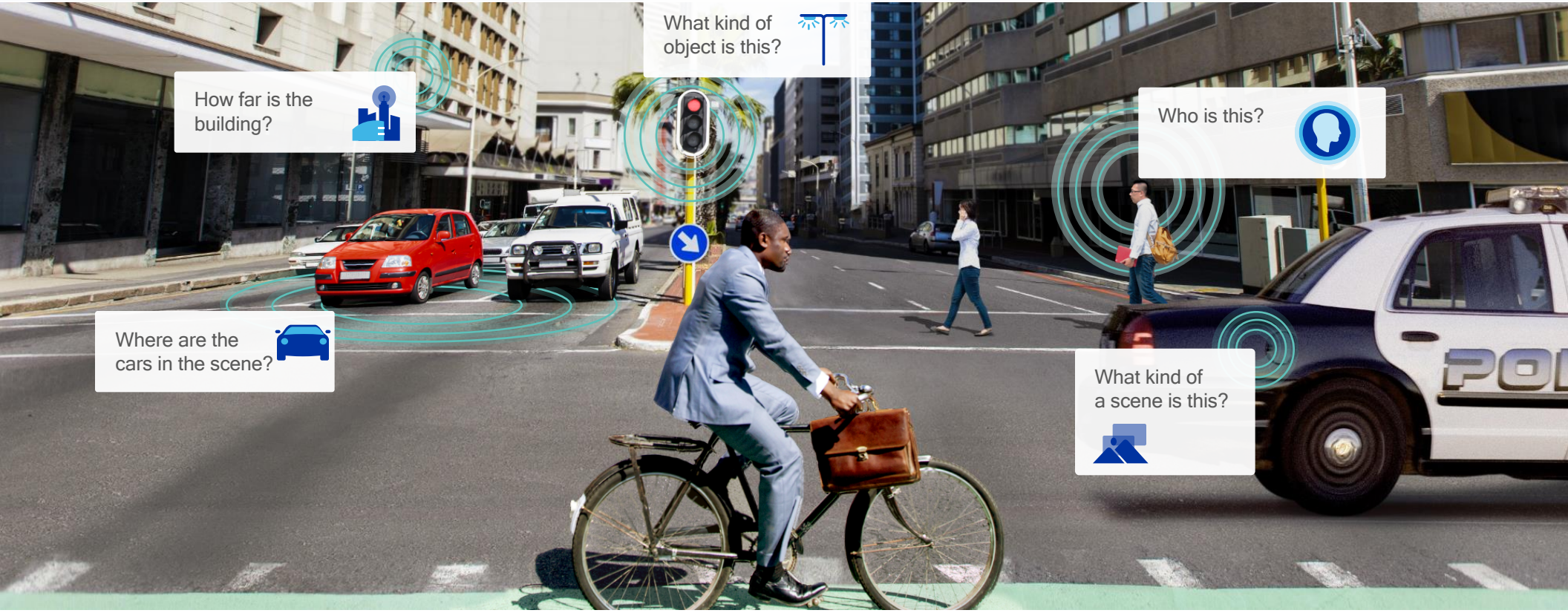
QUALCOMM®

Emerging Vision Technologies: Enabling a New Era of Intelligent Devices

Computer vision overview

Computer vision is being integrated in our daily lives

Acquiring, processing, and understanding visual data in images, videos and the real world



Enabling vision applications for different ecosystems

Example use cases

	Mobile	Virtual Reality	IP Camera	Robotics/ Drones	Automotive
Ecosystems					
Use Cases	Touch to focus	HMD position location	Face recognition	Obstacle avoidance	Pedestrian detection
Computer vision features	Touch-to-track	6-DOF positional tracking	Face detection and recognition	Simultaneous localization and mapping	Object detection

Mobile use case

Touch-to-focus

- User selects which part of the image the camera should focus on
- Computer vision identifies key points within the selected region of interest
- Camera tracks the key points and informs the AF algorithms where to focus even while the region of interest moves

Touch-to-track

Robust, low power, multiple object tracker

Track — Lucas-Kanade optical flow tracking

Detect — Decision Forest and NCC detection

Learn — Online learning bootstrapping binary classifier

Benefits

- Multiple object tracking up to 4 objects
- Robust tracking algorithm tightly integrated with 3A and auto-zoom algorithms
- Low power and thermal object tracking by using hardware accelerated object tracker
- High performance tracking on 1080p @ 30fps

VR use case

VR will be the new paradigm for how we interact with the world

Offering unprecedented experiences and unlimited possibilities

Play

Learn

Communicate

Experiences in VR

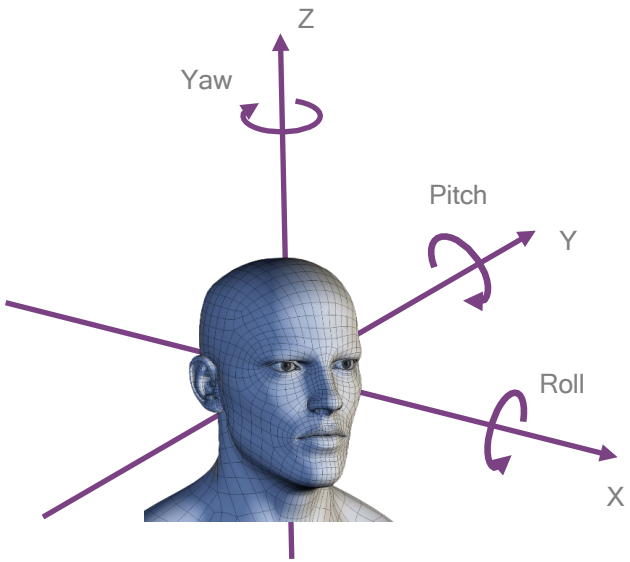
- Immersive movies and shows
- Live concerts, sports, and other events
- Interactive gaming and entertainment

- Immersive education
- Training and demos
- 3D design and art

- Social interactions
- Shared personal moments
- Empathetic storytelling

Precise motion tracking of head movements

Through 6-DOF positional tracking



3 degrees of freedom (3-DOF)

- “In which direction I look”
- Detect rotational movement
- Main benefit: Look around the virtual world from a fixed point

6 degrees of freedom (6-DOF)

- “Where I am and in which direction I look”
- Detect rotational movement and translational movement
- Main benefit: Move freely in the virtual world and look around corners

Achieving precise head motion tracking on the device

Visual inertial odometry (VIO) for rapid and accurate 6-DOF pose

Monocular camera data

Captured from tracking camera image sensor at ~30 fps

Accelerometer & gyroscope data

Sampled from external sensors at 800 / 1000 Hz

Snapdragon "VIO" subsystem

Camera feature processing

Inertial data processing

Hexagon™ DSP algorithms

- Camera and inertial sensor data fusion
- Continuous localization
- Accurate, high-rate "pose" generation & prediction

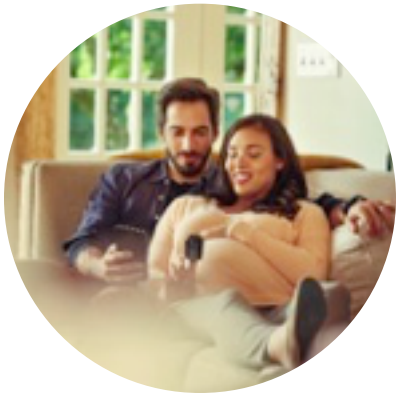
6-DOF position & orientation

(aka "6-DOF pose")

IP camera use case

Face recognition

A key vision use case for IP camera



Keeping our homes safer

Capturing our important moments

Keeping our communities safer

Home surveillance

Family member Recognition and intruder detection

Action camera

Face recognition and visual tracking

Professional surveillance

Traffic/parking monitoring

Face recognition

Using face detection and recognition

Detection

Detects faces in real time at low power

Recognition

Involves accurately comparing detected face with a library of known faces

Benefits

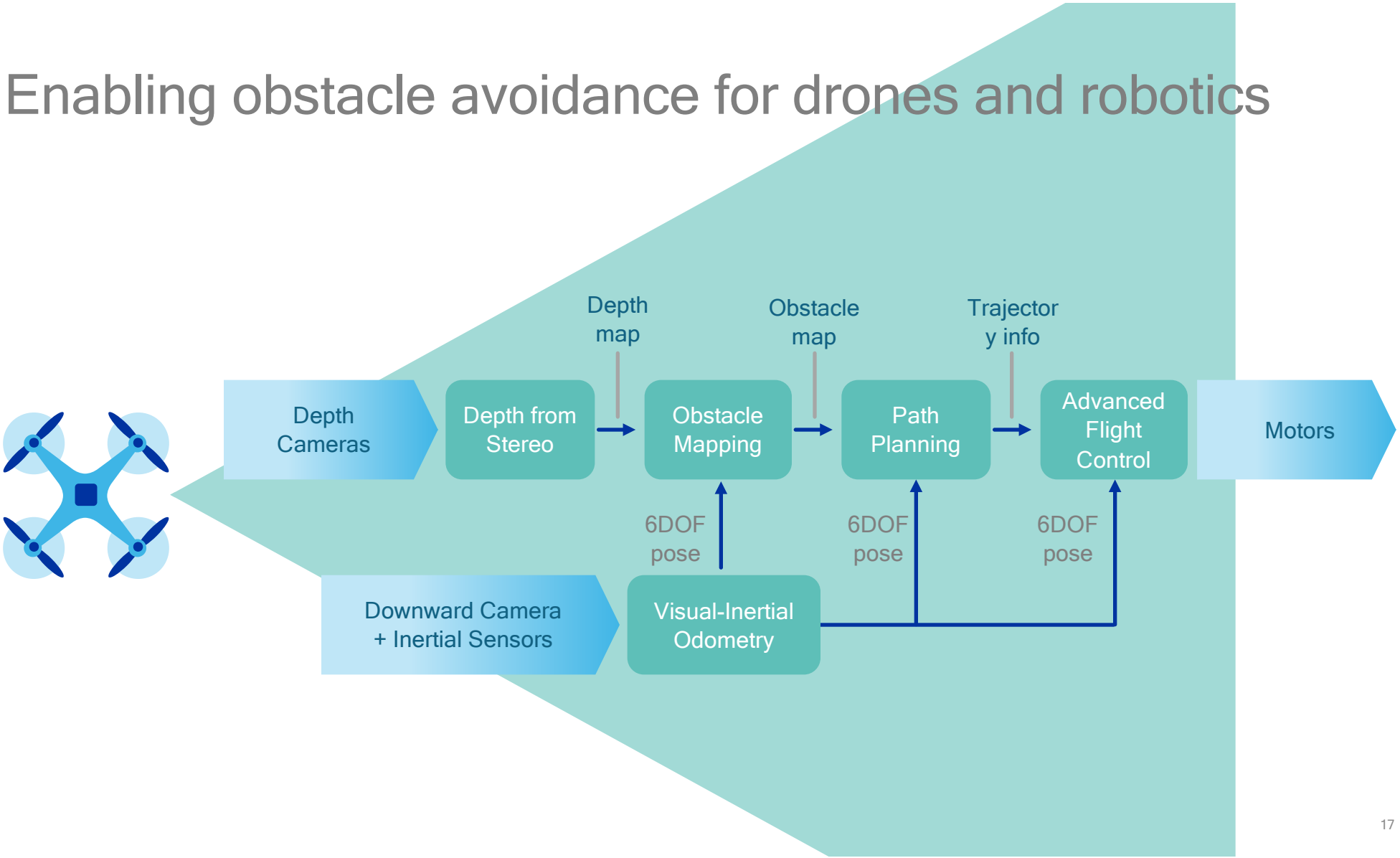
- Robust recognition at a distance
- Real time identification (no cloud needed)
- Better off-angle recognition
- Take action by just touching the photos in the image

Drone use case

Obstacle avoidance

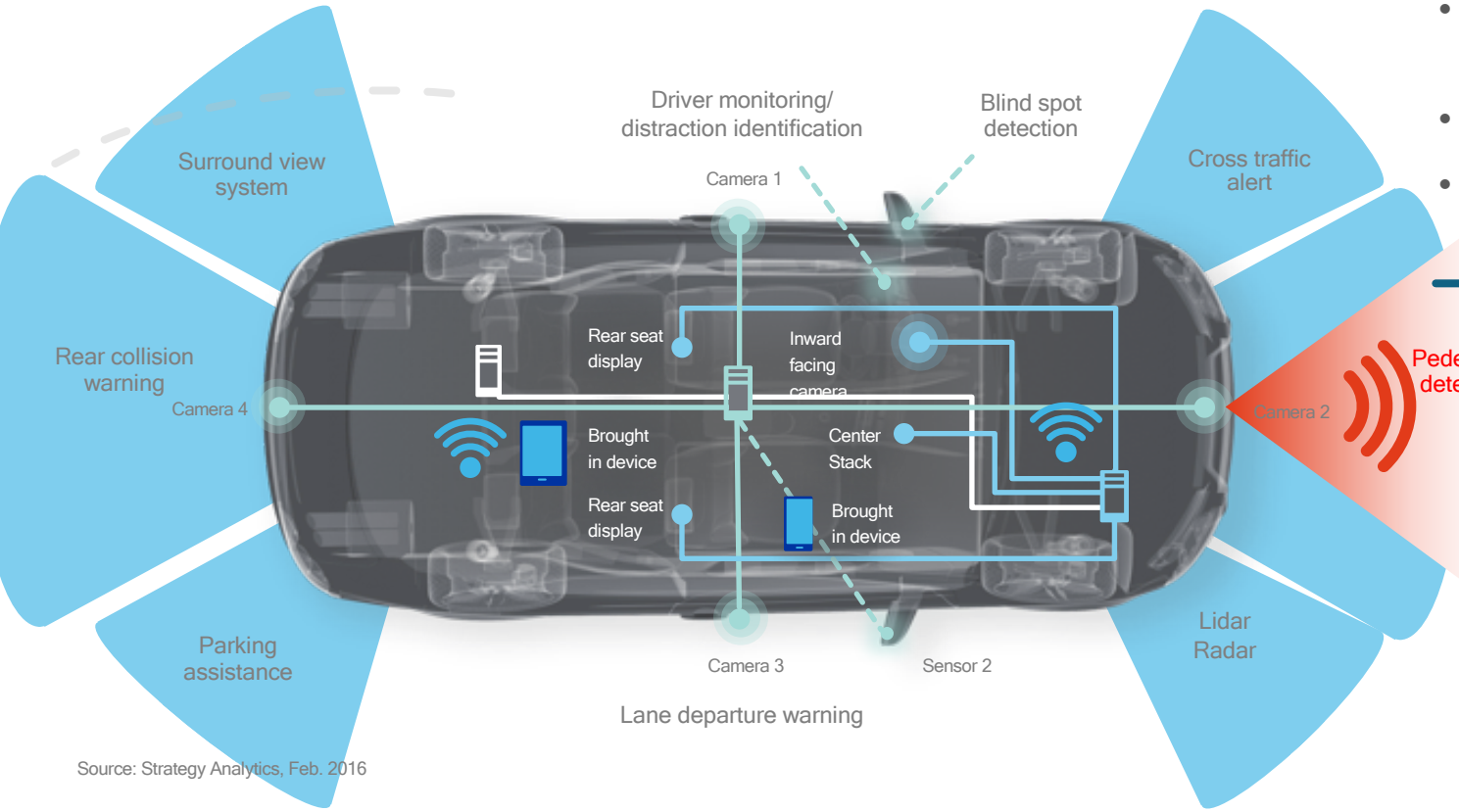
- Computer vision allows drones to map and avoid obstacles in their path, making navigation safer
- Key for many drone applications, including flying cameras, delivery, agriculture and public safety drones

Enabling obstacle avoidance for drones and robotics



Automotive use case

Vision is enabling ADAS today and autonomous driving in the future



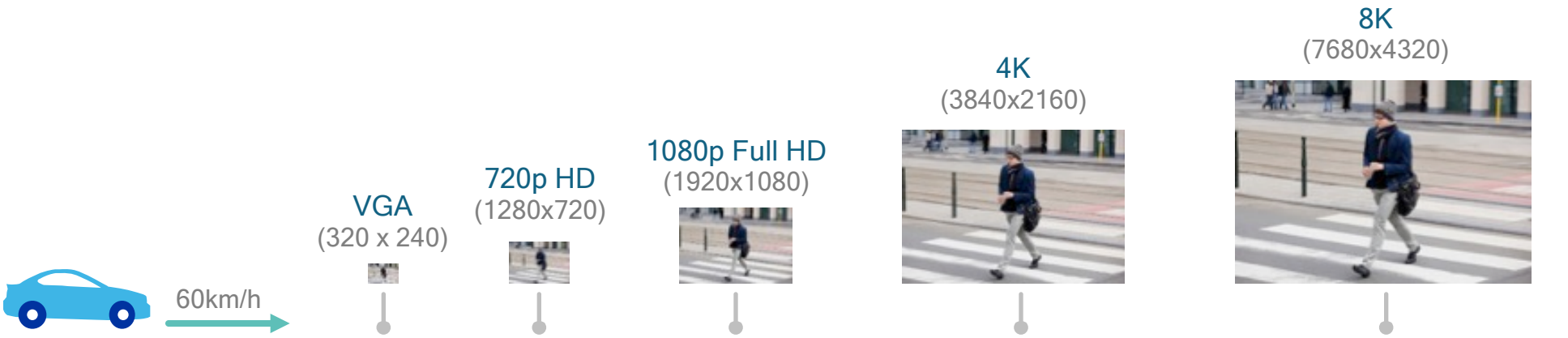
Pedestrian detection

- Uses robust object detection and tracking
- Looks for specific patterns
- Required to work at a distance

Source: Strategy Analytics, Feb. 2016

Pedestrian detection

Using robust feature detection at a distance

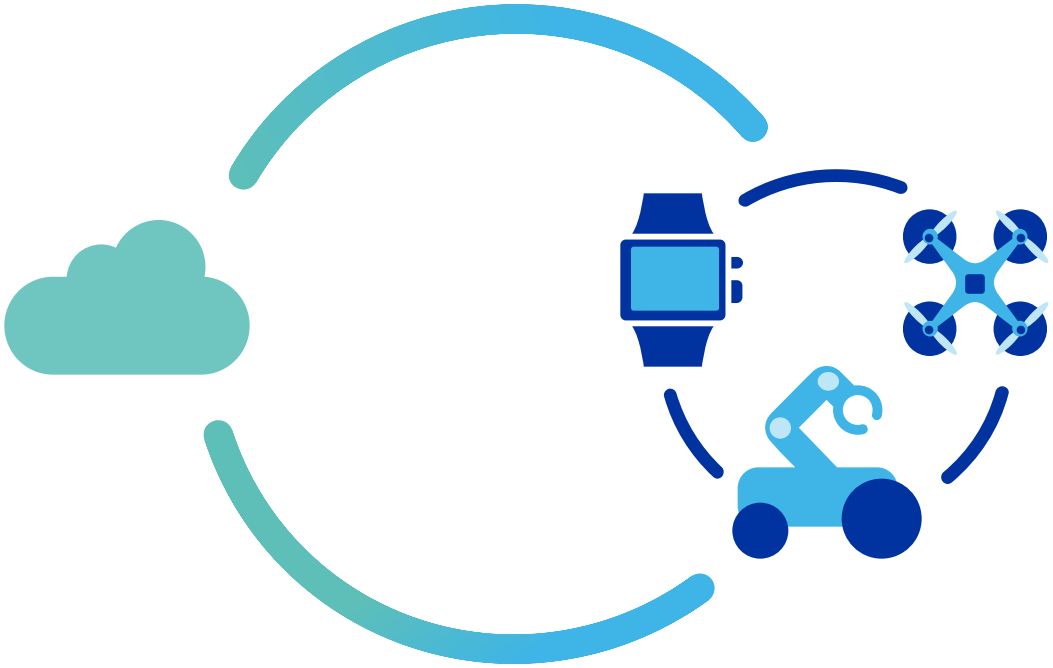


Alert time	0.6s	1.2s	2.1s	4.2s	8.4s
Distance from Camera	10m	20m	35m	70m	140m
Increase in compute requirements	1X	3X	6.75X	27X	108X

Bringing CV to mobile devices is challenging

On-device processing for vision workloads is key

Process data closest to the source, complement cloud



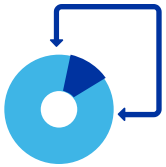
- Efficient use of network bandwidth
- Security and user privacy
- Reliability
- Low latency

Processing vision on mobile devices is a key challenge

Compute intensive work loads in mobile constrained environment

Visual perception workloads

Compute intensive



Evolving requirements

Constrained mobile environment

Storage and memory bandwidth limitations

Battery powered

Thermal efficiency

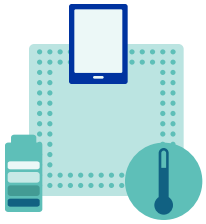
Qualcomm Technologies is tackling mobile vision challenges

Example: High-resolution 3D Reconstruction on a mobile processor

How Qualcomm Technologies is solving mobile vision challenges

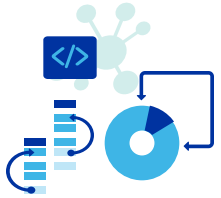
Superior camera support with Spectra ISP

- Efficient image processing
- Flexible interfaces for 2D and 3D sensors



Powerful Heterogeneous Snapdragon Processors

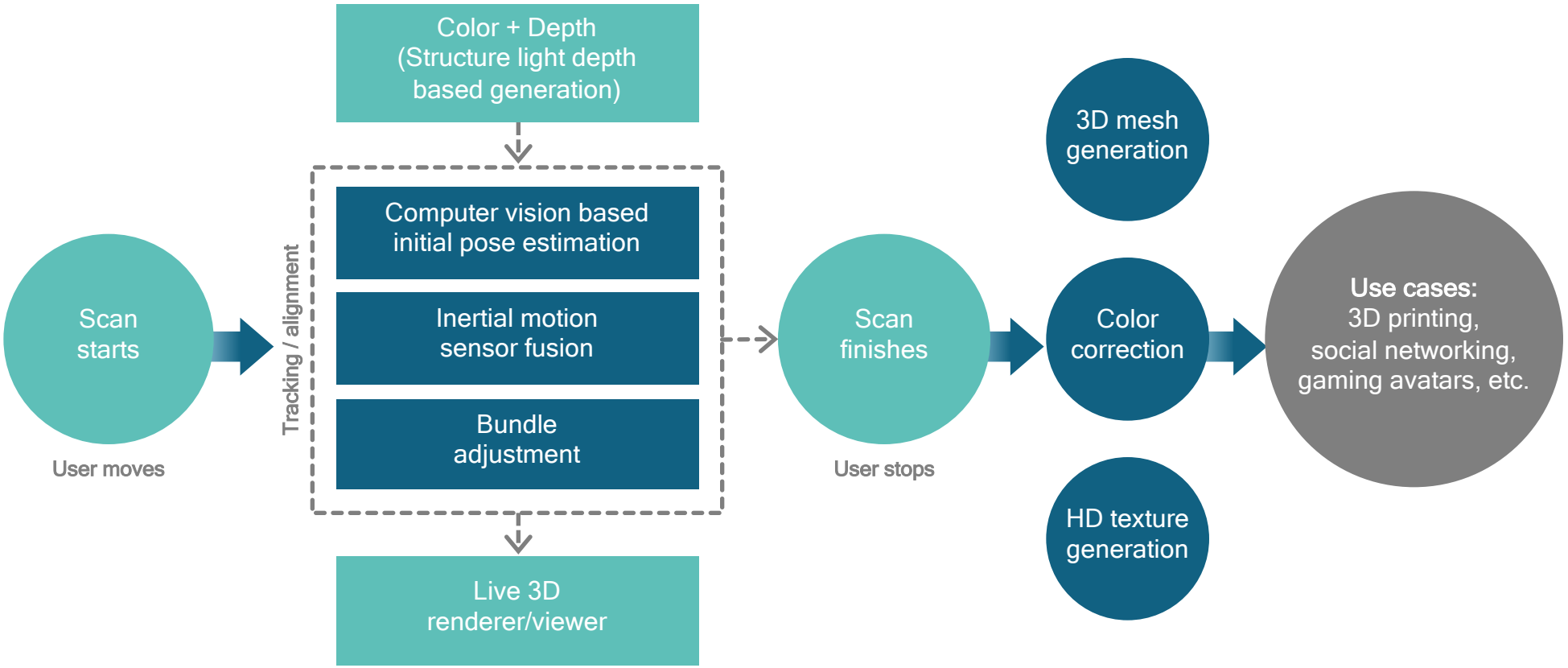
- Running the right algorithm on the right processing engine
- Process compute intensive CV features within power and thermal limits



Optimized algorithmic support and availability

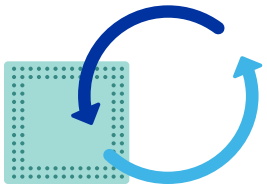
- Access to top tier CV algorithms with OpenCV and FastCV libraries
- FastCV provides a mobile optimized library for key CV functions
- Profiling tools to quickly identify performance bottlenecks

3D reconstruction block diagram



Spectra™ ISP enables 3D Reconstruction on mobile

1. Great interface support



- Connects 2D color sensors to Snapdragon via MIPI, enabling color information to be applied to 3D-reconstructed objects
- Allows various kinds of 3D depth sensors to connect to Snapdragon

2. Camera Synchronization

- Supports tight hardware and software synchronization of camera frames, facilitating multi-sensor frame alignment

Powerful heterogeneous Snapdragon processors

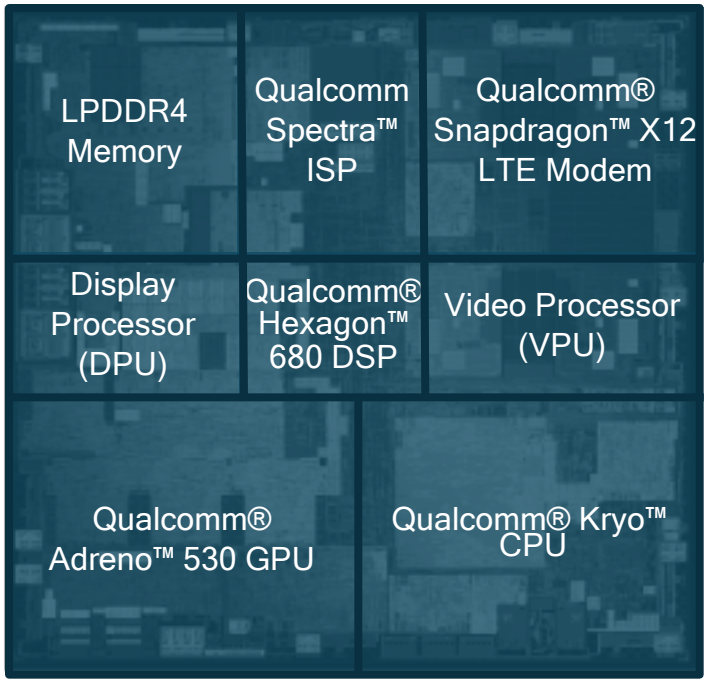
Enabling high performance at low power and thermal

Parallelism

- Partitioning 3DR algorithms across our heterogeneous engines

Matching CV algorithms to appropriate processing engine

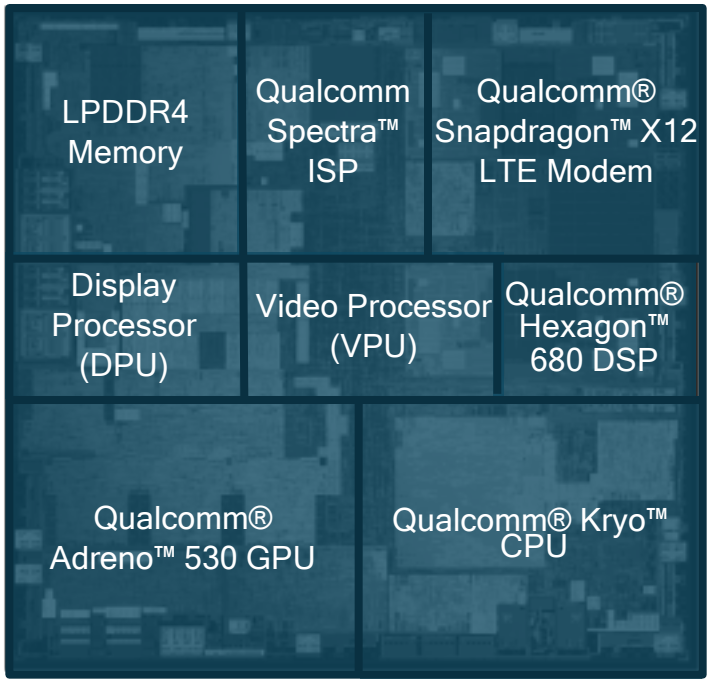
- Achieving more work to be done per clock cycle, power savings, and reduced latency



Example: 3D Reconstruction on Snapdragon 820

Using heterogeneous computing framework to do all of this at 15 FPS

- RGB sensor processing
- Depth sensor interface

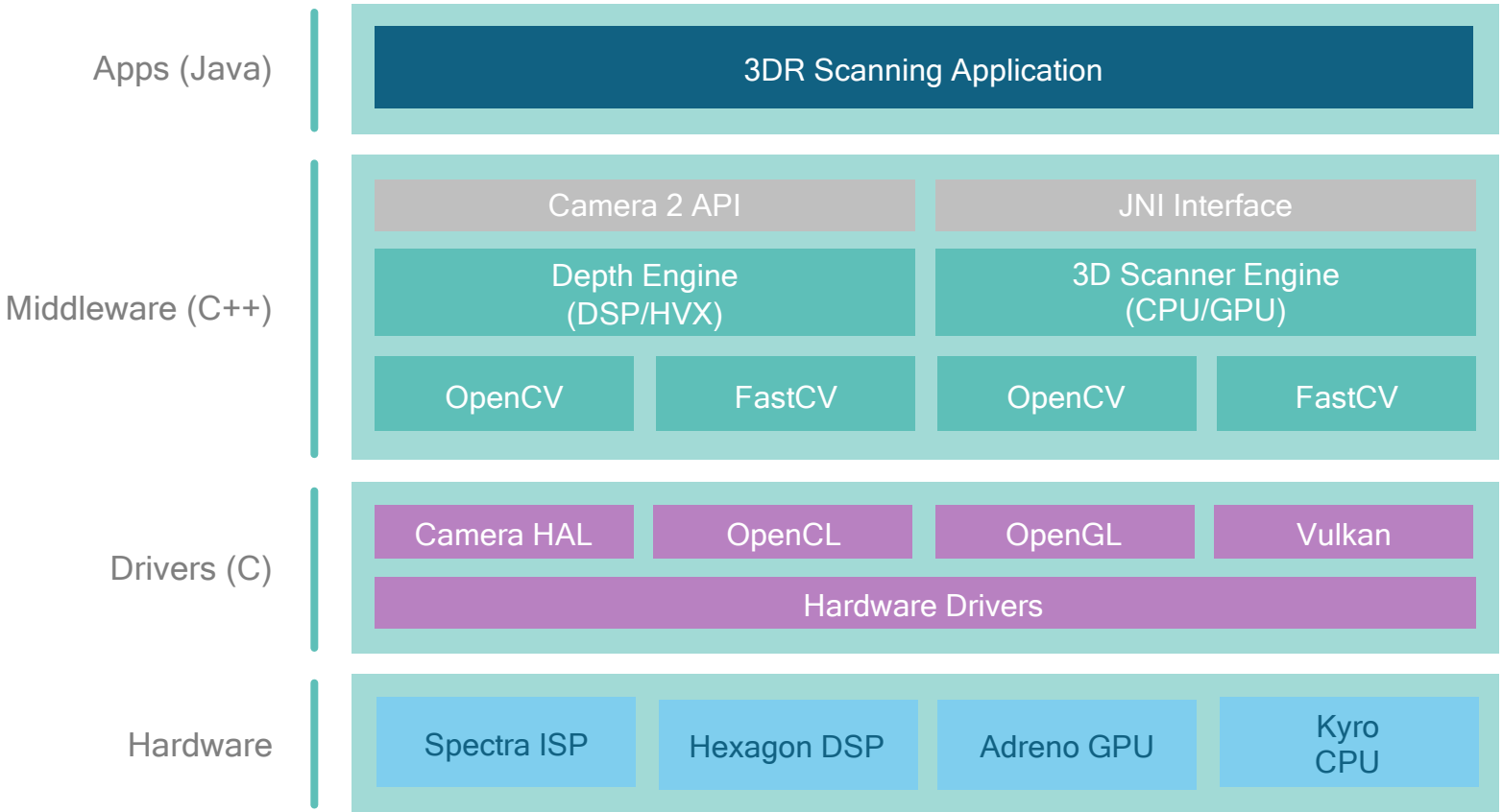


- Depth extraction from structured light

- Point cloud rendering
- Texture mapping
- Shading

- Pose estimation and tracking
- Bundle adjustment
- Visual and inertial sensor data fusion
- Mesh generation

3D Reconstruction stack diagram



Ubiquitous deployment of visual intelligence

Bee-sized flying cameras

Intelligent cameras

Adaptive self-driving cars

Aerial, 360 virtual reality

We enable ubiquitous deployment of visual intelligence

1. Computer vision enables a broad range of applications for different market segments
2. On-device processing is key to ubiquitous adoption of vision in our daily life
3. Qualcomm® Snapdragon™ brings CV to mobile devices at low power and thermal
4. Qualcomm Technologies is bringing our mobile vision to different ecosystems

Thank you

Follow us on:

For more information, visit us at:
www.qualcomm.com & www.qualcomm.com/blog