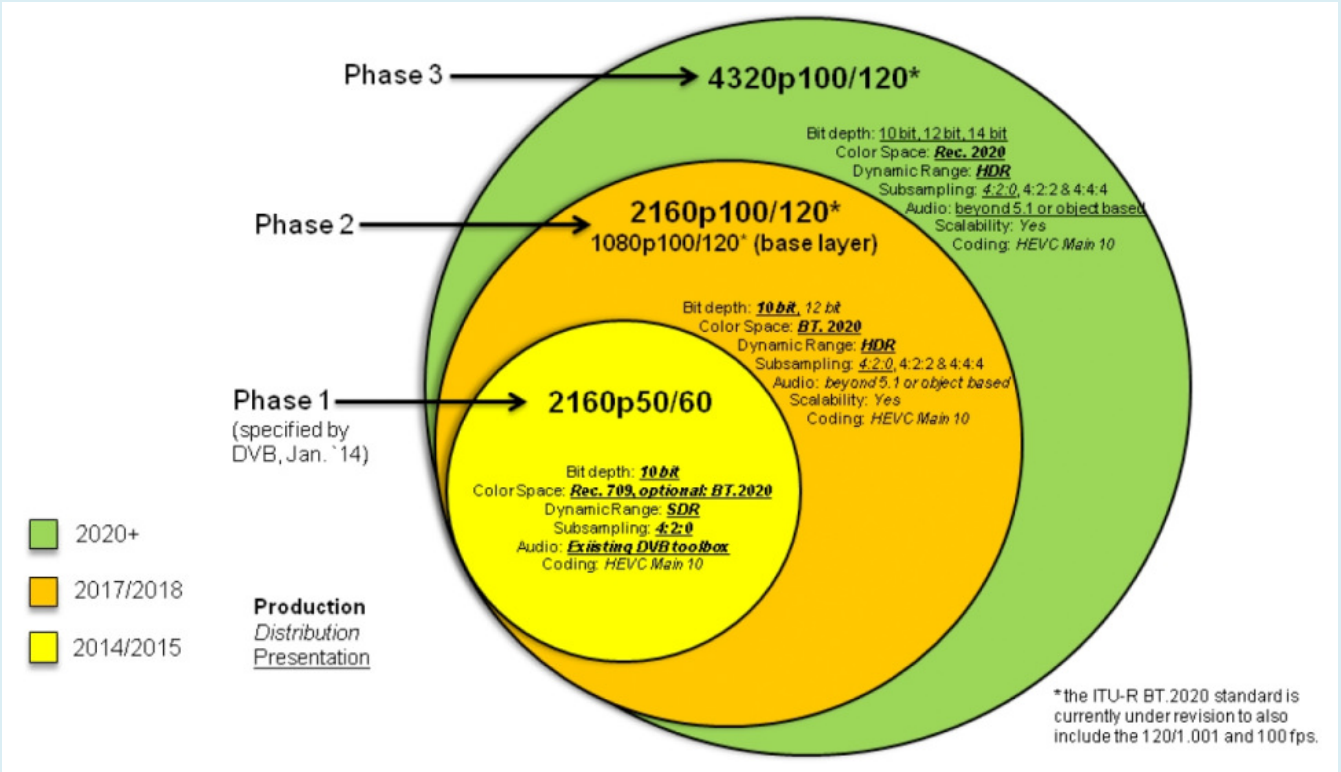
Machine Learning driven Variable Frame-Rate for Production and Broadcast Applications

VidTrans20, Benoît Le Ludec

b com

VSF
VIDEO SERVICES FORUM

Introduction : VFR for what ?



B-com/BLL , VidTrans20

VSF
VIDEO SERVICES FORUM

VFR versus HFR

- Variable Frame Rate is naturally applied to High Frame Rate (100/120fps and above) in order to reduce HW requirements to improve QoE
- But VFR can also be used in the case of standard Frame (50/60fps)

High Frame Rate context

- Benefits of HFR:
  - Better Quality of Experience coupled with UHD resolution
  - Increase of the perceived video quality, motion fluidity

- But
  - Increase of the encoder complexity (around 40%)
  - Increase of the compression bit-rate (6 to 7%)
  - Increase of the storage capacity (+100%)

VFR applied to HFR

- To face these challenges, VFR is a powerful tool that allows to use the minimum frame rate necessary to maintain the quality of experience

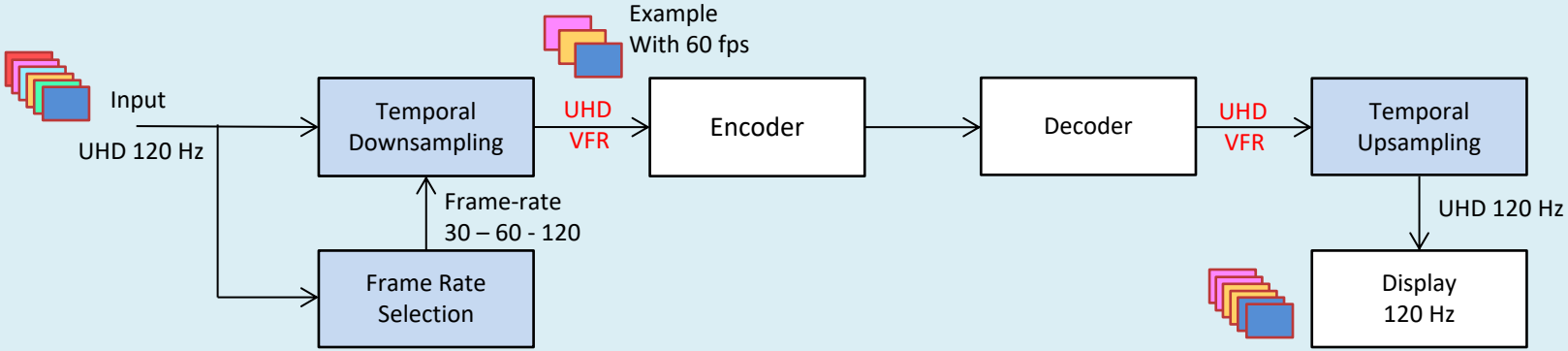
- **The problem boils down to a content aware determination of the minimum frame-rate**

System level considerations

- Must be compatible with live broadcast, content production
- Must have no visual impact

→The proposed approach is a real time system capable of

  - Frame decimation as a temporal downsampling
  - Frame duplication as a temporal upsampling



B-com/BLL , VidTrans20

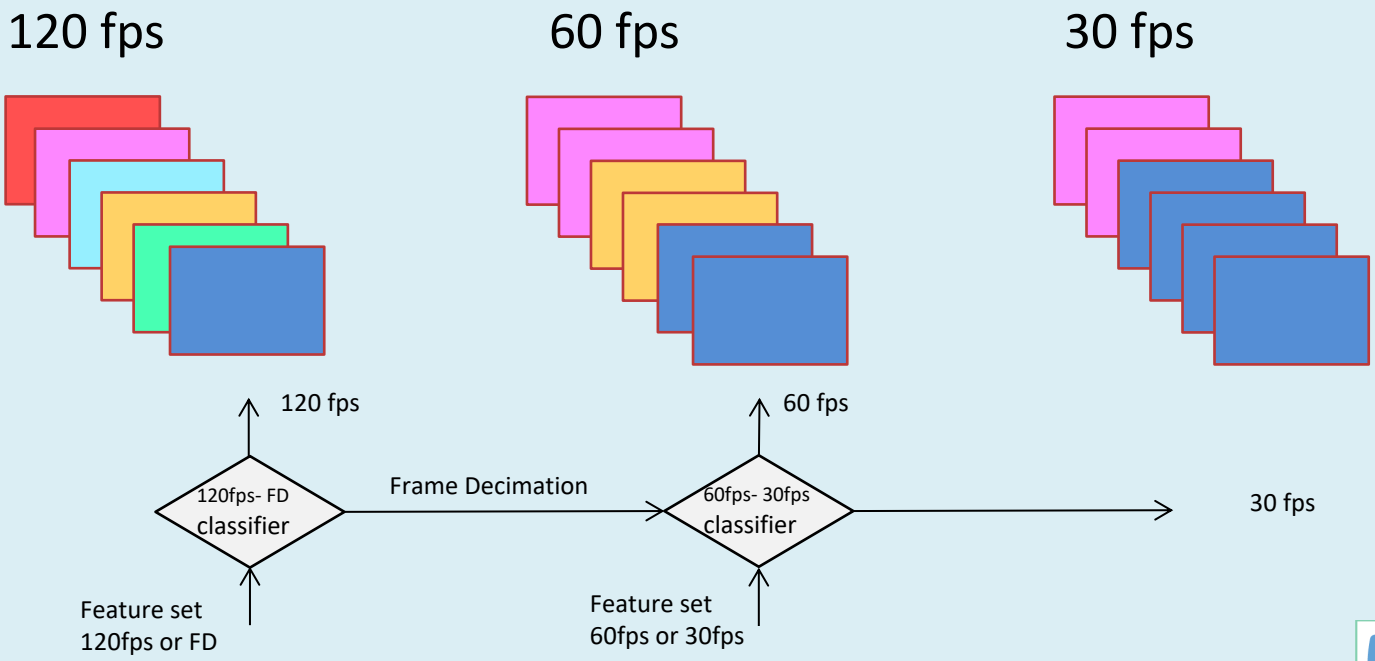
VSF
VIDEO SERVICES FORUM

Critical Frame-Rate Prediction

- The problem to be solved is a classification of the frame-rate F:
  - F or F/2 or F/4
- Machine Learning is used to predict the critical frame-rate:
  - 2 binary Random Forest classifiers to predict the critical frame-rate
- Definition of 2 feature-sets to feed the classifiers:
  - Feature set to select F or Frame Decimation (FD)
  - Feature set to select F/2 or F/4

VFR classification

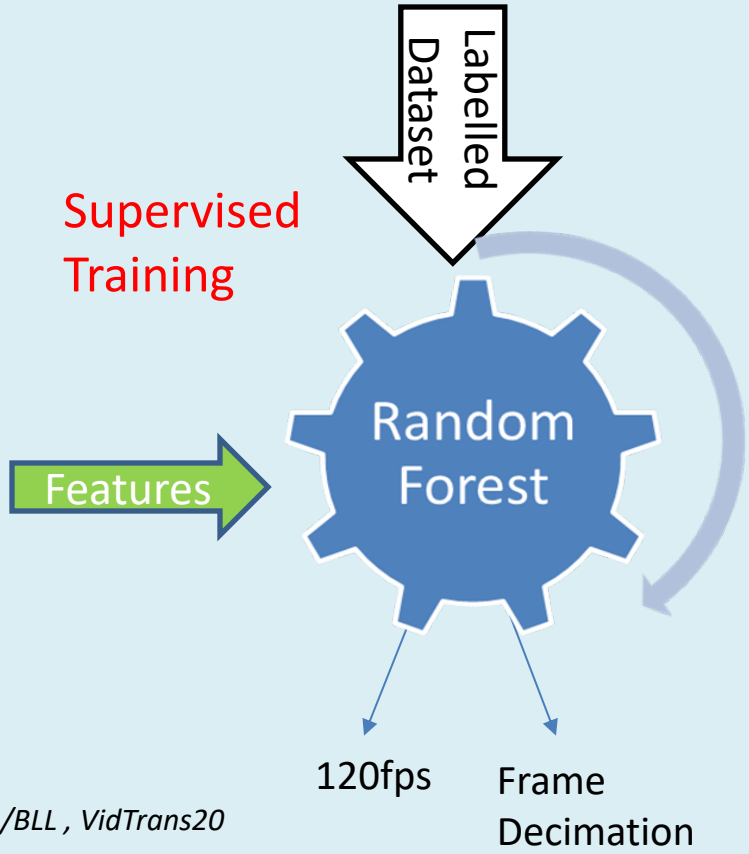
- Definition of 3 classes of VFR frame-rate (example with F= 120fps):

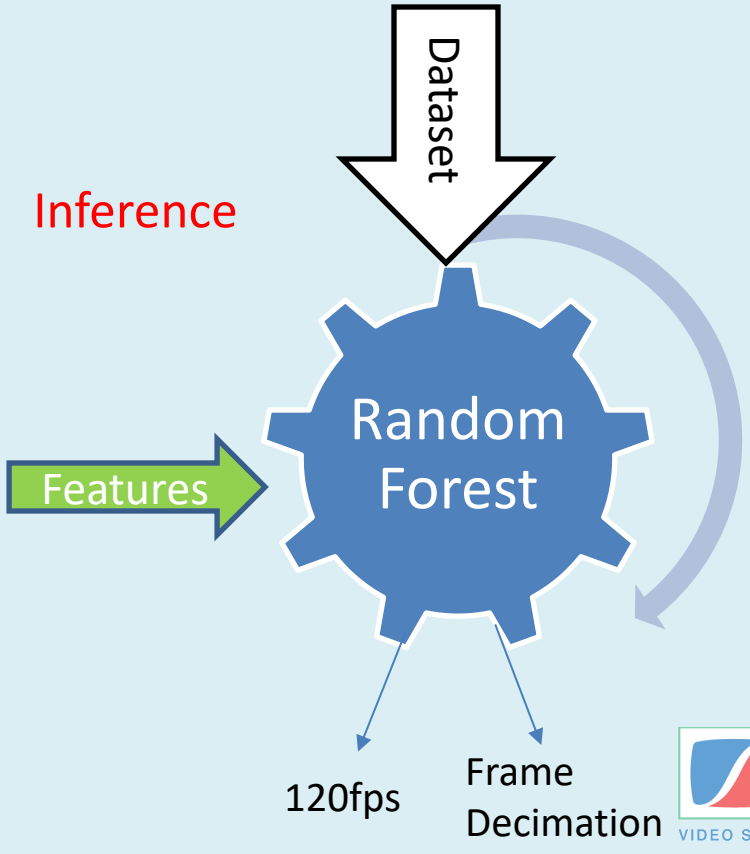


Methodology with F=120fps

- Creation of 2 dataset sequences to train the RF classifiers:
  - One dataset for 120fps-FD classifier
  - One dataset for 60fps-FD classifier
  - → HFR database: 360 video clips have been manually annotated, equivalent to 50000 chunks that have been individually labeled.
- Extraction of the features set gathering different metrics:
  - Motion vector (movement)
  - Pixel luminance (flickering)
  - Etc..
  - → A set of 32 features
- Training of the model with the datasets and the features

Training and Inference





B-com/BLL , VidTrans20

VSF
VIDEO SERVICES FORUM

Random Forest Characteristics

- The Random Forest has been optimized for each classifier:

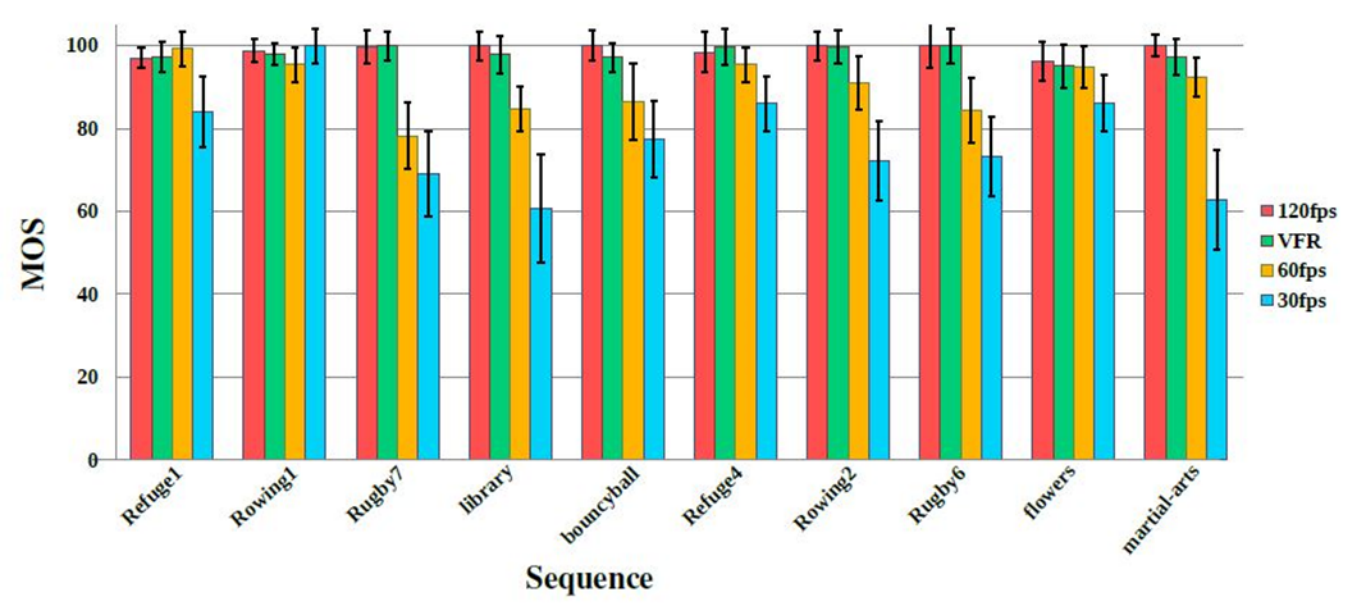
Classifier	120fps-FD	60fps-30fps
Number of features	26	13
Number of trees	200	100
Tree depth	7	7

Validating the trained Random Forest model

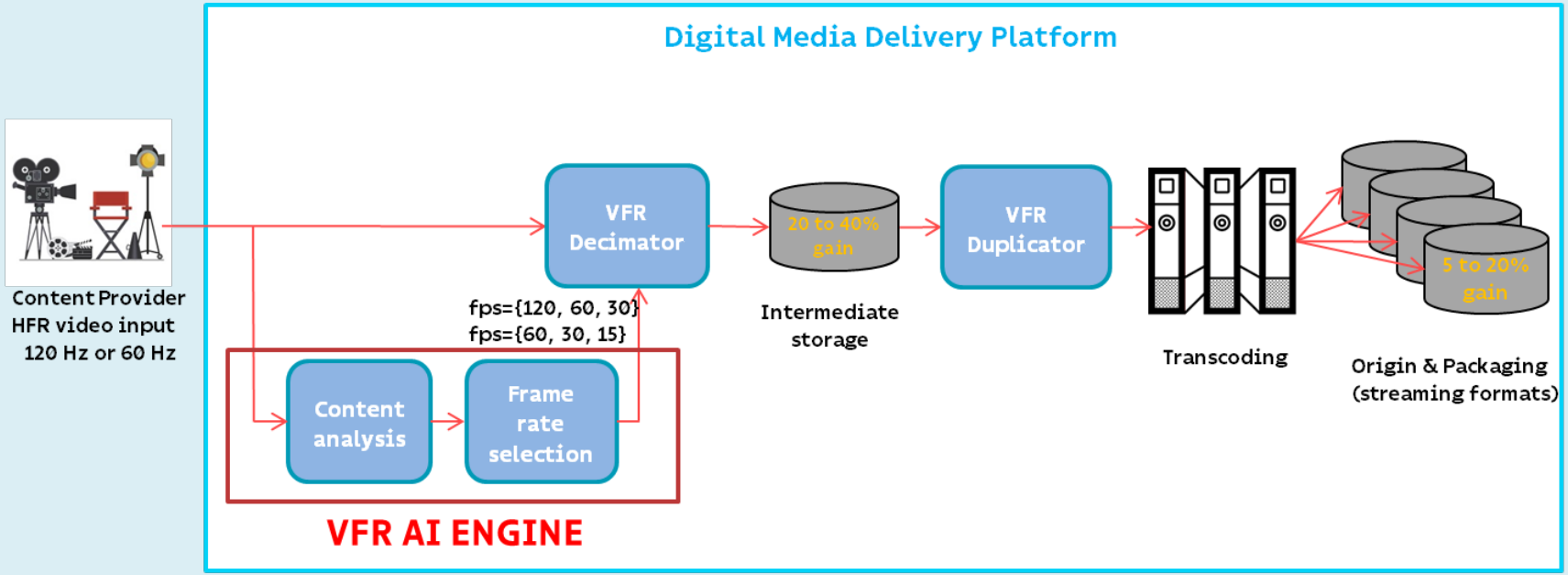
- Validation of the Random Forest models with sequences unknown from the model:
  - 15 sequences at 120 fps frame-rate
  - Duration between 9 and 13 seconds

Subjective visual quality of the VFR

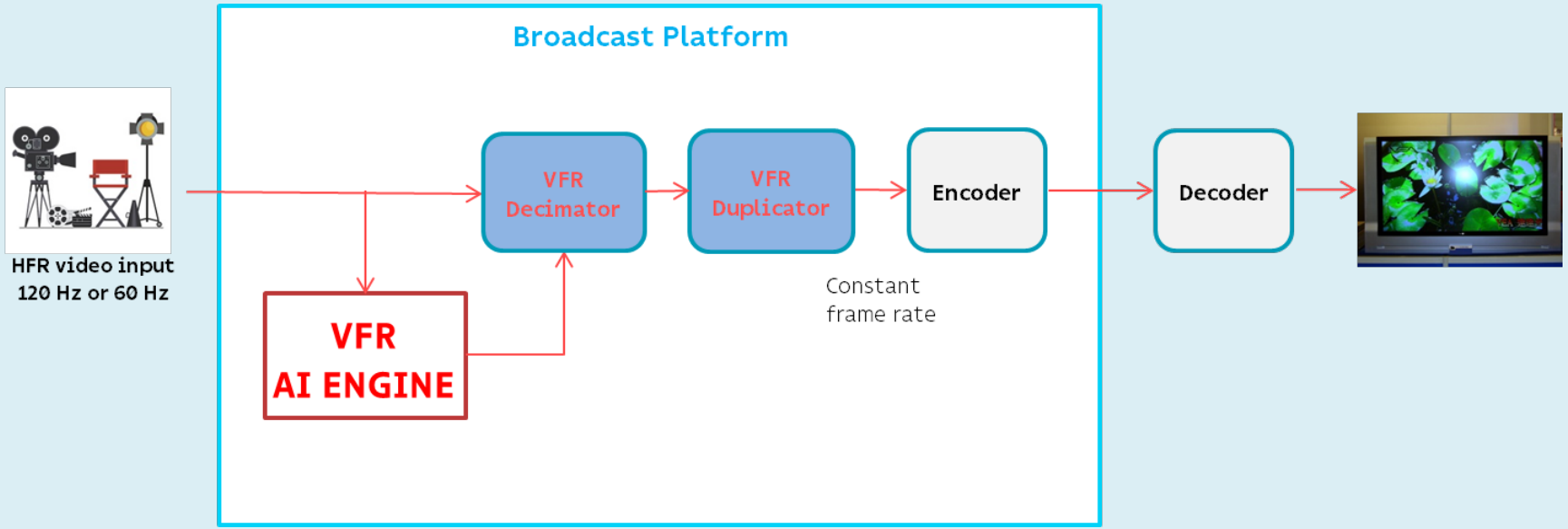
- 19 participants
- Criteria: MOS (Mean Opinion Score)



Production use case



Broadcast use case



B-com/BLL , VidTrans20

VSF
VIDEO SERVICES FORUM

Performance in HFR 120fps

- Performance in UHD resolution between HFR 120fps and VFR (120-60-30):

Broadcast use case		
VFR 120-60-30fps Compared to UHD 60fps	**HFR 120fps +encoder**	**HFR+AI engine+ VFR decimator/duplicator +encoder**
Encoding Complexity*	+40%	< 10%
Bit-rate*	+5 to +20%	<10%

Production use case		
VFR 120-60-30fps Compared to UHD 60fps	**HFR 120fps +encoder**	**HFR+AI engine+ VFR decimator/duplicator +transcoder**
Storage after transcoding	+5 to +20%	< 10%
Intermediate storage	+100%	< 60%

B-com/BLL , VidTrans20

VSF
VIDEO SERVICES FORUM

Performance in 60fps

- Performance in UHD resolution between 60fps and VFR (60-30-15):

Broadcast use case	
VFR 120-60-30fps Compared to UHD 60fps	**HFR+AI engine+ VFR decimator/duplicator +encoder**
Encoding Complexity*	-20% to -30%
Bit-rate*	-5% to -20%

Production use case	
VFR 120-60-30fps Compared to UHD 60fps	**HFR+AI engine+ VFR decimator/duplicator +transcoder**
Storage after transcoding	-5% to -20%
Intermediate storage	-20% to -40%

B-com/BLL , VidTrans20

VSF
VIDEO SERVICES FORUM

Conclusion

- Machine Learning is a major tool to improve Quality of Experience without dramatically increasing computing resources or program bit-rate
- SW implementation allows to process HD in real time at 120fps
  - UHD implementation in progress
  - FPGA implementation under consideration
- Several patents applied
- Demonstration at NAB 2020
  - With an exclusive preview here at VidTrans20 !!

Thank you for your attention!

Benoît Le Ludec – HW Engineering manager – benoit.leludec@b-com.com

- These persons have contributed to the study:

  Jean-Yves Aubié and Franck Chi {b<>com}

  Glenn Herrou, Charles Bonnineau {b<>com, INSA Rennes}

  Wassim Hamidouche and Luce Morin {b<>com, INSA Rennes}

  Patrick Dumenil {b<>com, Harmonic Inc.} Jérôme Fournier {b<>com, Orange Labs}

- Conference Reference:

  Jean-Louis Diascorn (Harmonic Inc): SMPTE 2019 “How AI Technology is Improving Video Compression for Broadcast & OTT Content Delivery”

b<>com
1219 avenue des Champs Blancs
35510 Cesson-Sévigné / France

VIDEO SERVICES FORUM