

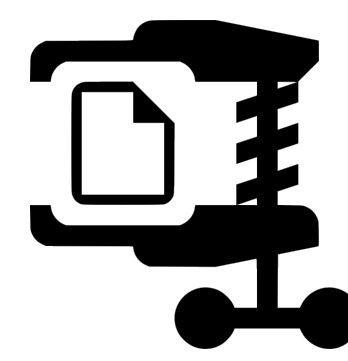
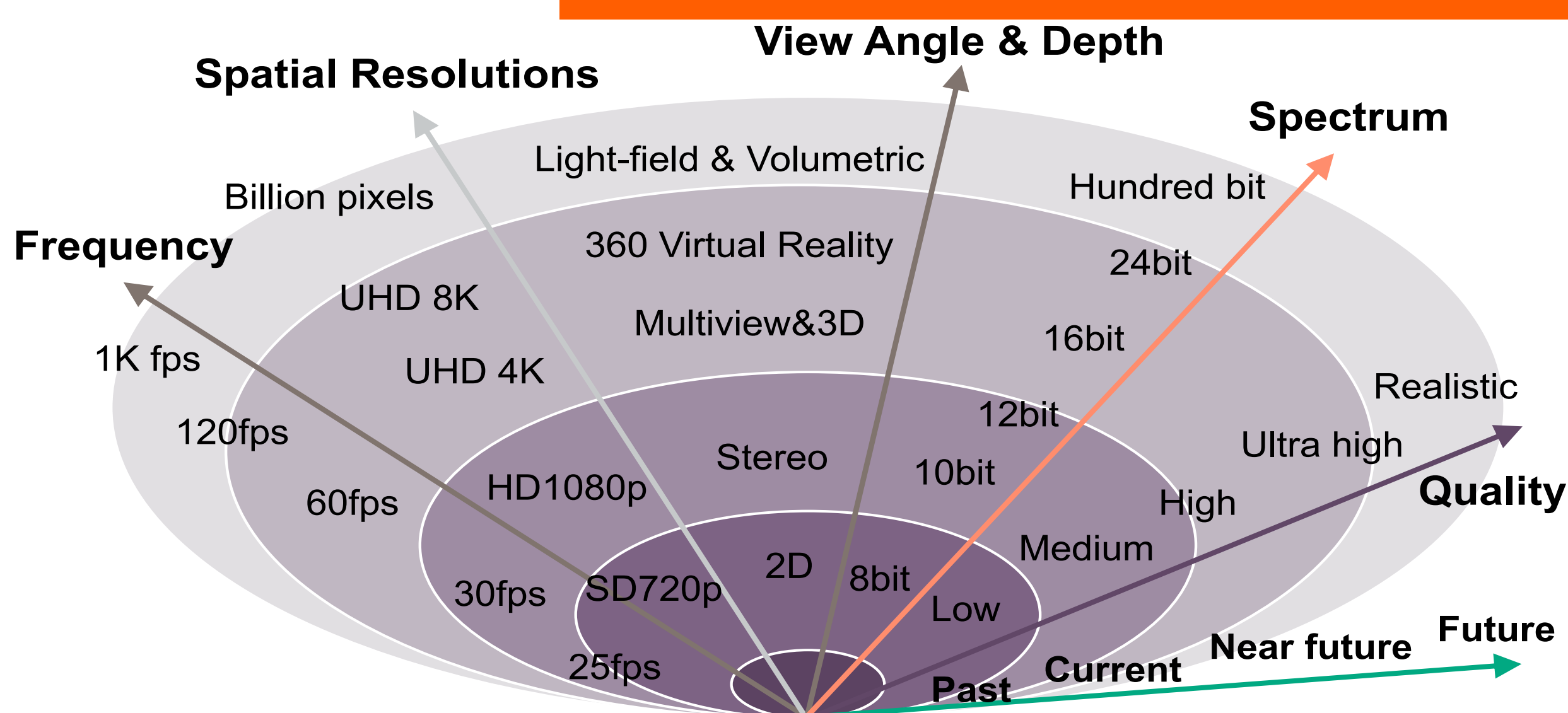
Towards AI-assisted Video Compression



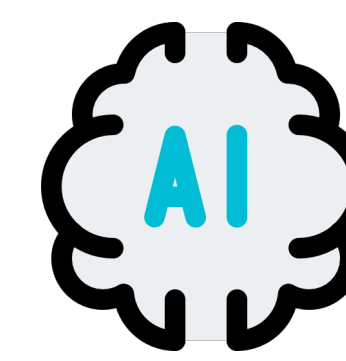
Sicheng Li, Di Xu, Jian Lou, Fu Wang, Yen-Kuang Chen
Alibaba Group
{sicheng.li, di.x, j.lou, yunfan.wf, yk.chen}@alibaba-inc.com



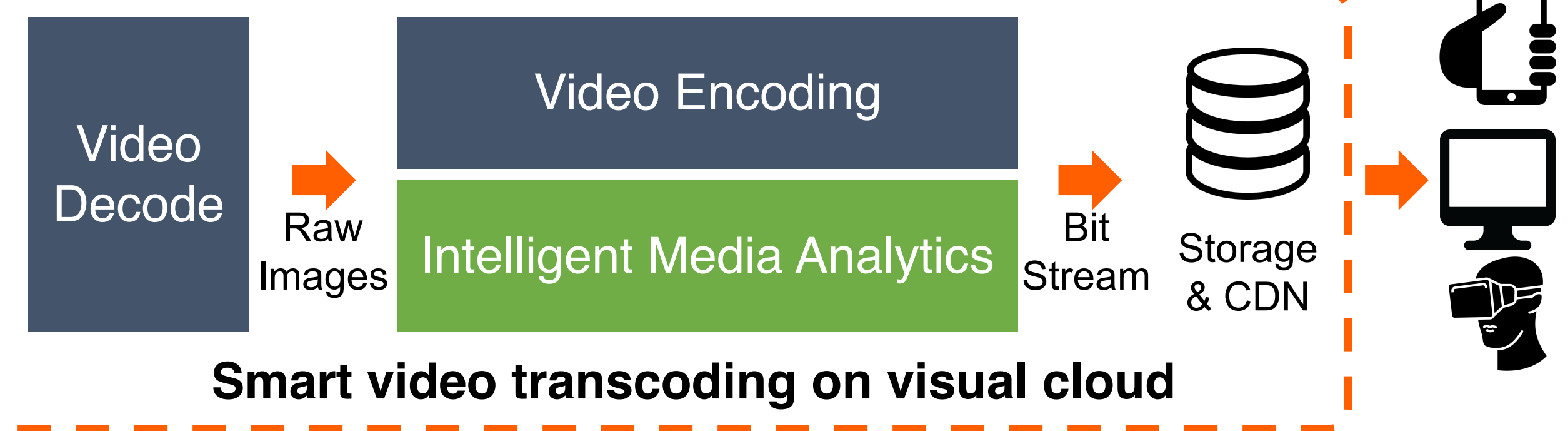
TREND OF VIDEO DATA REPRESENTATION TO ENHANCE USER EXPERIENCES



Better compression
MPEG-2/H.264/H.265/VCC
VP9/VP9/AV1/AV2
JPEG 2000/TICO/JPEG-XS



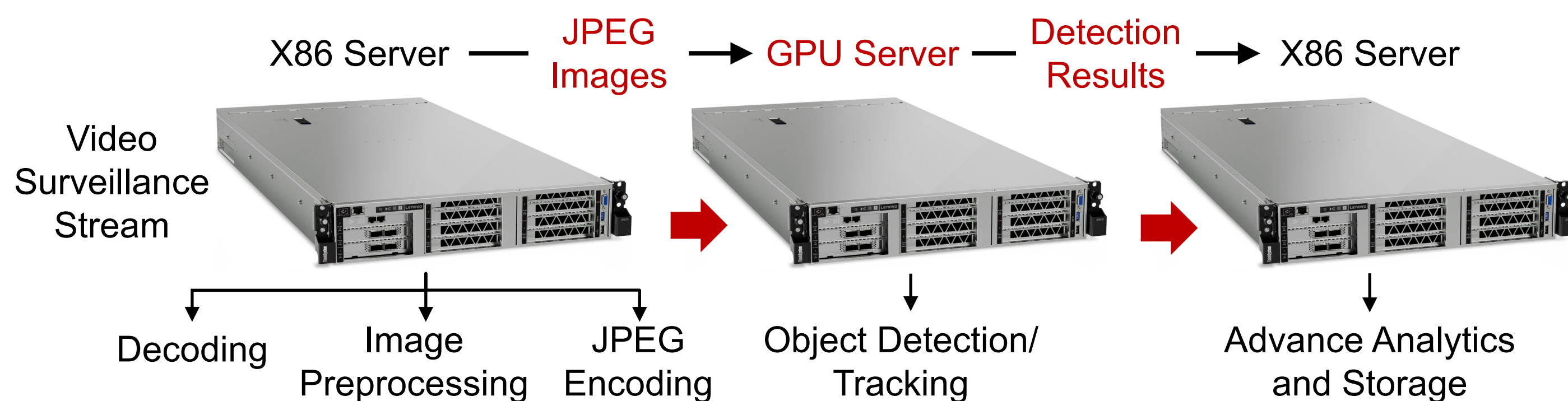
Intelligent Analytics
Event detections, face recognition,
translations for video conferences,
codec optimizations



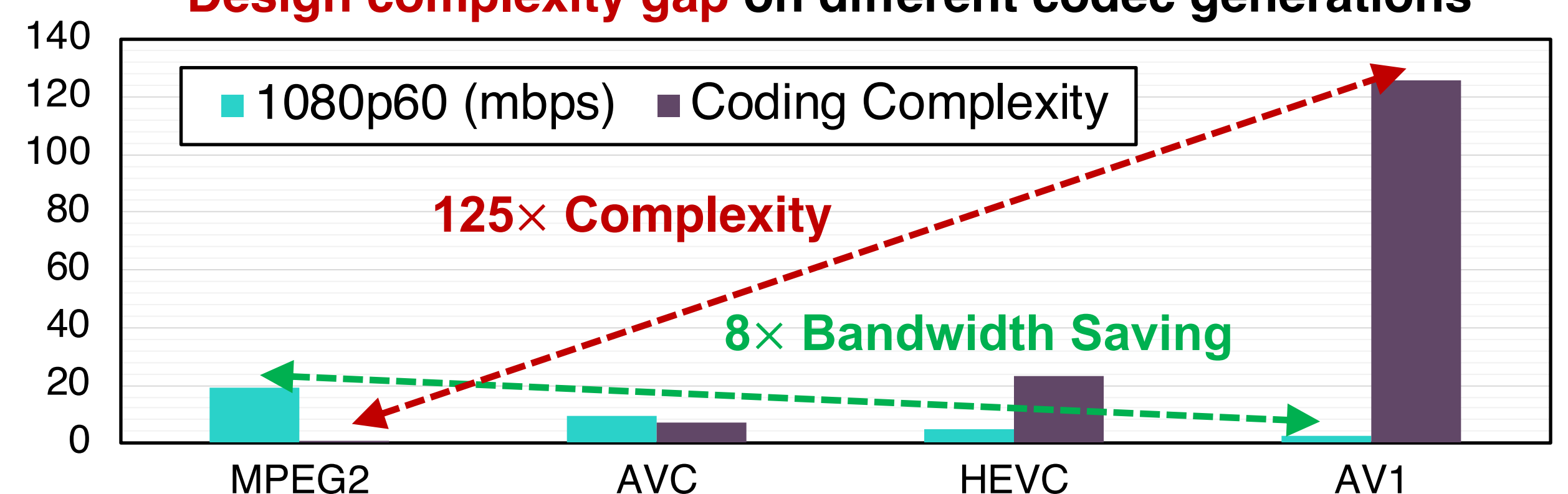
Significantly large data volume is required to represent high fidelity in video content [1]

DESIGN CHALLENGES ON VIDEO COMPRESSION AND INTELLIGENT ANALYTICS

Traditional solution brings **high bandwidth cost and latency overhead** [2]

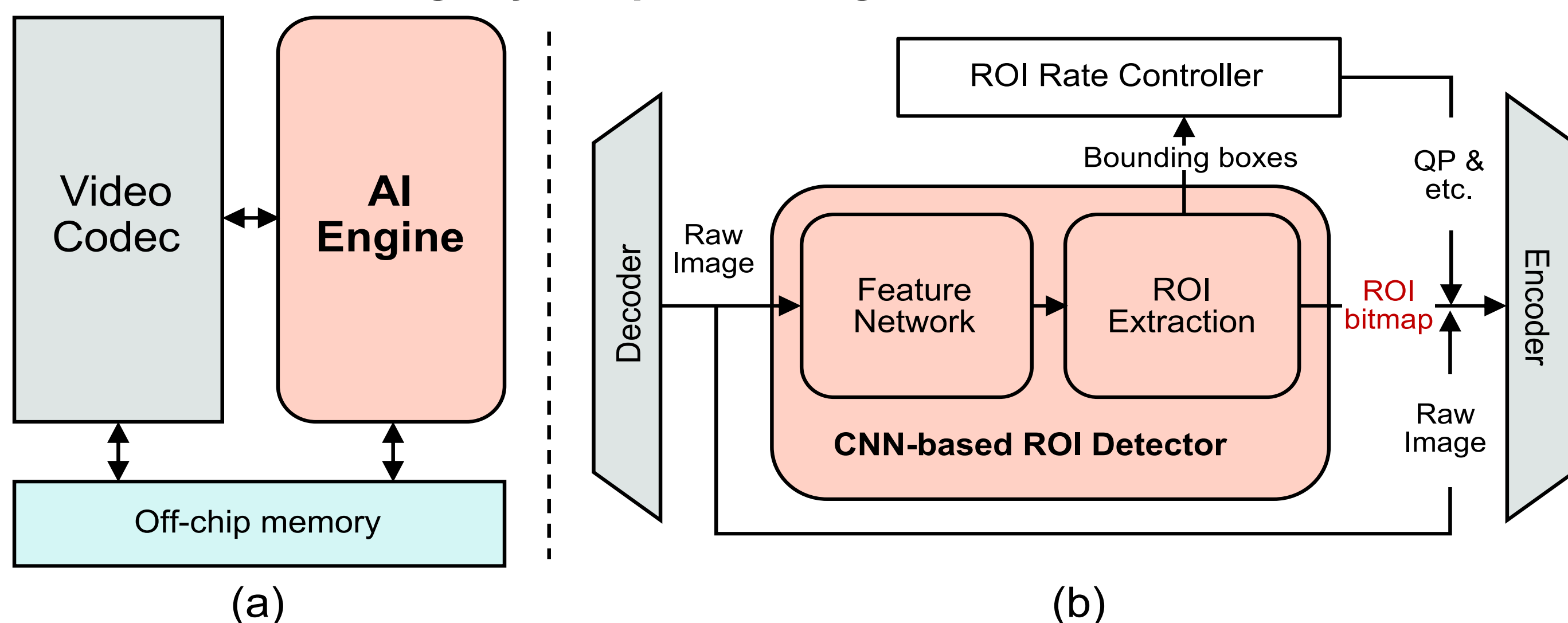


Design complexity gap on different codec generations



AI-ASSISTED VIDEO COMPRESSION: ALL-IN-ONE SOLUTIONS

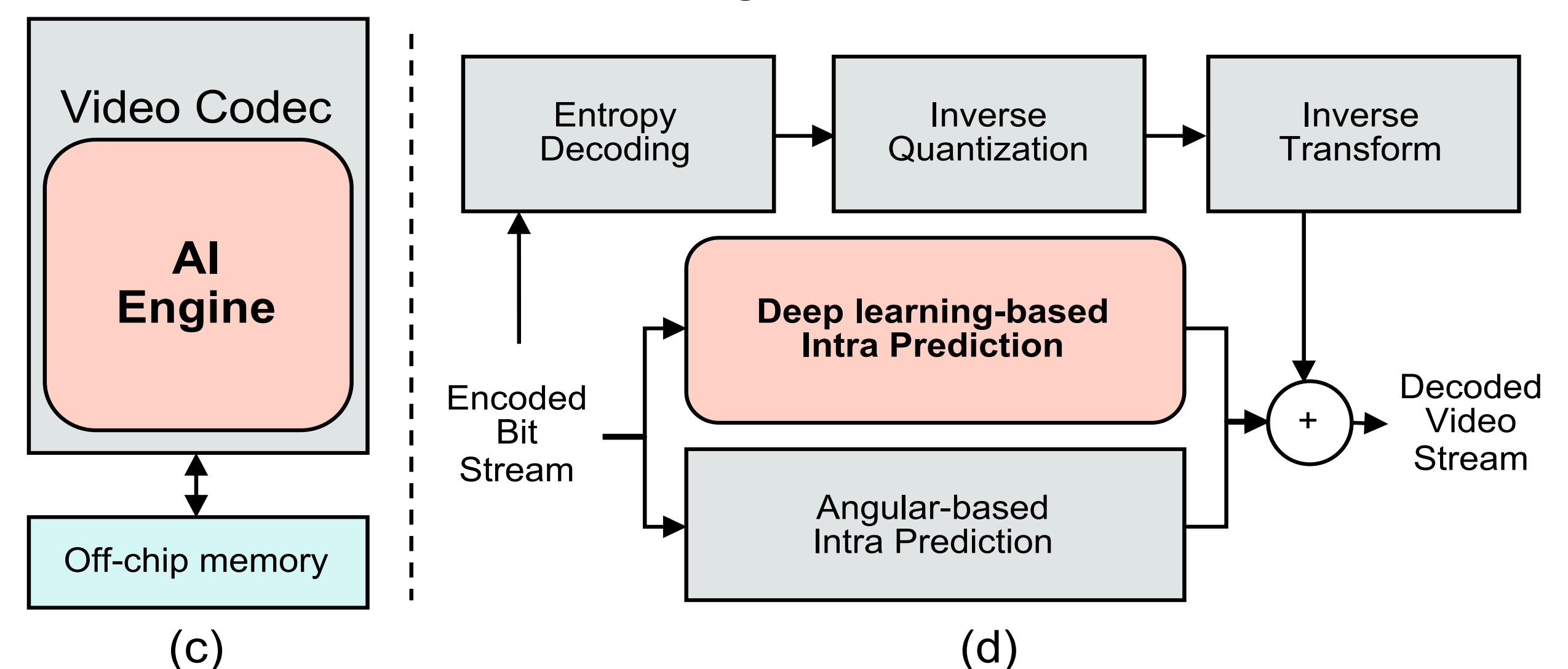
Tightly-coupled AI engine and Codec



(a) Merging analytics and recognition tasks with video coding, not only facilitates the reuse of video information, but also reduces encoding bitrate while maintaining the same visual quality.

(b) demonstrates an CNN-base ROI detector which receives raw images from decoder and generates ROI bitmap to indicate salient regions for encoding bitrate saving.

Fused AI engine and Codec

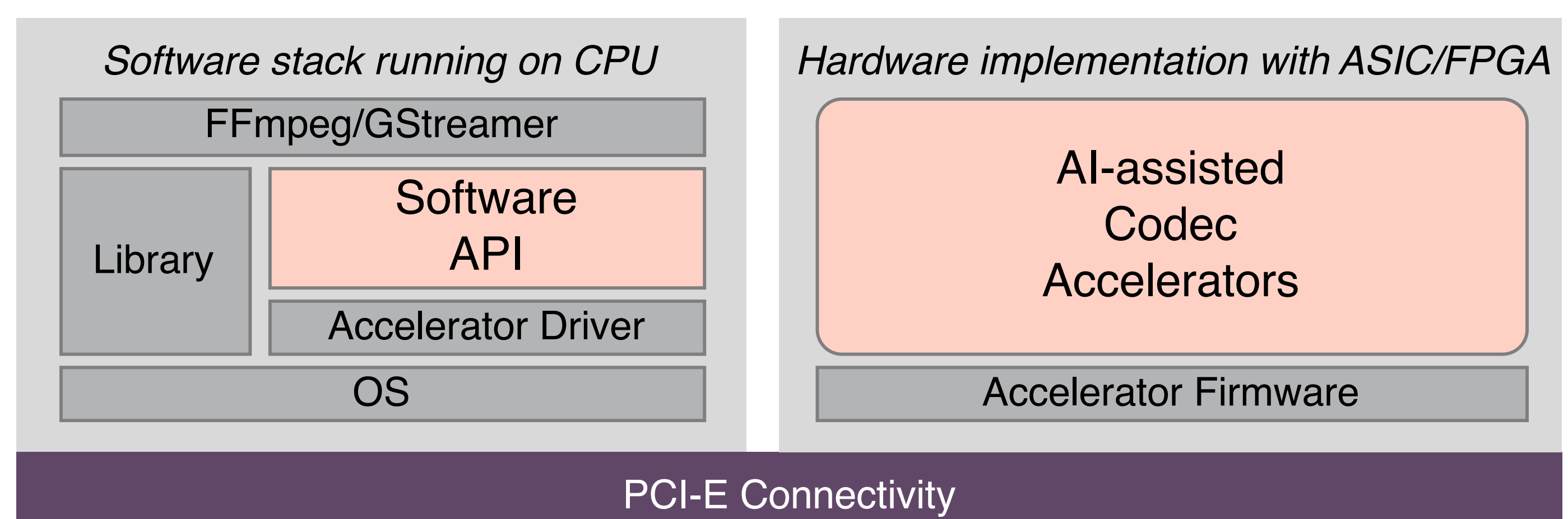


(c) This architecture enables two research directions: hybrid schemes that contain AI-based building blocks within the traditional Codec and end-to-end coding solutions that are built primarily upon AI.

(d) illustrates a deep learning based intra prediction module to remove spatial redundancy and generates flexible prediction patterns for coding efficiency improvements.

FUTURE WORK: OPEN-SOURCE SW/HW CODESIGN FRAMEWORK

- We propose a codesign framework to facilitate the design trade-off between compression ratio, computational cost and design flexibility.
- OCP is a great platform for us to look for partners who are interested in this concept and would like to build the hardware accelerator together. The accelerator could be in the form of a FPGA or a discrete ASIC depending on performance and TCO targets. Even though all parties who participate in this project may have diverse video applications, we believe that they can all benefit from the hardware accelerator while being differentiated by their software stack.



[1] Dai, Q., Wu, J., Fan, J., Xu, F. and Cao, X., 2019. Recent Advances in Computational Photography. Chinese Journal of Electronics, 28(1), pp.1-5. [2] Zhang, Y., Kwong, S. and Wang, S., 2020. Machine learning based video coding optimizations: A survey. Information Sciences, 506, pp.395-423.



2020 OCP Global Summit