Decoder-side Intra Mode Derivation for next generation Video coding

Anthony Nasrallah

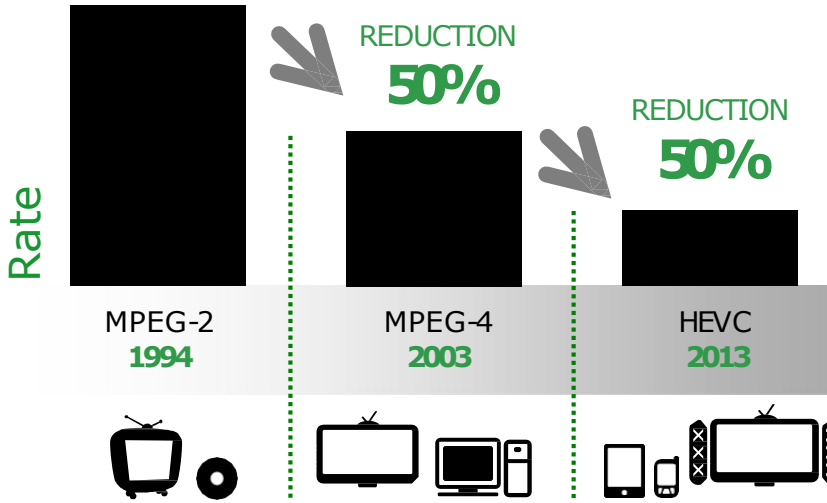
Plan

1. Introduction
2. State of the Art
3. Presentation of the problem
4. Description of the method
5. Results
6. Conclusion

Introduction

HEVC: A 50% gain in compression efficiency

Evolution of the rates obtained by the different standards



HD content Distribution on mobiles and content UHD at home and at the cinema

Video = 82% of global Internet traffic in 2021

*Source :*Cisco Corp., VNI Global Fixed and Mobile Internet Traffic Forecasts

Continued growth in global TV content consumption

Multiplicity of catalogues of video On demand

Increased resolutions (HD then UHD)

Multiplicity of broadcast FORMATS, MULTI-screens (tablets, smartphones,...)

Ateme × anevia

Stronger Together

Context

- **HEVC** would also present **limitations** that will prevent it from **meeting** the **requirements** of the market to the 2020 horizon

- Need to increase the efficiency of compression compared to HEVC with the introduction of VVC (Versatile Video Coding).

- Each 3 months, JVET (MPEG+ITU) organizes meetings between experts in the video field to proceed the development of the work until the standardization of the VVC which ends up in July 2020.

- A collaboration between ATEME and Telecom Paris was created in order to start a PhD thesis that aims to propose techniques to VVC.

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

ateme × anevia

Stronger Together

State of the art

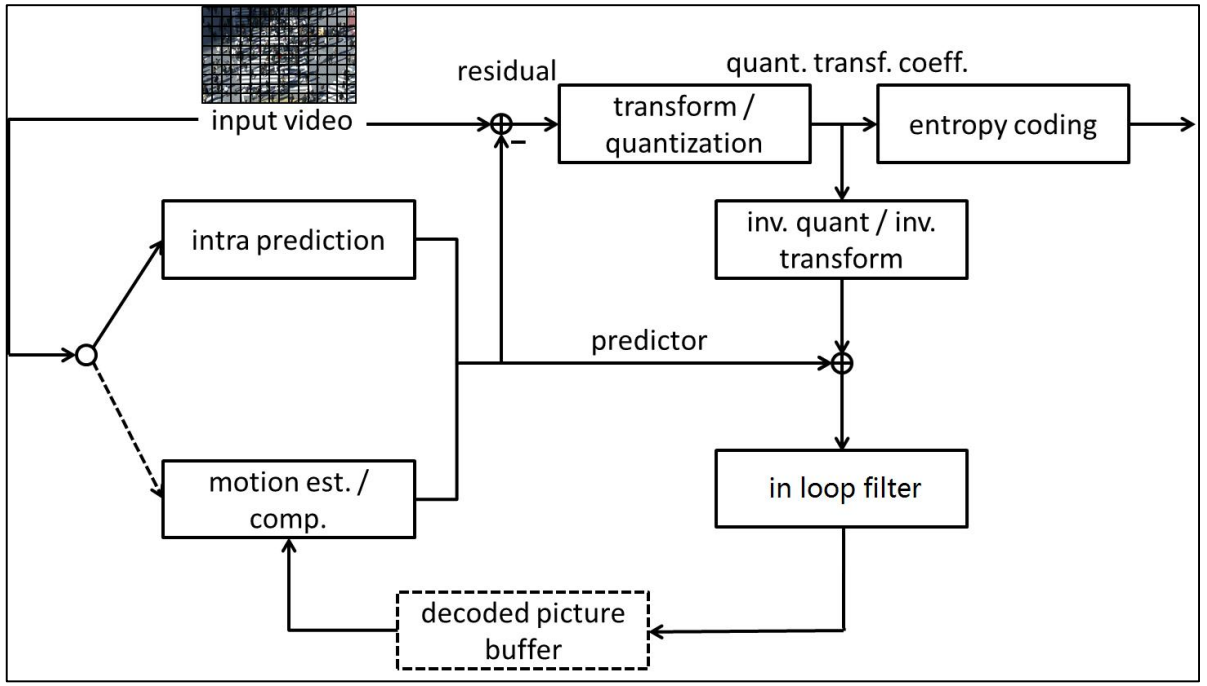
TELECOM Paris

IP PARIS

INSTITUT POLYTECHNIQUE PARIS-SACLAY

General notions

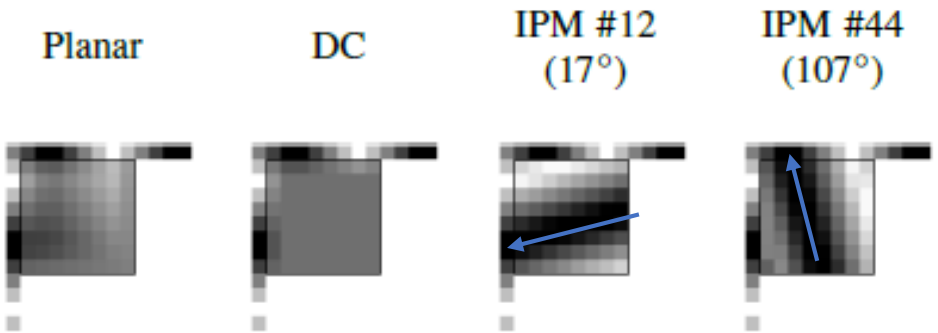
- Frames are divided into blocks.
- Intra Prediction : Spatial redundancy, in the same frame.
- Inter Prediction : Temporal redundancy, between frames.

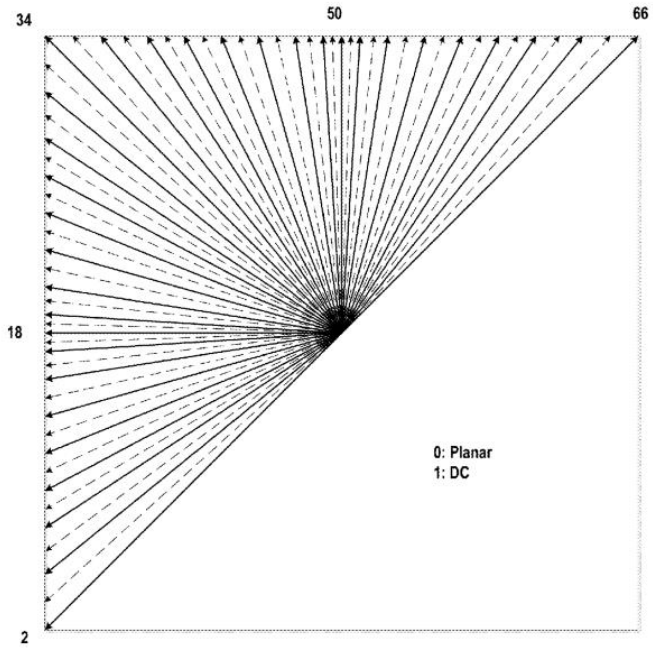


Generic Hybrid Video Encoder

Intra prediction

- Spatial extrapolation of a set of reference pixels is used for pixel prediction in VVC.
- 67 Intra Prediction modes (IPMs) in VVC, instead of 35 in HEVC. It has been reported that, on average, the IPM signaling takes about 9% of rate in VVC.





Stronger Together

ateme × anevia

Stronger Together

Presentation of the problem

TELECOM Paris

IP PARIS

INSTITUT POLYTECHNIQUE PARIS-SACLAY

Presentation of the problem

- One of the solutions is the increase in the number of intra-prediction modes, thus allowing more efficient predictions and less energetic residues
- Problem: Increasing number of prediction modes ➔ Signaling increased

- **Problematic: How to add intra prediction modes without increasing the cost of coding?**

- **Solution: Deriving the intra-angular mode at the decoder, thus avoiding its transmission in the bitstream.**

- Something interesting to know is that the IPM signaling takes on average about 9% of rate in VVC

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

ateme × anevia

Stronger Together

Description of the method

TELECOM Paris

IP PARIS

INSTITUT POLYTECHNIQUE PARIS-SACLAY

Summary of the method

- Our method **reduces the encoding rate** of a constant-quality video by adding a **new intra coding mode** called DIMD.

- The operation referred consists of treating the **DIMD identically to other coding modes** and in case this mode presents the **cheapest encoding cost**, it will be chosen among the modes and its choice will be **reported to the decoder** by a flag.

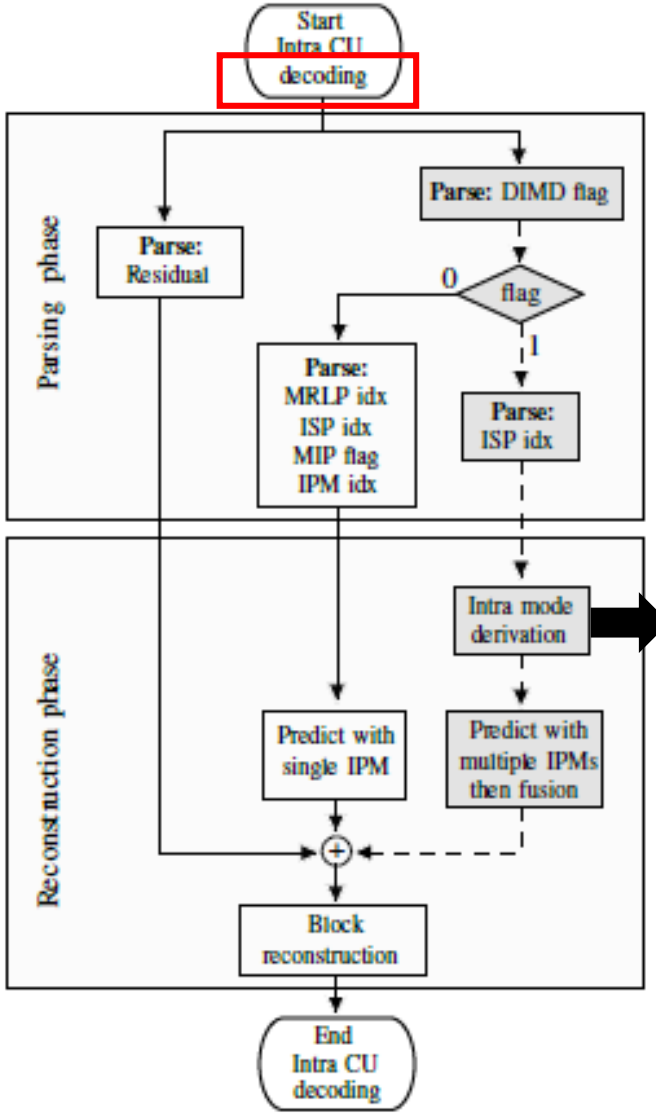
- **Instead** of explicitly **signaling** the intra-prediction mode, the **decoder** implicitly **derives one or more** potential intra-prediction directions by a **low-complexity texture analysis.**

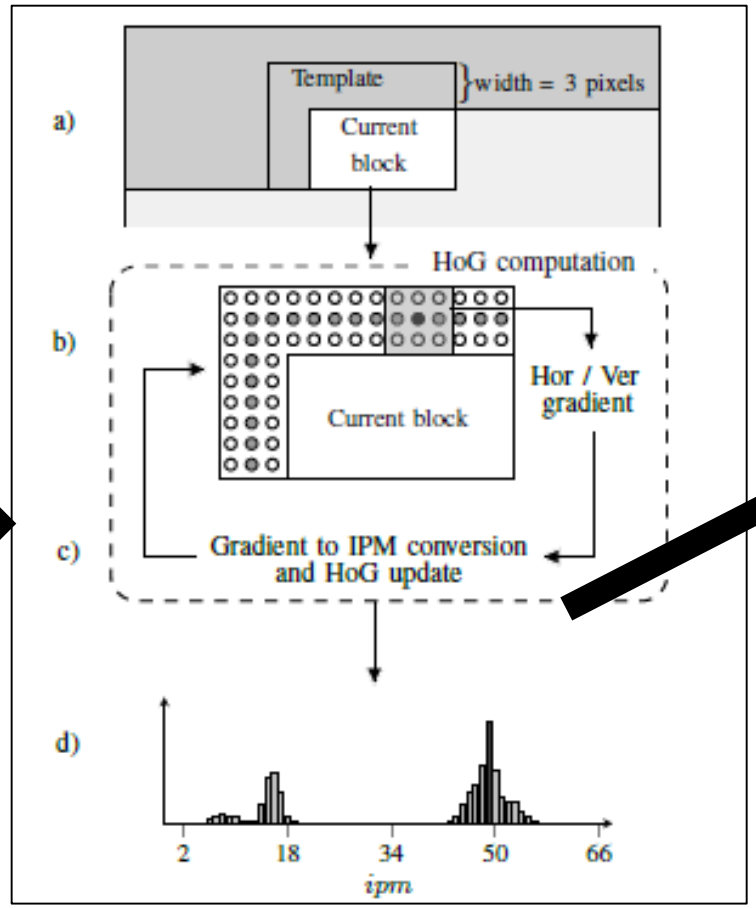
- The **different** inferred **intra-prediction modes** can then be **combined** by **means** of **prediction fusion**, to obtain the final intra-prediction for a given block. This fusion property allows DIMD to **more accurately predict areas with complex textures**, which in conventional coding schemes would instead require a finer partitioning and/or transmission of high energy residuals

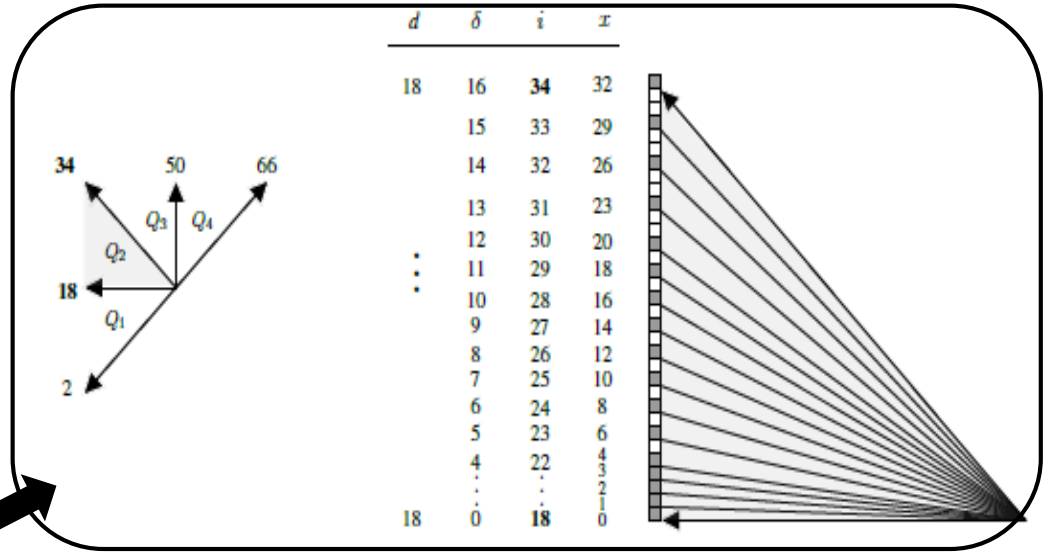
ateme × anevia
Stronger Together

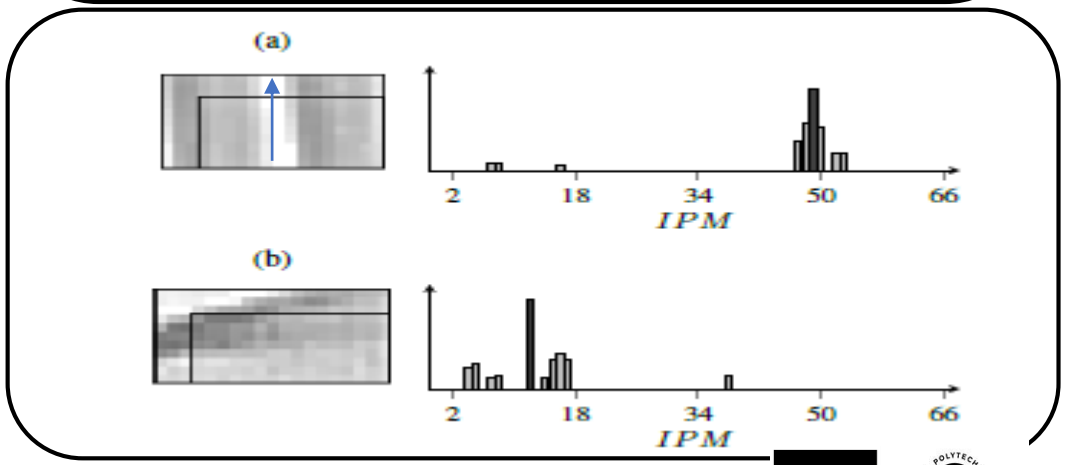
TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

Decoder-side Intra Mode Derivation





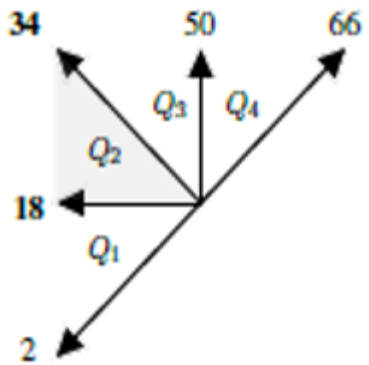


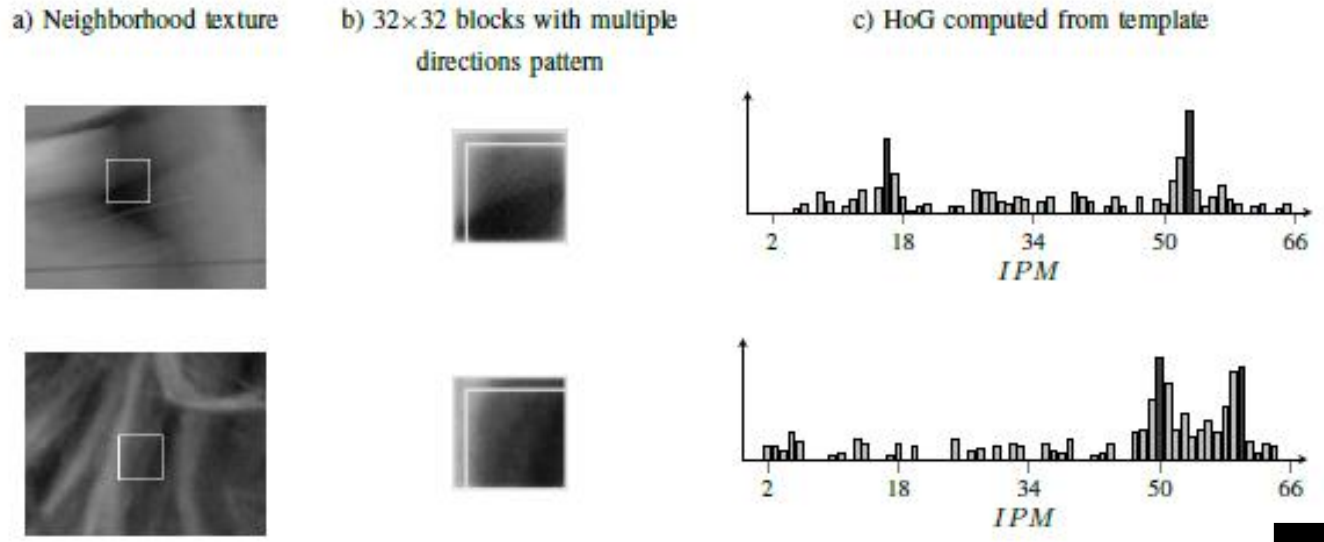


Stronger Together

Prediction fusion

- Total number **of 3 fusion modes** are used.
- The **first and second** fusion modes are always the HoG modes corresponding to the two **highest histogram amplitudes.**
- The choice of the **third** fusion mode **depends on the relative index of the two selected HoG modes.**





ateme × anevia
Stronger Together

Fusion modes choice

- If the **absolute difference between the IPM index** of the two HoG modes **is more than a threshold** → block with **multiple dominant local angular textures** → a **smooth transition** between them → the third fusion mode is selected to be **planar mode**.

- **Otherwise**, the block is assumed to have simple directional pattern that corresponds to the best HoG modes → the **third fusion** mode is selected to be the **MIP mode equivalent to the first HoG mode.**

- In our experiments, **threshold** is selected to be **10** and the three fusion modes are combined with equal weights.

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

ateme × anevia

Stronger Together

Results

TELECOM Paris

IP PARIS

INSTITUT POLYTECHNIQUE PARIS-SACLAY

Experimental settings

- Integrated in VTM-5.0 (C++); The aim of this work was to forge the VVC standard, that's why we contributed to JVET.

- Tests
  - Launched with All Intra(AI), Random Access (RA) and Low delay (LD-B) configurations.
  - On 4 QPs 22, 27, 32 and 37 (CTC)
  - All the video sequences defined in the CTC
  - On a server hosting 2 CPUs running at 2.50GHz.

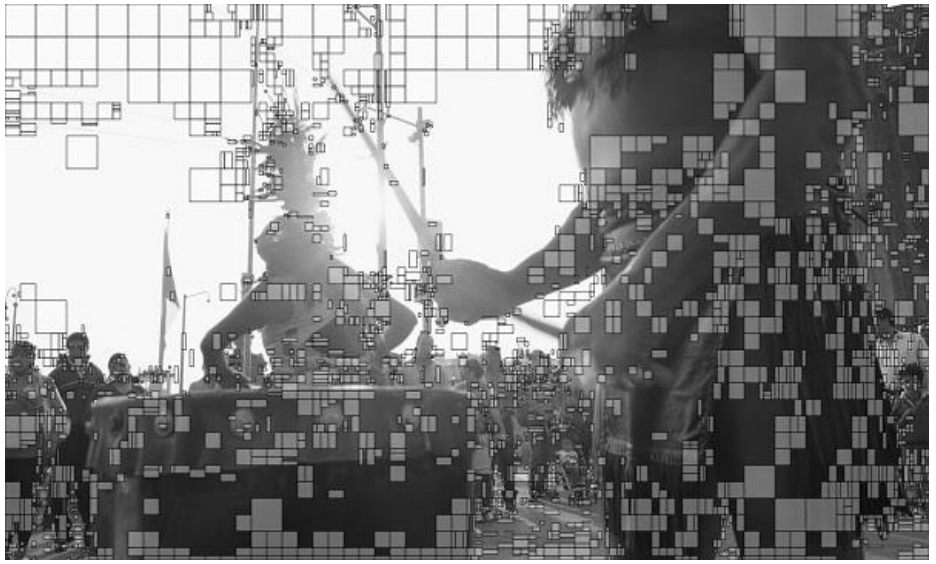
- The development of this work was strongly influenced by feedback gathered during the VVC standardization process, in order to satisfy hardware friendliness requirements, decoder complexity issues, and general implementation constraints.

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

Spatial distribution of DIMD blocks

in Tango (3840×2160)



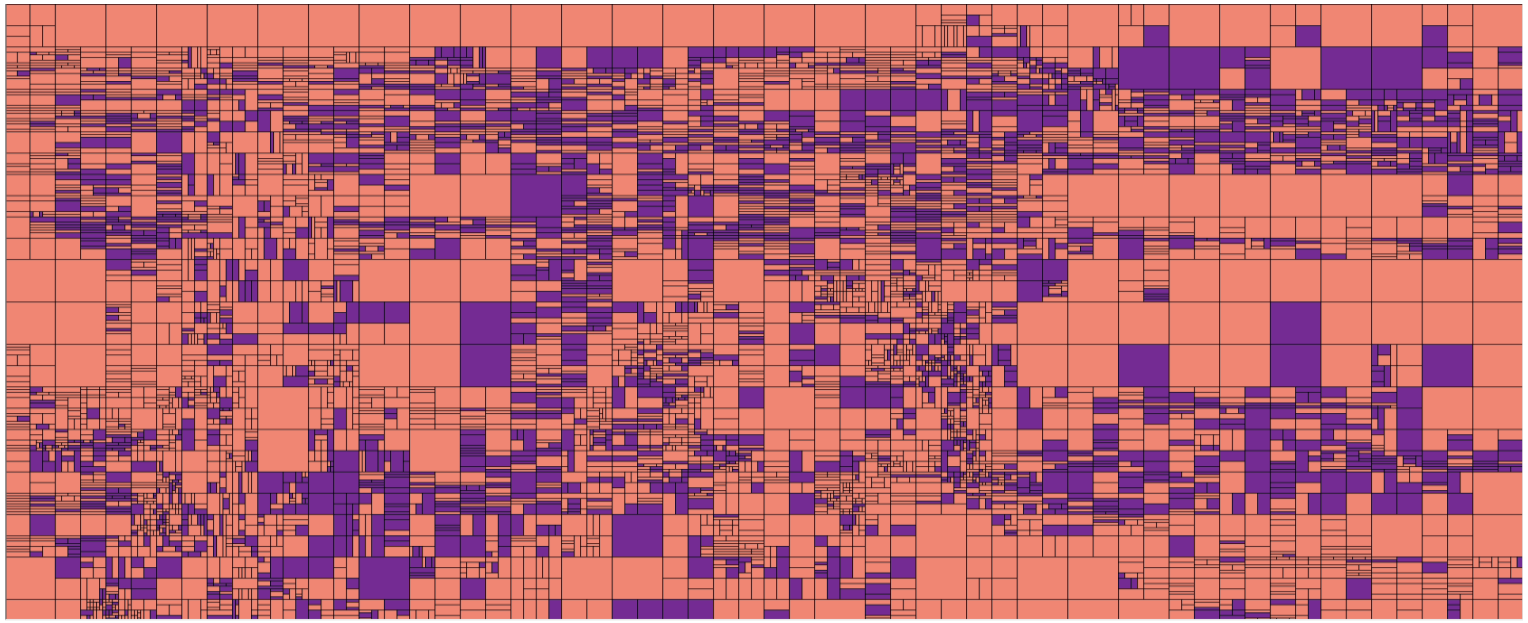
and RitualDance (1920×1080).

			Block width				
			4	8	16	32	64
Block height	QP22	4	36%	39%	36%	31%	–
		8	34%	32%	29%	24%	–
		16	32%	29%	27%	21%	–
		32	34%	29%	28%	24%	–
		64	–	–	–	–	17%
	QP37	4	41%	31%	34%	31%	–
		8	28%	27%	28%	26%	–
		16	21%	24%	26%	25%	–
		32	23%	23%	26%	26%	–
		64	–	–	–	–	28%

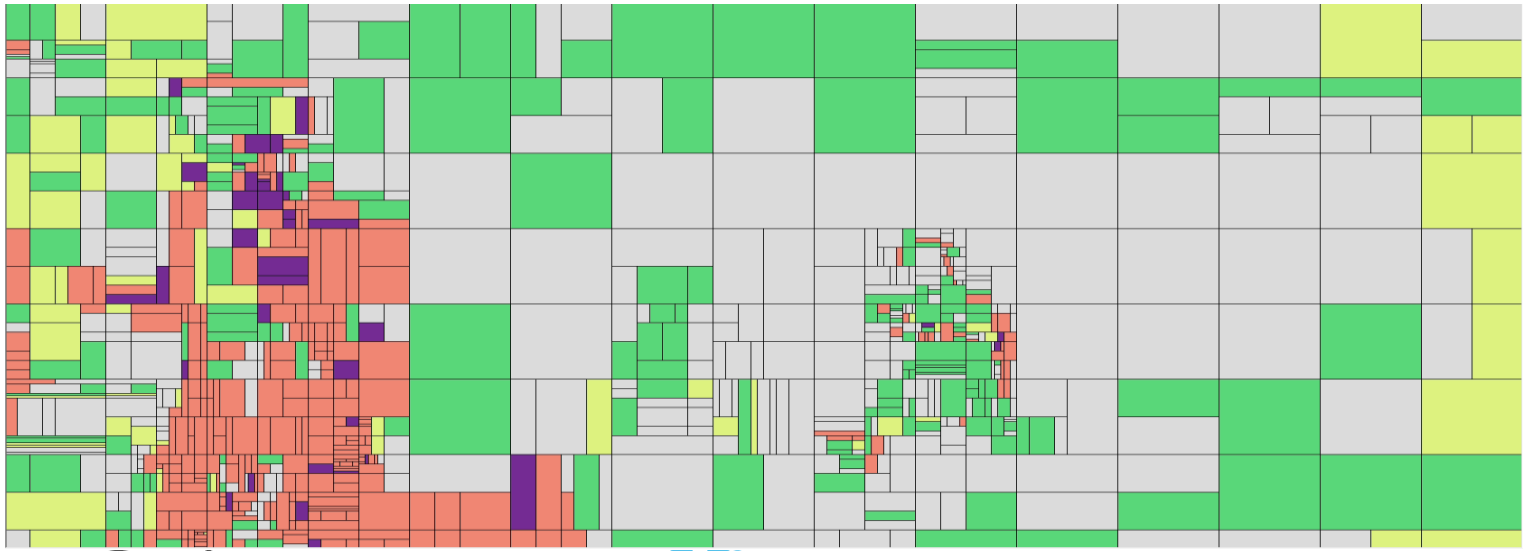
ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

Final selection DIMD mode



Block coded in DIMD mode
Block coded in intra mode



Block coded in Inter mode
Block coded in DIMD mode
Block coded in merge mode
Block coded in Intra mode
Block coded in skip mode

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

Class	Sequence	All Intra			Random Access			Low Delay B		
		BDR	ET	DT	BDR	ET	DT	BDR	ET	DT
A1	Tango2	-0.97%	108%	104%	-0.51%	102%	102%	-0.18%	103%	101%
	FoodMarket	-0.97%	108%	104%	-0.37%	102%	102%	-0.29%	102%	100%
	Campfire	-0.59%	109%	104%	-0.44%	102%	102%	-0.25%	103%	101%
	Average	**-0.84%**	**108%**	**104%**	**-0.44%**	**102%**	**102%**	**-0.24%**	**103%**	**101%**
A2	CatRobot	-0.55%	109%	103%	-0.26%	102%	102%	-0.17%	102%	101%
	DaylightRoad2	-0.43%	106%	104%	-0.20%	102%	102%	-0.26%	102%	101%
	ParkRunnig3	-0.47%	108%	104%	-0.17%	102%	102%	-0.32%	102%	101%
	Average	**-0.48%**	**108%**	**104%**	**-0.21%**	**102%**	**102%**	**-0.25%**	**102%**	**101%**
B	MarketPlace	-0.47%	107%	104%	-0.21%	101%	101%	-0.24%	102%	101%
	RitualDance	-0.85%	109%	104%	-0.43%	101%	102%	-0.29%	101%	102%
	Cactus	-0.51%	108%	104%	-0.34%	102%	102%	-0.28%	102%	101%
	BasketballDrive	-0.68%	109%	105%	-0.46%	102%	102%	-0.25%	102%	100%
	BQTerrace	-0.25%	108%	104%	-0.14%	102%	102%	-0.13%	102%	101%
	Average	**-0.55%**	**108%**	**104%**	**-0.31%**	**102%**	**102%**	**-0.24%**	**102%**	**101%**
C	BasketballDrill	-0.32%	109%	104%	-0.34%	102%	101%	-0.25%	101%	101%
	BQMall	-0.73%	108%	104%	-0.38%	102%	102%	-0.18%	102%	102%
	PartyScene	-0.61%	109%	105%	-0.36%	102%	102%	-0.23%	102%	101%
	RaceHorses	-0.70%	108%	105%	-0.37%	102%	101%	-0.27%	102%	101%
	Average	**-0.59%**	**109%**	**105%**	**-0.36%**	**102%**	**102%**	**-0.23%**	**102%**	**101%**
D	BasketballPass	-0.67%	109%	104%	-0.36%	102%	102%	-0.04%	102%	101%
	BQSquare	-0.42%	110%	104%	-0.18%	102%	101%	-0.29%	102%	101%
	BlowingBubbles	-0.89%	110%	104%	-0.25%	103%	102%	-0.29%	103%	102%
	RaceHorses	-0.68%	111%	104%	-0.28%	102%	101%	-0.14%	103%	102%
	Average	**-0.67%**	**110%**	**104%**	**-0.27%**	**102%**	**102%**	**-0.19%**	**103%**	**102%**
E	FourPeople	-0.74%	108%	104%	-0.41%	102%	102%	0.04%	102%	101%
	Johnny	-0.79%	108%	104%	-0.48%	103%	102%	-0.11%	103%	101%
	KristenAndSara	-0.72%	109%	104%	-0.47%	102%	103%	-0.45%	102%	101%
	Average	**-0.75%**	**108%**	**104%**	**-0.45%**	**102%**	**102%**	**-0.17%**	**102%**	**101%**
F	BasketballDrillText	-0.28%	111%	104%	-0.30%	102%	102%	-0.27%	102%	101%
	ArenaOfValor	-0.80%	111%	104%	-0.39%	102%	102%	-0.28%	102%	101%
	SlideEditing	-0.16%	110%	104%	-0.17%	102%	102%	-0.23%	102%	102%
	SlideShow	-0.13%	110%	104%	-0.24%	103%	103%	-0.21%	102%	101%
	Average	**-0.34%**	**111%**	**104%**	**-0.28%**	**102%**	**102%**	**-0.25%**	**102%**	**101%**
All		**-0.6%**	**109%**	**104%**	**-0.31%**	**102%**	**102%**	**-0.23%**	**102%**	**101%**

DIMD module	Encoder run-time overhead
Full RD-check	89%
Intra mode derivation	6%
Fusion	5%
Total encoder overhead	**100%**

DIMD module	Decoder run-time overhead
Intra mode derivation	28%
Fusion	72%
Total decoder overhead	**100%**

Conclusion

- This work represents an invention. This offers me the chance to write my first **patents (3)** declaring an invention and 2 publications (IEEE: DCC-Utah-USA and ICIP--Taipei-Taiwan).

- The gains presented and the accepted complexity added to the original codec encouraged us to present this technique as a **contribution** to the MPEG meeting.

- DIMD passed several steps and was so close from being added to the VVC standard, but unfortunately this did not happen finally. However, we have been contacted lately by QUALCOMM to have our DIMD in their codec.

- DIMD is being developed by my colleague/supervisor at ateme, while I moved to another project.

ateme × anevia
Stronger Together

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY

Thank You

TELECOM Paris
IP PARIS
INSTITUT POLYTECHNIQUE PARIS-SACLAY