Alternative Image Sensors for Intelligent In-cabin Monitoring, Home Security and Smart Devices

Dr. Petronel Bigioi
XPERI

XPERI

About XPERI

XPERI

Portfolio of Trusted Brands

dts	HD Radio	IMAX ENHANCED	invensas	TESSERA	TiVo
Sound, Connected Media, Imaging, and Computer Vision Solutions	Digital Broadcast Radio and Connected Car Services	Creation and Delivery of Immersive Cinematic Entertainment	Semiconductor and Interconnect Packaging Technology and Solutions	Semiconductor Intellectual Property Licensing	Entertainment Software, Consumer Hardware, Data and Advertising Solutions

Introduction

- Visible light constitutes a small portion of the electromagnetic spectrum, and image frames capture an instant in time rather than a continuous sequence.
- This motivates exploring beyond the visible and finding alternatives to the use of frame-based imaging.
- What important information do we miss with frame based visible imaging?

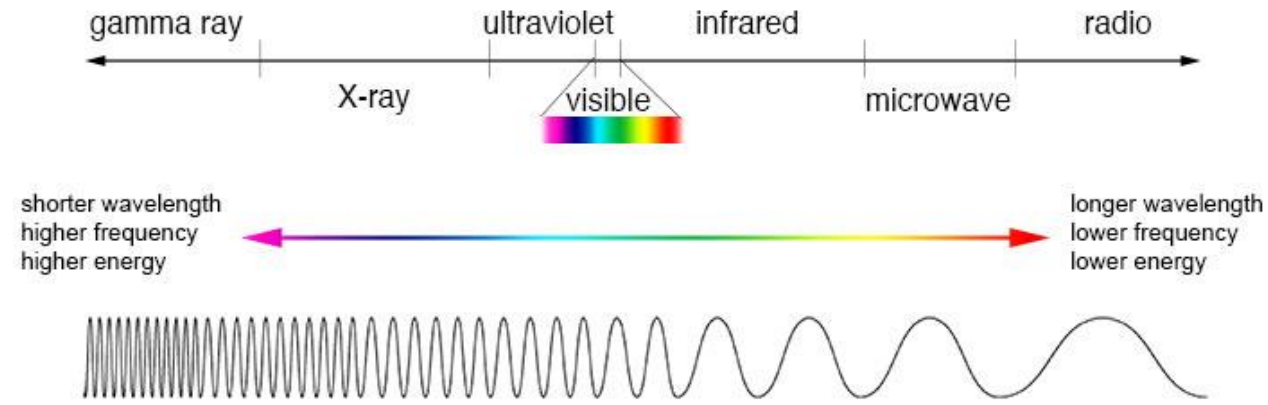


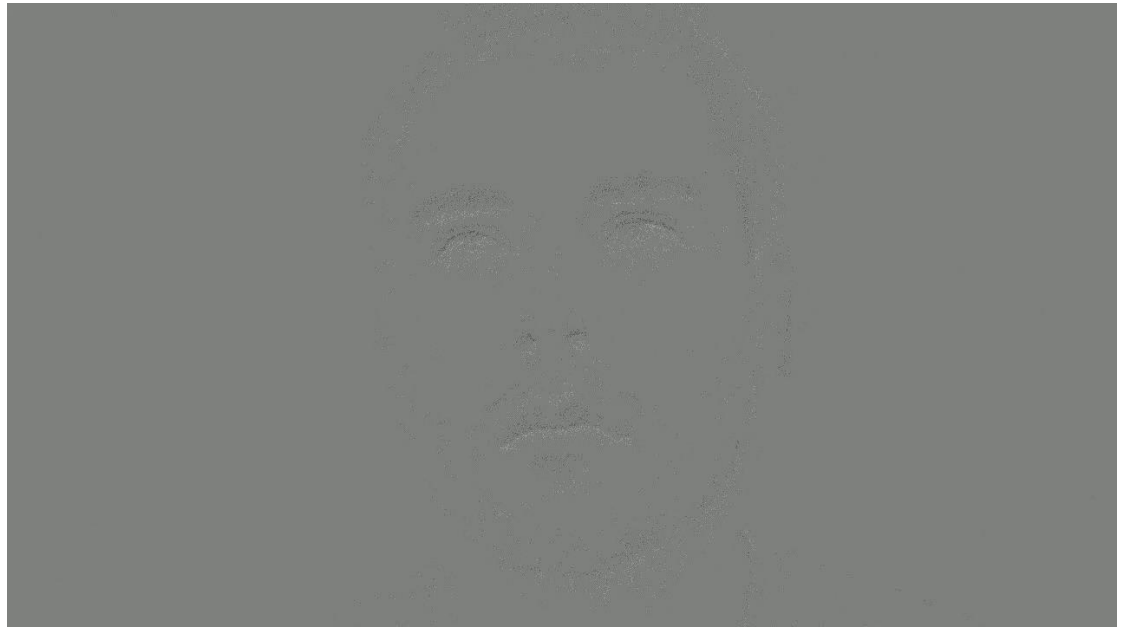
Image Credit NASA Explore

Event Cameras

- Event cameras report light intensity changes
- Each pixel responds only to motion or changes in the scene
- Asynchronous “events” instead of frames

Event Stream

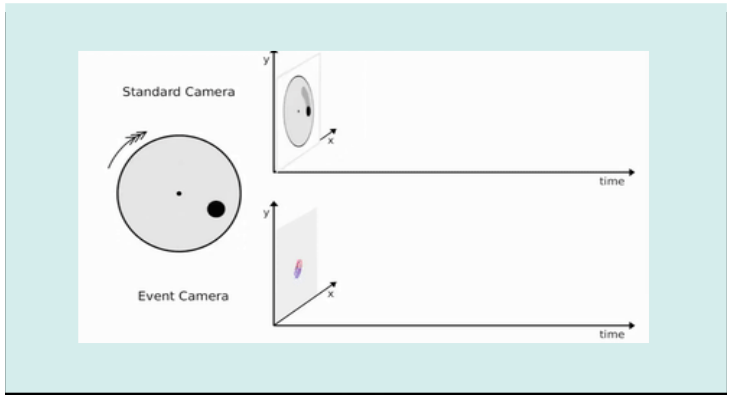
time	x	y	polarity
1.1177435000000000e+07	7.0600000000000000e+02	1.3600000000000000e+02	1.0000000000000000e+00
1.1177437000000000e+07	4.8500000000000000e+02	5.9100000000000000e+02	0.0000000000000000e+00
1.1177446000000000e+07	7.6600000000000000e+02	3.0400000000000000e+02	0.0000000000000000e+00
1.1177465000000000e+07	6.2900000000000000e+02	1.7900000000000000e+02	0.0000000000000000e+00
1.1177485000000000e+07	8.6100000000000000e+02	1.8600000000000000e+02	0.0000000000000000e+00
1.1177489000000000e+07	6.5100000000000000e+02	1.2200000000000000e+02	0.0000000000000000e+00
1.1177489000000000e+07	7.0200000000000000e+02	4.8400000000000000e+02	0.0000000000000000e+00
1.1177489000000000e+07	7.3000000000000000e+02	1.5900000000000000e+02	0.0000000000000000e+00
1.1177494000000000e+07	7.6500000000000000e+02	4.9300000000000000e+02	0.0000000000000000e+00
1.1177494000000000e+07	7.9800000000000000e+02	1.9900000000000000e+02	0.0000000000000000e+00
1.1177495000000000e+07	6.9800000000000000e+02	3.9500000000000000e+02	0.0000000000000000e+00
1.1177499000000000e+07	7.8800000000000000e+02	1.0800000000000000e+02	0.0000000000000000e+00
1.1177509000000000e+07	6.5700000000000000e+02	1.1700000000000000e+02	0.0000000000000000e+00
1.1177510000000000e+07	6.8400000000000000e+02	1.8700000000000000e+02	0.0000000000000000e+00
1.1177525000000000e+07	6.9100000000000000e+02	3.8400000000000000e+02	1.0000000000000000e+00
1.1177550000000000e+07	6.9900000000000000e+02	1.4400000000000000e+02	0.0000000000000000e+00
1.1177558000000000e+07	7.0000000000000000e+02	1.6900000000000000e+02	0.0000000000000000e+00
1.1177559000000000e+07	8.3300000000000000e+02	9.5000000000000000e+01	0.0000000000000000e+00
1.1177562000000000e+07	6.7600000000000000e+02	2.7800000000000000e+02	0.0000000000000000e+00
1.1177569000000000e+07	6.7300000000000000e+02	3.9600000000000000e+02	0.0000000000000000e+00



Why Event Cameras

Much higher temporal resolution (microseconds)

- Event cameras can operate at the equivalent of > 10,000 fps
- Can capture fast moving objects with no motion blur or loss of information
- Video shows a reconstruction from an event-based camera compared with one from an NIR camera. Note the motion blur in the visible.



Reconstructed Events

Frames from NIR camera

XPERI

Why Event Cameras

Higher dynamic range (> 120 dB)

- Event camera can operate in extreme lighting conditions
- Operation under low lighting conditions is crucial for in-cabin monitoring, home surveillance and smart cameras
- Closer to dynamic range of the eye

Copyright Prophesee inc. Used with permission.

Low-lighting examples

Why Event Cameras

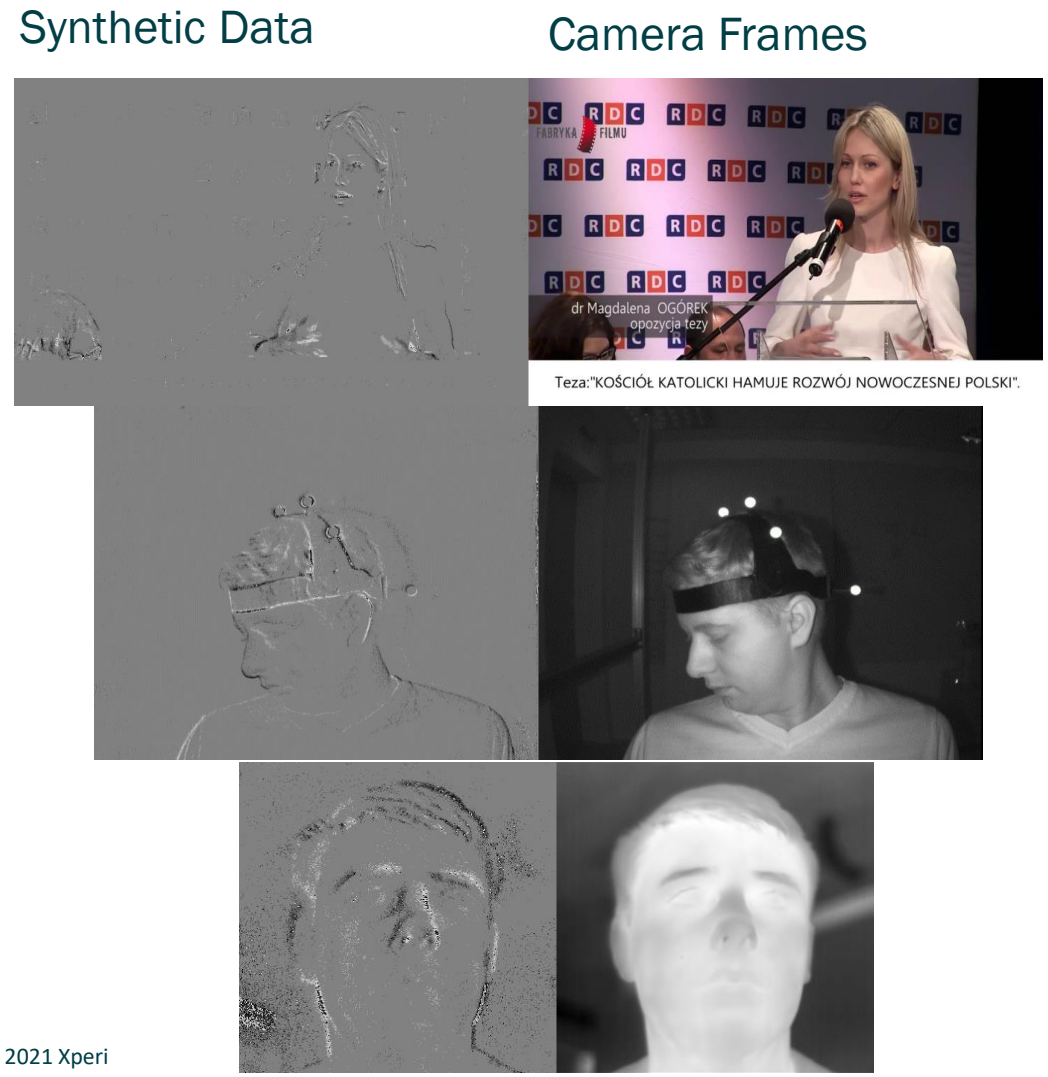
Responds to motion (asynchronous pixels)

- Enable adaptive framerates i.e. data is processed only when a scene has changed
- Ignores “unwanted” information i.e. static, unchanging objects. This is often an advantage but is also a disadvantage when static information is required.
- Reduces computational costs and memory usage
- We are normally only interested in changes in the state of the driver (driver monitoring) or changes in the scene (surveillance)

Surveillance example

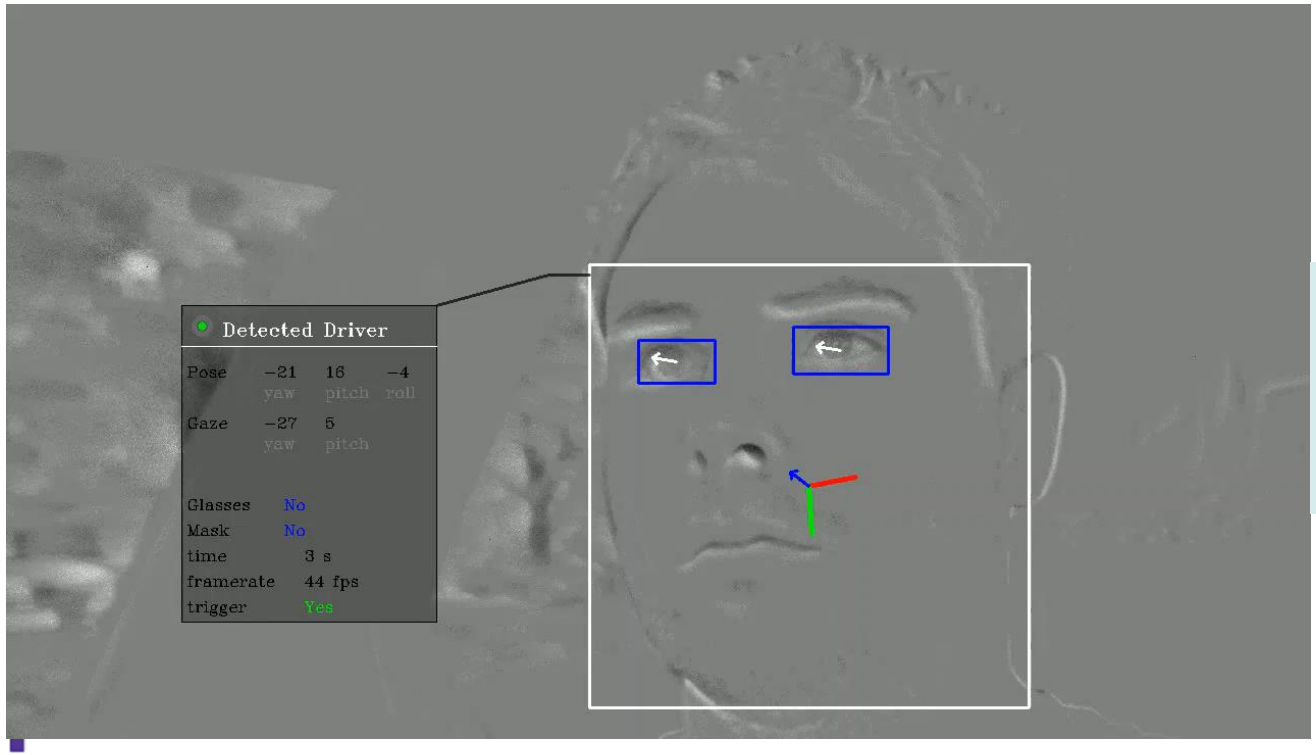
Dataset re-use – Training machine learning models

- Traditional frame-based images can be used to create synthetic event data for DNN training
- Both real and synthetic datasets can be used for neural network training
- Synthetic data simulators are more effective in comparison to frame-based cameras methods
- Event can be simulated using visible (RGB), near-infrared and even thermal images



Facial Analytics

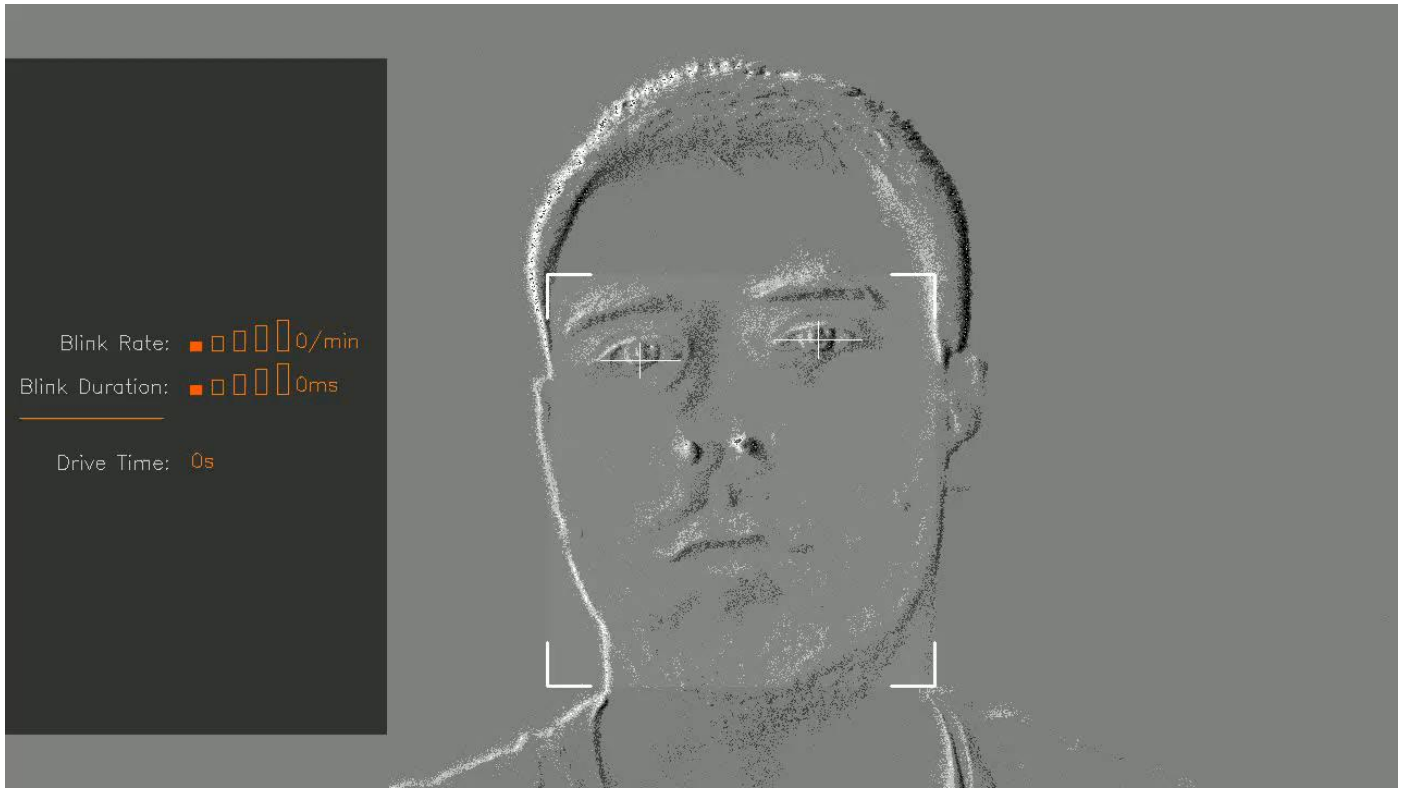
- Driver facial analytics
- Adaptive framerate based driver motion (video slowing effect with faster motion)



- Face/Eye detection
- Head pose estimation
- Eye gaze estimation
- Occlusion detection

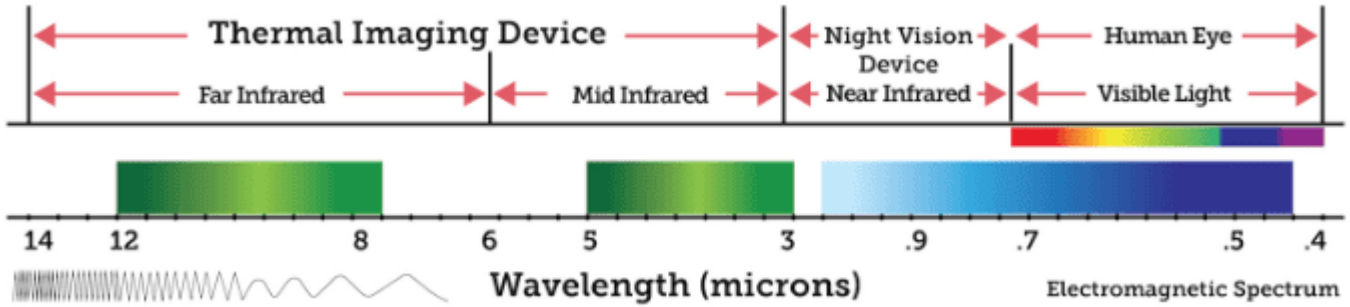
Advanced Analytics

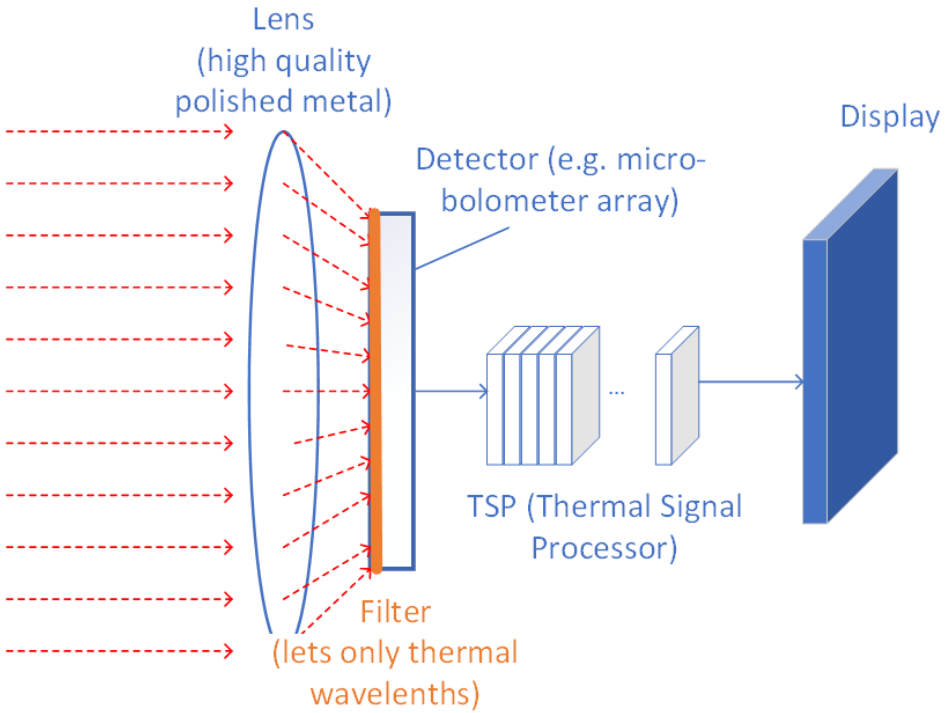
- Eye blinking patterns can be analysed at a desired rate
- Possible features include blink frequency, duration, speed, downward/upward duration and speed, time closed etc.



Thermal Cameras

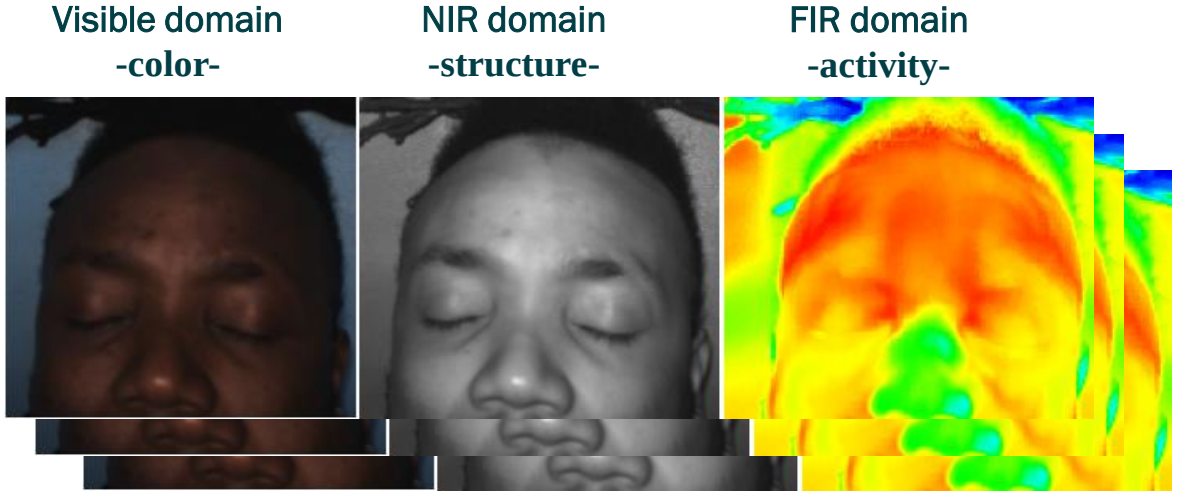
- Thermal cameras detect far infrared energy i.e. heat
- They typically capture wavelengths between 8-14 µm
- Microbolometer array sensors are used and respond to heat changes
- Thermal reflections can be a problem and different materials have different emissivity.





Thermal Cameras

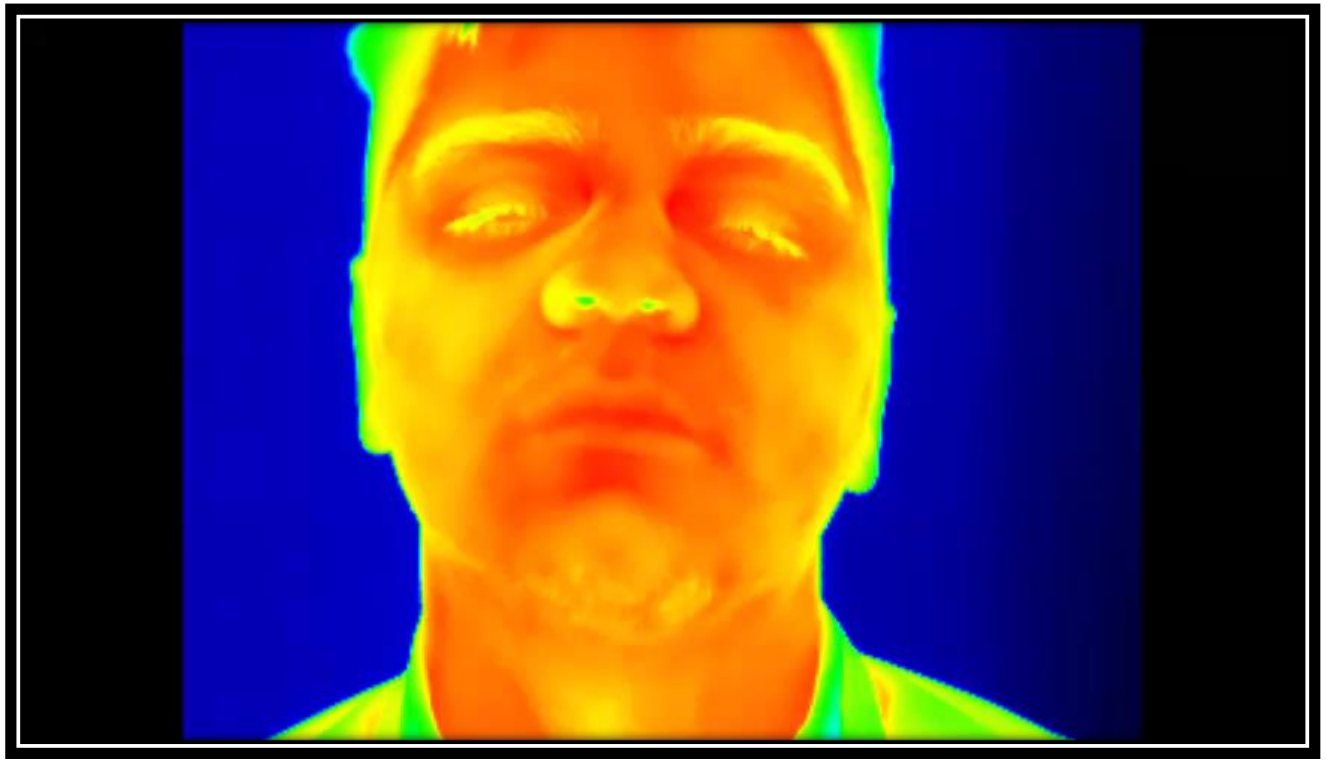
- Provide information beyond the visible and near-infrared light spectrums
- They can operate under any lighting conditions
- Disadvantages include low resolution, increased cost, and the difficulty of determining certain ocular parameters such as eye gaze.



Thermal Cameras

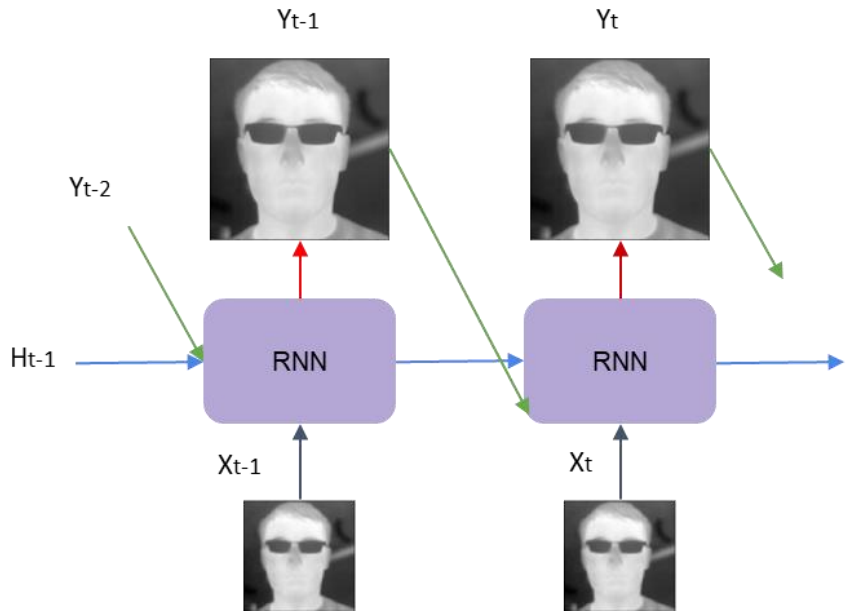
- Thermal cameras enable us to capture physiological features beyond the capabilities of visible or NIR cameras

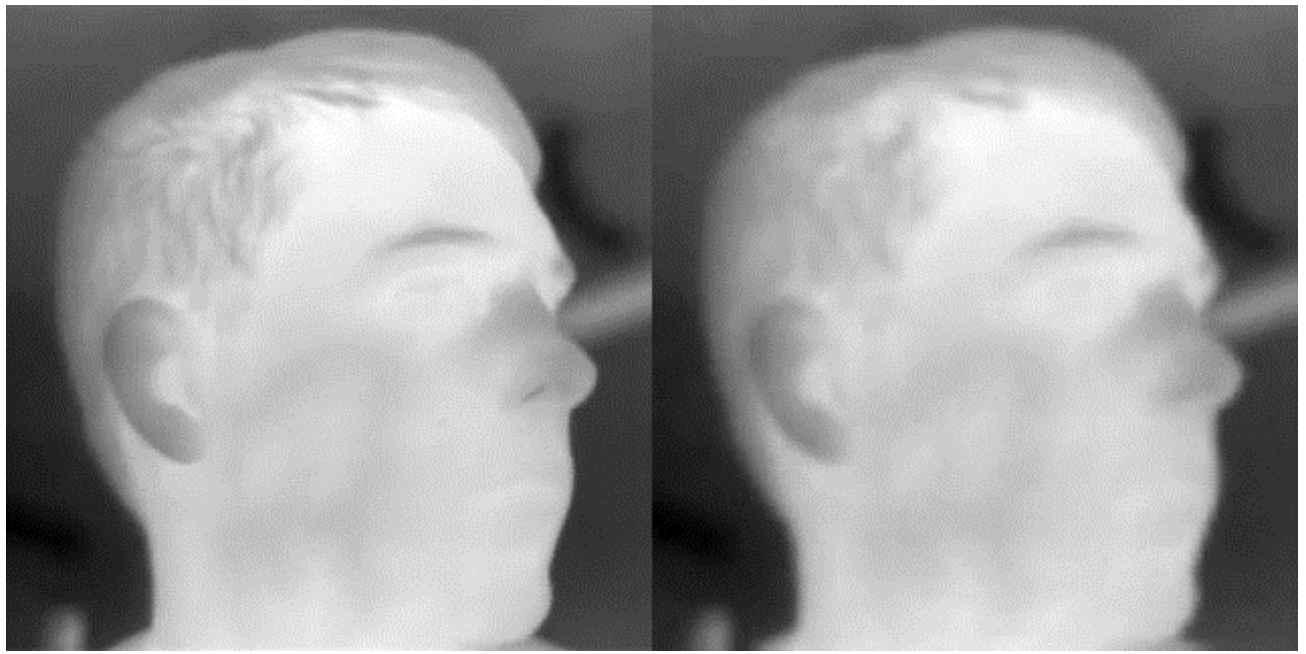
- Advanced vital sign monitoring:
  - Breathing rate
  - Heart rate
  - Temperature
  - Enhanced Drowsiness or attentiveness
  - Skin health



Thermal Cameras
Thermal Super-Resolution

- Recovering high-resolution images from single or multiple low-resolution images
- Super-Resolution includes de-noising, de-blurring and up-sampling
- Supports the use of cheaper, lower-resolution thermal cameras
- Convolutional Neural Networks provide state-of-the-art and real-time performance

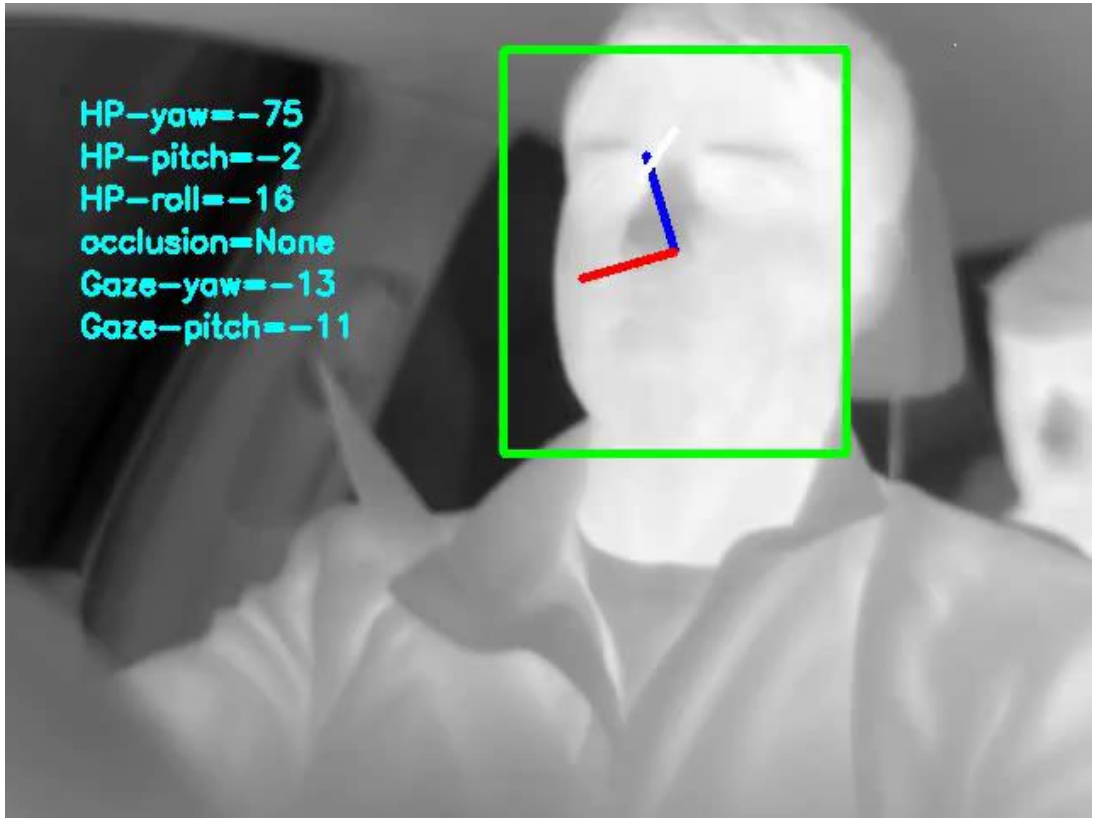




Thermal Cameras

Early driver monitoring system prototype from Xperi built with a thermal camera supporting features such as:

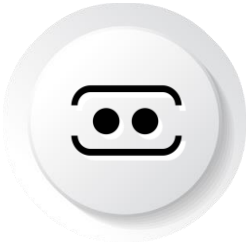
- Head Pose estimation
  - Yaw, pitch and roll angles
- Occlusion detection
  - glasses, masks, phones etc.
  - Some objects are easier to detect in thermal compared to visible or near-infrared
- Liveness detection
  - Anti-spoofing



The Future of Sensing

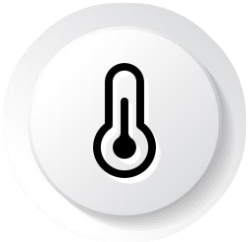
2021 embedded VISION summit

ALTERNATIVE SENSING



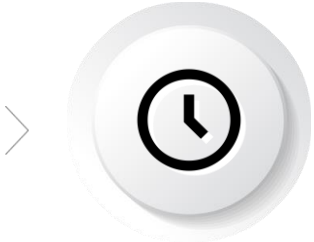
DUAL SYSTEM CAMERA

- narrow + wide configuration
- Face recognition + IRIS (Bank and payment grade biometrics)



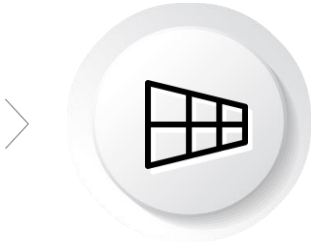
THERMAL SENSING

- vital sign monitoring
- driver comfort monitoring



TIME EVENT CAMERA

- precise eye analytics
- NIR camera substitute for DMS



DEPTH SENSING

- airbag deployment + body estimation
- 3D Facial Recognition, payment-grade + anti-spoofing

MMW RADAR

- reliable presence sensing

XPERI

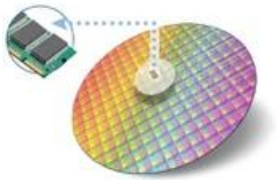
Resources

- https://www.prophesee.ai/
- https://www.prophesee.ai/category/dataset/
- https://www.heliaus.eu/
- https://doi.org/10.1016/j.neunet.2021.03.019
- https://doi.org/10.1109/QoMEX48832.2020.9123135
- https://github.com/uzh-rpg/rpg_e2vid
- https://github.com/uzh-rpg/rpg_vid2e
- https://www.flir.eu/oem/adas/adas-dataset-form/
- https://dts.com/autosense/

Thank You!

XPERI

XPERI Market Presence

DTS® and HD Radio™ audio and broadcast data technologies		Over 2.65 billion devices and vehicles
TiVo® and DTS® enabling premium entertainment		Over 5 billion hours of video consumed
DTS® imaging hardware and software		Over 3.65 billion devices and vehicles
Invensas ZiBond® and DBI® interconnect technologies		Over 1 billion smartphones
Tessera packaging/ interconnect IP		Over 100 billion chips