Putting Eyes and ARMs on Machines

Sun Huo
Product Marketing
Mass Market Group
Marketing Strategy Center

Agenda

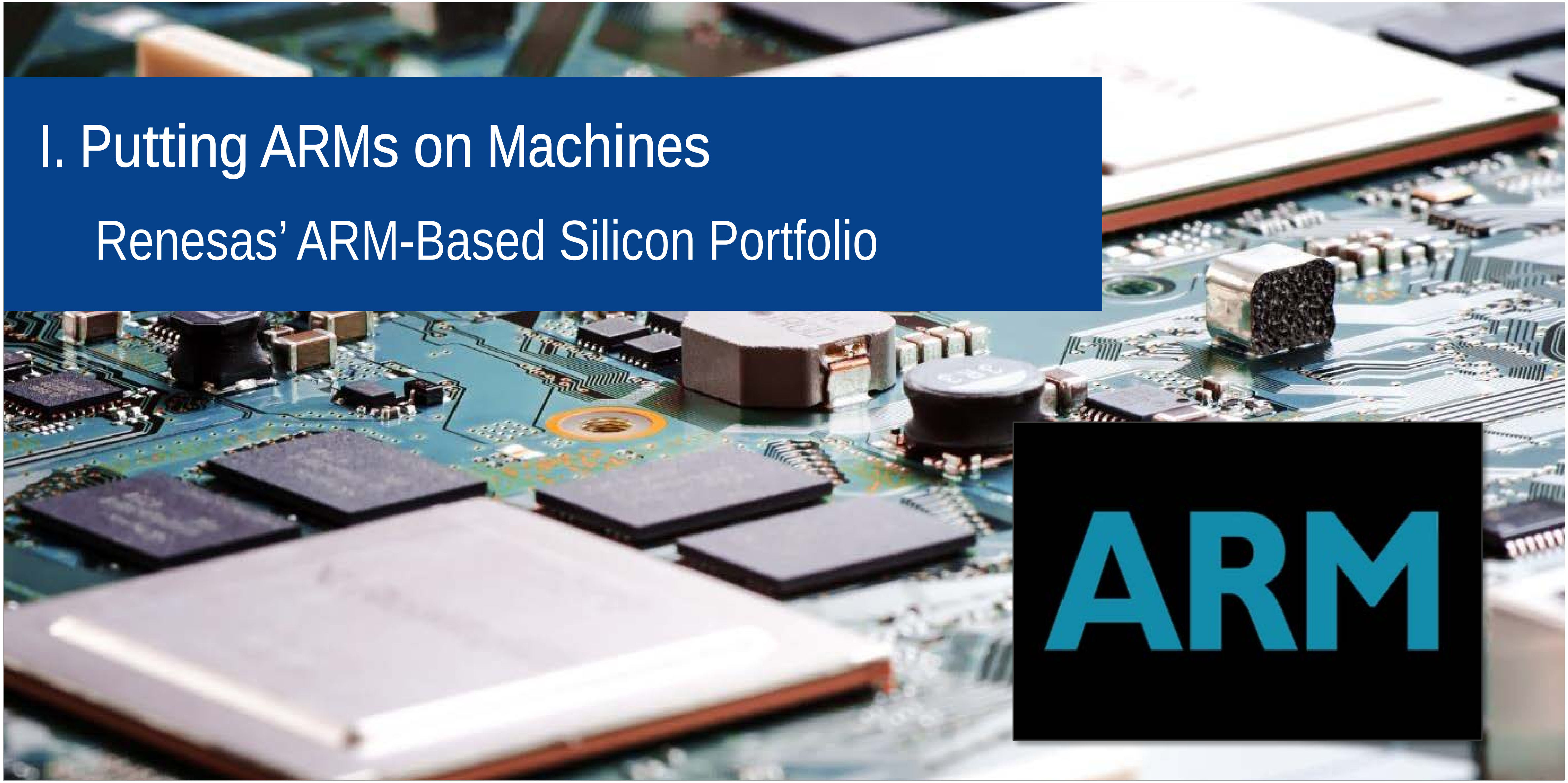
I. Putting ARMs on machines
   - Renesas' ARM-based silicon portfolio

II. Putting eyes on machines
   - Embedded vision concepts and applications

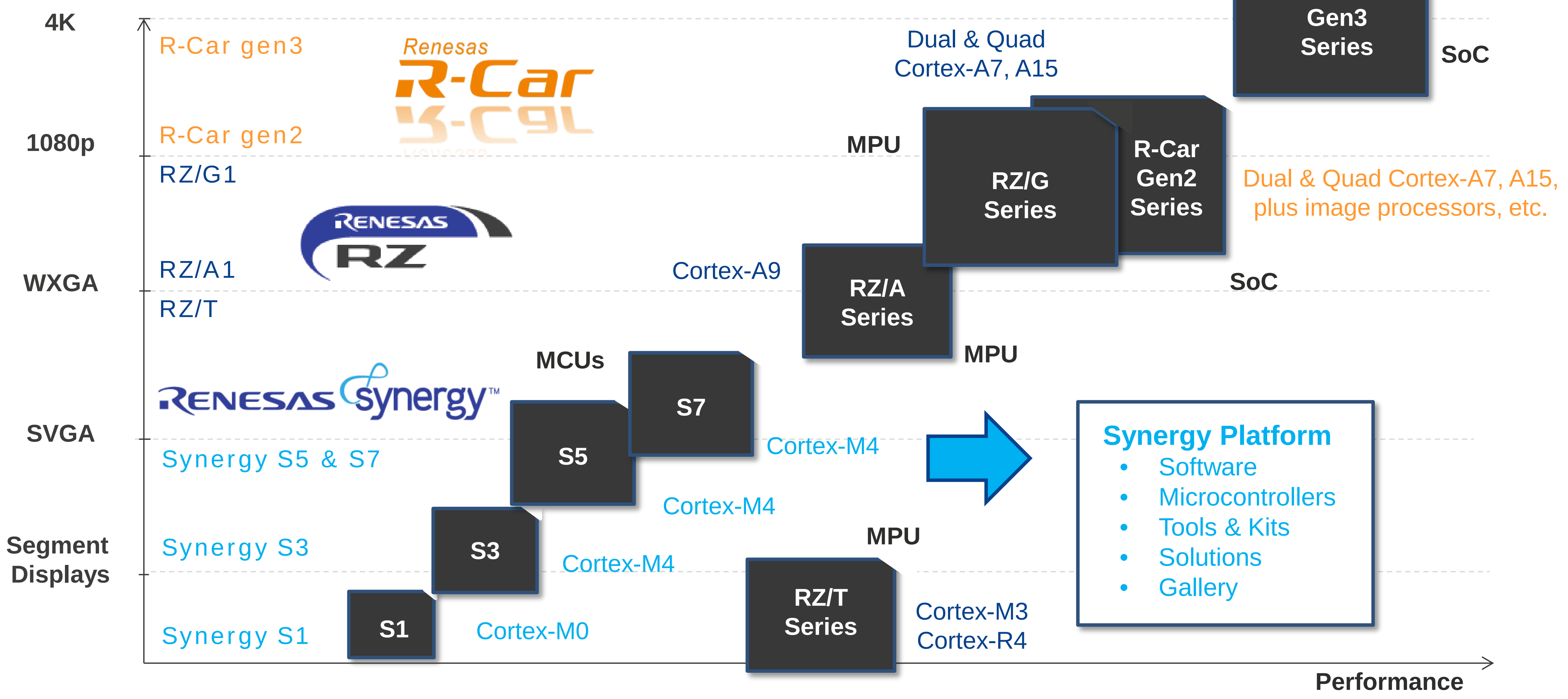
III. Future outlook and summary

I. Putting ARMs on Machines

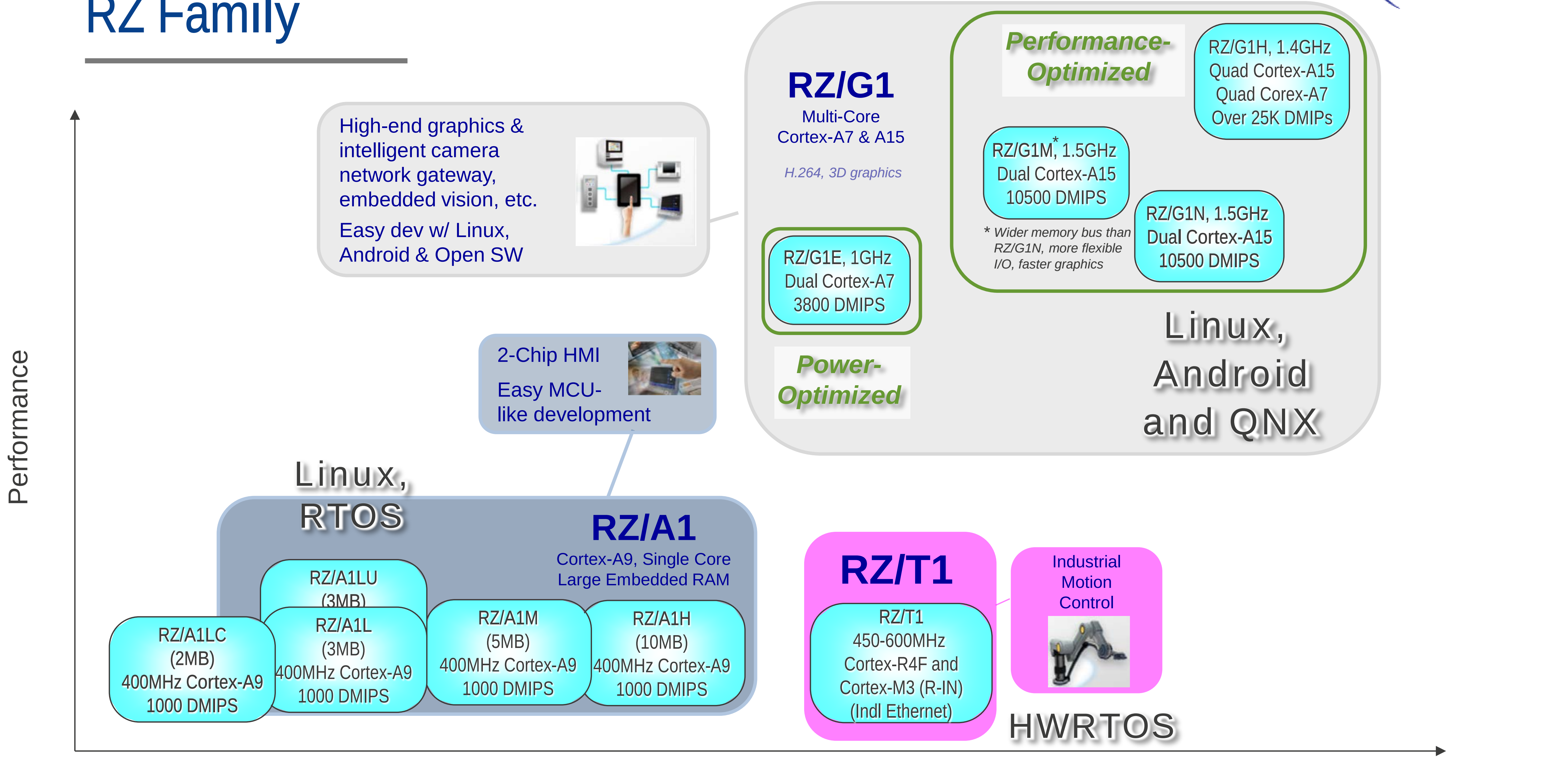
Renesas' ARM-Based Silicon Portfolio



Renesas ARM-Based MCUs and MPUs

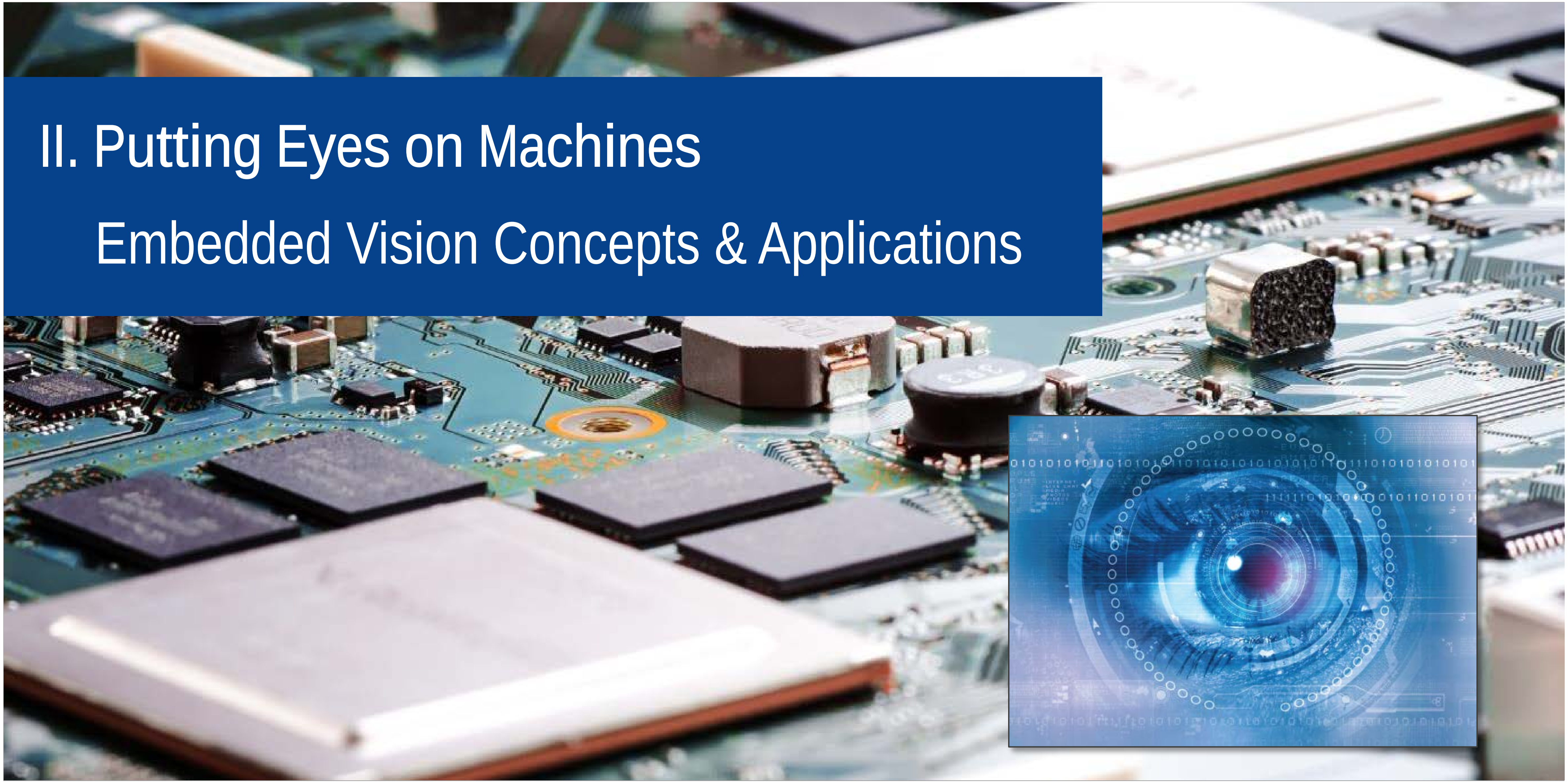


RZ Family


Performance
High-end graphics &
intelligent camera
network gateway,
embedded vision, etc.
Easy dev w/ Linux,
Android & Open SW
RZ/G1
Multi-Core
Cortex-A7 & A15
H.264, 3D graphics
Performance-
Optimized
RZ/G1H, 1.4GHz
Quad Cortex-A15
Quad Corex-A7
Over 25K DMIPs
RZ/G1M*, 1.5GHz
Dual Cortex-A15
10500 DMIPS
* Wider memory bus than RZ/G1N, more flexible I/O, faster graphics
RZ/G1N, 1.5GHz
Dual Cortex-A15
10500 DMIPS
RZ/G1E, 1GHz
Dual Cortex-A7
3800 DMIPS
Power-
Optimized
Linux,
Android
and QNX
2-Chip HMI
Easy MCU-
like development
Linux,
RTOS
RZ/A1
Cortex-A9, Single Core
Large Embedded RAM
RZ/A1LU
(3MB)
RZ/A1LC
(2MB)
400MHz Cortex-A9
1000 DMIPS
RZ/A1L
(3MB)
400MHz Cortex-A9
1000 DMIPS
RZ/A1M
(5MB)
400MHz Cortex-A9
1000 DMIPS
RZ/A1H
(10MB)
400MHz Cortex-A9
1000 DMIPS
RZ/T1
RZ/T1
450-600MHz
Cortex-R4F and
Cortex-M3 (R-IN)
(Indl Ethernet)
Industrial
Motion
Control
HWRTOS

II. Putting Eyes on Machines

Embedded Vision Concepts & Applications



From the Electric-Eye to Modern Security Camera Systems

An **electric eye** is a photodetector used for detecting obstruction of a light beam

Modern photodetector systems use an infrared light-emitting diode modulated at a few kilohertz, which allows the detector to reject stray light and improves the range, sensitivity and security of the device.

Popular Science 1937

AUTOS ARE COUNTED BY ELECTRIC EYES

VEHICLES passing a given point are automatically counted by a new traffic-recording device just introduced. Two infra-red lamps, housed a short distance apart and mounted on one side of a road, cast invisible beams across the highway to a photo-electric receiving unit on the other side. Interruption of the two beams by an auto actuates an electrical counting device, which can be set to total the number of passing vehicles by the hour, day, week, or month. Pedestrians are not counted, since the apparatus is so constructed that it registers only when both beams are blocked at the same instant.

Passing vehicles intercept parallel beams of light and are counted by a photo-electric recording device

36 POPULAR SCIENCE MONTHLY

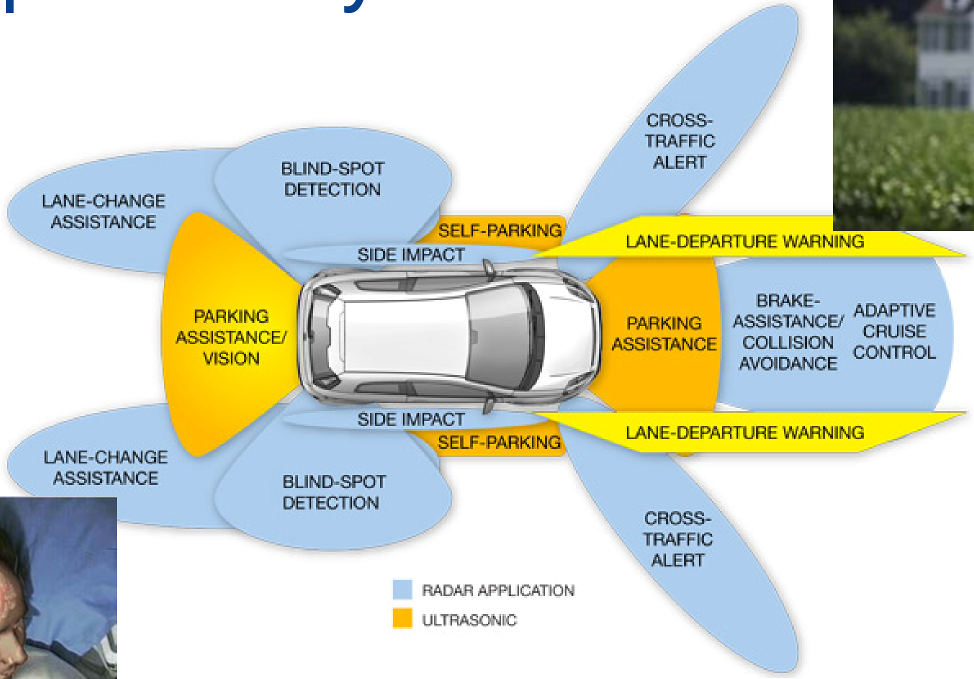
Outdoor Motion-Sensitive Light

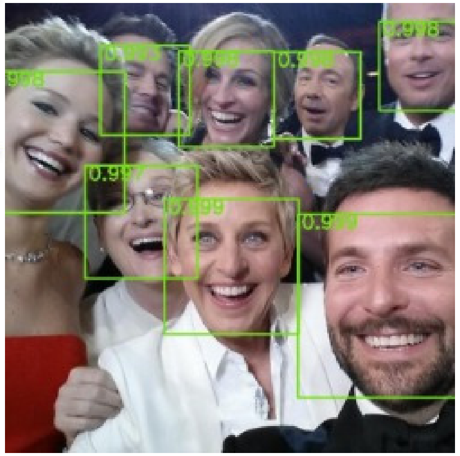
Laser Hallway

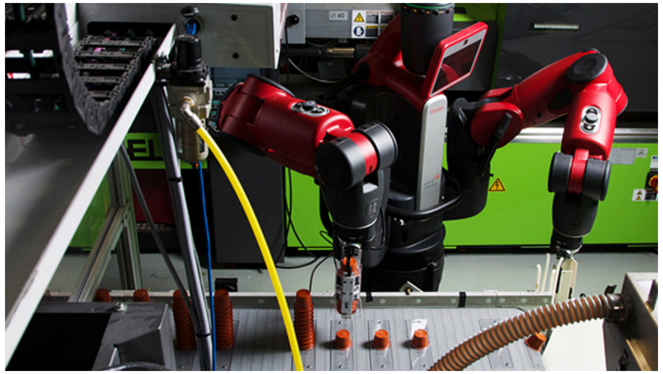
Garage Door Safety Sensors

Networked Security Systems

What Embedded Vision Implies Today







Embedded Computer Vision

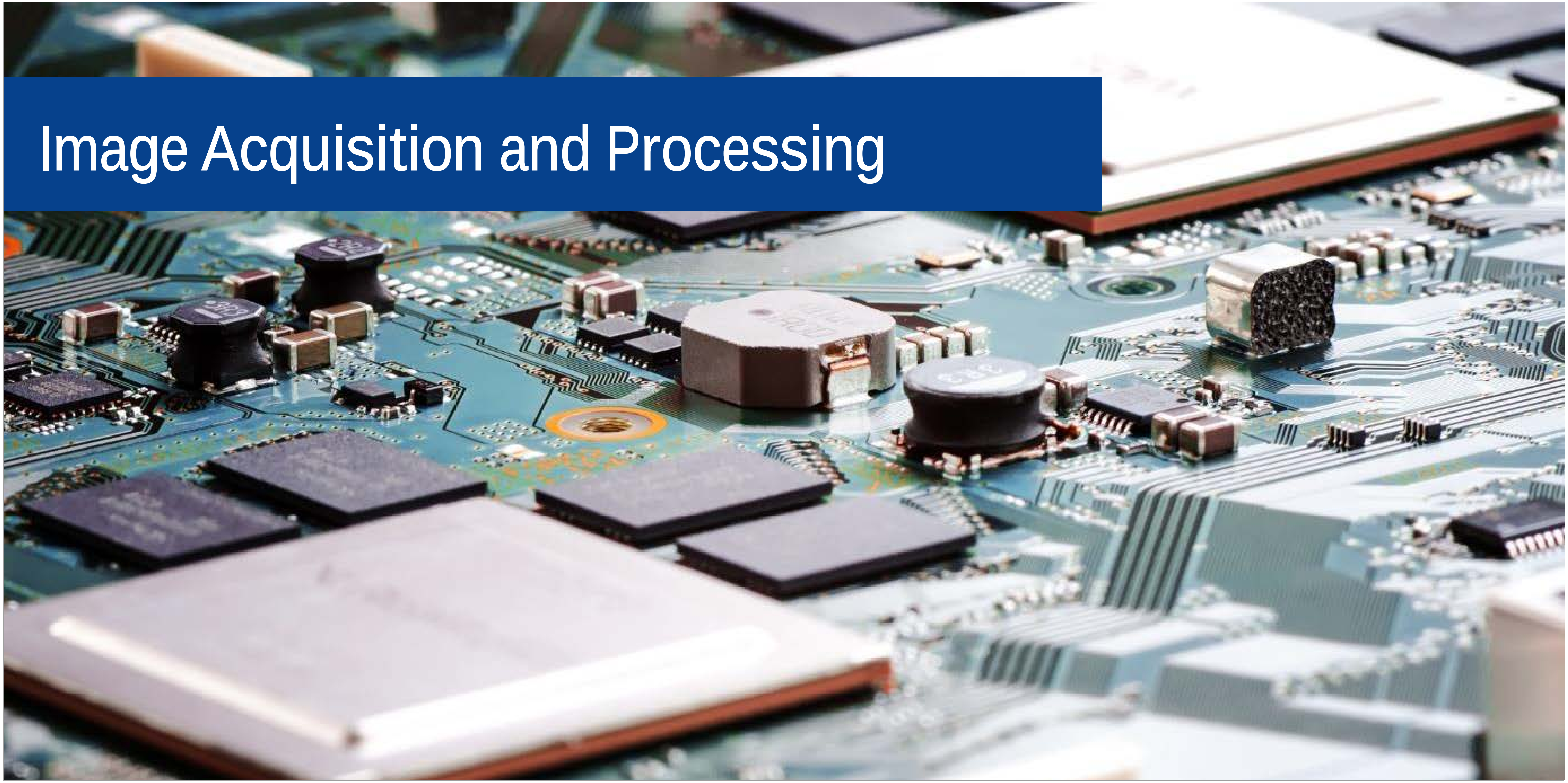
According to BDIT (URL)

- "Embedded Vision" refers to the practical use of computer vision in machines that understand their environment through visual means, and the use of digital processing and intelligent algorithms to interpret meaning from images or video.

Some key concepts in computer vision

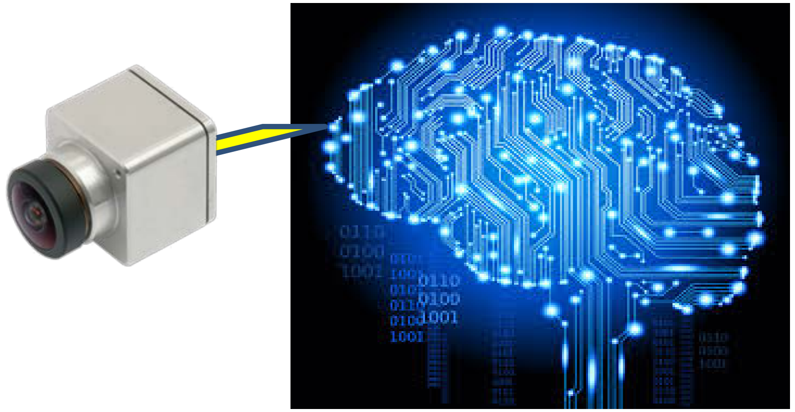
- Image acquisition
- Image processing
  - **Motion detection**
  - **Feature detection**
  - **Object identification**

Image Acquisition and Processing



RENESAS

Image Acquisition



Analog Cameras

- Used in cars, drones, FPV (first person view) controllers, etc.
- Max resolution ~720p due to wire bandwidth

Digital camera sensors - **CCD** and **CMOS**

(2D arrays of tiny cells that convert light into electrons and often send raw digital data to semiconductor device for processing)

CCD (Charge-Coupled Device)

- Move charge btwn. capacitive bins
- Very sensitive, very high-quality

CMOS (Complementary Metal-Oxide Semiconductor)

- Simpler manufacturing and lower power
- Several transistors / cell, so less detection and more noise

Serial digital cameras (MIPI CSI or LVDS)

- Send raw camera sensor data into MPU in high-speed serial format

Serial digital camera module

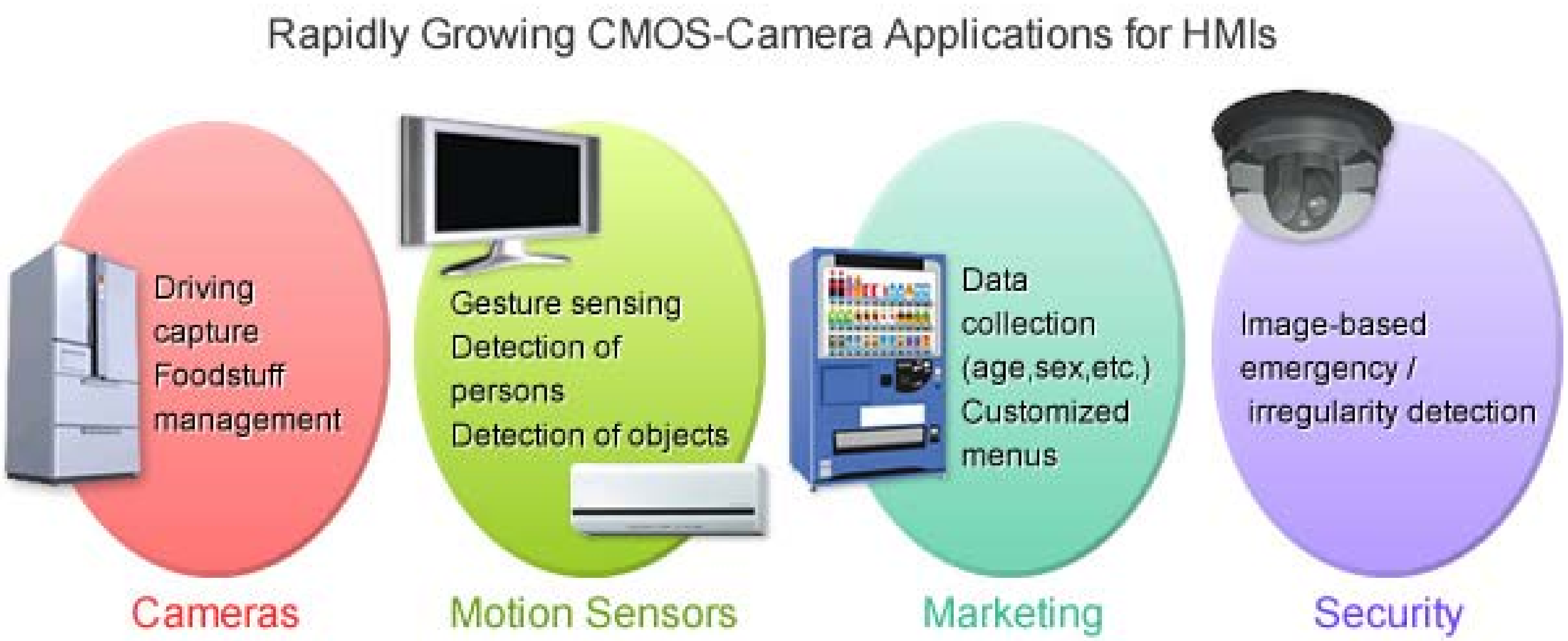
- Send processed data in serial format to semiconductor device

Parallel digital camera

- Send camera data to semiconductor device in RGB or YCbCR format

Camera Modules Used in Industrial and Consumer Products

For GUI & Up to 4
Digital Camera Inputs

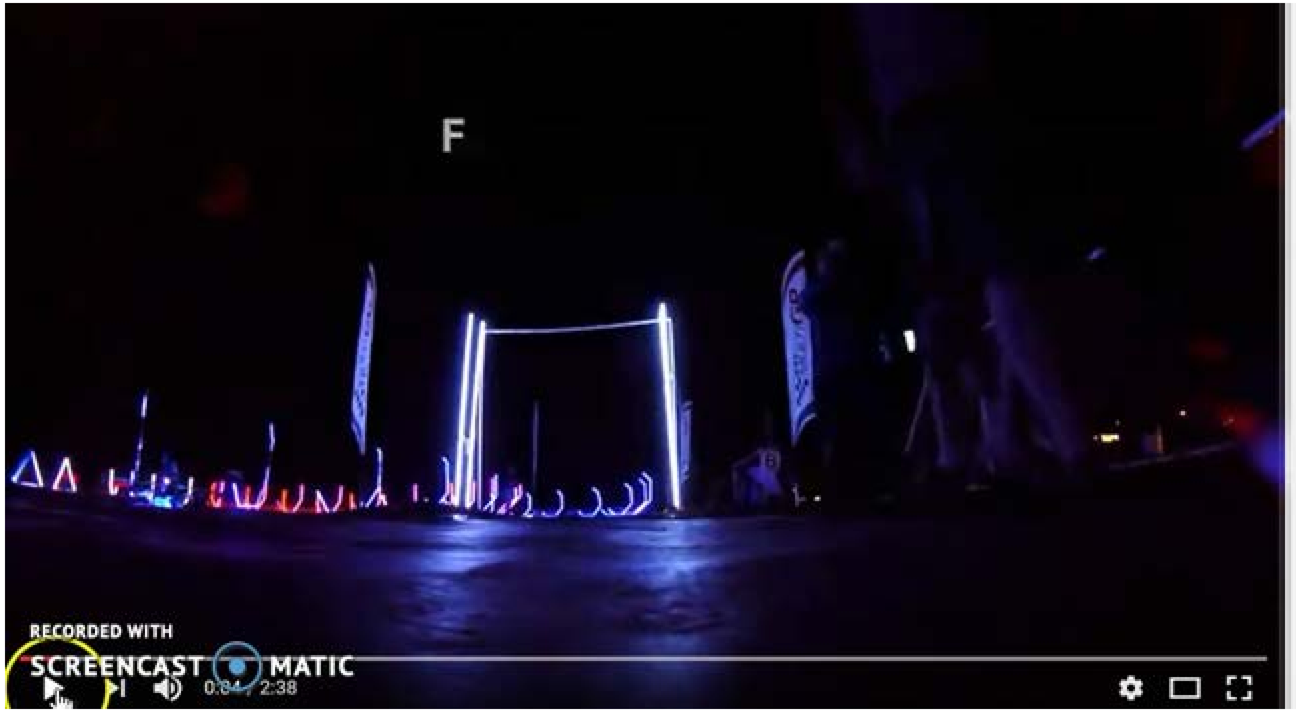


Cameras Used in Consumer Drones

Drones with auto-tracking & stabilization

Self-Flying Lily Drone (URL)

Drone racing



GoPro Awards: Epic Drone Race at Night (URL)

Motion Detection

Miscellaneous Techniques Used in Motion Detection

Background subtraction

Temporal difference detection

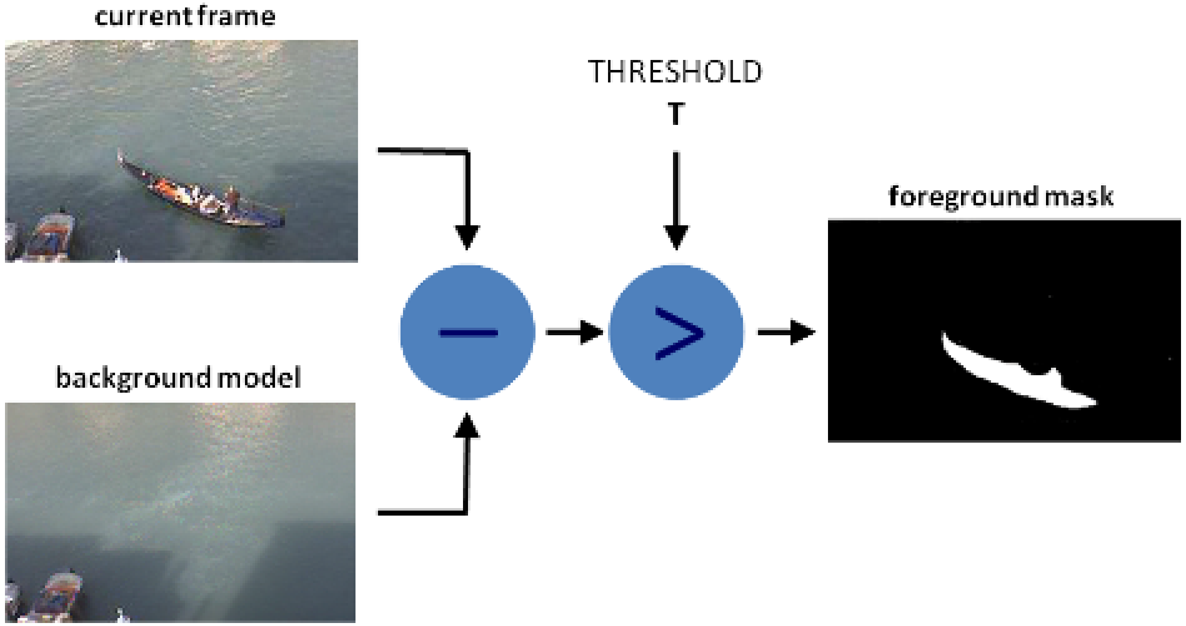
Optical flow

Background Subtraction

Background Subtraction

Motion is detected by comparing incoming frames to a known background image

- Foreground mask, which contains the pixels belonging to moving objects in the scene, is compared to a model of the background reference to determine if motion occurred



Pros

- Reduces the amount of information needed to process in the algorithm
- Flexible and fast
- Good for stationary environments

Cons

- It is only useful when camera is stationary
- With indoor scenes, reflections can lead to background changes.
- With outdoor scenes, wind, rain or lighting changes can trigger false detection

Temporal Difference Detection

Temporal Difference Detection (AKA Frame Differencing)

Algorithm that checks for the **difference** between two consecutive video **frames**

- If the pixels have changed, then something changed in the image

Often incorporates blur and threshold, to distinguish real movement from noise (e.g. when lighting conditions change, camera auto-focuses or corrects for brightness)

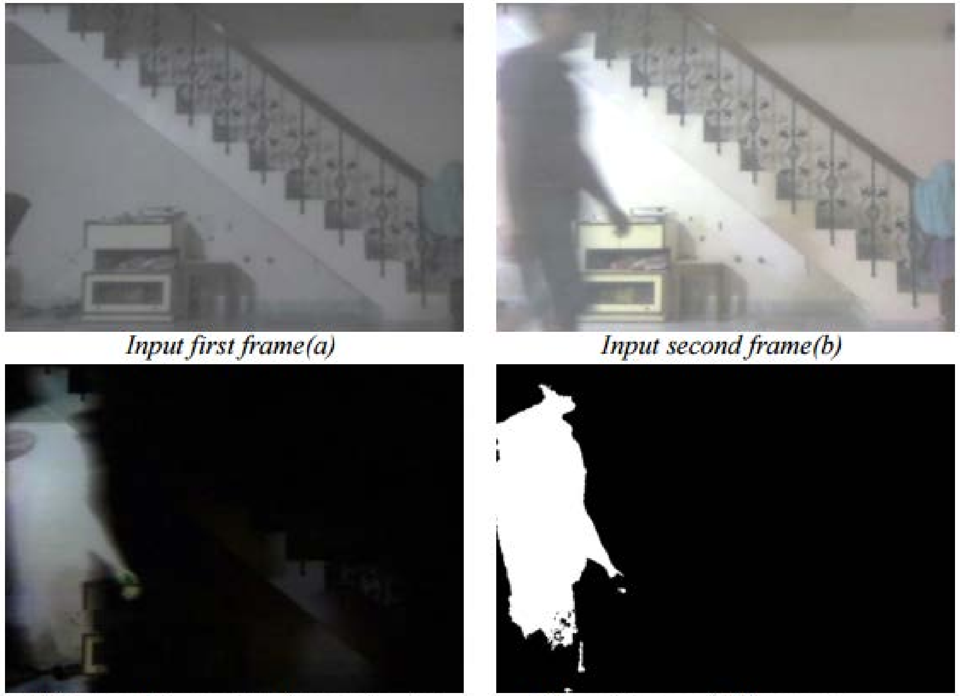
- Motion tracking occurs when change exceeds a certain threshold within "region of awareness"

Pros

- This method is highly adaptive to dynamic (i.e. moving) environments

Cons

- Does poor job of extracting complete shapes of moving objects

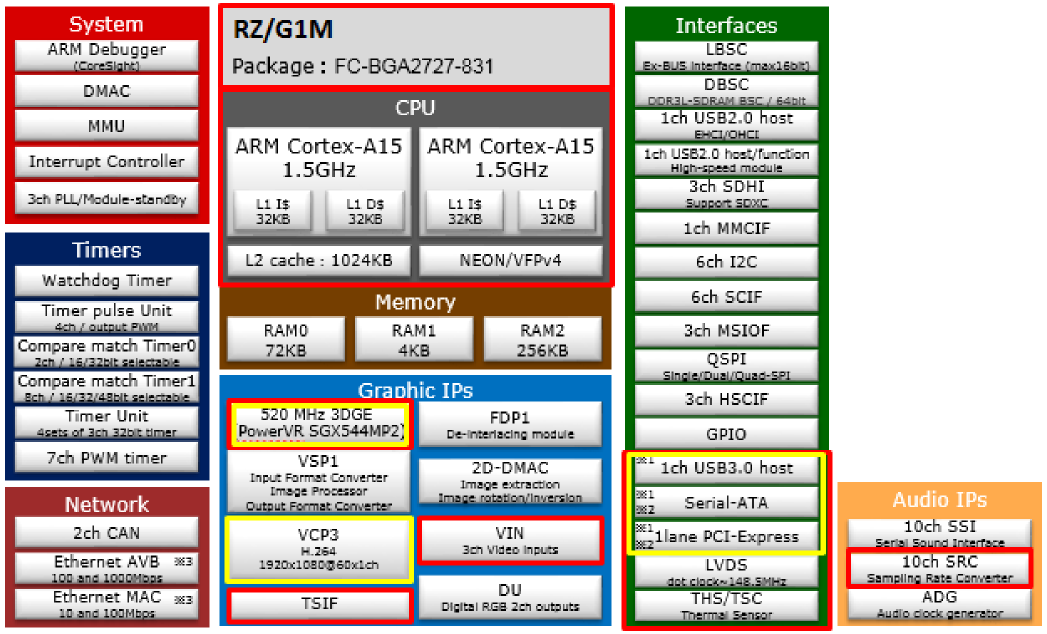


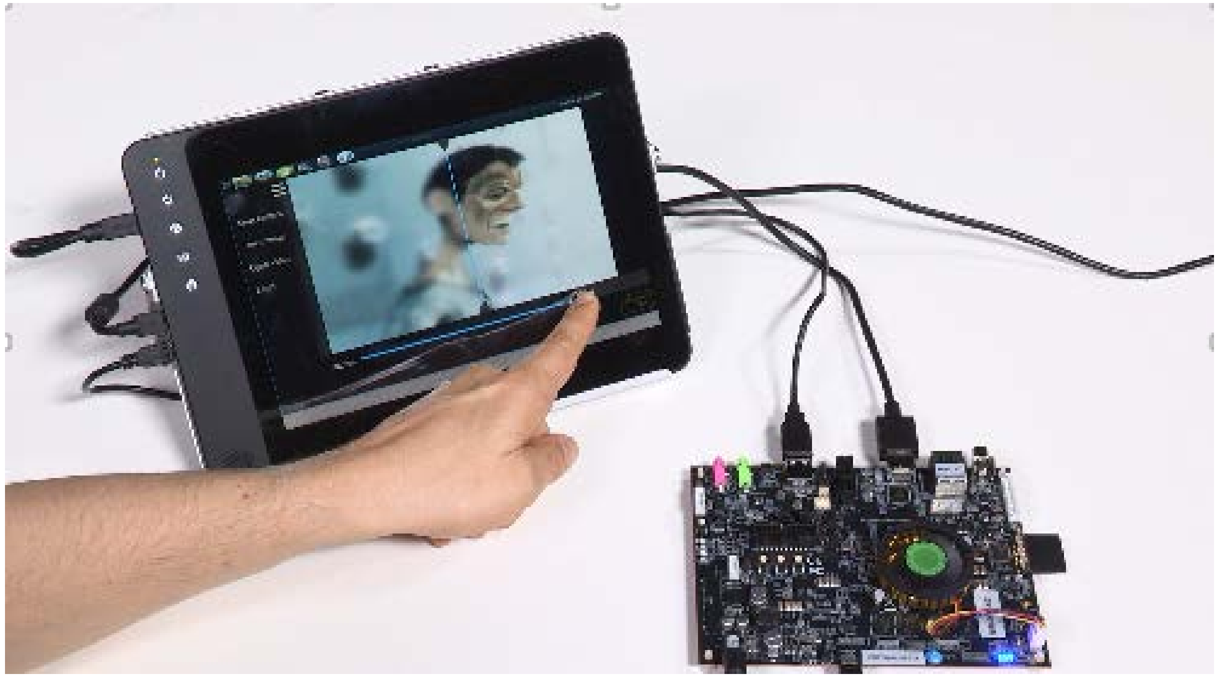
Input first frame(a) *Input second frame(b)*

Difference between two frame showing moving object *Binary image of difference image.*

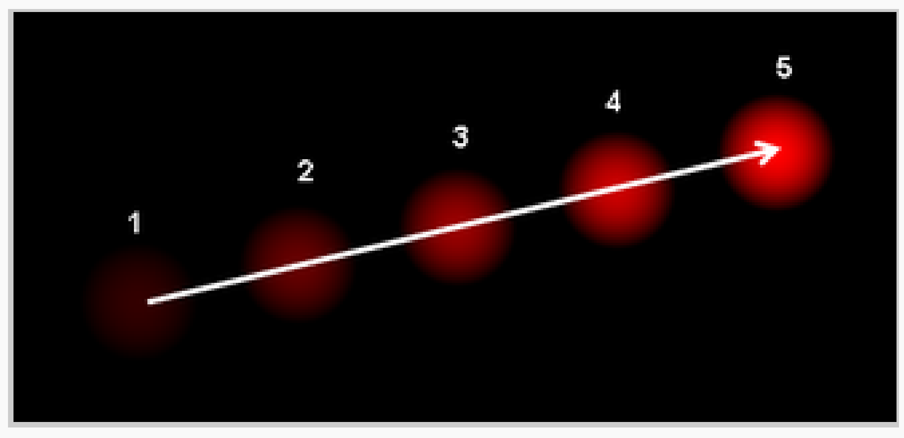
Renesas RZ/G1 Edge Blurring Example

Deming the technique for making the image more stable, not the motion detection





Optical Flow



Optical Flow

Form of temporal difference (i.e. it looks at successive frames)

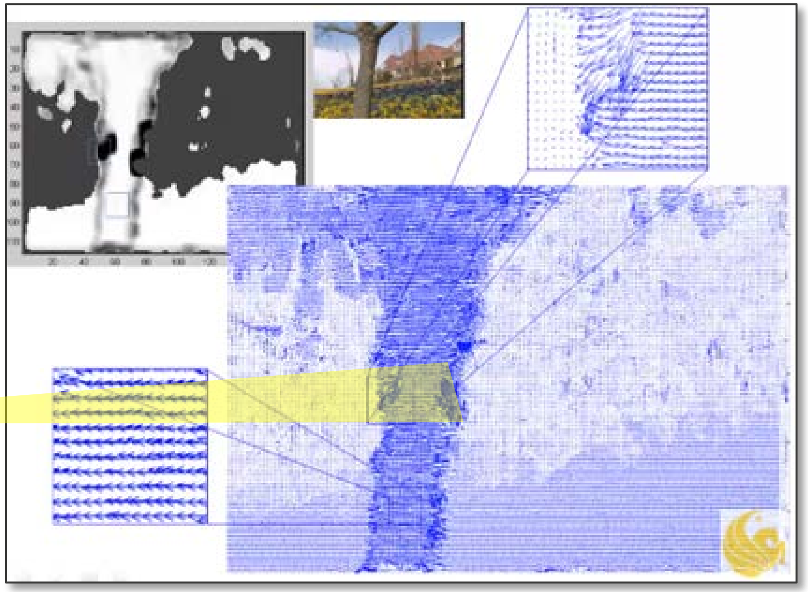
Algorithm determines displacement or “optical flow” vector (i.e. direction, gradient, and magnitude) for every pixel

- Gives you displacement of each pixel vs. previous frame.
- Lets you interpolate and extrapolate motion to predict position

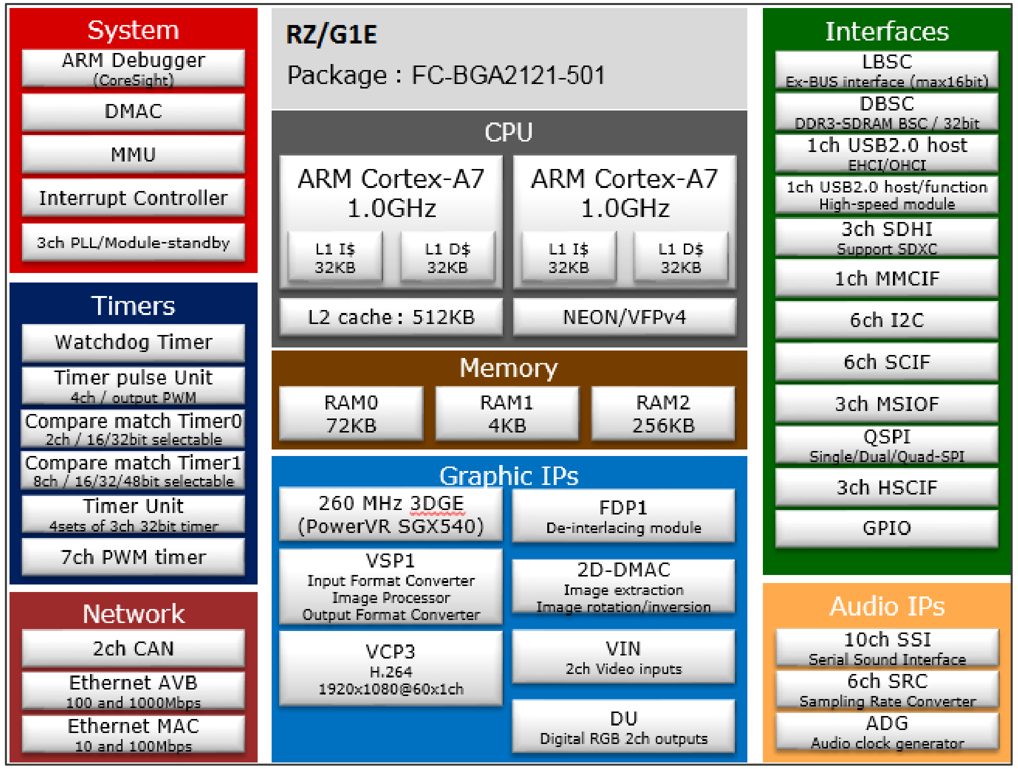
Relies on “Brightness and Constancy” assumption

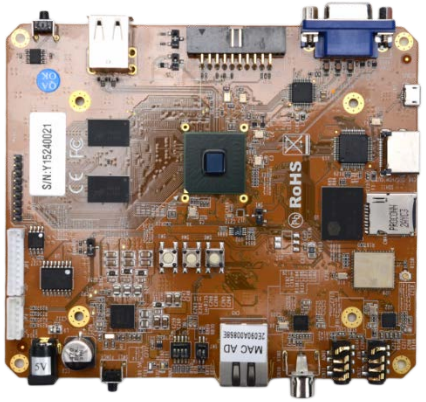
- Corresponding points in two successive frames should not move more than a few pixels
  - Not great for very fast moving video

Interfaces between objects can cause confusion



Example – Optical Flow





Localize region of interest around the ball to minimize computational load, and achieve optical flow on low-cost, low-power RZ/G1E MPUs

(see also URL)

Feature Detection

Miscellaneous Techniques and Tools Used in Feature Identification

Calculation-based techniques

- Like edge detection

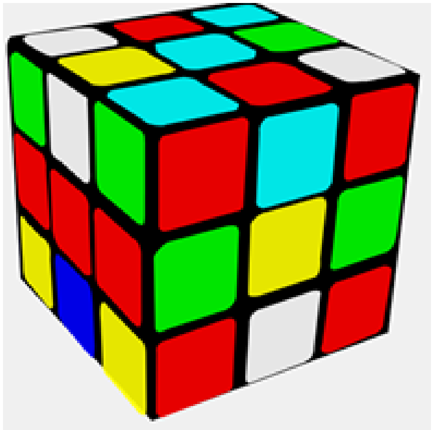
Template matching

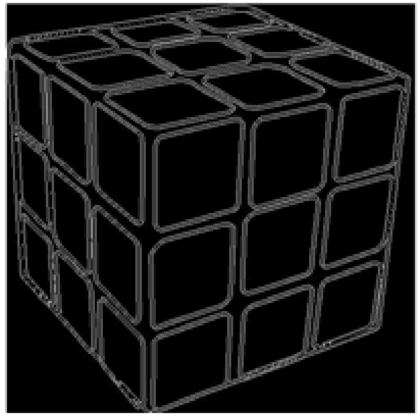
- Simple shapes and more complex cascade classifiers

What do I mean by “feature”?

- Edges, corners, points, faces, codes, text, etc.

Edge Detection





Edge Detection

Edge detection is an image processing technique for finding the boundaries of objects within images

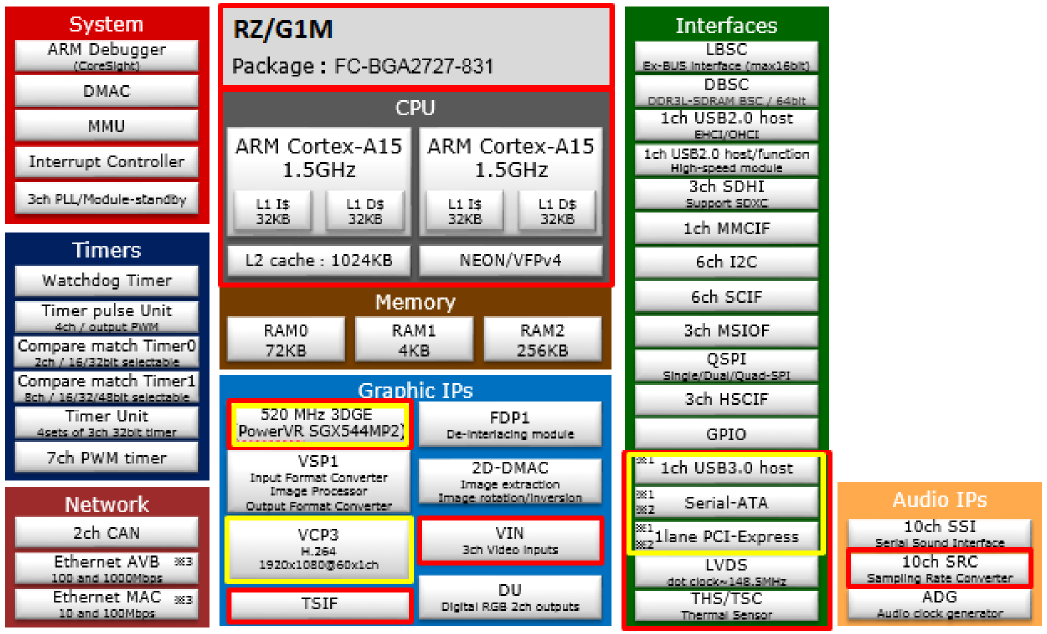
- It works by **detecting** discontinuities in brightness
- Used for image segmentation and data extraction in image processing and computer vision
- Often used an intermediate step to reduce the amount of data you need to process

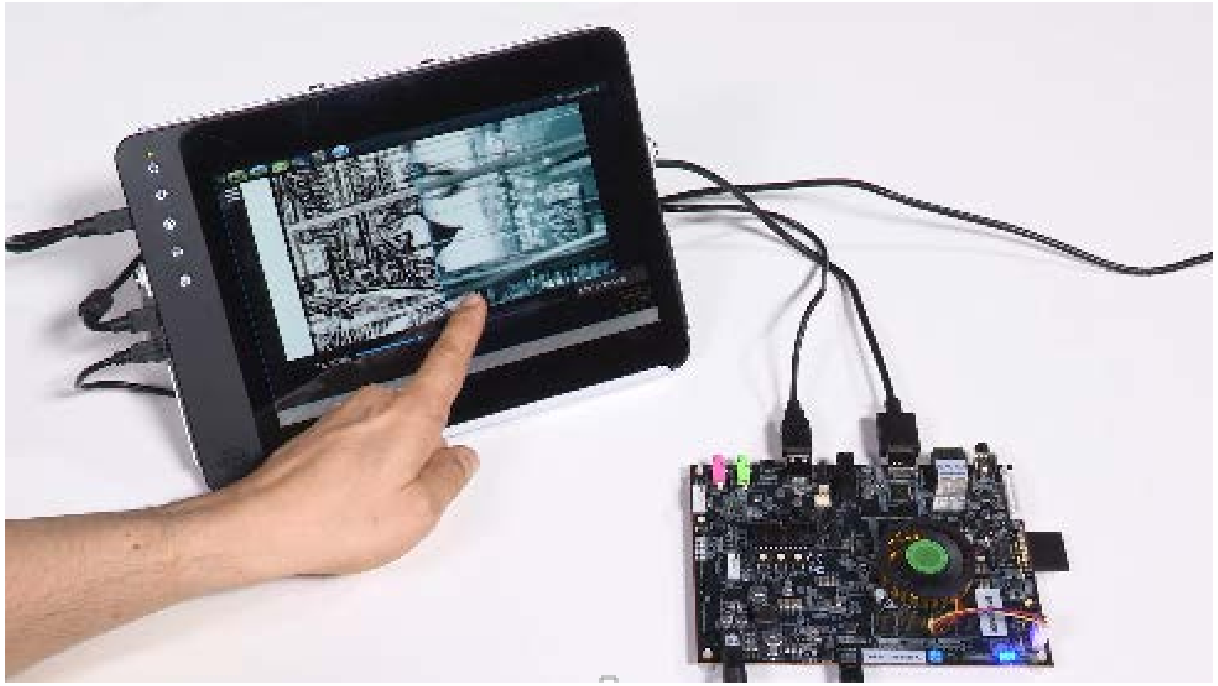
Common edge detection algorithms include Sobel, Canny, Prewitt, Roberts, and fuzzy logic

- Canny (multi-stage) operation shown below

Edge Detection Renesas RZ/G1M

Edge detection running on 1080p 30
real-time video playback on RZ/G1M

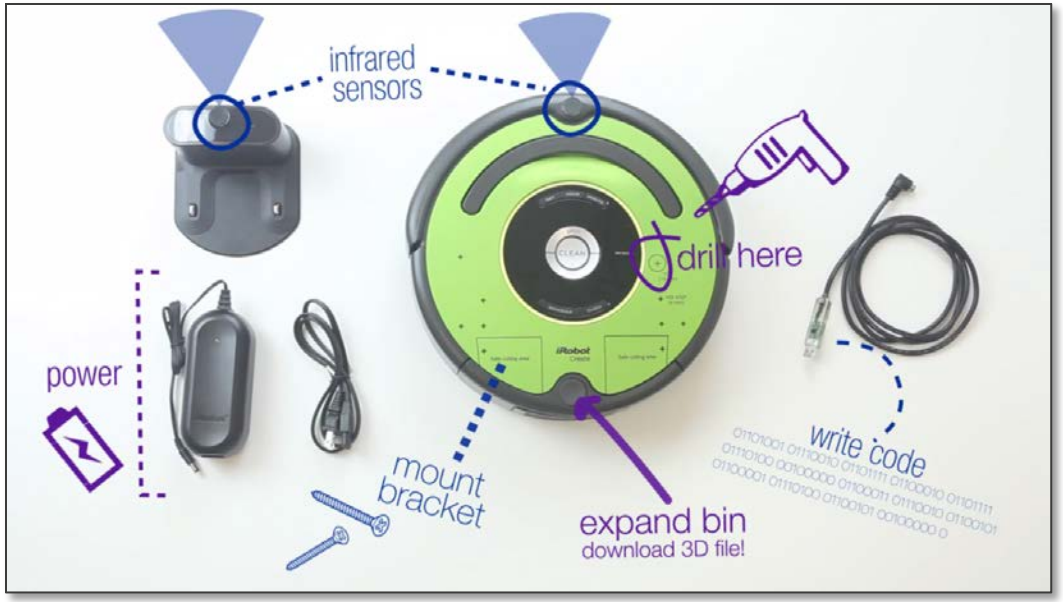




vSLAM - Visual Simultaneous Localization and Mapping

Autonomous robots must be able to explore their environment without user-intervention or external data, build a reliable map, and localize themselves within it

- No initial map needed
- Determine appropriate reference points, correct for slippage, collisions, relocation
- Handles dynamic changes in the environment (lighting, moving objects and/or people)

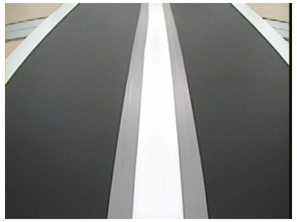


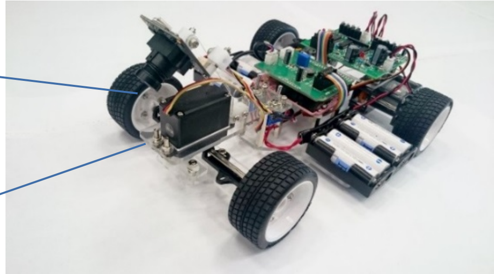
Roomba SDK available for hobbyists and inventors

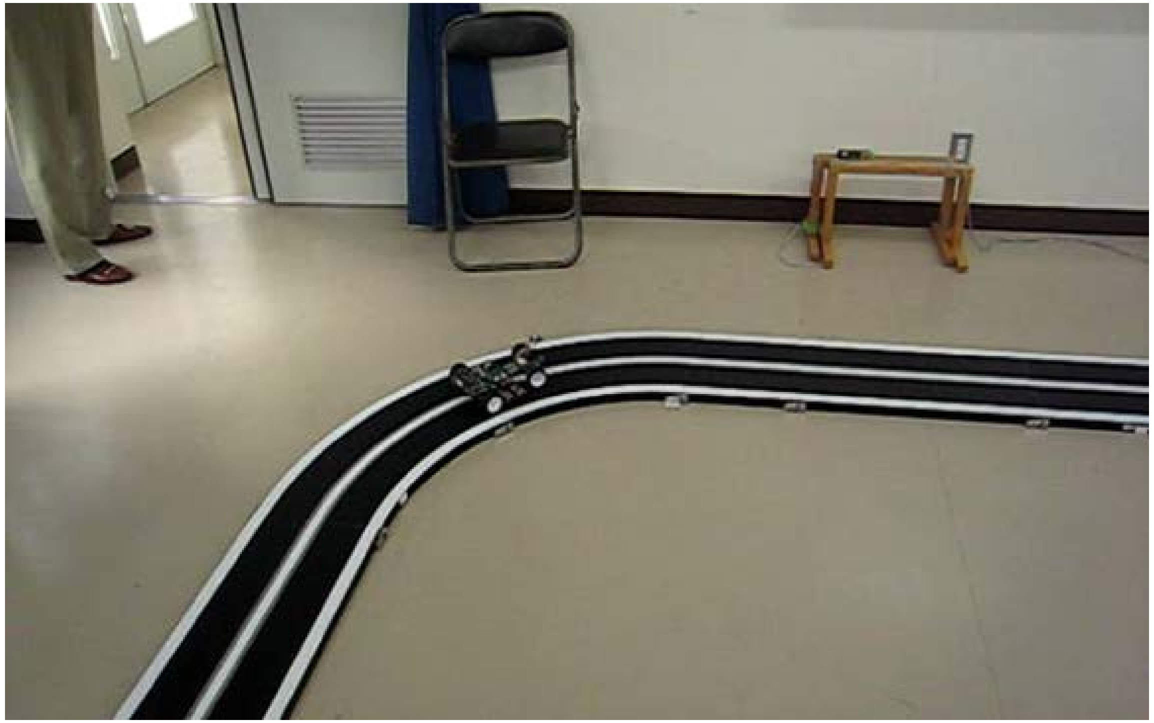
vSLAM utilizes edge detection for localization for this purpose

- Vision and odometry-based, enabling low-cost navigation in cluttered / populated environments

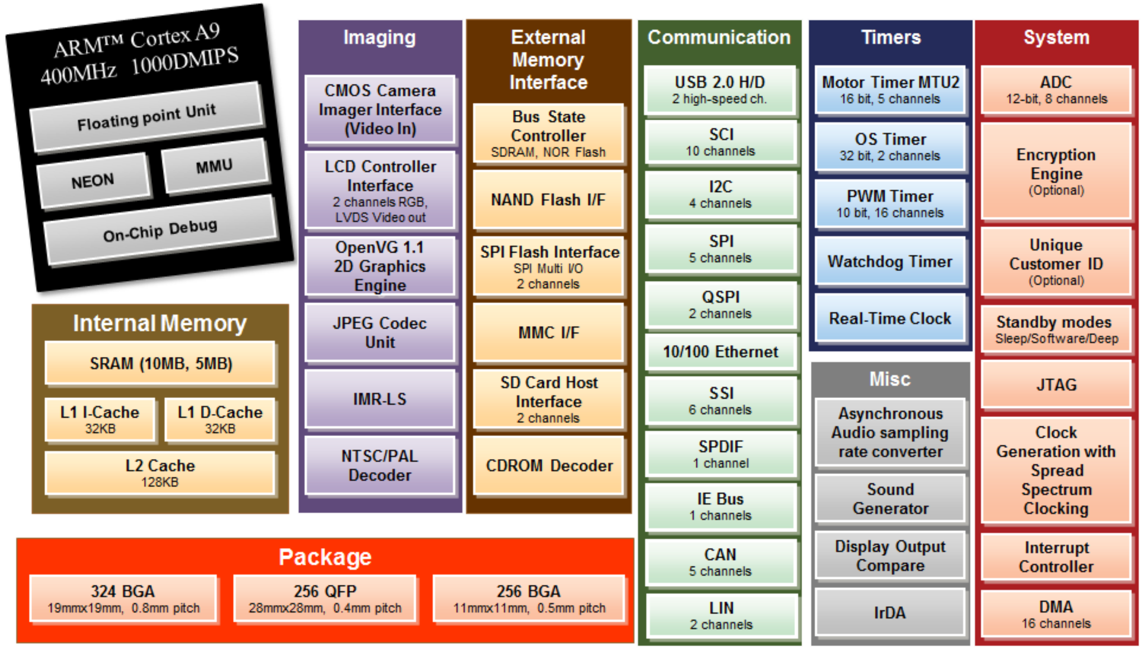
Renesas RZ/A Autonomous Car





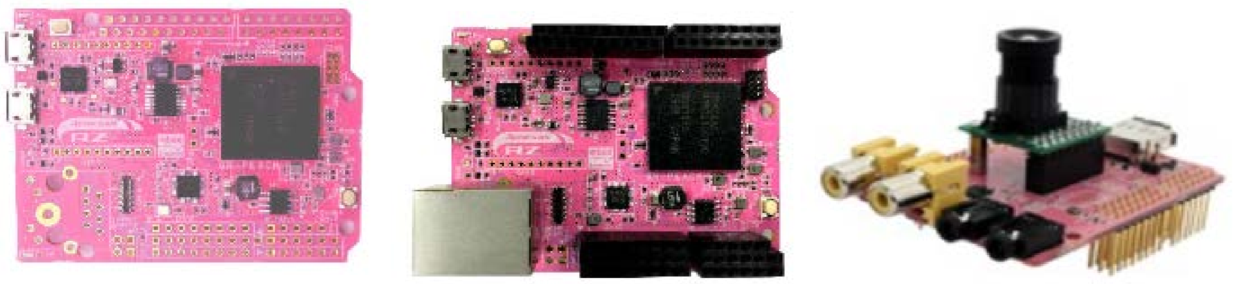


RZ/A1



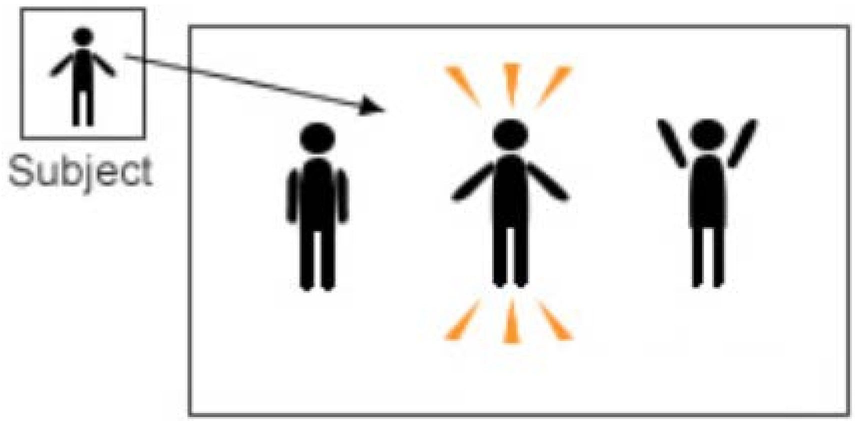
(See also URL)

ARM MBED "GR Peach" Board



World's only Cortex A-Class MBED product

Template Matching



Template Matching

Digital image processing technique for finding small parts of an image, which **match** a pre-defined **template**

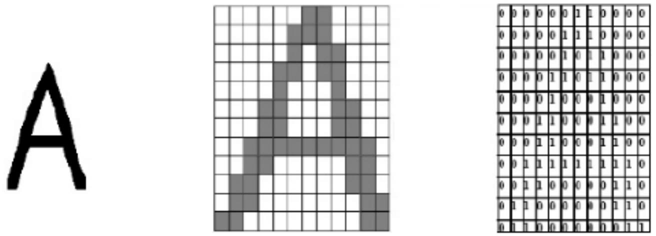
- A template is a small image inside a larger image.
- There is a similarity criteria (metric) – an equation of correlation that your algorithm follows to perform that matching

Flexible and relatively easy to use, which makes them very popular for object localization, especially in machine vision

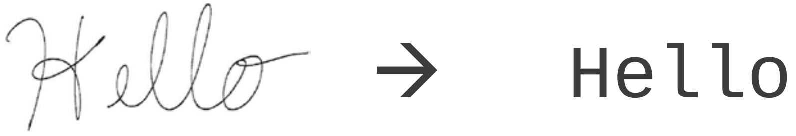
Can be used in motion detection, robot navigation, object-detection and in manufacturing as a part of quality control

- Applicability is limited by computational power available
- Big and complex templates can be time-consuming to identify

Character Recognition



Handwriting Recognition



UPC Code

9 87654 32109 8

QR Code

RENESAS

Running Z-Bar on RZ/A

Renesas RZ/A running Z-Bar Code

(See also URL – from 1 min, 30 sec onward)

"Open source software suite for reading bar codes from various sources, such as video streams, image files and raw intensity sensors

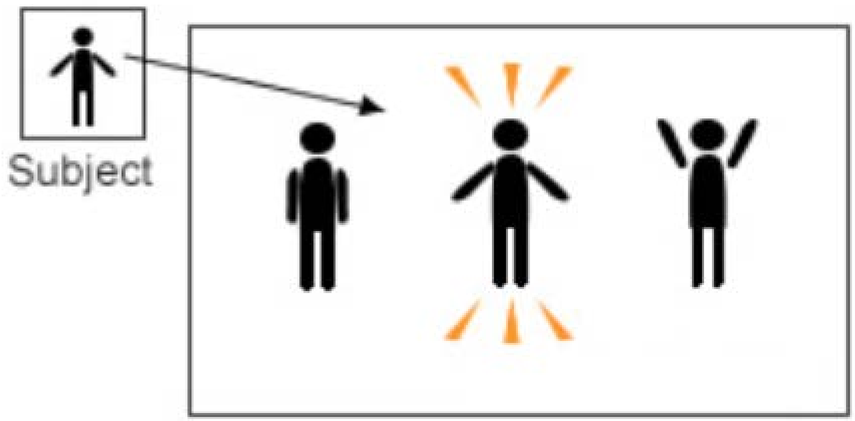
It supports many popular symbologies (types of bar codes) including EAN-13/UPC-A, UPC-E, EAN-8, Code 128, Code 39, Interleaved 2 of 5 and QR Code

ZBar is licensed under the GNU LGPL 2.1 to enable development of both open source and commercial projects."

http://zbar.sourceforge.net/

Object Identification

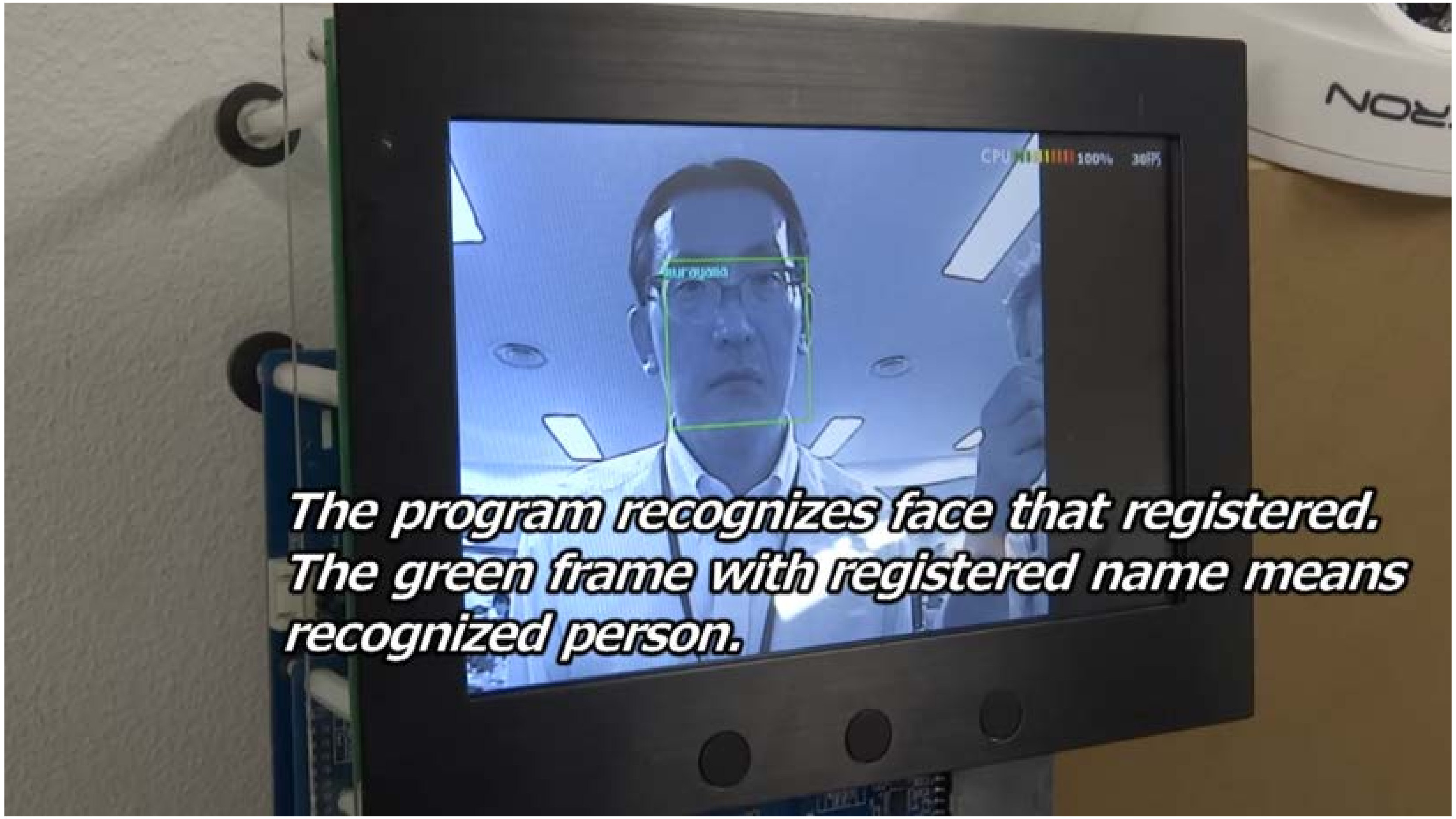
More Complex Template Matching



Renesas RZ/A Face, Gender and Age Identification Example

Face detection by template-matching

Limited to frontal view, relatively stationary subject



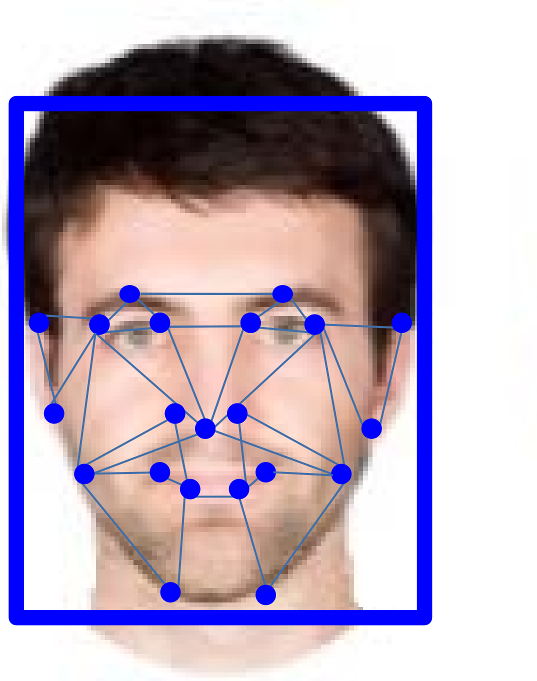
(See also URL)

Face Identification Using Constellation Model

Constellation model performs category-level object recognition

A “**part-based**” model

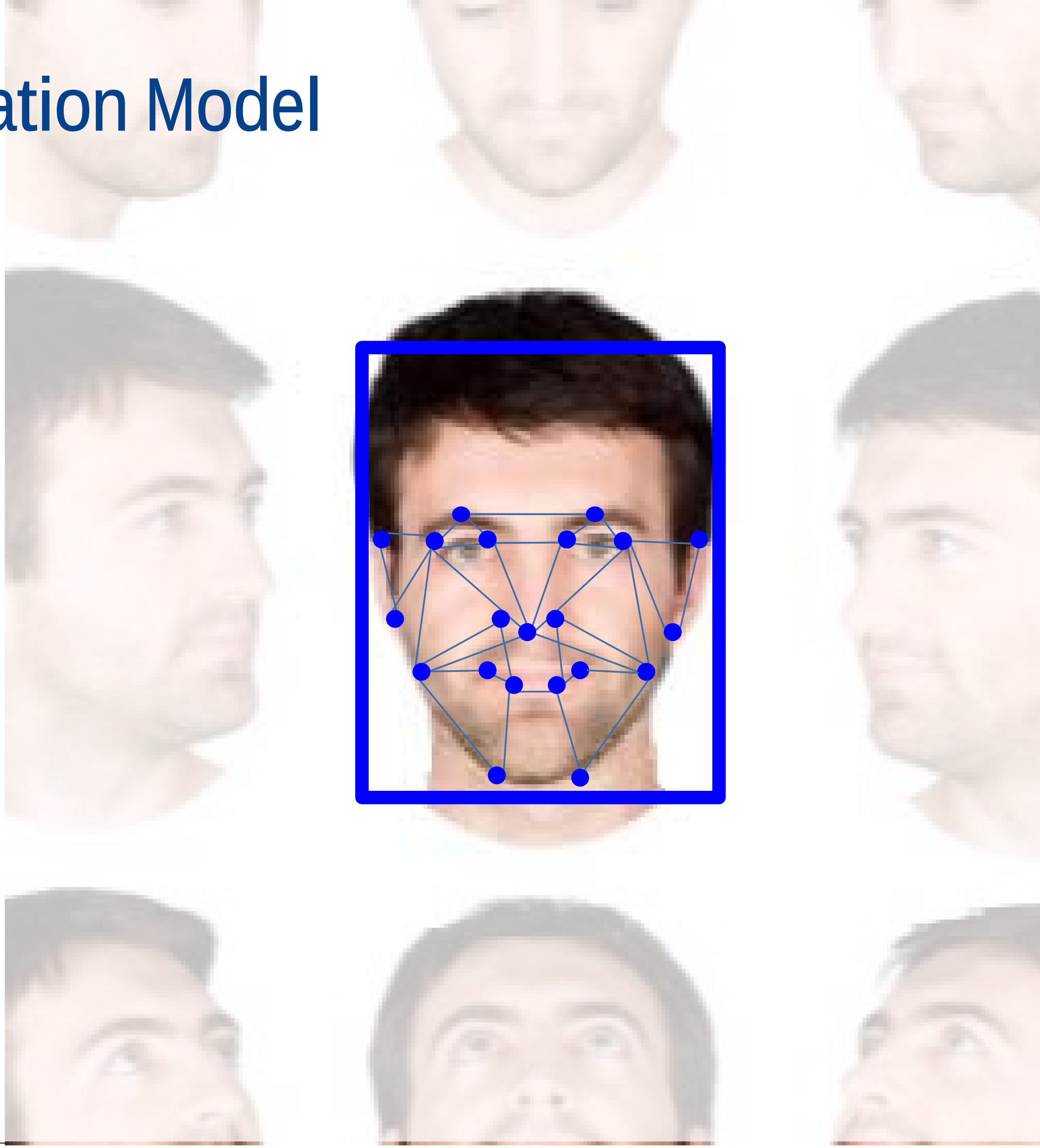
- It attempts to represent classes of objects by a set of parts that are related by geometric constraints
- It models relative location, relative scale, and appearance of the parts in a given object category



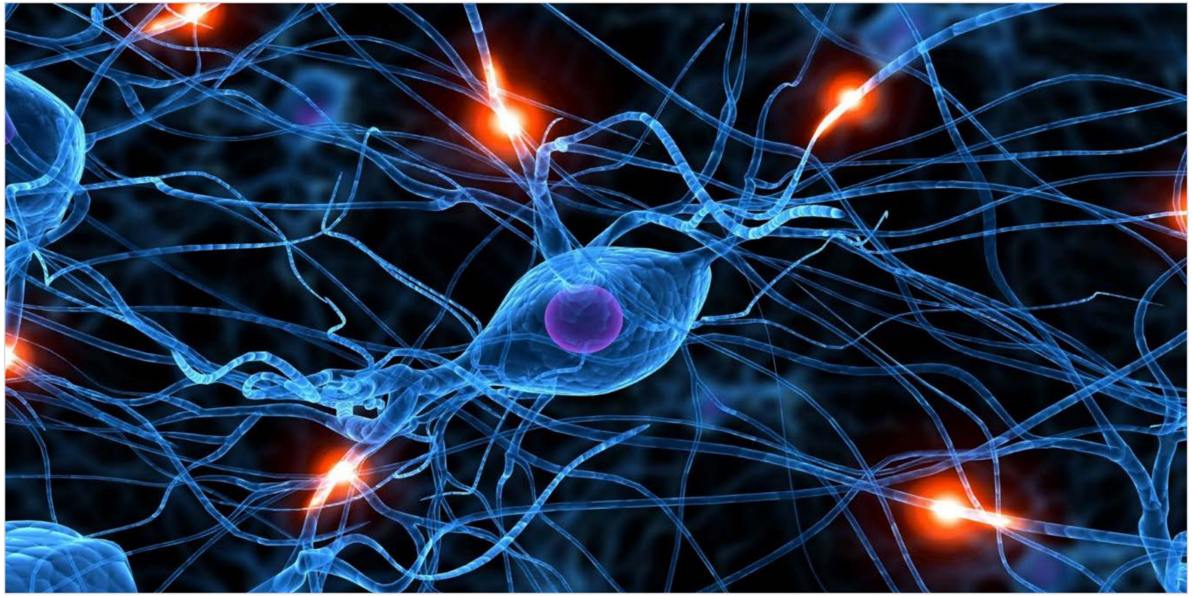
Model parameters are estimated using **unsupervised learning** algorithm, so visualization of an object class can be extracted from an unlabeled set of training images

Face Identification Using Constellation Model

But how do we make it rotation-invariant and less sensitive to temporal fluctuations in subjects' appearance?



Neural Networks



Artificial Neural Networks (ANN)

“Deep Learning Machine Teaches Itself Chess in 72 Hours. Plays at International Master Level!” (URL)

A type of machine learning algorithm

Loosely modelled after how the brain works

A framework for solving predictive and analytical problem

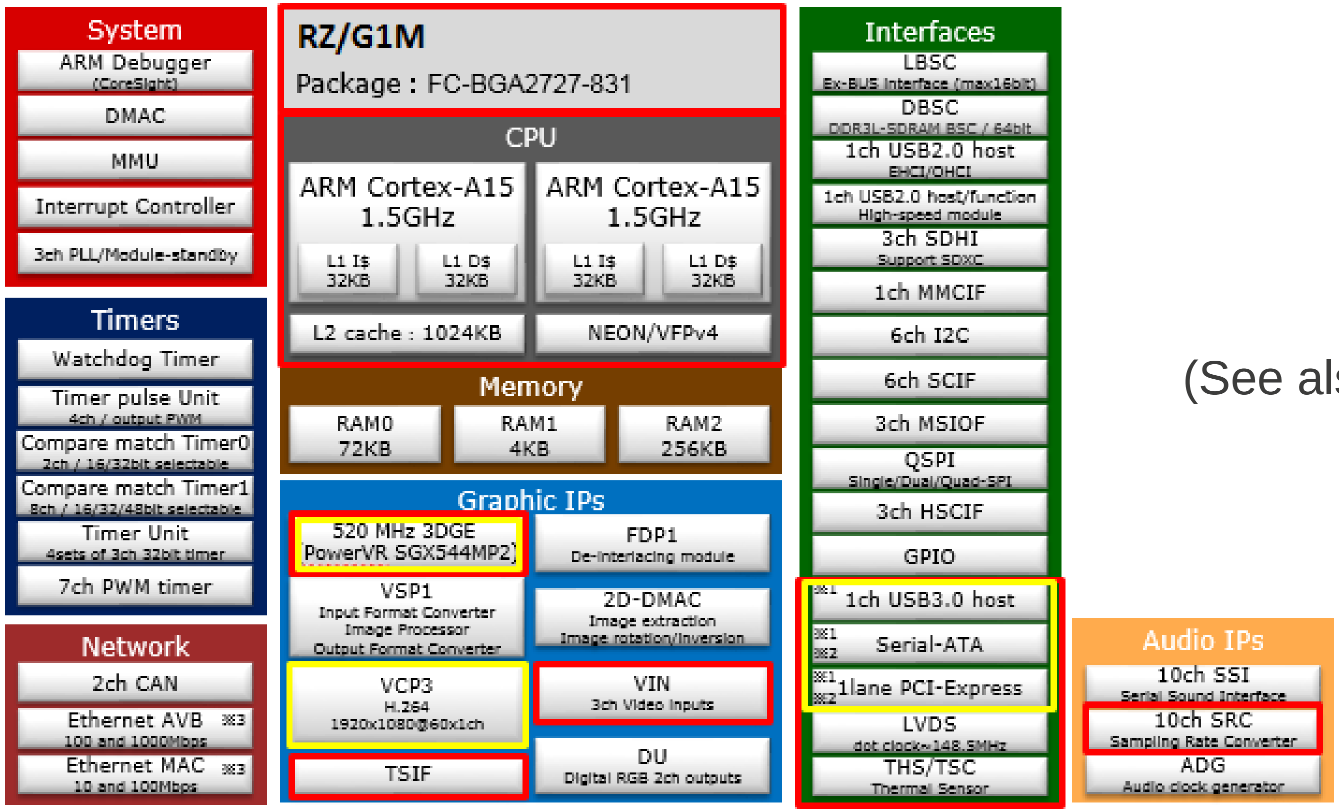
Has exhibited success in solving hard, big problems that rely on a lot of data, including games, natural language interpretation, Big Data trend analyses, and vision

It uses **supervised** regressions (vs. unsupervised regressions). This process is called training

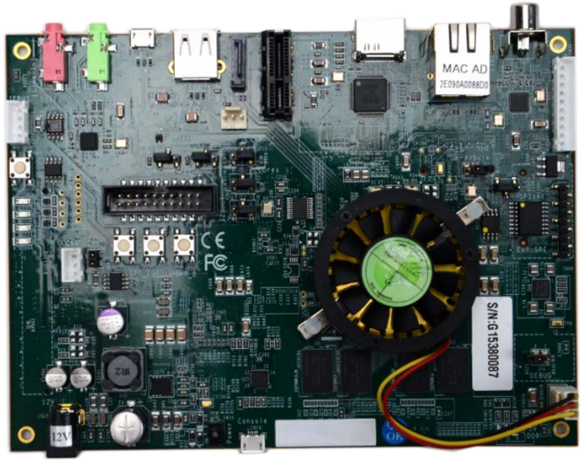
RENESAS

Example – Image Identification

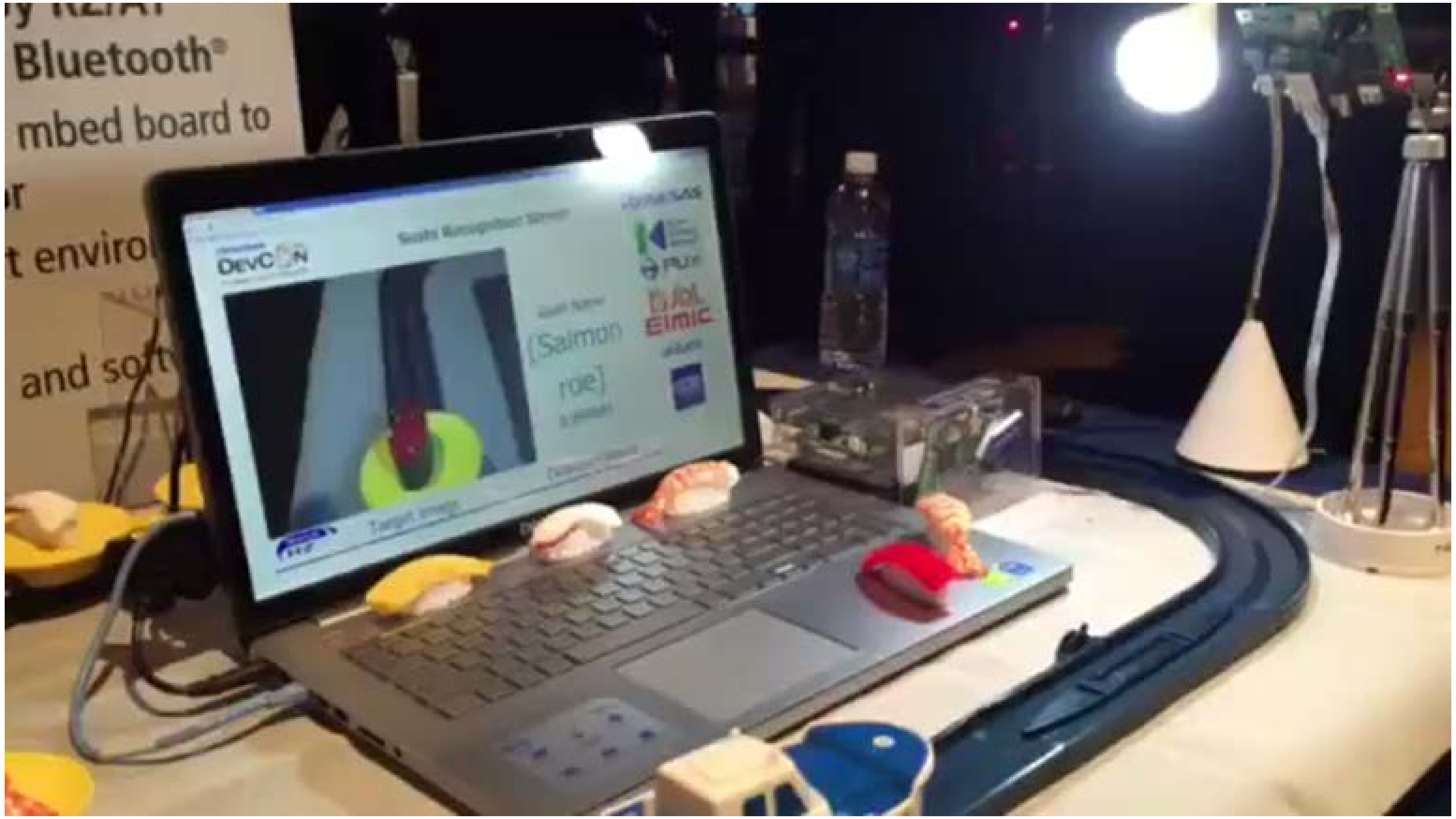
Uses partner software



(See also URL)



“Fog Computing” Approach – Sushi Demo



(See also URL)

Example IoT application

- Object detection at endpoint
- CNN in cloud (host)

By partner **Uncanny Vision** on **GR Peach Board**

III. Future Outlook and Summary

Summary and Future Outlook

Interest in embedded machine vision is growing - it is dragging us into the realm of AI

Motion, feature, and gesture recognition can fit into industrial and automotive products

Image **detection** vs. **identification** have different computational requirements

Renesas can enable your embedded vision needs

What we do with this technology can help make the world a better, safer, more efficient place

Use it! But use it responsibly! ☺

vs.

Renesas Electronics America Inc.