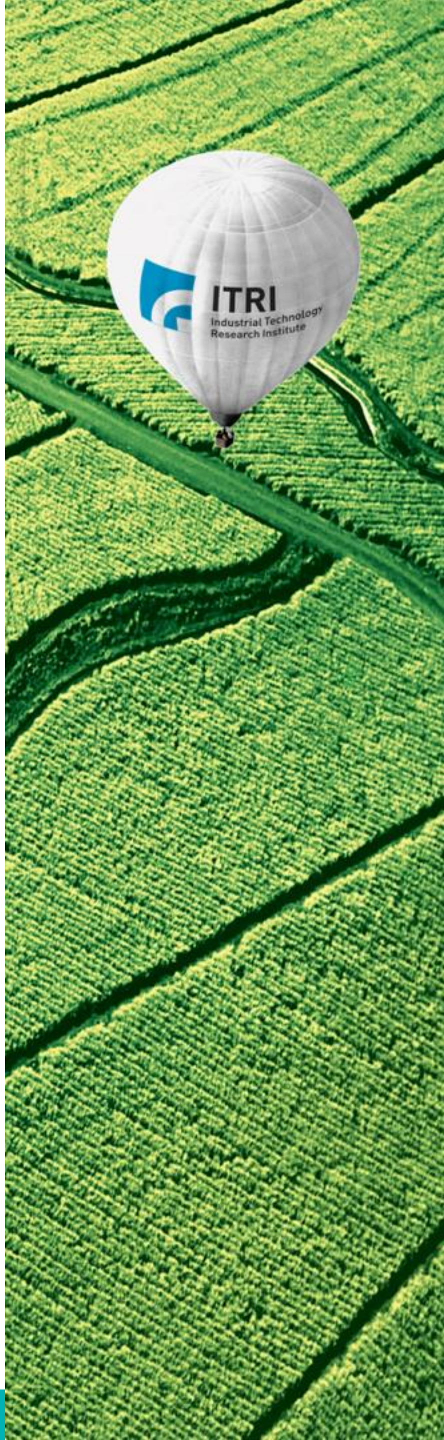
ITRI
Industrial Technology
Research Institute



MPEG標準會議分享

王聖博

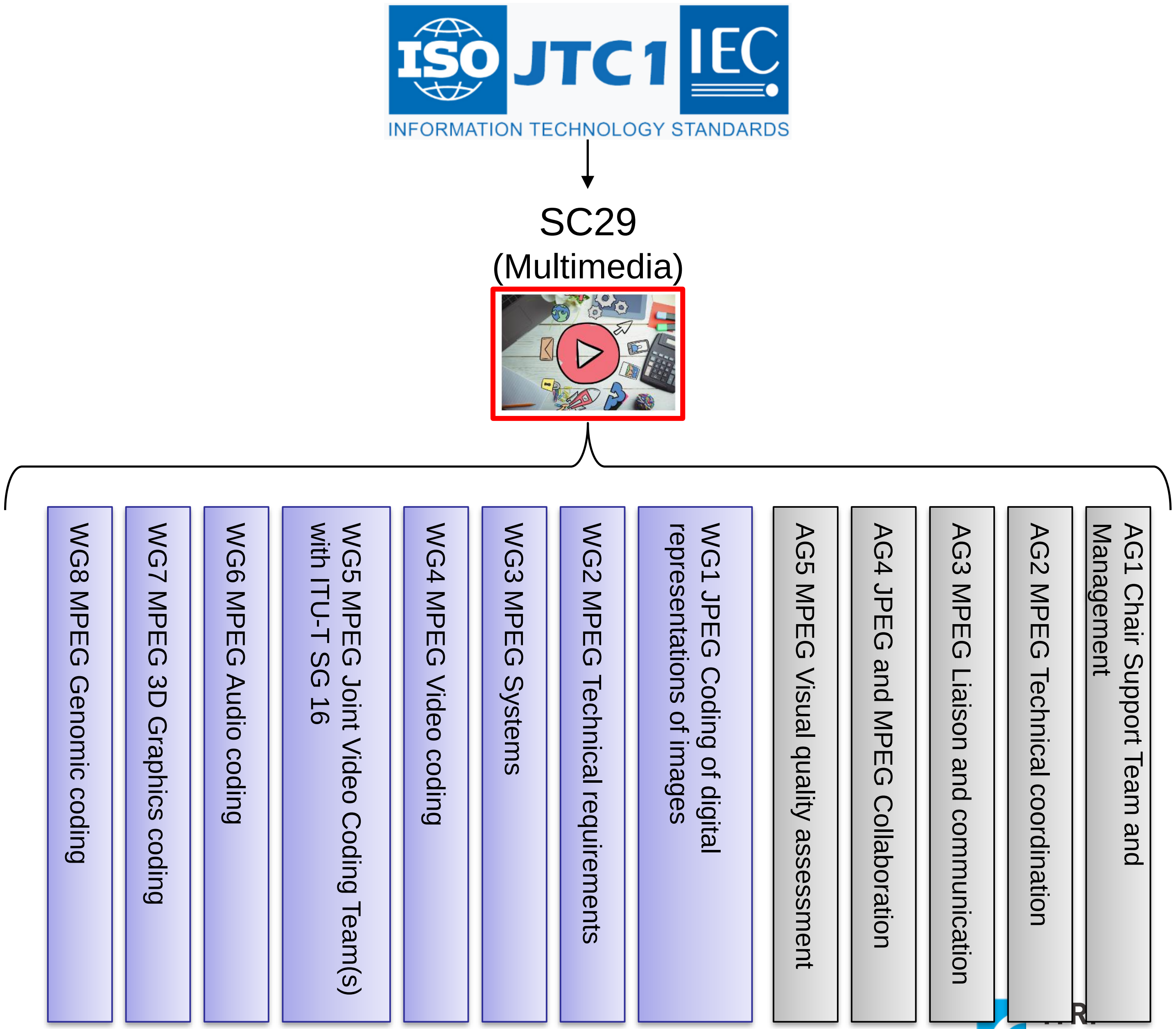
29th October 2021

Outline

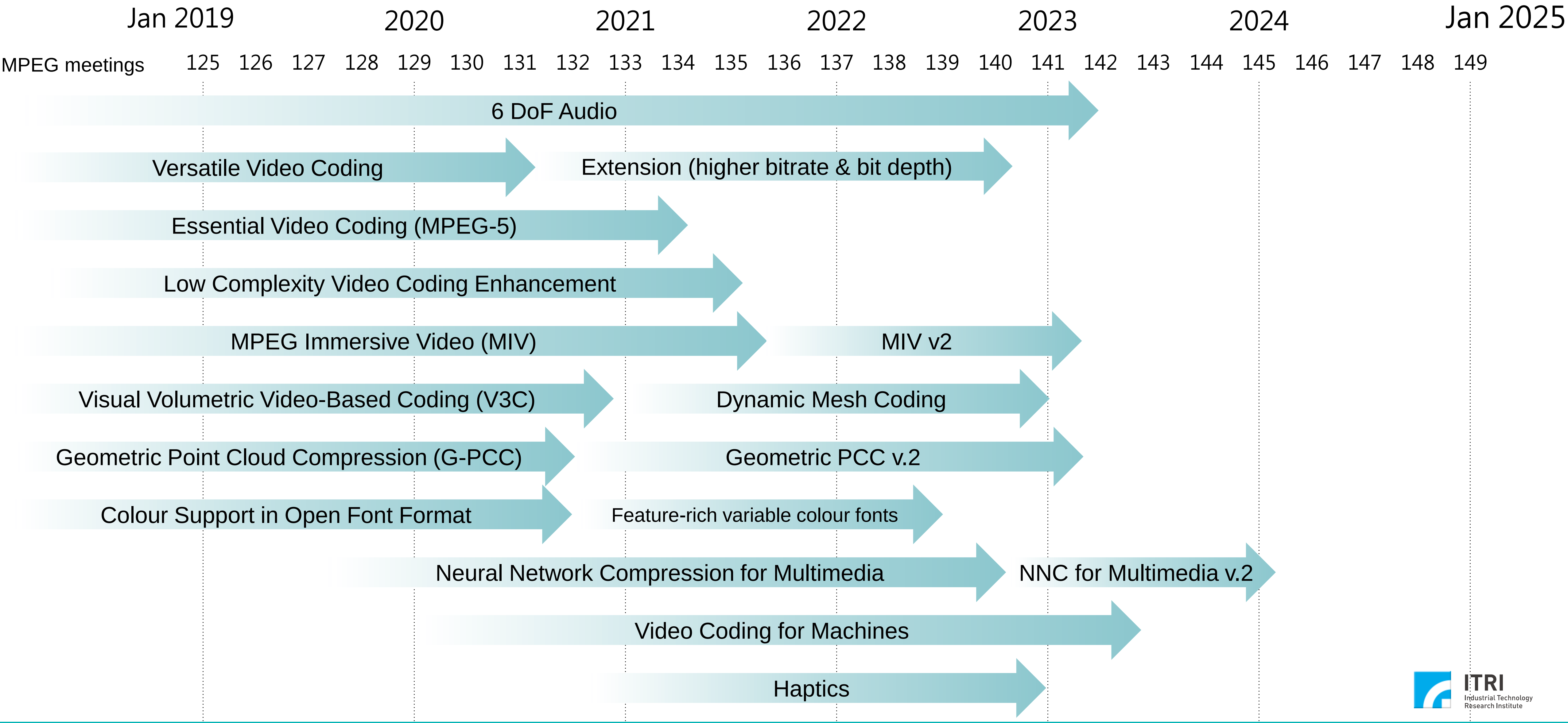
- Introduction of MPEG
- 3D Graphics Coding (3DG)
- MPEG Video Coding
- Video Coding for Machine (VCM)

MPEG標準組織架構

- MPEG(Moving Picture Experts Group)
  - ISO/IEC Joint Technical Committee 1, Subcommittee 29
  - 5個Advisory group與8個working group
    - Working group 針對特定主題進行技術討論與標準制定工作
    - Advisory group 負責組織管理、聯繫與共通性議題的處理
  - 每年開會四次，最近七次會議皆採遠端會議形式
  - 預計2022/04開始採用遠端/實體複合式會議的形式舉辦



MPEG Roadmap



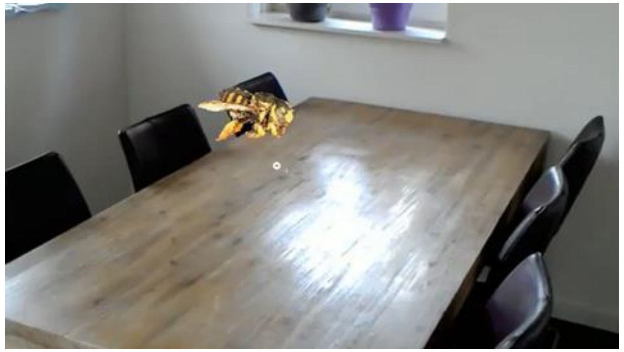
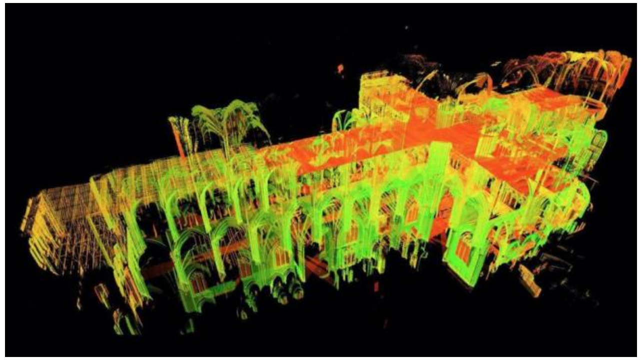
ITRI Industrial Technology Research Institute

Source: m55611 Draft Roadmap after MPEG 132

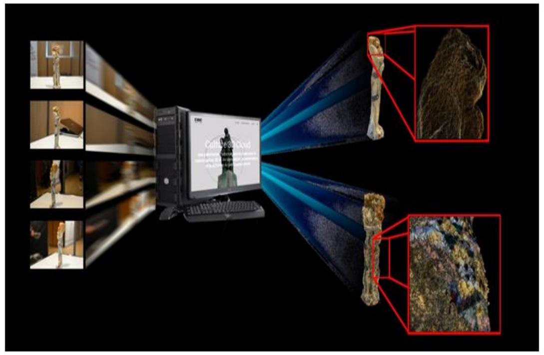
3D GRAPHICS CODING (3DG)

Applications and use cases

- Real-time telepresence, VR, Sport Broadcasting, Cultural heritage, Geographic Information, Autonomous Navigation



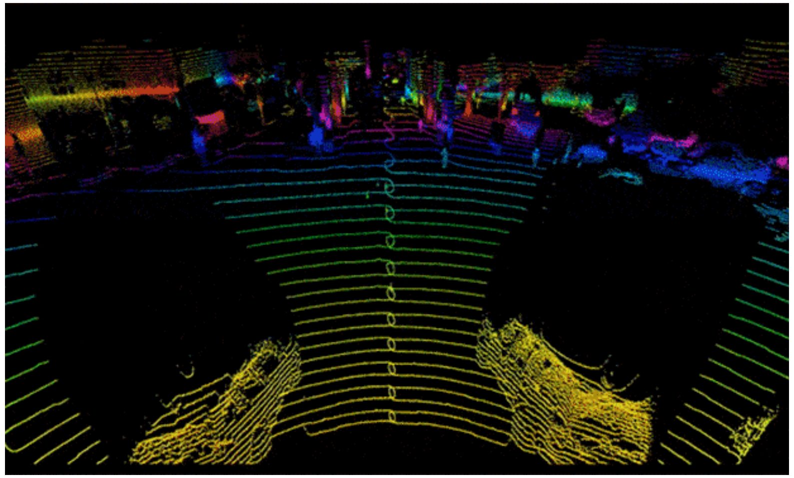


ITRI
Industrial Technology Research Institute

MPEG 3DG Report

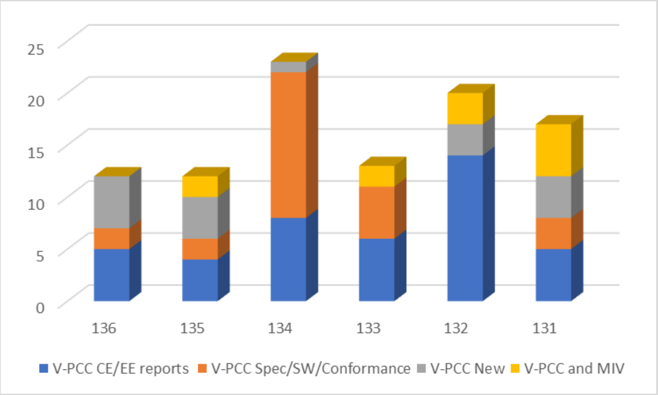
Graphics Compression

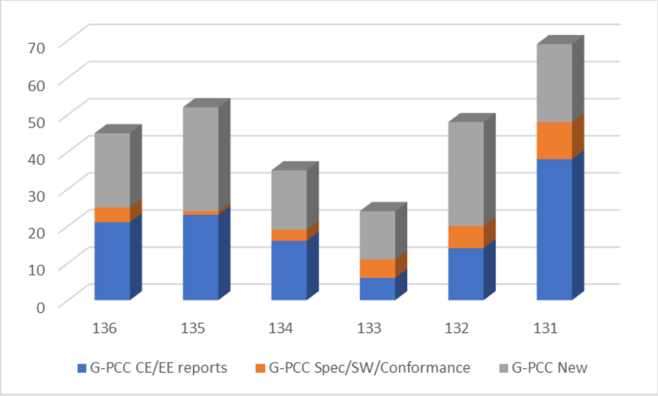
V-PCC

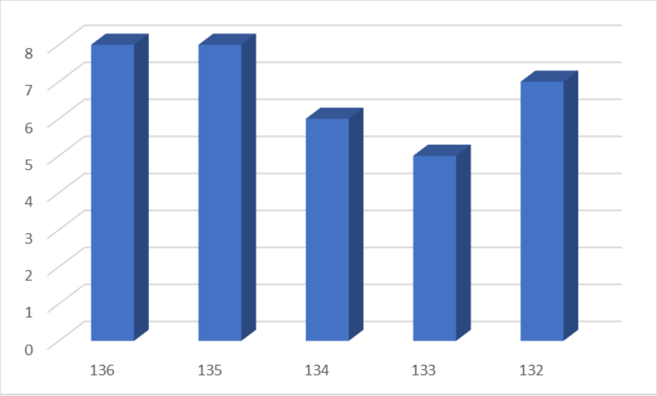


G-PCC

Dynamic Mesh

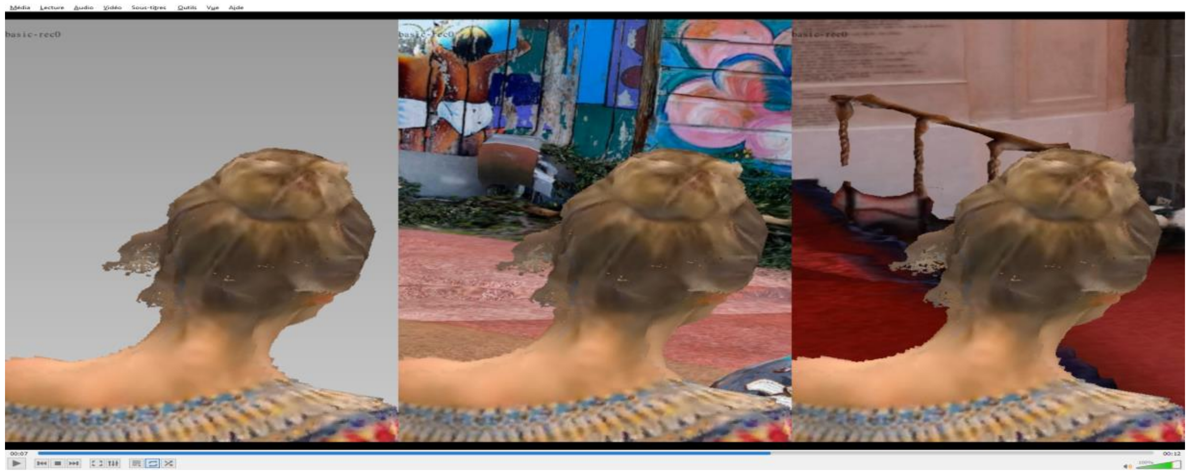






V-PCC activities

- V-PCC related activities
  - RDO model, parallel processing
  - Spatially Scalable Video-Based Point Cloud Compression
  - Evaluation of VVC coding tools for V-PCC software
  - Conformance and Ref. S/W on track (promoted DIS)
  - V-PCC Verification Tests



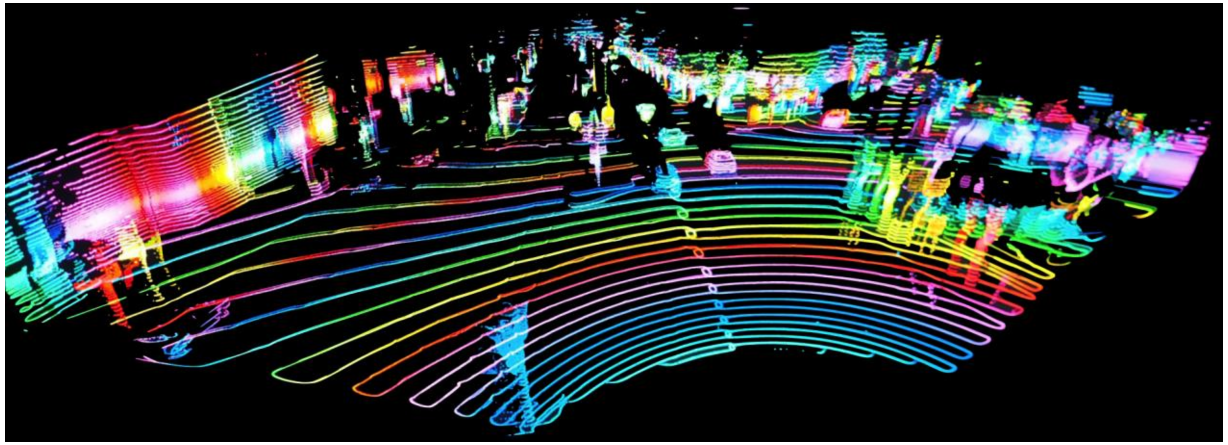
Dynamic Mesh Coding

- Previous mesh standards did not take into account time varying attribute maps and connectivity information
- It could reuse the very promising framework set up by V-PCC that is heavily relying on video-compression
- Video-based mesh (V-mesh) ➔ Dynamic Mesh Coding CfP
- CfP issued at October 2021!
- Plan to review responses at April 2022 and finalize standardization in 2024

ITRI Industrial Technology Research Institute

G-PCC main topics

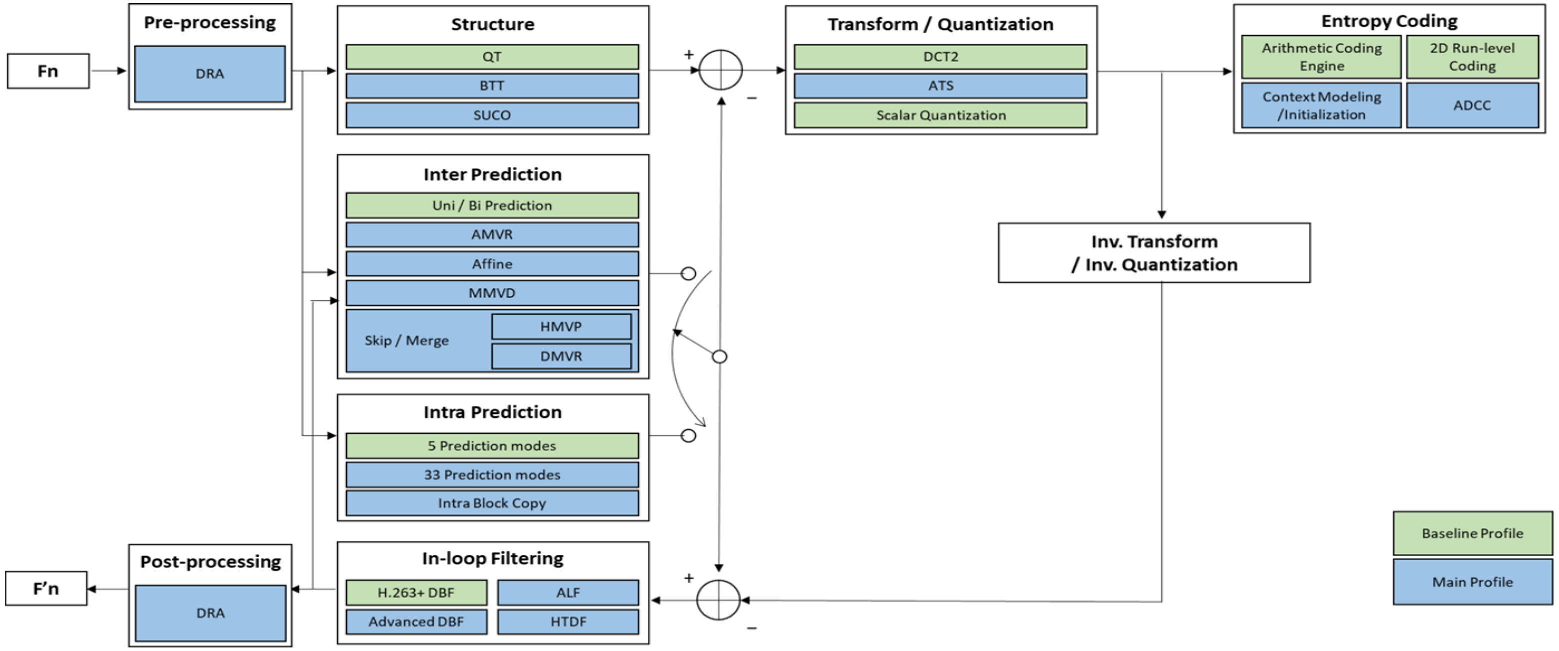
- Activities in G-PCC
  - Revision of the FDIS Editors comments
  - Conformance for G-PCC
- Technical topics
  - Inter-frame coding
  - Point cloud coding with very low complexity and latency
  - AI technologies for point cloud coding



MPEG Video Coding

Essential Video Coding

- Goal:
  - Encourage timely publication of licensing terms to allow reliable business plans to be created
  - Coding efficiency at least as good as HEVC
  - Complexity suitable for practical real time encoding
- Profiles:
  - Baseline profile: Only royalty free technologies
  - Main profile: Small number of additional tools, each can be switched off independently



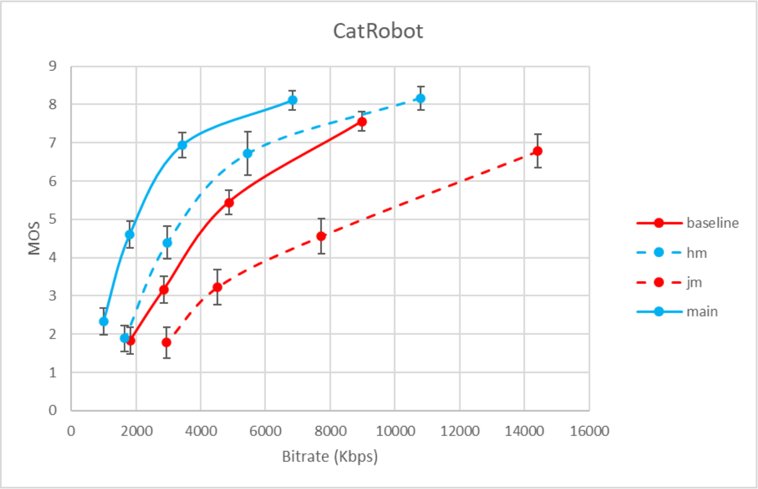
Source: MDS20735 Draft of White paper on Essential Video Coding (EVC), ISO/IEC SC29 WG4 Online July 2021

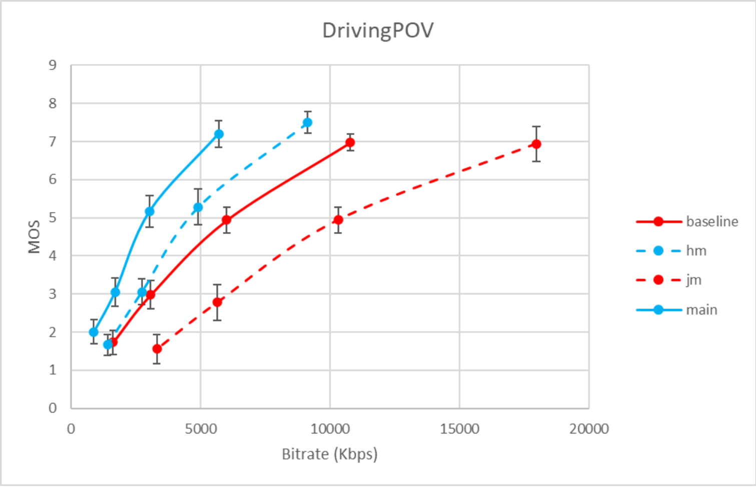
Essential Video Coding

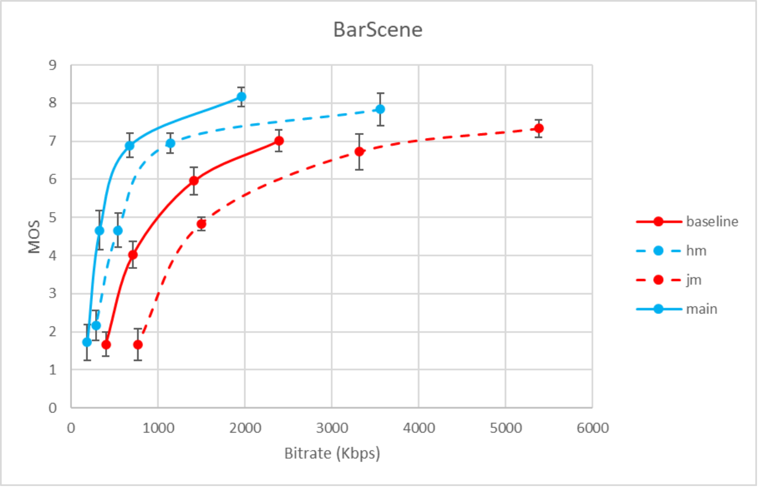
- FDIS on conformance and reference SW
  - ETM 7.3 as reference SW
  - Public available of reference SW and conformance bitstreams
- Implementation & market support
  - Real-time Mobile Video Player based on ETM 6.1 and optimized for ARM (by Solveig Multimedia & Huawei)
    - Tested on Huawei P40 Pro
    - 30+ FPS on 1080p
  - Zond 265 video bitstream analyzer include EVC

Summary of EVC Verification Test Results

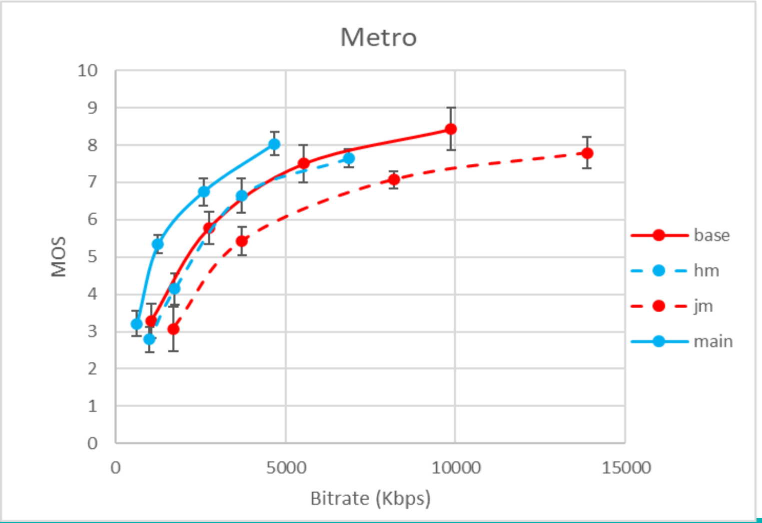
Random Access, UHD

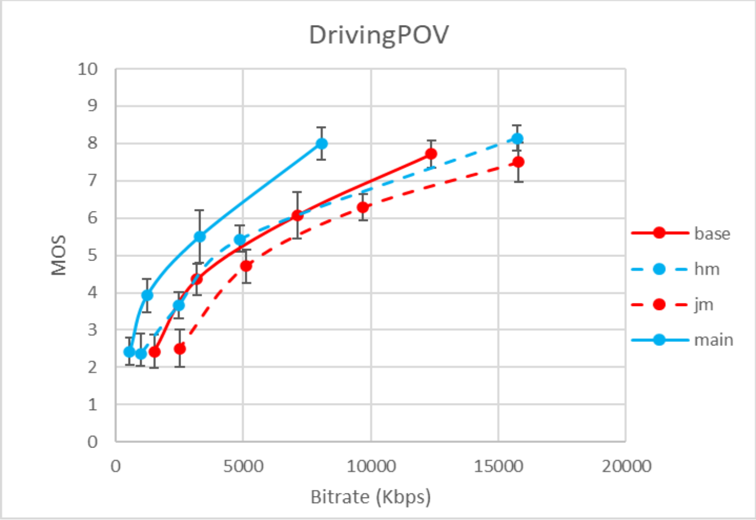


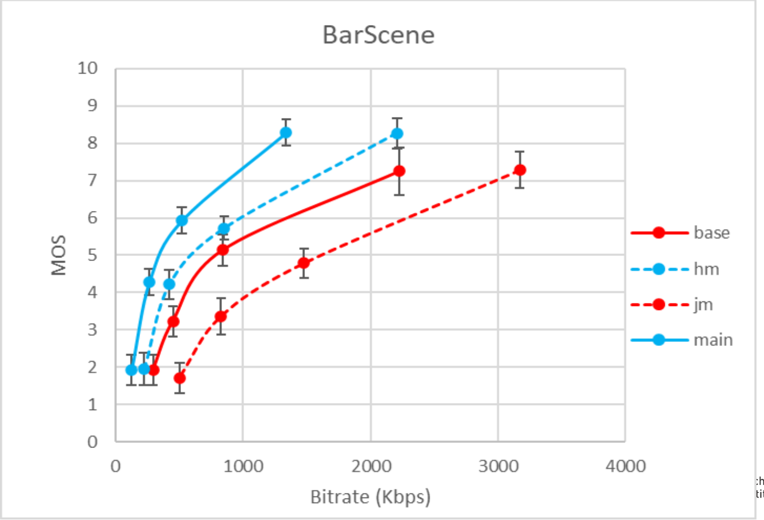




Low delay, HD







Source: w20000 Report on Essential Video Coding compression performance verification testing for SDR content, ISO/IEC SC29 WG4 Online January 2021

Summary of EVC Verification Test Results

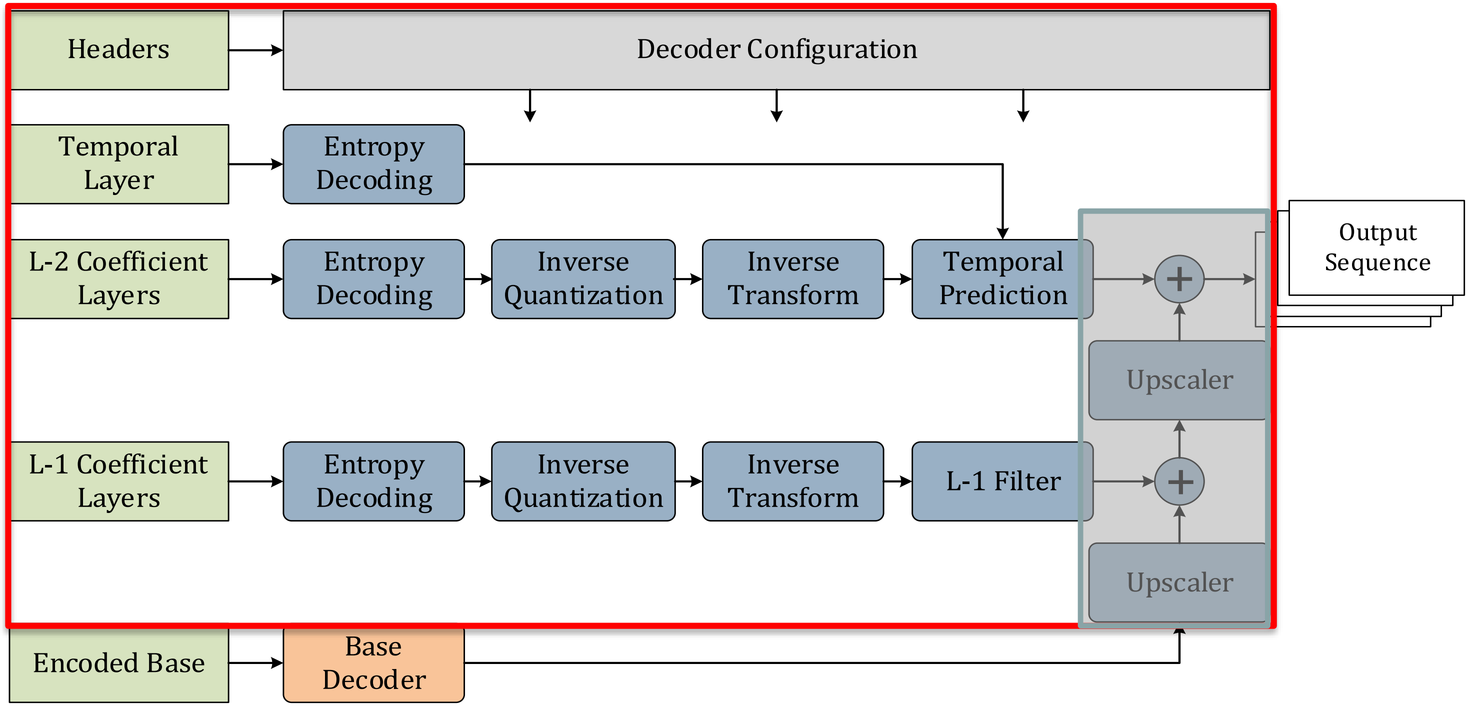
- EVC Main profile vs. HEVC Main10 profile:
  - SDR-RA (UHD): 39% gain
  - SDR-LD (HD): 41% gain
- EVC Baseline profile vs. AVC High10 profile
  - SDR-RA (UHD): 39% gain
  - SDR-LD (HD): 34% gain

PS: All in MOS BD-rate

Source: w20000 Report on Essential Video Coding compression performance verification testing for SDR content, ISO/IEC SC29 WG4 Online January 2021

Low Complexity Enhancement Video Coding

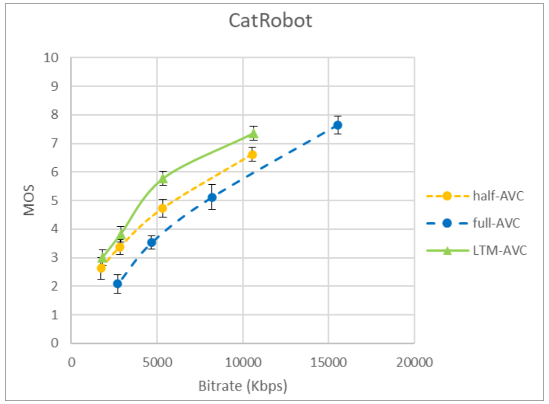
- 23094-2 LCEVC
  - FDIS approved
- 23094-3 Conformance and Reference Software
  - DIS ballot on going

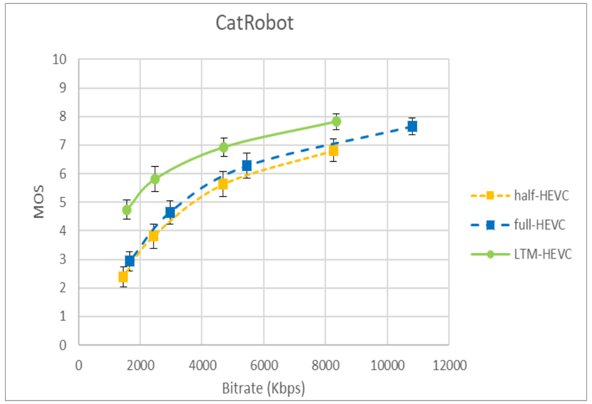


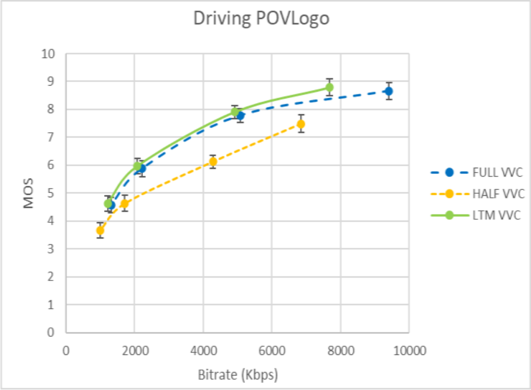
Summary of LCEVC Verification Test Results

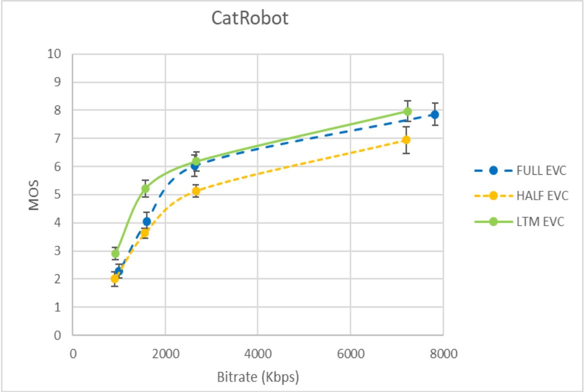
- Anchor: AVC, HEVC, VVC and EVC with full and half resolutions.
- Test: LTM5.1

Anchor	Condition	Full Resolution Avg. BD-rate	Half Resolution Avg. BD-rate
AVC / H.264	SDR_UHD	-45.90%	-27.52%
	SDR HD	-28.47%	-27.04%
HEVC / H.265	SDR_UHD	-30.87%	-33.65%
	SDR HD	-24.14%	-25.80%
VVC / H.266	SDR_UHD	-15.66%	-33.36%
	SDR HD	-14.14%	-20.53%
EVC	SDR_UHD	-17.77%	-37.86%
	SDR HD	-8.55%	-21.09%







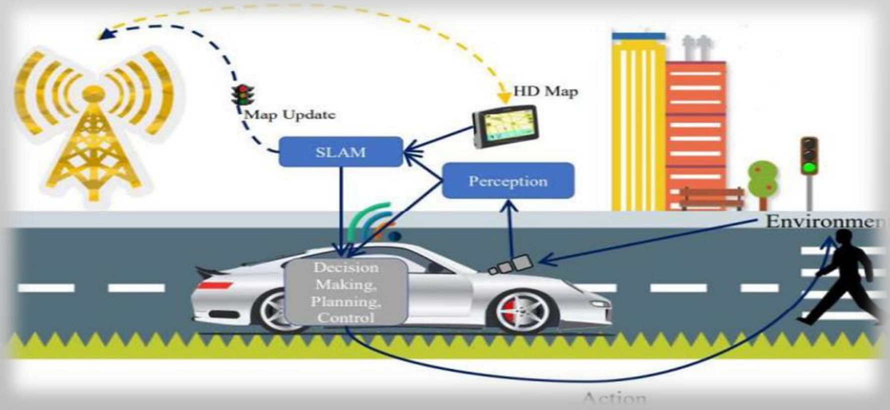


Video Coding for Machine(VCM)

Video Coding for Machine

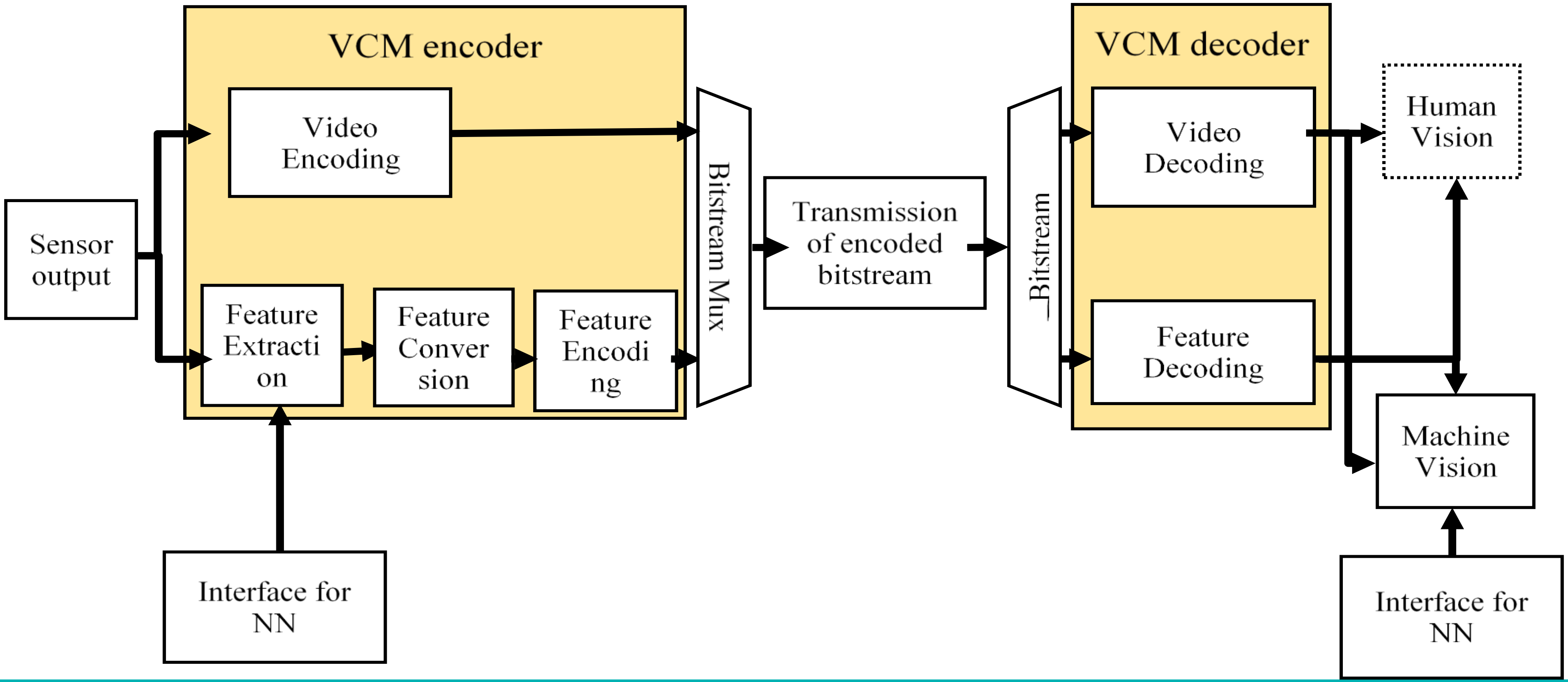
- Traditional coding methods aim for the best video under certain bit-rate constraint for human consumption.

- Machines will communicate amongst themselves to perform tasks without a human.
- Transmission and archive systems require a more compact data representation and low latency solution for machine vision.



Potential VCM architecture

- The VCM codec could be video codec, feature codec, or both.
- Machine vision tasks could be split into two stages and being implemented in the encoder side and decoder side respectively.



Status and progress

- Main topics in VCM
  - Feature compression
  - End-to-end image/video compression for machine
  - Hybrid machine / human compression
- Dedicated to improve anchor and define evaluation framework
- First call for evidence was issued in January 2021 and completed in April 2021
  - Received 5 responses, 2 are accepted as evidences
  - Both accepted responses are regarding object detection task
- Based on the result of CfE, the AhG starts to discuss EE establishment and prepare draft CfP

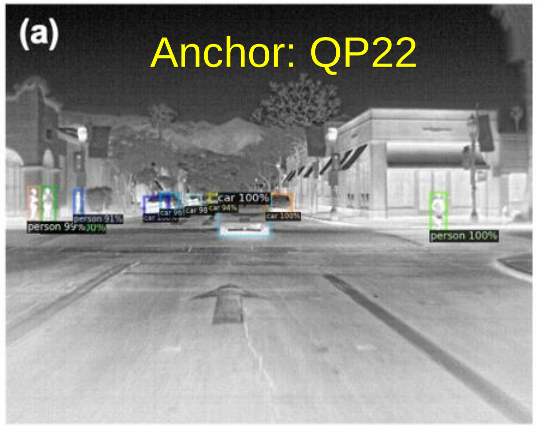
ITRI Industrial Technology Research Institute

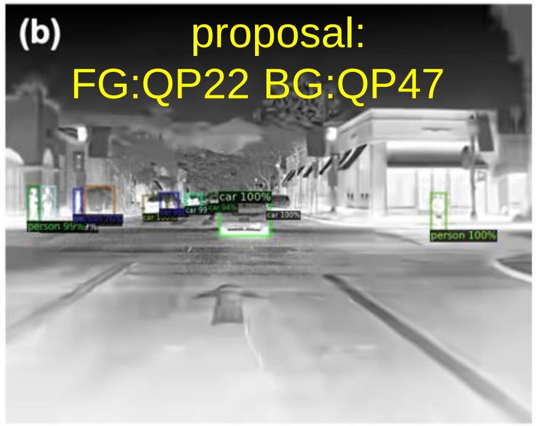
Responses to Call for Evidence

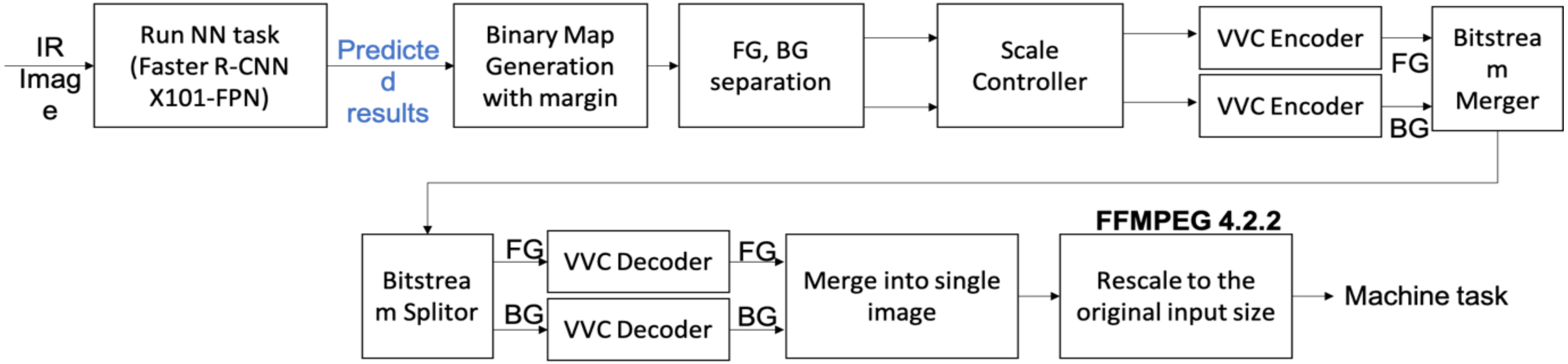
Company	Title	Description	Conclusion
ZJU	[VCM] ZJU response to cfe: deep learning-based compression for machine vision	End-to-end codec for object detection	Accepted
Konkuk Univ. / ETRI	[VCM] Response to CfE: Object detection results with the FLIR dataset	Separate image into FG and BG and compress them with different resolution/QP.	Accepted
Tencent	[VCM] Response to CfE: Investigation of VVC Codec for Video Coding for Machine	Investigate performance of VVC tools in terms of machine vision.	Not recognized as evidence
USTC	[VCM] Evidence of VCM: Object Detection Evaluation on Semantically Structured Image Compression (SSIC)	learning-based semantically structured image coding framework for high level and low level features	Not recognized as evidence
China Telecom / HQT	[VCM] Response to Call for Evidence of Video Coding for Machine: K-means and BAC based feature compression	Use K-means algorithm to analyze feature data for adaptive quantization.	Insufficient results

Response to CfE (ETRI, Konkuk Univ.)

- Creates the binary maps from the predicted results by adding margin pixels and separate the original image into FG and BG.
- Selectively down-scale FG and BG and encoded with different QPs and merged to a single bitstream
- About 18.89% BD-rate saving



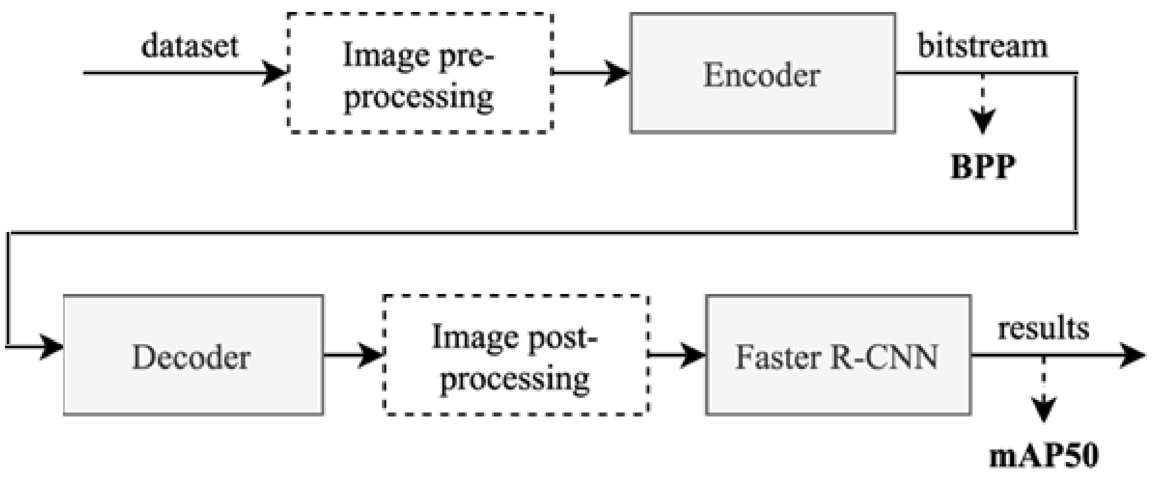




BD-rate (%)		
mAP	100%	-18.89
	75%	-36.04
	Average	-27.45
Pareto mAP		-30.76

Response to CfE (ZJU)

- End-to-end deep learning-based image coding networks (Cheng2020's networks) optimized for machine vision tasks.
- The re-trained model (optimized for mAP) can achieve 22.8% BD-rate saving compared to VTM8.2 (optimized for PSNR)



QP	Anchor (VTM-8.2)		re-trained Cheng2020's networks	
	BPP	mAP	BPP	mAP
22(1)	0.83601788	0.76517548	0.5556372957	0.76346890
27(2)	0.496323715	0.76032878	0.4531612348	0.76153704
32(3)	0.277115735	0.749032079	0.3147077104	0.75533322
37(4)	0.146083233	0.722172818	0.1958046374	0.74696396
42(5)	0.073909964	0.672795607	0.1167023101	0.72394438
47(6)	0.035294544	0.566341802	0.0746458029	0.70362475
BD-rate	-		-22.80%	

Source: m56445 [VCM] ZJU response to cfe: deep learning-based compression for machine vision, ISO/IEC SC29 WG2 Online April 2021

Thank You

INNOVATING A BETTER FUTURE!

ITRI
Industrial Technology Research Institute