# ENCODING COMPLEXITY ANALYSIS THROUGH THE MEASUREMENT OF THE SEARCH SPACE

## SVCP 2022 - Summer School on Video Coding and Processing

*Presenter:*

Adam Wieckowski

*Video Communications and Application Dep.*
*Fraunhofer Heinrich Hertz Institute (HHI)*
*Berlin, Germany*




Fraunhofer
HHI

# Introduction

- Based on the paper:
  - **VVC Search Space Analysis including an Open, Optimized Implementation**
  - In IEEE Transactions on Consumer Electronics, vol. 68, no. 2, pp. 127-138, May 2022.

- Comparisons based on 3 encododers
  - HM 16.22
  - VTM 11.0
  - VVenC 1.0.0

# Agenda

- What is complexity in video encoding?
  - Encoding vs decoding
- Empirical search space quantification
  - Measurement
  - Application to partitioning
  - *Application to mode search*
- Conclusion

# Video coding complexity

- Video coding is complex
  - In literature: VVC encoding 10x more complex than HEVC, decoding 2x

- Coding complexity is measurable
  - Runtime
  - Energy consumption (electric bill, device heat)

- What causes the complexity and how can it be controlled?
- Does the complexity depend on the standard used?

# Encoding vs decoding complexity

- What is **video decoding**?
  - Fixed sequence of steps described in the standard specification
  - The **worst case** can be quantified, using e.g. number of Multiply-Add-Operations

- What is **video encoding**?
  - Video encoders incorporate the decoder
  - Beyond that its basically a **search problem**
    - Decoder is part of the **cost function**

# Video encoding as a search problem

- Usual formulation: find a sequence of **N bits**, such that ($2^N$ possibilities)
  - The sequence is **compliant** with a given standard
  - **Minimizing** the distortion between the **decoded video** and the original
- Components influencing enc. complexity: bitrate, complexity of decode, accuracy
- ➢ Contradiction
  - Assumption: VVC provides **50%** bitrate savings vs HEVC
  - N half as large for VVC as for HEVC, at double decode complexity
  - **VVC encoding** should actually be **less complex**?
- Smart search algorithms only evaluate fraction of overall search space
  - ➢ How big is the **actual visited search space**?
  - ➢ Why is this search space larger for each next generation codec?



Fraunhofer HHI

# Empirical search space quantification

- Recap
  - Decoder is part of the encoder cost function
- Idea
  - Quantify how many times a sample is decoded during encoding
- Problem
  - When is a sample decoded during encoding?
- Solution
  - During video decoding each sample is
    - Only contained in a single block
    - Dequantized maximally once
  - ➢ Measure the plurality of partitioning and quantization test per sample



Fraunhofer HHI

# Empirical search space quantification

## Partitioning search

- Encoder search visits a set of $i_P \in 1 \ldots N_P$ blocks during encoding of a frame
  - A frame has $N_s$ samples
  - Each block has a height of $W_P(i)$ and $H_P(i)$
- Partitioning overhead is thus:

$$S_P = \sum_{i=1 \ldots Np} W_P(i) \cdot H_P(i) \,/\, N_S$$

- Cumulative area of blocks visited during encoding normalized to frame size

# Empirical search space quantification

## Coding mode search

- In each visited block, the encoder tests a plurality of coding modes
  - Often most complex step is the rate-distortion-optimized quantization (RDOQ)
  - The RDOQ is applied to $i_Q = 1 \ldots N_Q$ blocks of size $W_Q(i) \cdot H_Q(i)$
- Quantization overhead is thus:

$$S_Q = \sum_{i=1 \ldots Nq} W_Q(i) \cdot H_Q(i) \, / \, ( N_S \cdot S_P )$$

- Cumulative quantized area, normalized to the partitioning area
- Overall encoding overhead (i.e. **empirical search space**): $S = S_Q \cdot S_P$



Fraunhofer HHI

# Partitioning search space

## Exemplary for HEVC

- Assuming the G-BFOS search algorithm
  - *Forward-only* search
  - Each sample once in each config.
- For HEVC assuming CTU of 64x64
  - 4 different CU sizes
  - PU splits also count as partitioning
  - Easy to enumerate
- Simple formulation of upper bound

| Depth / CU size | Allow PU splits |
|---|---|
| 0 / 64x64 | NxN, N/2xN/2, NxN/2, N/2xN, 4x assymetric modes |
| 1 / 32x32 | |
| 2 / 16x16 | |
| 3 / 8x 8 | NxN, N/2xN/2, NxN/2, N/2xN |

| Partitionings per sample | | |
|---|---|---|
| Intra | 4 | 5 |
| Inter | 4 | 28 |

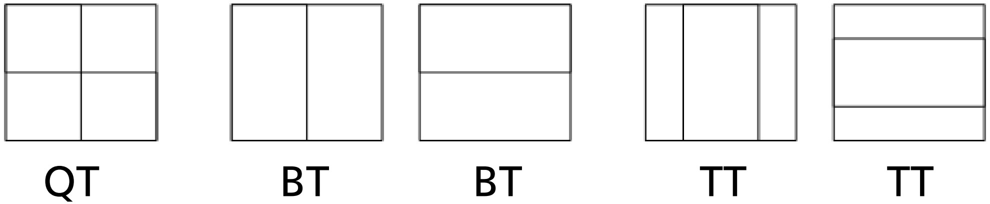


Fraunhofer HHI

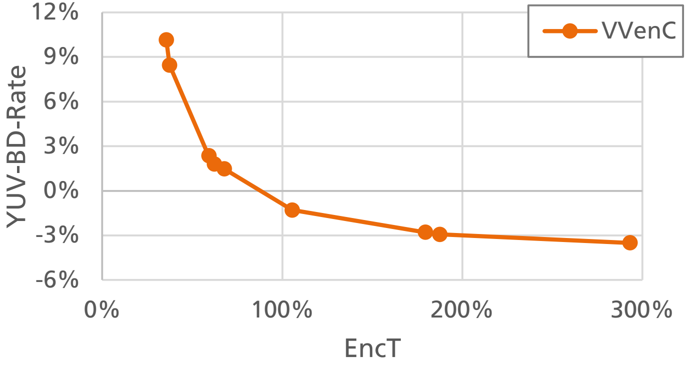
# Partitioning search space

## Recap partitioning in VVC

- Five recurisve splits
  - QT only with a QT
- Exponential growth
- Empirical upper bound




  - Assuming full traversal with G-BFOS, including
  - Chroma separate tree
  - Local dual tree, mode restrictions
  - Depends on high-level partitioning parameters
- Figure
  - CTU128, QT: 128x128 to 8x8
  - BT and TT splits: 0 to 4 recursion levels



# Partitioning search space

## Upper bounds for VVC search space

- QT depth is simply the difference between min and max block size
  - BTT depth has to be kept much lower to limit complexity
- VVC partitioning space can be much larger than HEVC
  - For practical encoders it can be very well kept lower

| | Intra Frames | | | | | | P/B Frames | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CTU | Max size | | Max depth | | bound | CTU | Max size | | Max depth | | bound |
| | | QT | BTT | QT | BTT | | | QT | BTT | QT | BTT | |
| **faster** | 64 | 64 | N/A | 4 | 0 | **4.67** | 64 | 64 | N/A | 4 | 0 | **5.00** |
| **fast** | 64 | 64 | 32 | 4 | 1 | **14.33** | 64 | 64 | N/A | 4 | 0 | **5.00** |
| **medium** | 128 | 128 | 32 | 4 | 2 | **37.50** | 128 | 128 | 128 | 4 | 1 | **24.00** |
| **slow** | 128 | 128 | 32 | 4 | 3 | **85.42** | 128 | 128 | 128 | 4 | 2 | **75.00** |
| **slower** | 128 | 128 | 32 | 4 | 3 | **85.42** | 128 | 128 | 128 | 4 | 3 | **220.75** |
| **VTM-11.0** | 128 | 128 | 32 | 4 | 3 | **85.42** | 128 | 128 | 128 | 4 | 3 | **220.75** |



Fraunhofer HHI

# Partitioning search space

## Empirical results for Inter frames

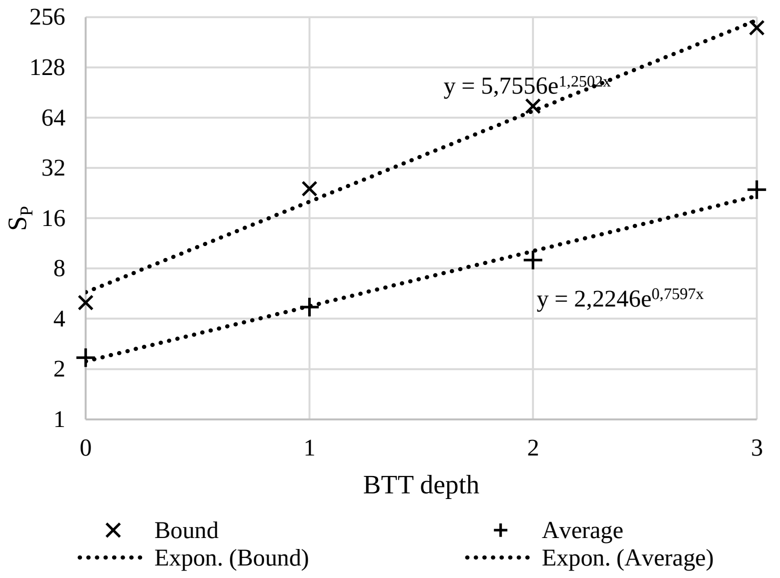
- Measured on JVET CTC Classes A1, A2, and B, in CTC encoding conditions
- Observations for HM
  - HM has no CU depth speedup
  - ... but has PU split speedups
- VVenC slower visits less partitions than VTM, bcs of more aggresive optimization
- The larger the search space, the easier it is to limit
- Search space varies, e.g. with target quality (more so for VVencC)

| | Partitioning $S_P$ | | | |
|---|---|---|---|---|
| | avg | min | max | bound |
| VVenC faster | 2.39 | 1.71 | 3.31 | 5.00 |
| VVenC fast | 2.34 | 1.68 | 3.23 | 5.00 |
| VVenC medium | 4.69 | 3.32 | 6.75 | 24.00 |
| VVenC slow | 8.98 | 6.01 | 14.12 | 75.00 |
| VVenC slower | 23.71 | 13.49 | 42.04 | 220.75 |
| VTM | 29.18 | 17.54 | 49.60 | 220.75 |
| *HM CU* | *3.95* | *3.95* | *3.95* | *4.00* |
| HM CU+PU | 15.51 | 13.97 | 17.77 | 28.00 |

# Partitioning search space

## Exponential growth, theoretical and measured for P/B frames

- Fast algorithm can limit the growth
  - Both base and exponent reduced
  - Exponent bound: ~1.25 x *Depth*
  - Exponent average: ~0.76 x *Depth*
- The curves diverge
  - The larger the search space, the more opportunities for limitation






Fraunhofer HHI

# Overall search space

## First observations

- Why is partitioning such a popular topic for optimization in literature?
  - $S_P >> S_Q$ for both VTM and HM
  - In VVenC, which was Pareto-Optimized, $S_P \sim S_Q$ (except for *slower*)
- It is hard to define an upper bound for $S_Q$

| | Partitioning $S_P$ | | | | Quantization $S_Q$ | | |
|---|---|---|---|---|---|---|---|
| | avg | min | max | bound | avg | min | max |
| VVenC faster | 2.39 | 1.71 | 3.31 | 5.00 | 3.15 | 2.92 | 3.55 |
| VVenC fast | 2.34 | 1.68 | 3.23 | 5.00 | 4.46 | 4.13 | 5.05 |
| VVenC medium | 4.69 | 3.32 | 6.75 | 24.00 | 5.81 | 5.29 | 6.59 |
| VVenC slow | 8.98 | 6.01 | 14.12 | 75.00 | 10.08 | 9.37 | 10.99 |
| VVenC slower | 23.71 | 13.49 | 42.04 | 220.75 | 13.08 | 11.85 | 14.84 |
| VTM-11.0 | 29.18 | 17.54 | 49.60 | 220.75 | 14.28 | 12.35 | 16.91 |
| *HM-16.22 CU* | *3.95* | *3.95* | *3.95* | *4.00* | *20.93* | *17.18* | *27.27* |
| HM-16.22 CU+PU | 15.51 | 13.97 | 17.77 | 28.00 | 5.32 | 4.84 | 6.05 |



Fraunhofer HHI

# Overall search space

## Further observations

- VVenC *medium* has a search space similar to HM, but provides **~30% BD-rate gain**
- Average $S_Q$ ~3x larger in VTM than HM, $S_P$ ~2x larger in VTM than HM
  - Partitioning adds complexity to VVC, but new encodings modes add even more
  - E.g. a lot of new merge modes

| | Partitioning $S_P$ | | | | Quantization $S_Q$ | | |
|---|---|---|---|---|---|---|---|
| | avg | min | max | bound | avg | min | max |
| VVenC faster | 2.39 | 1.71 | 3.31 | 5.00 | 3.15 | 2.92 | 3.55 |
| VVenC fast | 2.34 | 1.68 | 3.23 | 5.00 | 4.46 | 4.13 | 5.05 |
| VVenC medium | 4.69 | 3.32 | 6.75 | 24.00 | 5.81 | 5.29 | 6.59 |
| VVenC slow | 8.98 | 6.01 | 14.12 | 75.00 | 10.08 | 9.37 | 10.99 |
| VVenC slower | 23.71 | 13.49 | 42.04 | 220.75 | 13.08 | 11.85 | 14.84 |
| VTM-11.0 | **29.18** | 17.54 | 49.60 | 220.75 | **14.28** | 12.35 | 16.91 |
| *HM-16.22 CU* | *3.95* | *3.95* | *3.95* | *4.00* | *20.93* | *17.18* | *27.27* |
| HM-16.22 CU+PU | **15.51** | 13.97 | 17.77 | 28.00 | **5.32** | 4.84 | 6.05 |

# Conclusions

## Empirical search space measurement

- Recap of the presentation
  - Empirical measurement of search space was presented
  - Data was presented and discussed, with focus on partitioning
- Conclusions
  - Partitioning is complex, but is already very effectively reduced
  - Is partitioning in VTM over-dimensioned?
- Shortcomings
  - Only measures the CU-loop search space
- Outlook
  - Measure distortion calculations per sample (i.e. prediction overhead)
  - Evaluate the ratio of prediction overhead to quantization overhead

# ENCODING COMPLEXITY ANALYSIS THROUGH THE MEASUREMENT OF THE SEARCH SPACE

Questions?

Fraunhofer
HHI