UHD Panoramic Video Coding for Multi-Camera and Multi-Processor Acquisition Systems

João Duarte, João Carreira and Pedro Assunção

Instituto de Telecomunicações
Politécnico de Leiria / ESTG
Leiria, Portugal

UBICOMM 2019, Porto, Portugal, September 24, 2019

Summary

- Introduction
- Simulation study procedure
- Experimental results
- Conclusions and future work

Context and motivation

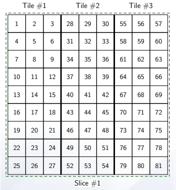
- High demand of Ultra High Definition (UHD) video captured by panoramic and omnidirectional cameras with wide Field-of-View (FoV).
- Improvements in the coding efficiency are mostly achieved at the expense of a great increase in computational complexity:
  - complex predictions;
  - high limitations using low-cost hardware.

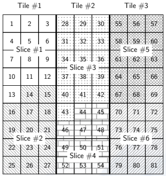
⇓

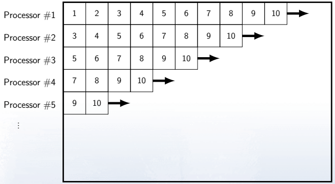
Take advantage of parallel processing to encode panoramic UHD video.

Related work

- Recent coding standard are already adapted for parallel processing.
- Parallel processing can be mainly achieved using:
  - Group of pictures and frame level;
  - Using frames divided into multiple slices or tiles;
  - Wavefront parallel processing (processing each row independently).







Problem formulation

Problem:

- Existing approaches require high memory storage or complex inter-process communication.
- What is the impact of using multi-processor architectures to compress narrow-FoV in the encoding times and compression efficiency?

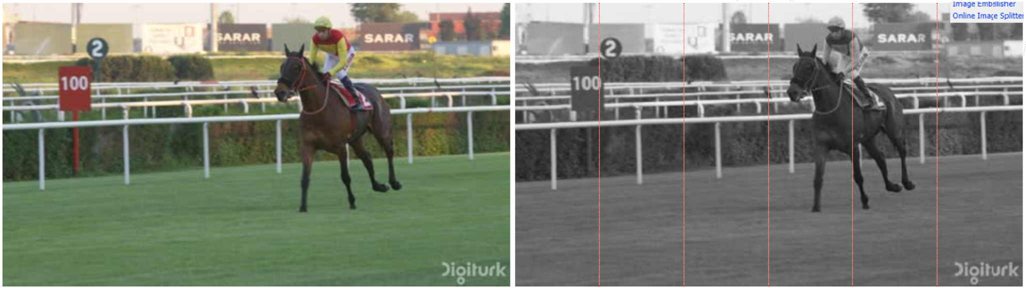
Solution:

- Study the use of multi-processor encoding under different FoV sizes and different coding parameters.

Encoding scheme

- Panoramic video acquired using multi-cameras;
  - partitioning the wide FoV into multiple narrow FoV videos.
- Use multiple encoders:
  - take advantage of parallel processing;
- Stitching operation can be performed as post-processing.

Field-of-View (FoV) partitioning



- Full-FoV UHD video frame (left) of 3840×2160 pixels.
- Partitioning into six FoVs of 640×2160 pixels each (right).

Simulation study procedure

FoV partitioning

- Six FoV partitioning schemes are adopted:

Type	Number of partitions	Spatial resolution of the FoV
Full-FoV	One	**3840**×2160
Sub-FoV	Six	**640**×2160
	Eight	**480**×2160
	Ten	**384**×2160
	Twelve	**320**×2160
	Fifteen	**256**×2160

- FoV partitioning is performed in the horizontal dimension.

Simulation study procedure

Test conditions

- UHD video sequences are encoded using the x265 software.
- Five different presents (0-ultrafast, 3, 5, 7 and 9-ultraslow):
  - control the partitioning schemes, slice types and search algorithms;
  - impact in the compression efficiency and encoding time.
- Four different Constant Rate Factors (CRFs):
  - control the Quantisation Parameter (QP);
  - impact in the reconstruction quality and bitrate.

Simulation study procedure

Test conditions

- Five UHD video sequences are used with a spatial resolution of 3840×2560;

Sequence	SI	TI	Description
Beauty	10.6	8.35	Very high spatial details in some regions (hair) and flat background
Bosphorus	13.4	3.75	Boat shipping at low motion with moderate complex background
HoneyBee	8.24	2.54	High spatial detail, with one low motion object
Jockey	11.5	16.2	High motion with one horse rider
ReadySteadyGo	18.0	19.0	Very high motion with several horse riders

- Each test condition is simulated a total of 20 times.

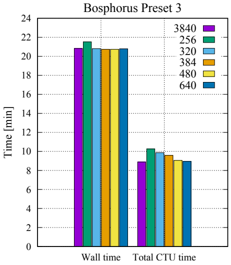
Simulation study procedure

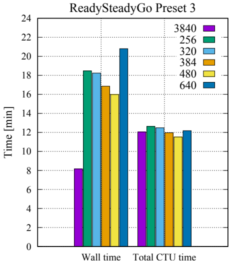
Studied variables

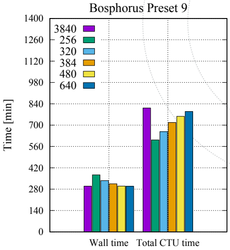
- Encoding times:
  - *Wall time:* difference between when the first CTU is ready to be compressed and the entire frame is outputed to the coded stream.
  - *Total CTU time:* the total time spent by working threads in compression and filtering operations of the CTUs of a given frame.
- Bitstream size and bitrate.
- Reconstruction quality using the PSNR metric.
- Results of each FoV are combined for a fair comparison.

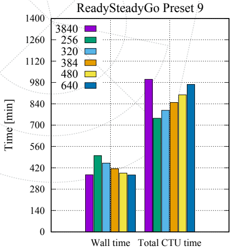
Encoding times evaluation

Encoding time for each FoV size









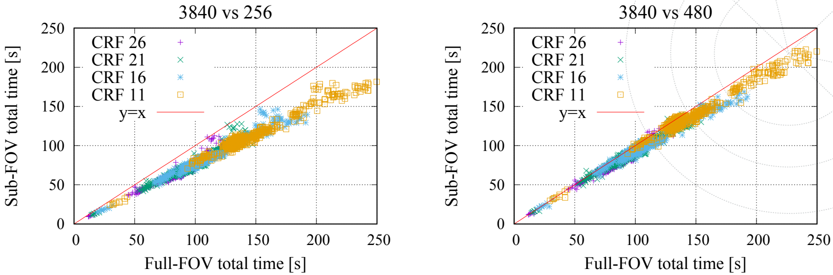
- Higher time variations for the sequence with higher motion (ReadySteadyGo).

instituto de telecomunicações

IPL instituto politécnico de leiria

Encoding times evaluation

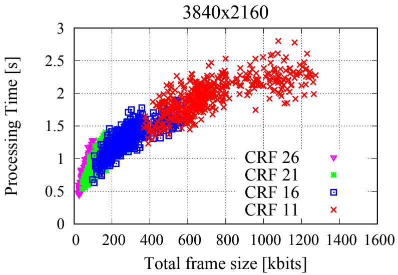
Full-FoV versus sub-FoV

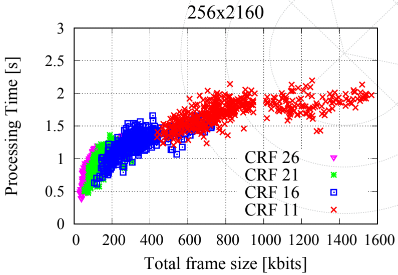


- These results reveal that quite linear correlation exists between the two encoding times.
- The overall time produced by the sum of sub-image videos is slightly lower than the single image video.

Coded frame size evaluation

Frame size versus encoding time

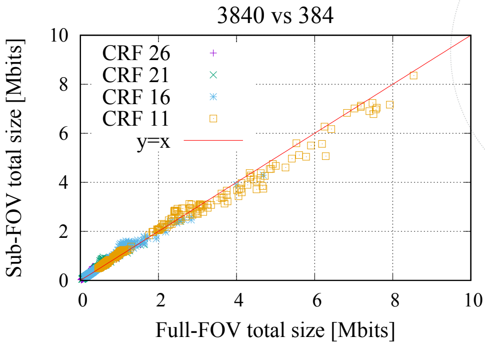




- Results reveal that video frames requiring higher amount of coded bits take longer time to be processed:
  - more coding modes are tested to achieve an efficient rate-distortion trade-off.
- Decreasing the CRF (i.e., lowering the QP) leads to higher encoding time.

Coded frame size evaluation

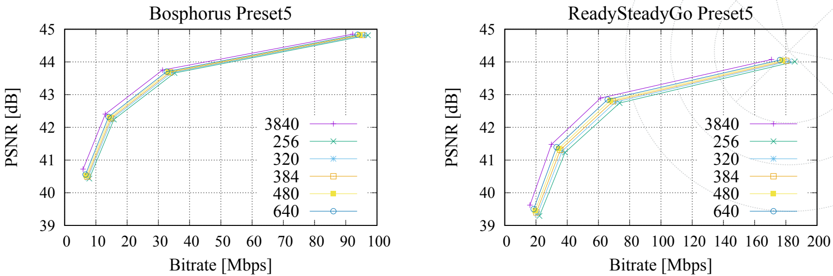
Full-FoV versus sub-FoV



- Encoding panoramic video using narrow FoVs videos leads to approximately the same amount of bits than encoding as a single video (Full-FoV).

Coding efficiency evaluation

Rate-distortion analysis



- Increasing the number of FoV partitions results in lower average video quality.
- For lower motion video the quality decreasing is less significant, revealing that this encoding approach does not have great impact in the overall performance.

instituto de telecomunicações

IPL instituto politécnico de leiria

Conclusions

- A coding scheme using for UHD panoramic video encoding was devised.
- The performance was evaluated using different sub-sequences of narrow FoV:
- Results show that the rate-distortion performance of multi-processor systems is quite similar to single processor one:
  - allows to distribute the huge computational requirements of HEVC encoding across several low-cost processors.
- This simulation study provides relevant insights on UHD panoramic video acquisition and coding systems using multiple cameras and processors.

Future work

- Extend this research from panoramic to omnidirectional video with ultra-high resolution.
- Evaluate this scenario using the upcoming Versatile Video Coding standard.
- Use machine learning techniques to further reduce the computational complexity.
- Combine multiple encoders architectures with scalable coding schemes.

Application example

Example of multiple low-complexity acquisition system

Professional omnidirectional camera

Thanks for your attention!

João Carreira