Tools for Creating Next-Gen Computer Vision Apps on Snapdragon

Judd Heape
VP Product Management for Camera,
Computer Vision and Video Technology

Computer Vision in Snapdragon

Three function levels to provide comprehensive CV solutions

CV hardware Acceleration blocks to support and enable hardware, software and system designs in Snapdragon platforms

CV Algorithms to demonstrate complete workflows that provide state-of-the-art solutions to certain perception problems

CV end-to-end Applications in mobile, XR, Automotive and IOT market segments to enable unique and enhanced user experiences

Engine for Visual Analytics (EVA): Computer Vision Hardware Blocks

Object/Face Detection	Optical Flow	Depth Estimation	Feature Extraction	Geometry Correction	XR & 3DR	
Object Detection HOG/SVM	Semi-dense OF GMO for video encode	DFS 1080p@30 Video Bokeh	Feature – video encode with 30% BR reduction HCD/NCC	Lens Distortion Correction		EVA1.x
ACF/RDF Face Detection			HCD/ORB – Centralized ME for camera	Exposure Compensation		EVA2.x
	Dense OF (SGM based) Dense motion map for multi-frame processing, sensor alignment	DFS (SGM based) • Bokeh better quality • Visual special effect • XR 3D reconstruction	Flow improvement for XR 6DoF/VIO	Motion and depth map warping LSR (in EVAa 3.5)	XRA - DoH,DoG FREAK, R-BRIEF (In EVAa/EVAv 3.5)	EVA3.x

Mobile Camera / Video XR Auto IoT

Qualcomm

Optical Flow

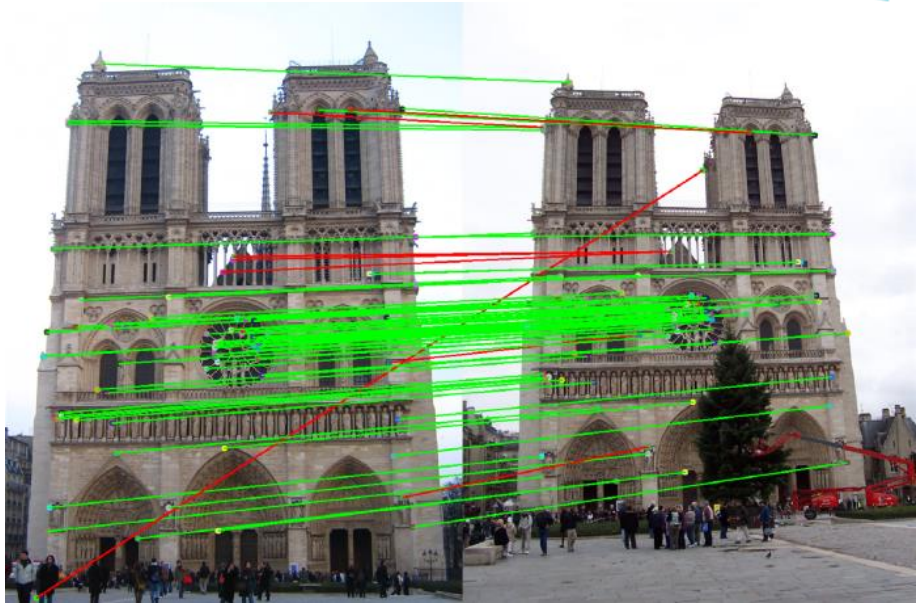
Sparse Motion

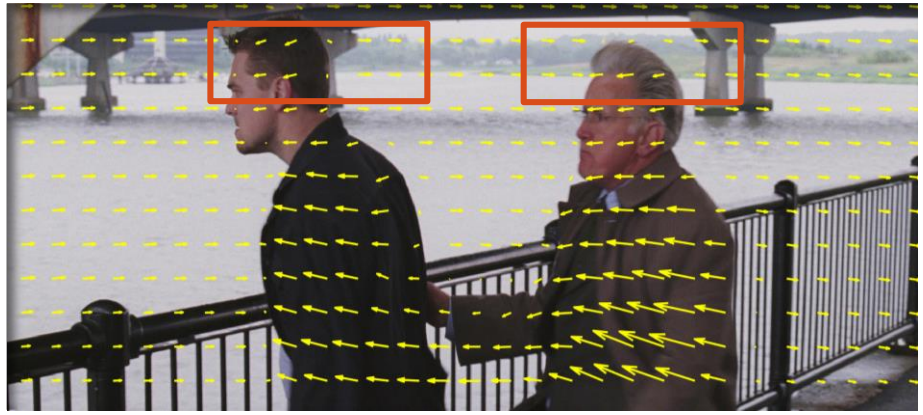
- Feature Point Detection, Local and Global Motion
- Various Detector and Descriptors (Harris, DoH, DoG, FREAK)

Dense Motion

- Semi-dense Optical Flow (sDOF)
- Dense Optical Flow (DOF)
- Hybrid Deep Learning based Motion + OF Segmentation Enhanced

	Semi-Dense OF	Dense OF
Motion Density	Every 2x2 block	Every pixel
Motion Accuracy	1/8 pixel	1/16 pixel
Motion Range (X,Y)	±128, ±64	±64, ±32
Max Resolution	1920x1080	1152x648
Confidence Map	8-bit	8-bit
Frames per Second	60	60



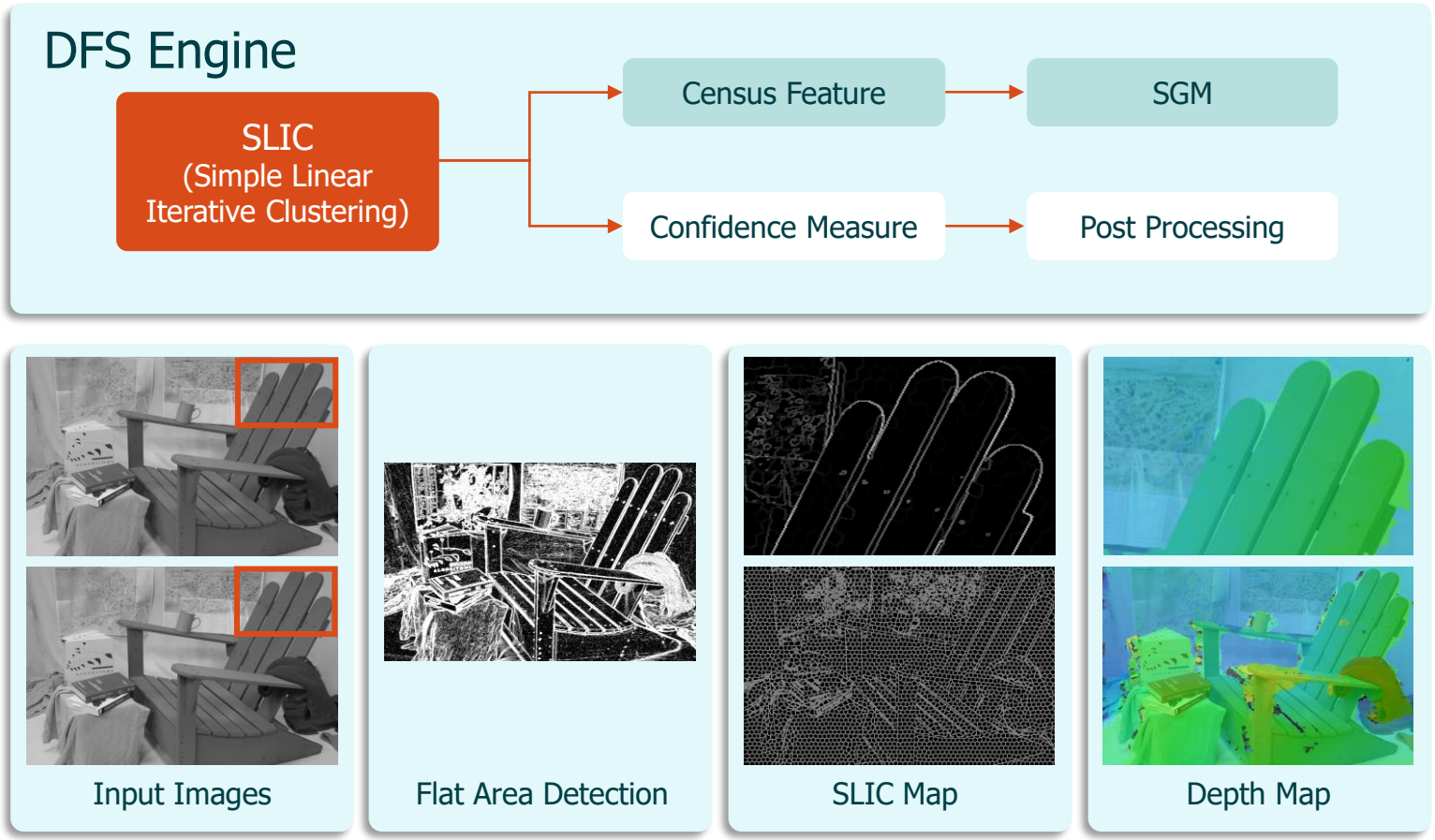


Depth from Stereo Estimation

Depth from Stereo (DFS)

- Super-pixel Segmentation on SLIC
- Feature Extraction and Matching
- Confidence Map and Post Processing

Depth Density	Every pixel
Disparity Accuracy	1/16 pixel
Disparity Level	[0,63]
Max Resolution	720P@60FPS



Geometric Correction Engine (GCE)

Low-power High-quality Warping

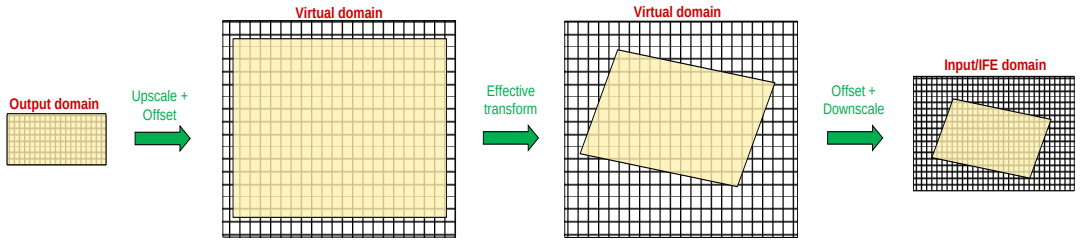
- ICA maps output pixels to input pixels

Effective Transformation

- Sparse grid transformation (35x27 or 67x51)
- Dense grid transformation (8 pixel grid)
- Perspective transformation (3x3 transform)

Effective Transform

Dense Grid Transformation → Perspective Transformation → Sparse Grid Transformation



GCE Use Cases

- Lens distortion correction
- Motion vector grid composition
- Rectification

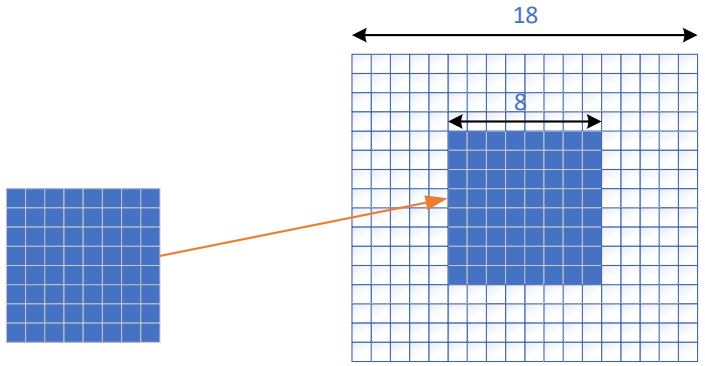
Lens Distortion Correction

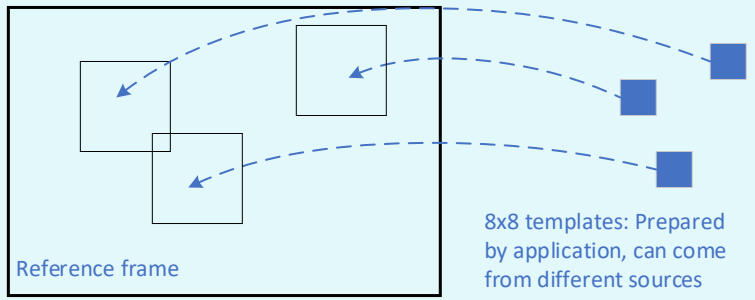
Rectification

Normalized Cross Correlation

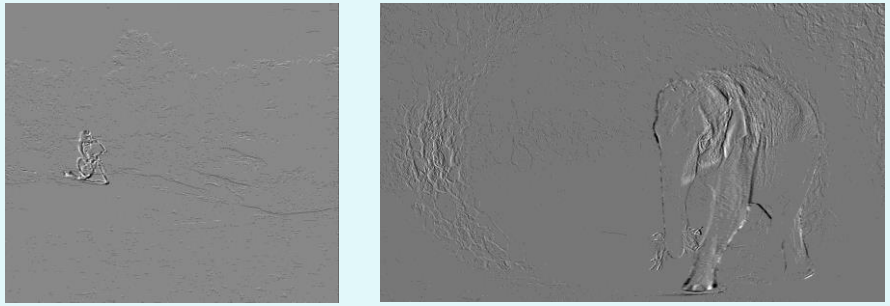
NCC Supports Two Modes

- Patch to Frame Mode
- Frame to Frame Mode

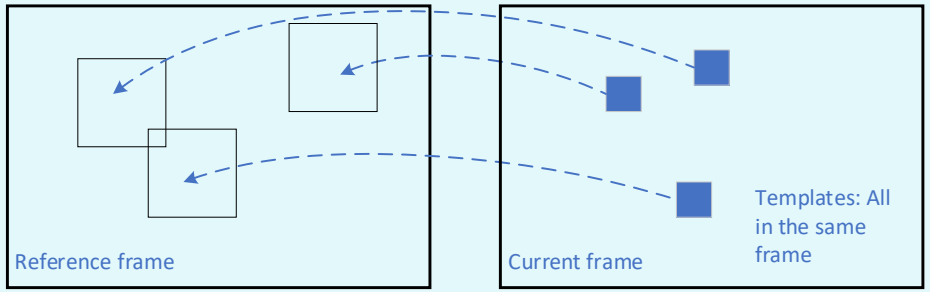




Patch-to-frame mode



Frame Matching Using Harris Corners and NCC



Frame-to-frame mode

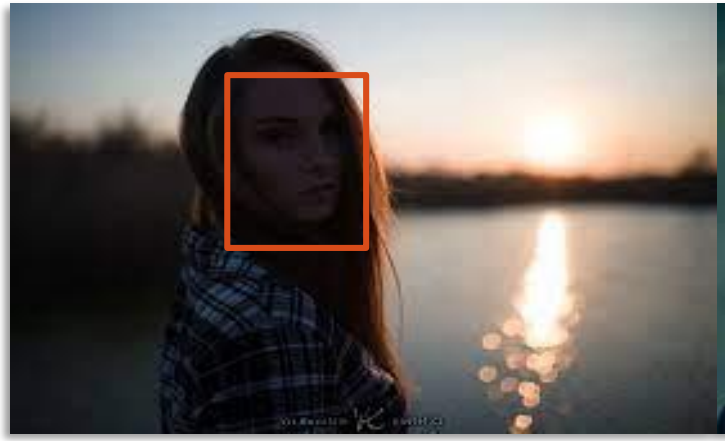
Face Detection

Deep Learning based Face Detection (FD)

- Min Face Size: 32x32
- Detection Accuracy: 95%
- 1080p@60FPS
- Multiple cameras supported

Under Non-Ideal Conditions

- Strong Backlight
- Full Profile
- Occlusions – Face Masks, Hats, Glasses, Sunglasses



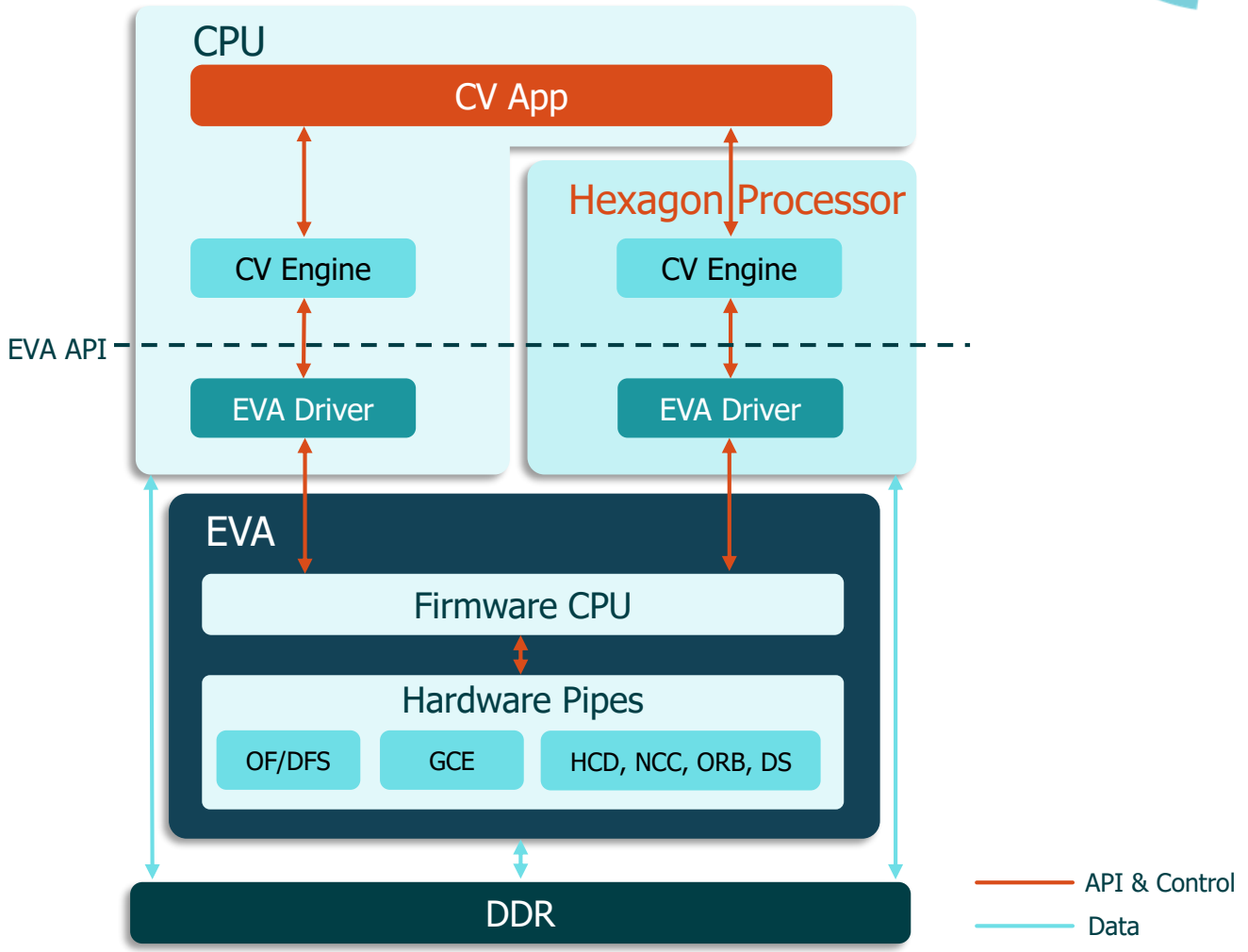
Strong Backlight

Full Profile

Occlusions

EVA Architecture and Access

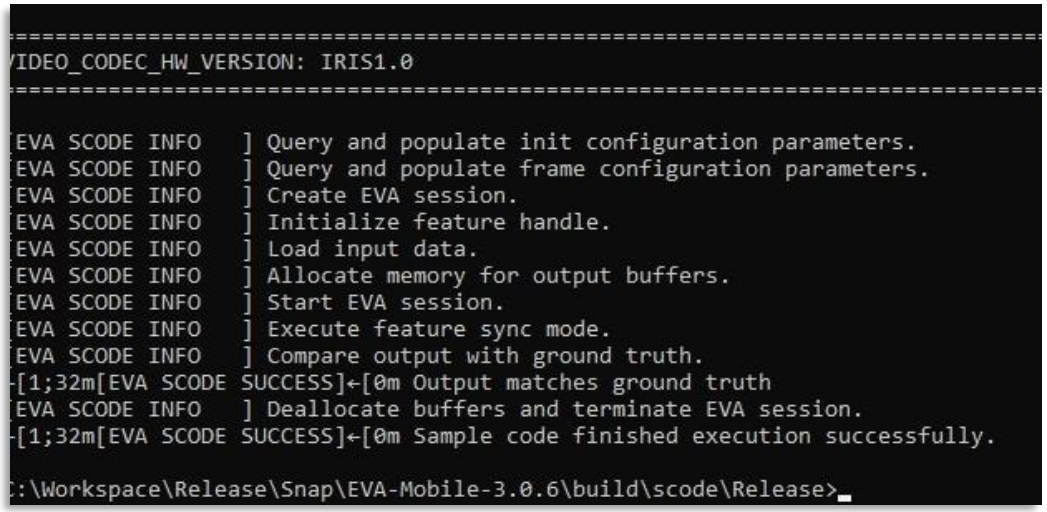
- The EVA APIs are exposed both from the CPU and Hexagon Processor sides
- It includes both synchronous APIs and asynchronous APIs
- There are direct interrupts between the Hexagon Processor and EVA cores for low latency communication
- EVA includes embedded CPU primarily for task scheduling and hardware pipes
- EVA hardware pipes are shared between certain functions

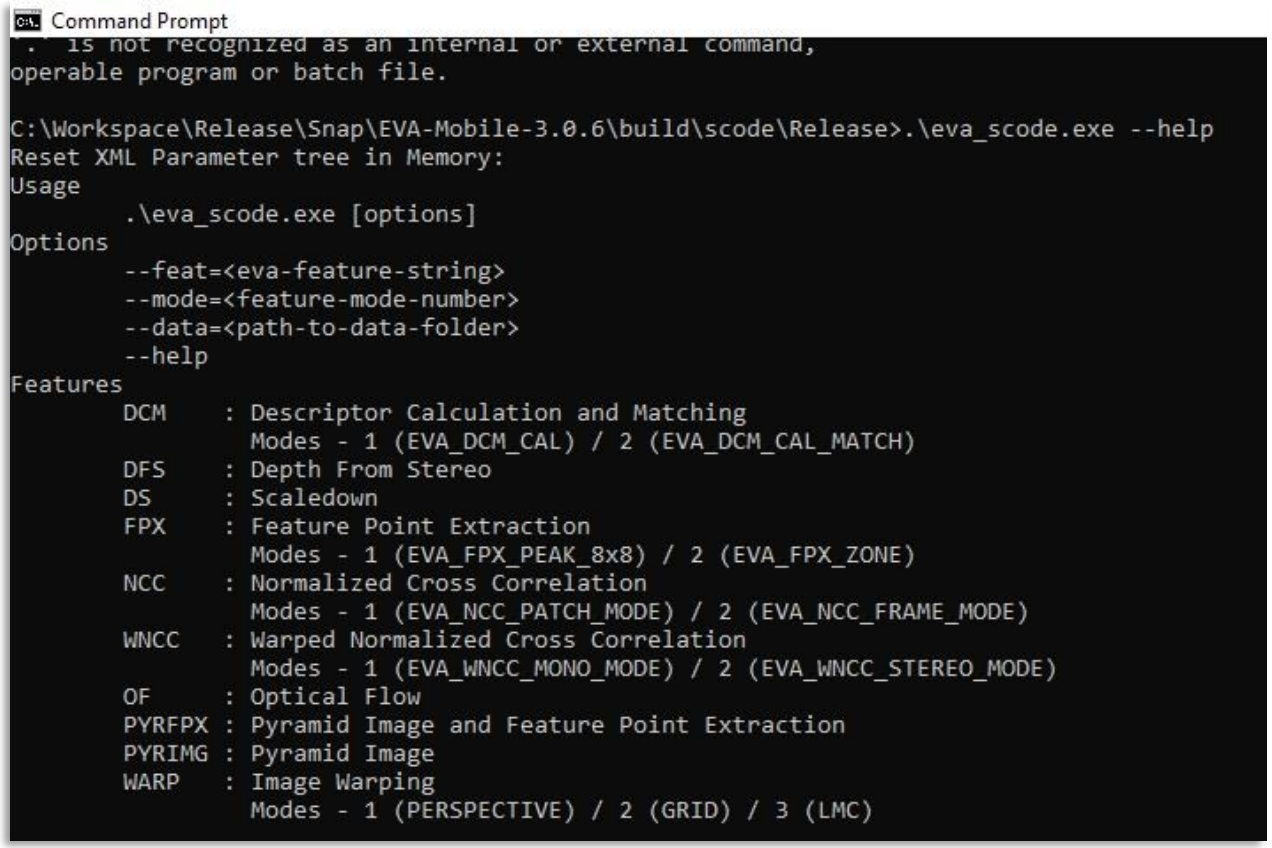


EVA Feature APIs

EVA3.0 Features	EVA API
Image Warping	evaWarp_Sync / evaWarp_Async
Depth from Stereo (DFS)	evaDfs_Sync / evaDfs_Async
Normalized Cross Correlation (NCC)	evaNccFrame_Sync / evaNccFrame_Async
Optical Flow (OF)	evaOF_Sync / evaOF_Async
Feature Extraction (HCD)	evaFeaturePoint_Sync / evaFeaturePoint_Async
Feature Descriptor Calc & Matching	evaDcm_Sync / evaDcm_Async
Downscaler	evaScaledown_Sync / evaScaledown_Async
Pyramid Image	evaPyramidImage_Sync / evaPyramidImage_Async

EVA SDK Simulator





CV Use Case 1

Depth Map from Stereo Cameras (DFS)

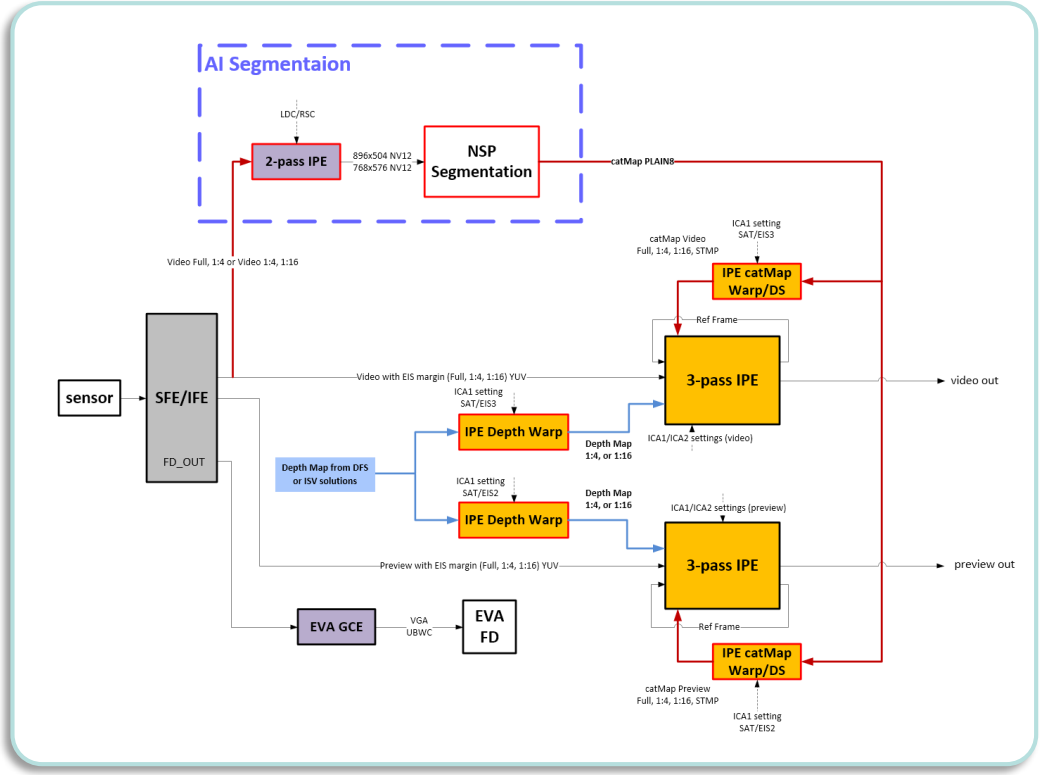
Applications

- Accurate Camera/Video Bokeh effect
- Background replacement in video recording or Zoom call
- AR/VR
  (3D Reconstruction, Video Passthrough, Occlusion)

Real Time Bokeh Effect using Depth Map from Stereo Cameras (DFS)

Applications

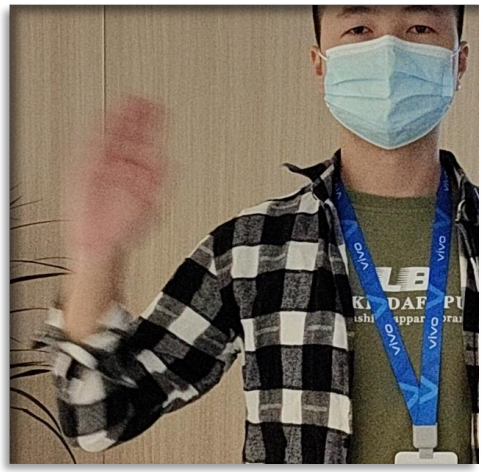
- Accurate Camera/Video Bokeh effect

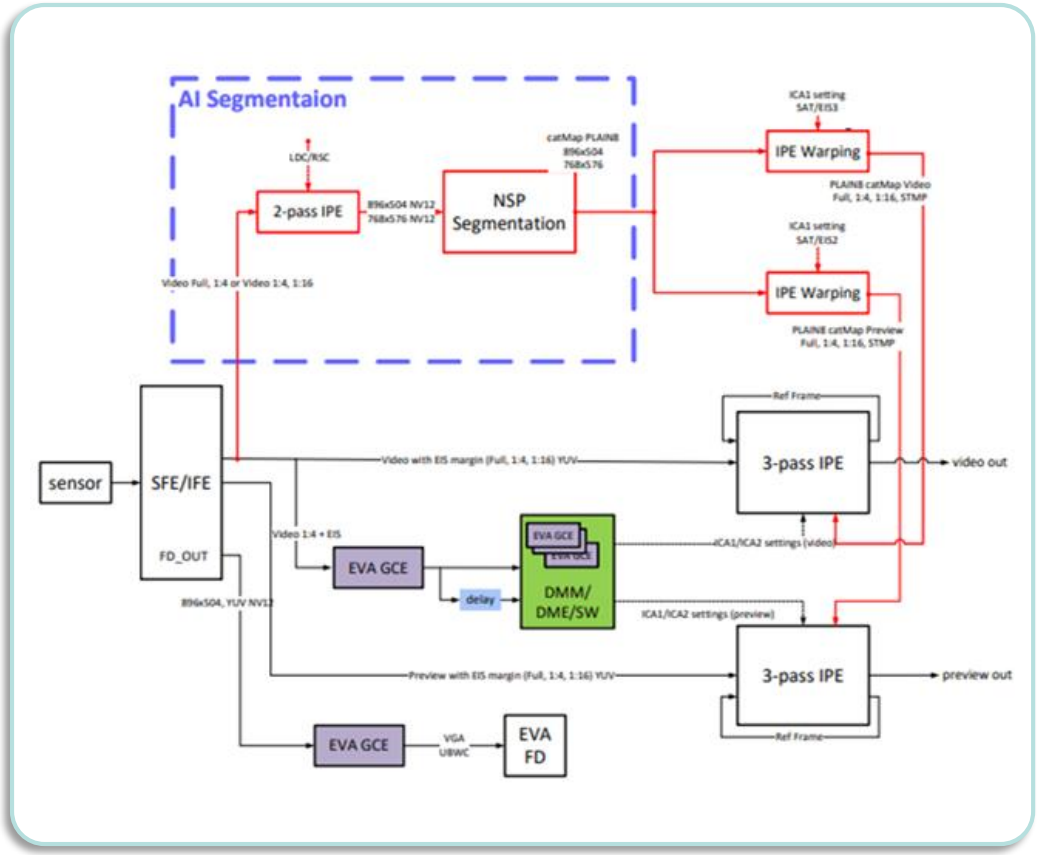


Dense Motion Map (DMM) for Video MCTF

Key Benefits of EVA

- Register multiple frames with local motion compensated
- Remove ghosting artifacts in combined video frames

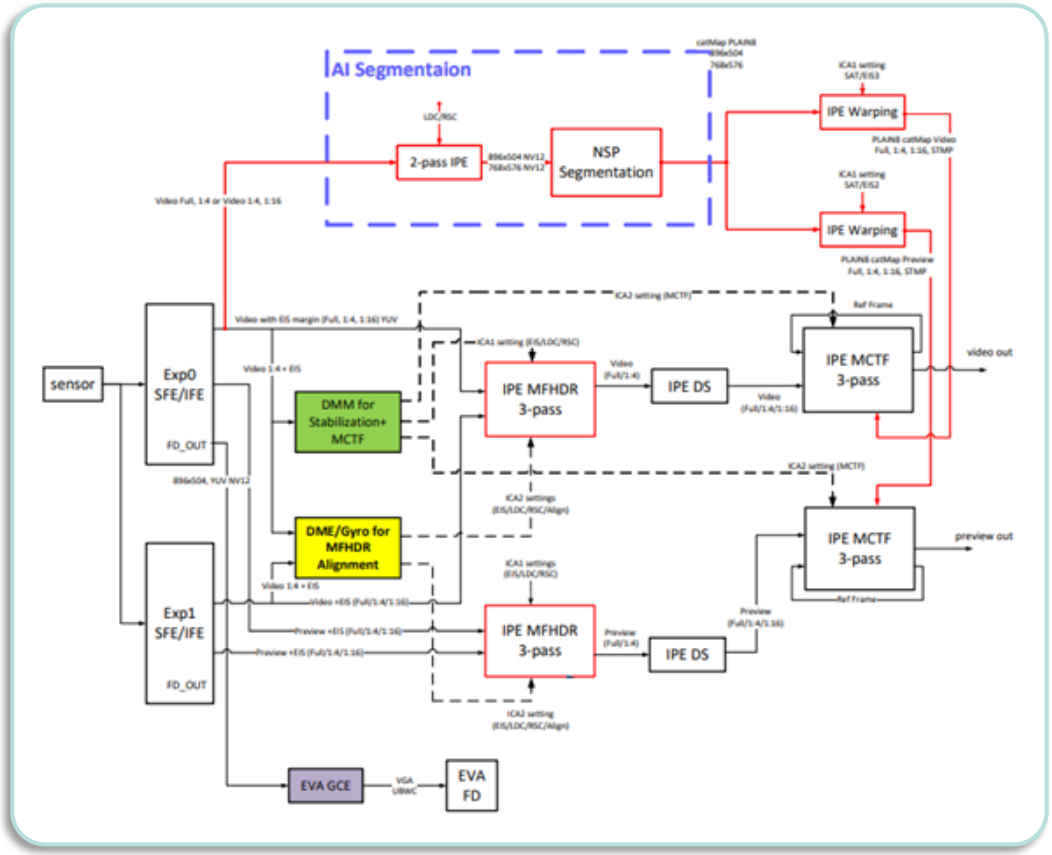




Dense Motion Map (DMM) for Video MFHDR

Key Benefits of EVA

- Estimating and compensating for motion is key to achieve high quality HDR video
- Remove ghosting artifacts in combined video frames
- Running global motion and local motion estimation simultaneously requires large amount of computation power

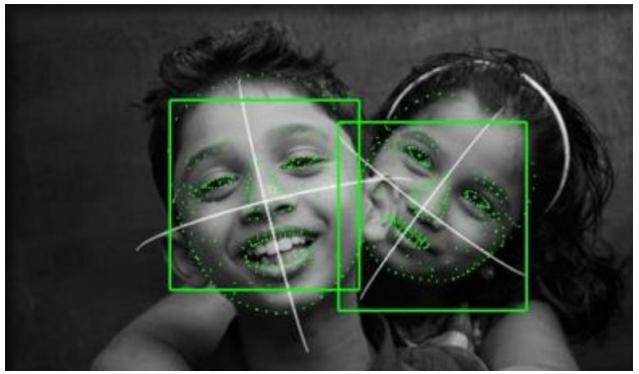


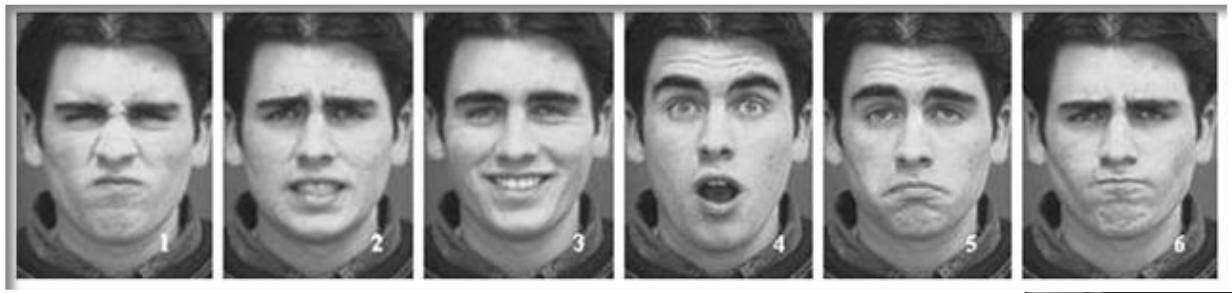
CV Use Case 5

Face Detection (FD) and Face Landmark Detection (FLD)

Applications

- Gender/Expression/ Emotion/Gaze detection
- Avatar animation
- Geometric personalization

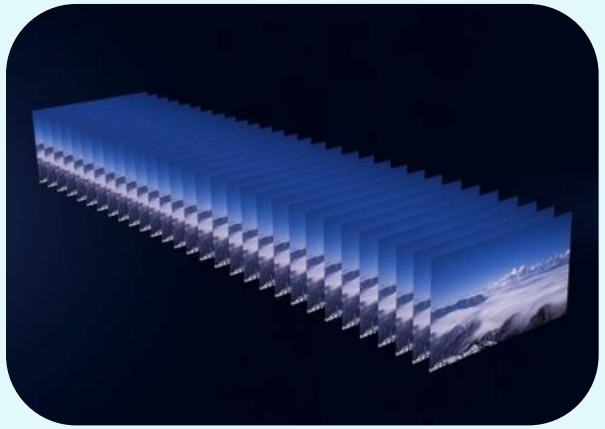




Qualcomm Deep Learning-based 3D face landmark detection reaches high accuracy in locating

115 or 300 facial landmarks

Start Developing on Snapdragon



Capture at higher FPS

Extend battery life

Tap into hardware-accelerated CV features with an SDK not previously available

Start Developing on Snapdragon

For access to the SDK contact:

Xin Zhong
Director, Product Management
xzhong@qti.qualcomm.com

Thank You

Qualcomm