.

- Step 3: Switch from the coarse search to the fine resolution inner search to find the best motion vector. With minimum point from previous search as center point, the search is done with small Diamond pattern having step size 1 .The resultant minimum point from this step will be used to calculate the motion vector.

4.1.2 Search Path Analysis

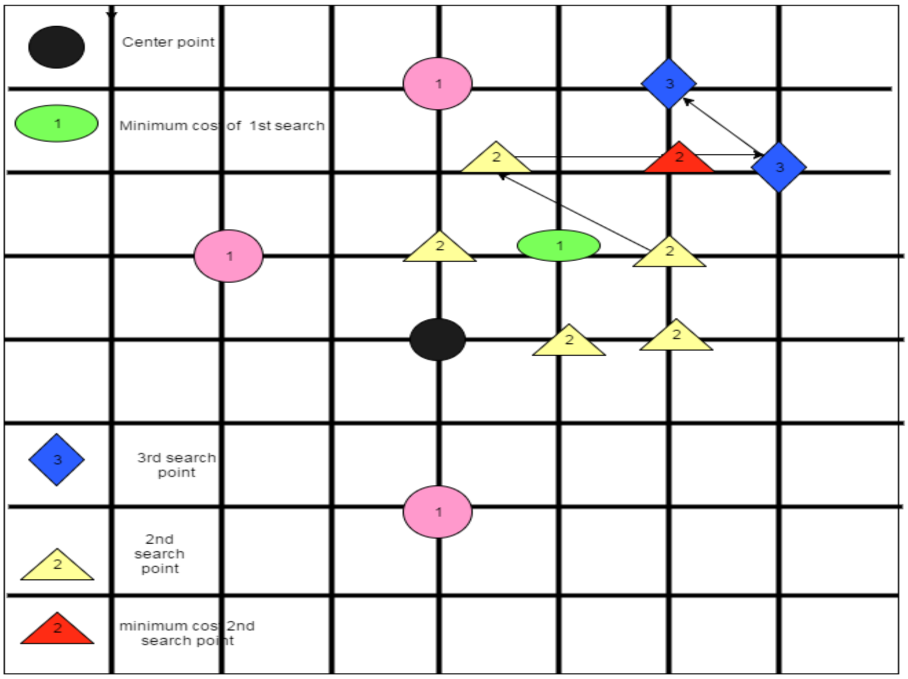


Figure 4.1: Proposed Algorithm

Proposed algorithm utilizes only 13 computations for the best match case,so the number of search candidates is decreased and which turns out to be the reducing motion estimation time and bandwidth as well.

Evaluation of the new CHDS algorithm(Fast search algorithm) having the featured inner search,compared with the existing methods in various aspects like search points , PSNR, average computational time saving motion estimation.The experimental setup is as follows: the block size of 16x16, the maximum displacement in the search area is 2 pixels and MAD (Mean absolute Difference) for distortion measurement.

Chapter 5

Implementation Results

The HM encoder provides two configurations of encoder tools, designated Main and Main10 in the HM test conditions and software reference configuration document. The coding tools for the Main configuration correspond to those supported in the Main profile of the HEVC specification. The coding tools for the Main10 configuration correspond to the Main 10 profile in the HEVC specification with the bit depth for luma and chroma both set to 10 bit.HEVC Test model uses three kinds of predictions depending on the experiments such as only intra prediction,low-delay and random access.[19]

5.1 Intra-only configuration

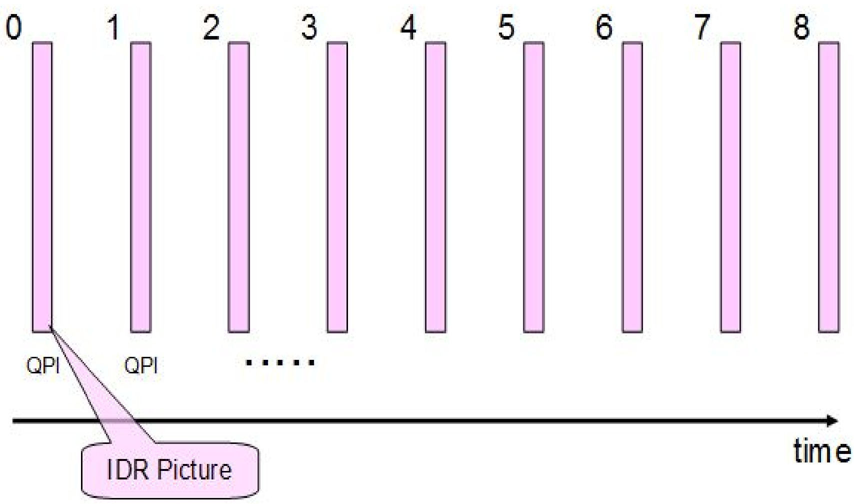


Figure 5.1: Graphical presentation of intra-only configuration

In the test case for intra-only coding, each picture in a video sequence is encoded as an IDR picture. No temporal reference pictures are used. QP does not change during a sequence within a picture. Figure gives graphical presentation of an intra-only configuration, where the number associated with each picture represents the encoding order.[9]

5.2 Low-delay configurations

For testing low-delay coding execution two coding design have been characterized. The primary picture in a video grouping is encoded as an IDR picture.alternate pictures are encoded as summed up P and B-pictures (GPBs). The GPB can utilize just reference pictures whose POC is shorter than the present picture.The substance of RefPicList0 and RefPicList1 are indistinguishable.[21]

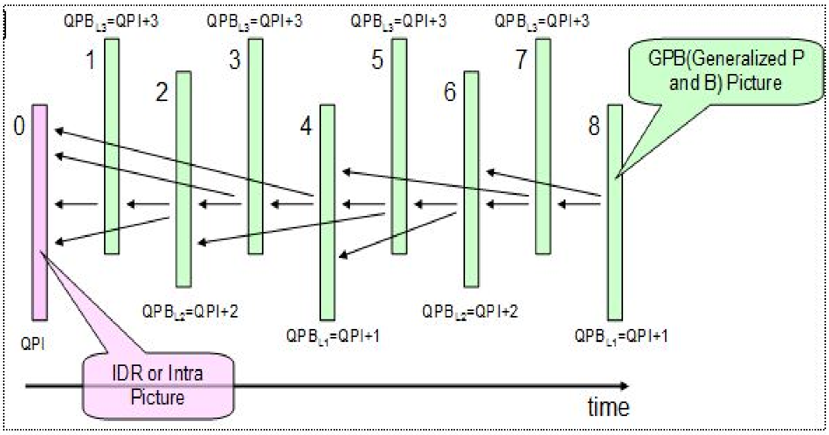


Figure 5.2: Graphical presentation of low-delay configuration

Entropy coding of the reference picture index is used for management by combination of reference picture.Graphical representation is shown in the figure representing the low delay configuration.The QP of each inter coded picture is derived by adding an offset to the QP of the intra coded picture depending on the temporal layer. In the alternative low-delay condition (which has not been mandatory in most CEs), all inter pictures are coded as P-pictures, where only the content of RefPicList0 is used for inter prediction.

5.3 Random-access configuration

For the random-access test condition, a hierarchical B structure is used for encoding. Graphical representation of a random-access configuration, each picture represents the encoding order associated with it. An intra picture is encoded at approximately one second intervals.

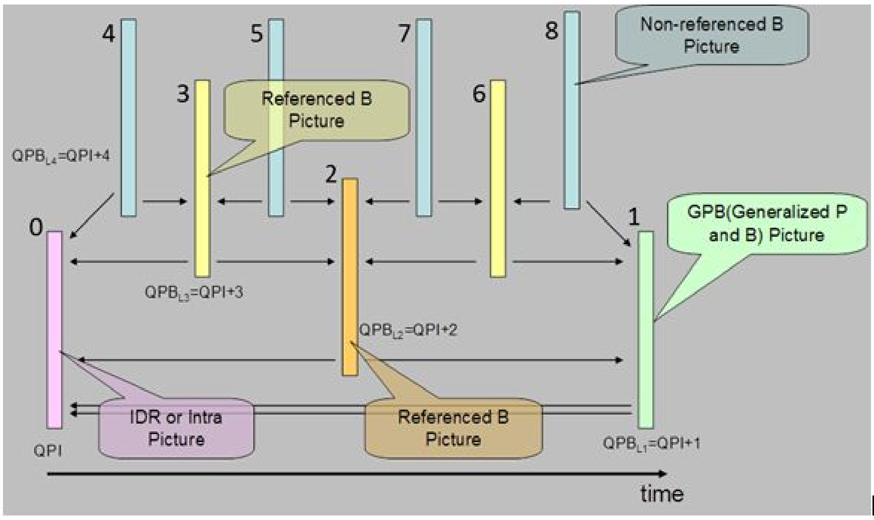


Figure 5.3: Graphical presentation of random-access configuration

5.4 Benchmark Videos

Benchmark videos are encoded by HM 16.8 software.Time stamp ,frame type are been traced from the video files by HM Software. Using HM Encoder we evaluate quality of various images that are transmitted. We firstly use the source YUV file to simulation file.We reconstruct YUV file at receivers end.We can download various benchmark videos from the site and use them for the evaluation process.[2]

Figure 5.4: Benchmark Videos

We use this YUV files and convert them into m4v format,this m4v files are then converted into mp4 files which are used by HM Encoder to get the image quality for transmitted file.this files are encoded and send for the transmission for receivers side .Thus this transmitted files are then converted to YUV files on the decoder side.They are also called reconstructed files which are different on decoder end as compared to encoder side.[2]

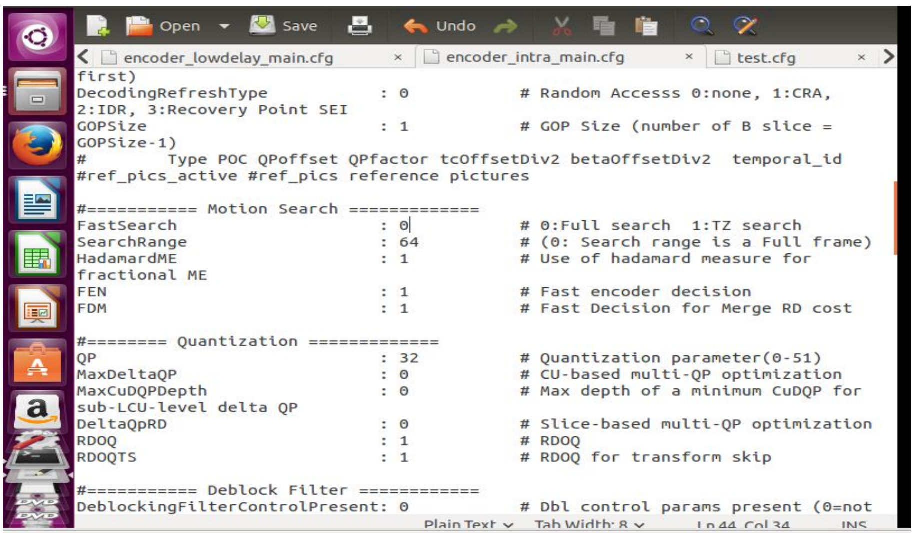


Figure 5.5: Configuration files for intra coding algorithm

The figure5.5 displays the Configuration files for intra coding algorithm.These configuration files are used in HM encoder to encode the video sequence.

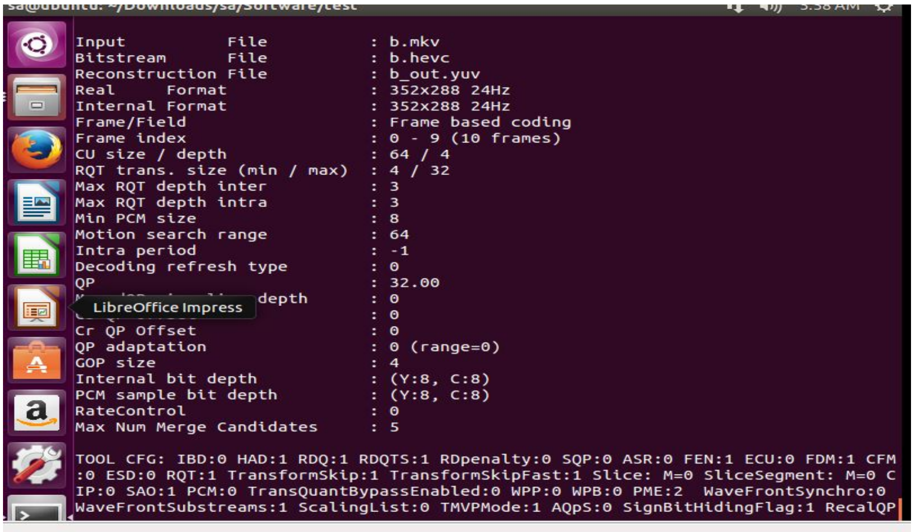


Figure 5.6: simulation for HEVC video files

The figure5.6 displays the simulation for various sequences.Detail describtionof various parameters like PSNR ratio, time stamp,quantization parameters,control rate,CU size etc.

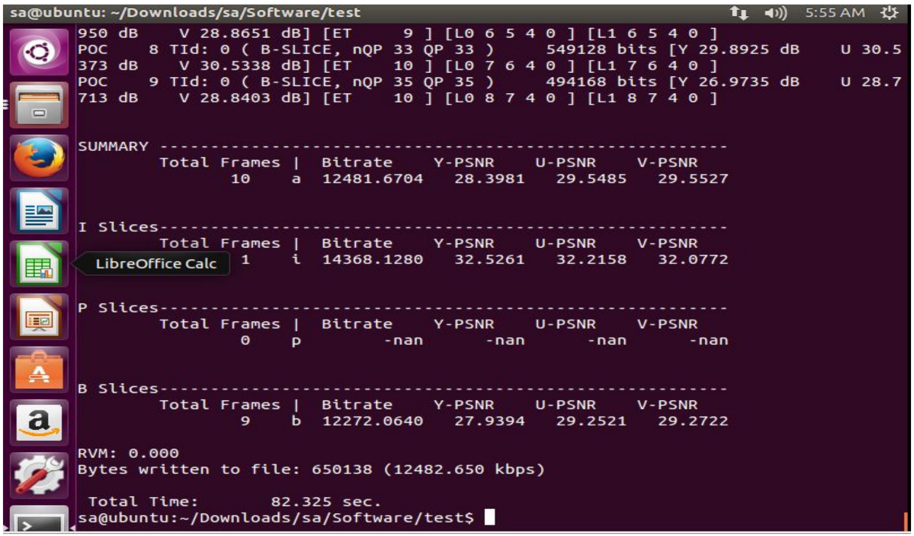


Figure 5.7: I,B,P slices

The figure5.11 displays the I,B,P slices.It shows different values of PSNR, total number of frames used and their slices.

RaceHorses_416x240_30.yuv, Number of frames encoded = 20						
	Random Access profile (FAST SEARCH)			Random Access profile (FULL SEARCH)		
QP	PSNR in dB	Bit rate in kbps	Encoding time in seconds	PSNR in dB	Bit rate in kbps	Encoding time in seconds
22	39.5858	1504.992	120.810	39.5969	1494.096	1399.795
27	35.9841	769.740	101.093	35.9926	762.0240	1508.792
32	32.7600	391.116	88.600	32.7769	389.296	1306.499
37	30.1072	202.572	71.437	30.1423	202.3080	1231.534

Figure 5.8: Results for RaceHorses.yuv sequence in Random Access Configuration

The figure5.8 displays Results for RaceHorses.yuv sequence in Random Access Configuration. The simulation is taken in consideration for four different quantization parameter, for two different algorithmshowing various result like psnr,bitrate,encoded time.

	BQMall_832x480_60.yuv , Number of frames encoded = 20					
	Random Access profile (FAST SEARCH)			Random Access profile (FULL SEARCH)		
QP	PSNR in dB	Bit rate kbps	Encoding time in seconds	PSNR in dB	Bit rate in kbps	Encoding time in seconds
22	40.6653	4538.160	308.840	40.7583	4528.440	4875.343
27	38.2347	2257.968	259.764	38.2446	2252.592	5214.150
32	35.4849	1200.864	244.682	35.4965	1196.352	4837.578
37	32.7343	656.088	224.326	32.7370	655.2240	5248.942

Figure 5.9: Results for BQMall.yuv sequence in Random Access Configuration

The figure5.9 displaysResults for KirstenAndSara.yuv ParkScene.yuv sequences in Random Access Configuration . The simulation is taken in consideration for four different quantization parameter, for two different algorithm showing various result like psnr,bitrate,encoded time.

	KristenAndSara_1280x720_60.yuv Number of fames encoded =20			ParkScene_1920x1080_24.yuv Number of fames encoded =20		
	Random Access profile (FAST SEARCH)			Random Access profile (FAST SEARCH)		
QP	PSNR in dB	Bit rate in kbps	Encoding time in seconds	PSNR in dB	Bit rate in kbps	Encoding time in seconds
22	44.3166	2676.456	533.038	40.7117	8421.072	1808.863
27	42.5646	1232.856	570.403	38.3703	3673.3632	1356.223
32	40.4650	667.344	481.128	36.0006	1697.606	1211.177
37	38.0715	381.720	444.487	33.7621	791.712	1107.963

Figure 5.10: Results for sequences in Random Access Configuration

The figure5.10 displays Results for sequences in Random Access Configuration. The simulation is taken in consideration for four different quantization parameter, for two different algorithm showing various result like psnr,bitrate,encoded time.

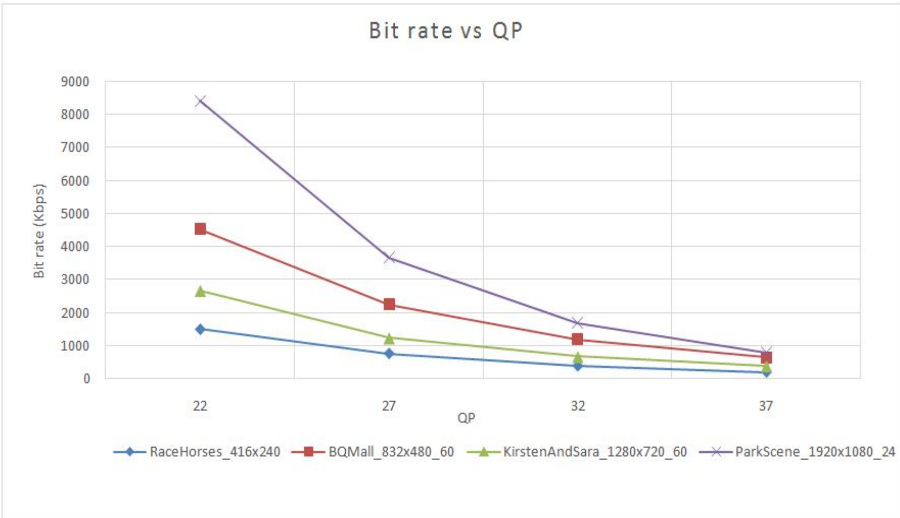


Figure 5.11: Bit rate vs QP for all test sequences

The figure5.11 displays the graph of Bit rate vs QP for all test sequences. The simulation is taken in consideration for four different quantization parameter.The graph represents bit-rate decreases with the increase in the quantization parameter for all test sequences.

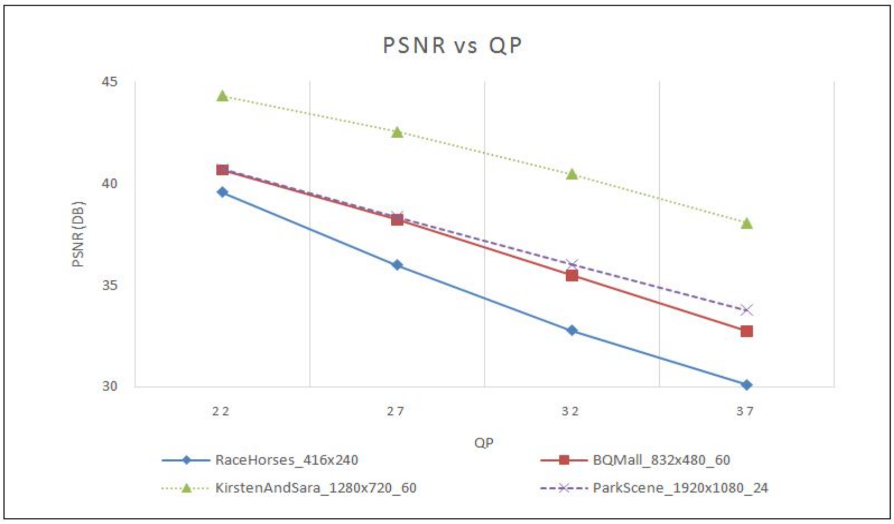


Figure 5.12: PSNR vs QP for all test sequences

The figure5.12 displays the graph of PSNR vs QP for all test sequences . The simulation is taken in consideration for four different quantization parameter. The graph represents values of psnr decreases with the increase in QP.

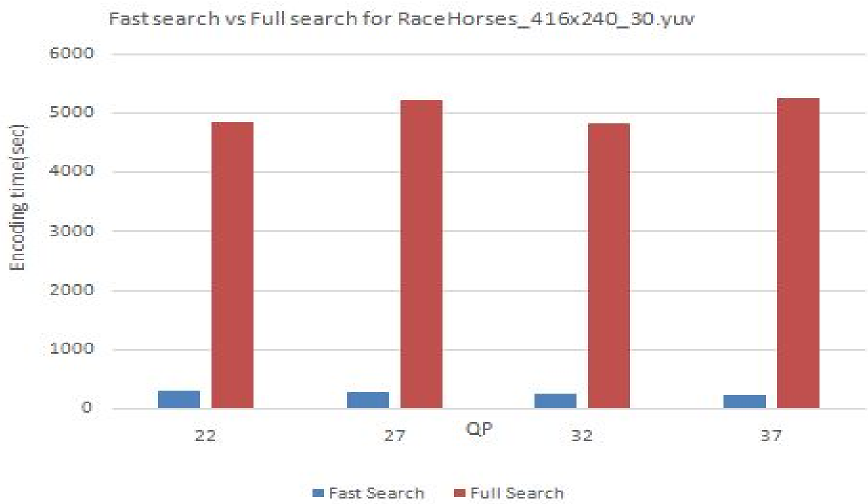


Figure 5.13: Encoding time comparison for RaceHorses 416×240.yuv sequence

The figure5.13 displays the graph of Encoding time comparison for RaceHorses.yuv sequence.The graph represents the time stamp versus quantization parameters.As from the graph we can see that two algorithms are used MCHD(fast search) and full search are been compared with each other fro Race horse video sequence.Graph shows the different time taken by full search algorithm is more compared to the fast search algorithms

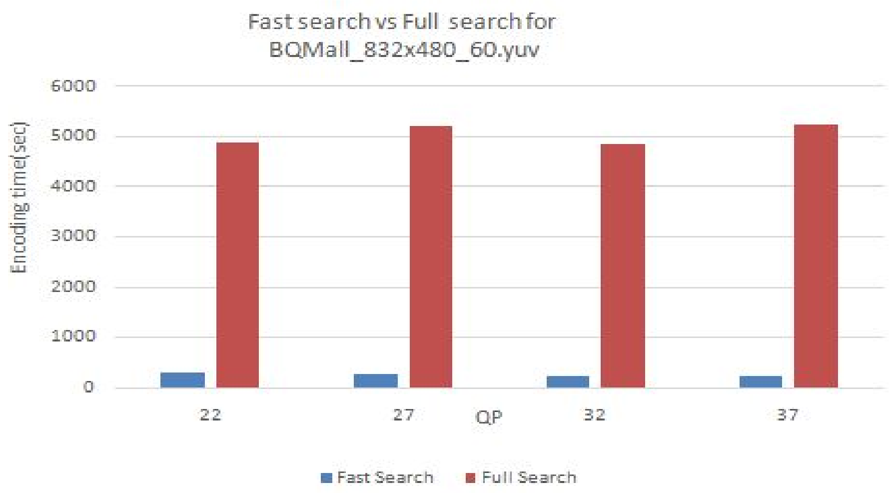


Figure 5.14: Encoding time comparison for BQMall832×480.yuv sequence

The figure5.14 displays the graph of Encoding time comparison for BQMall.yuv sequence. The simulation is taken in consideration for four different quantization parameter.The graph represents the time stamp versus quantization parameters.As from the graph we can see that two algorithms are used MCHD(fast search) and full search are been compared with each other fro Race horse video sequence.Graph shows the different time taken by full search algorithm is more compared to the fast search algorithms

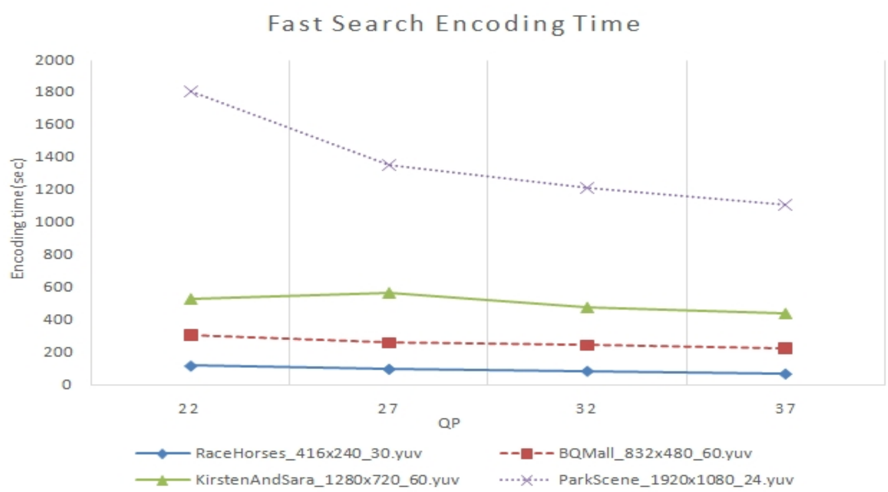


Figure 5.15: Encoding time comparison for all test sequences

The figure5.15 displays the graph of Encoding time comparison for all test sequences.For four different QP different encoding time has been plotted for four different sequences.Graph represents that the encoding time goes on decreasing with increase in the quantization parameter.

Chapter 6

Conclusion

- Implemented HM 16.8 software and analyzed how MCDH best match is selected. Simulated with different test sequences and obtained the results. Analyzed Full Search,Hexagonal search and Diamond Search and proposed MCDH algorithm. Compared full search and MCHD (Fast search) algorithm and found out that full search algorithm requires much time as well bandwidth increases but in modified cross hexagon diamond search requires less time and bandwidth.

  Developed functions like

  - To find PSNR w.r.t reference image.
  - To compute motion compensated image.
  - To find minimum distance among macro blocks.
  - To Compute Mean Absolute difference.

Bibliography

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] J. Klaue, B. Rathke, and A. Wolisz, "Evalvid–a framework for video transmission and quality evaluation," in *Computer Performance Evaluation. Modelling Techniques and Tools*, pp. 255–272, Springer, 2003.

[3] J. Dong, Y. Ye, and Y. He, "Cross-plane chroma enhancement for shvc inter-layer prediction," in *Picture Coding Symposium (PCS), 2013*, pp. 309–312, IEEE, 2013.

[4] D. S. Ragasudha, Narapaneni and S. Meher., ""block based motion estimation u sing hybrid hexagon kite cross diamond search algorithmv."," 2014.

[5] D. RAO, ""complexity reduction in h. 265 motion estimation using compute unified device architecture."," 2013.

[6] M. R. Hamidouche, Wassim and O. Deforges., ""multi-core software architecture for the scalable hevc decoder."," (2013).

[7] H. Amirpour and A. Mousavinia, ""motion estimation based on region prediction for fixed pattern algorithms."," 2013.

[8] C. Pandey and D. Pandey., ""literature review on block matching motion estimation algorithms for video compression."," 2013.

[9] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "Hevc complexity and implementation analysis," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1685–1696, 2012.

[10] N. Ragasudha, "Efficient video compression schemes. diss. department of electronics and communication engineering, national institute of technology rourkela," 2014.

[11] M. P. Podder, Pallab Kanti and M. Murshed., ""a novel motion classification based intermode selection strategy for hevc performance improvement."," 2014.

[12] L. Zhao, L. Zhang, S. Ma, and D. Zhao, "Fast mode decision algorithm for intra prediction in hevc," in *Visual Communications and Image Processing (VCIP), 2011 IEEE*, pp. 1–4, IEEE, 2011.

[13] ""the scientific world journal : http://dx.doi.org/10.1155/2014/267403"."

[14] e. a. Franois, Edouard, ""exploration of generalized residual prediction in scalable hevc."," (2013).

[15] X. S. Shi, Zhongbo and F. Wu., ""spatially scalable video coding for hevc." circuits and systems for video technology"," (2012).

[16] M. Domanski, A. Dziembowski, T. Grajek, A. Grzelka, L. Kowalski, M. Kurc, A. Luczak, D. Mieloch, R. Ratajczak, J. Samelak, O. Stankiewicz, J. Stankowski, and K. Wegner, "Methods of high efficiency compression for transmission of spatial representation of motion scenes," in *Multimedia Expo Workshops (ICMEW), 2015 IEEE International Conference on*, pp. 1–4, June 2015.

[17] M. T. P. Tohidypour, Hamid Reza and P. Nasiopoulos., ""adaptive search range method for spatial scalable hevc."," (2014).

[18] M. P. Podder, Pallab Kanti and M. Murshed., ""a novel motion classification based intermode selection strategy for hevc performance improvement."," 2014.

[19] M. P. Podder, Pallab Kanti and M. Murshed., "Inter-mode decision strategy for hevc on depth videos."," 2014.

[20] J. Boyce, Y. Ye, J. Chen, and A. Ramasubramonian, "Overview of shvc: Scalable extensions of the high efficiency video coding (hevc) standard," 2015.

[21] W. Hamidouche, M. Raulet, and O. Déforges, "Parallel shvc decoder: Implementation and analysis," in *Multimedia and Expo (ICME), 2014 IEEE International Conference on*, pp. 1–6, IEEE, 2014.

[22] P. B. Fortun, Denis and C. Kervrann, “”optical flow modeling and computation: a survey.” computer vision and image understanding,” 2015.

Sarthak

ORIGINALITY REPORT

8%	2%	9%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Öztekin, Abdulkerim, and Ergun Erçelebi. "An early split and skip algorithm for fast intra CU selection in HEVC", Journal of Real-Time Image Processing, 2015. Publication	1%
2	Pandian, S. Immanuel Alex. "A Study on Block Matching Algorithms for Motion Estimation", International Journal on Computer Science & Engineering/09753397, 20110101 Publication	1%
3	www.videosense.eu Internet Source	1%
4	Chen, Gaoxing, Zhenyu Pei, Lei Sun, Zhenyu Liu, and Takeshi Ikenaga. "Fast intra prediction for HEVC based on pixel gradient statistics and mode refinement", 2013 IEEE China Summit and International Conference on Signal and Information Processing, 2013. Publication	1%
5	Su, Jia. "Content based low complexity algorithms for high efficiency video coding",	1%

DSpace at Waseda University, 2013.
Publication

6	Integrated Circuits and Systems, 2014. Publication	1%
7	Han, G.J., J.R. Ohm, Woo-Jin Han, Woo-Jin Han, and T. Wiegand. "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Transactions on Circuits and Systems for Video Technology, 2012. Publication	1%
8	Signals and Communication Technology, 2015. Publication	1%
9	Miyazawa, K., T. Murakami, A. Minezawa, and H. Sakate. "Complexity reduction of in-loop filtering for compressed image restoration in HEVC", 2012 Picture Coding Symposium, 2012. Publication	1%

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	< 1%
EXCLUDE BIBLIOGRAPHY	ON		