Smart Imaging: from Silicon to Vision Challenges and… Specifications?

Jorge Fernández Berni

Institute of Microelectronics of Seville (IMSE – CNM)
CSIC – Universidad de Sevilla (Spain)
berni@imse-cnm.csic.es

Alpen-Adriá-Universität
Pervasive Computing Group

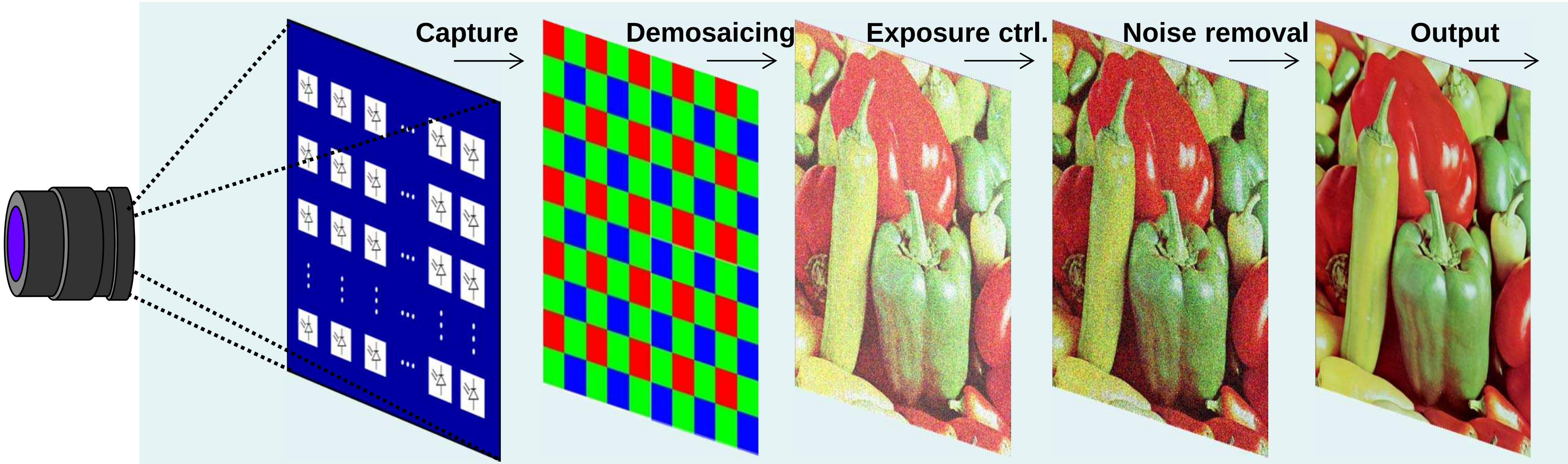
Klagenfurt
Austria

Outline

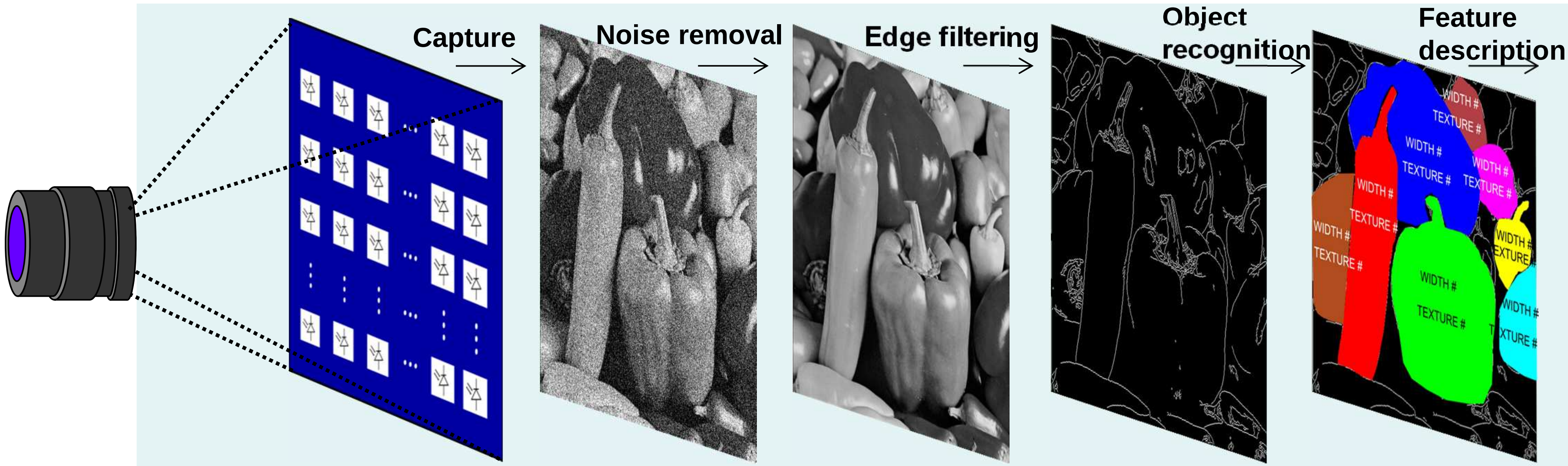
- Smart imaging: concept and rationale
- Early vision based on focal-plane massively parallel mixed-signal processing
- Smart CIS based on Focal-Plane Processing (FPP) @ IMSE
- Industry… is openVX the long expected framework?
- Specifications?

Concept of Imaging and Vision Sensors

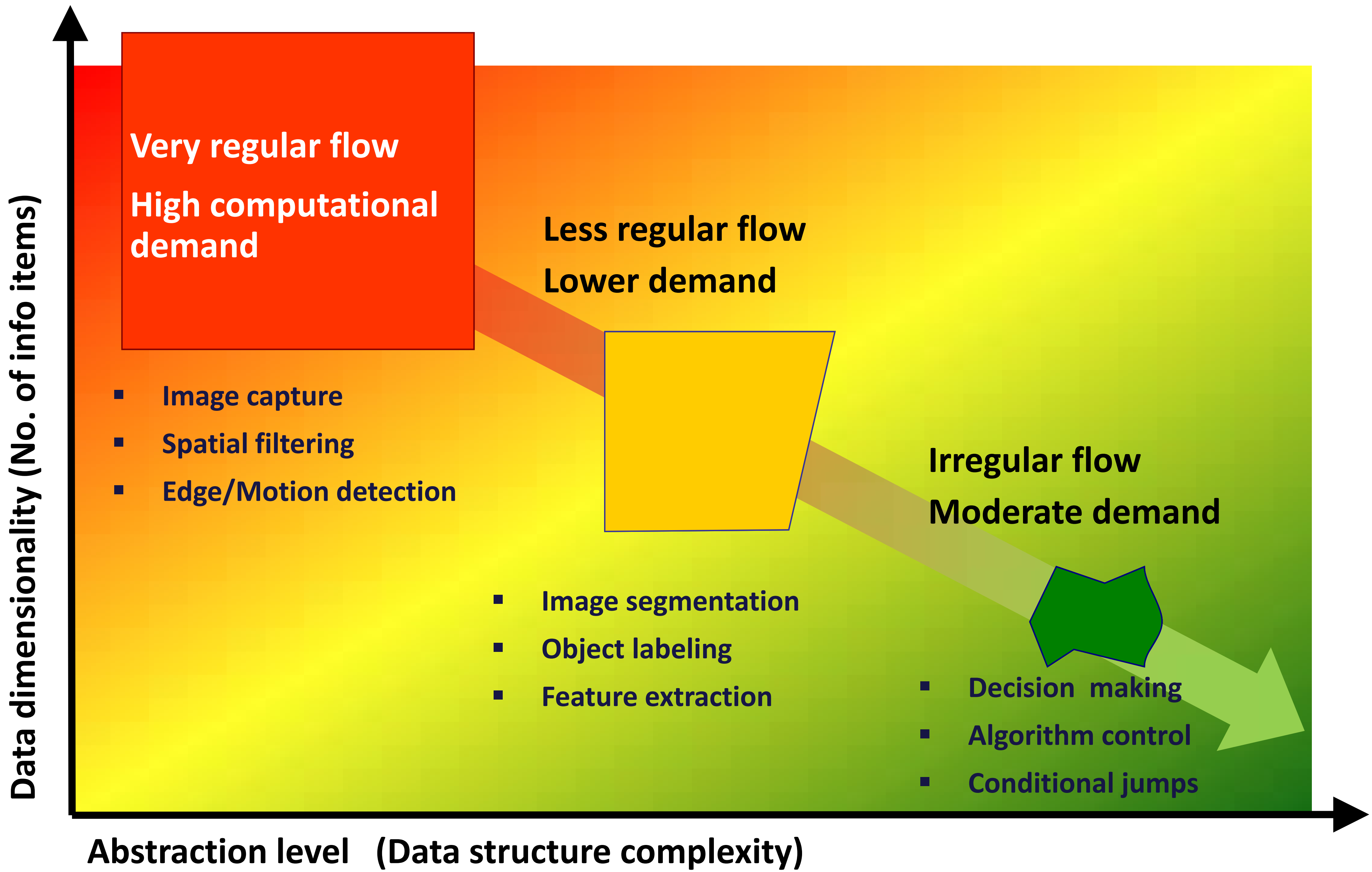
IMAGING **Main objective: image quality**



VISION **Main objective: scene understanding**

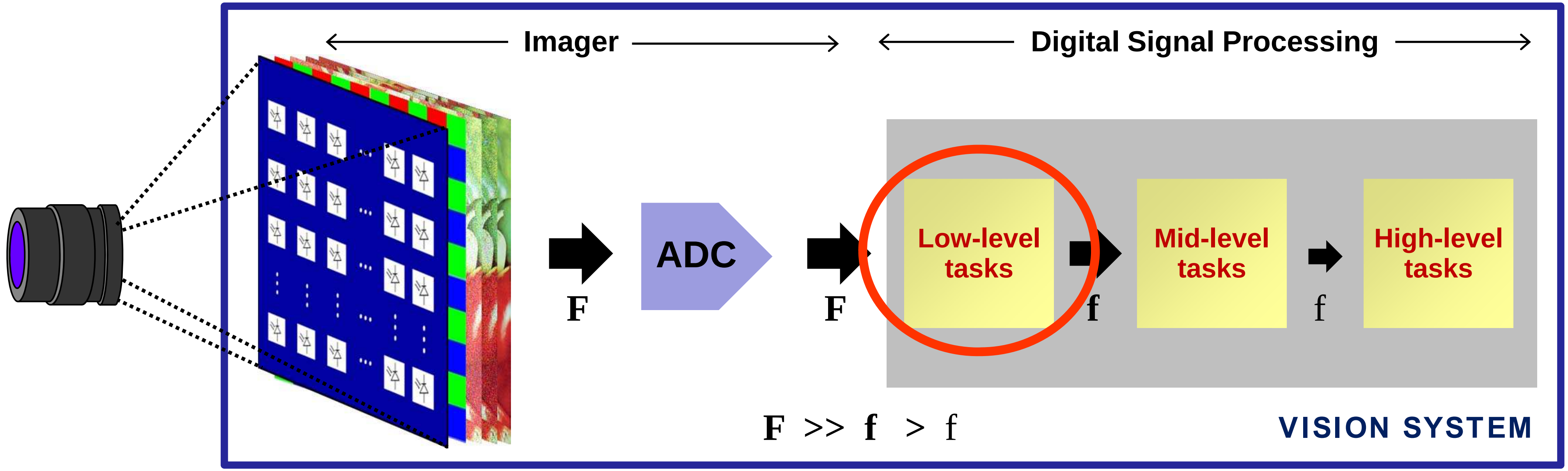


Computational demand in artificial vision



Vision processing chain

Conventional Processing Scheme



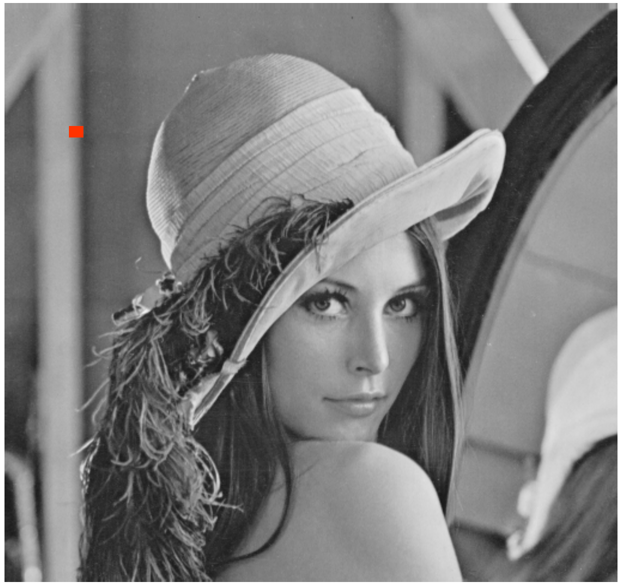
- Imaging stage delivering a **serialized raw digital data flow**
- Subsequent computing from this **serial flow** performed by a digital processor
- Inherent **bottleneck** associated with **early vision tasks**
- **No exploitation** at all of the **parallel nature** of early vision

Example of early vision task

INSTRUCTION FLOW ↓

COMPUTATIONAL LOAD ↑

Edge Detection



$$\mathbf{G}_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \mathbf{G}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$p_{ij_x} = p_{i-1,j-1} + 2p_{i,j-1} + p_{i+1,j-1} - p_{i-1,j+1} - 2p_{i,j+1} - p_{i+1,j+1}$$

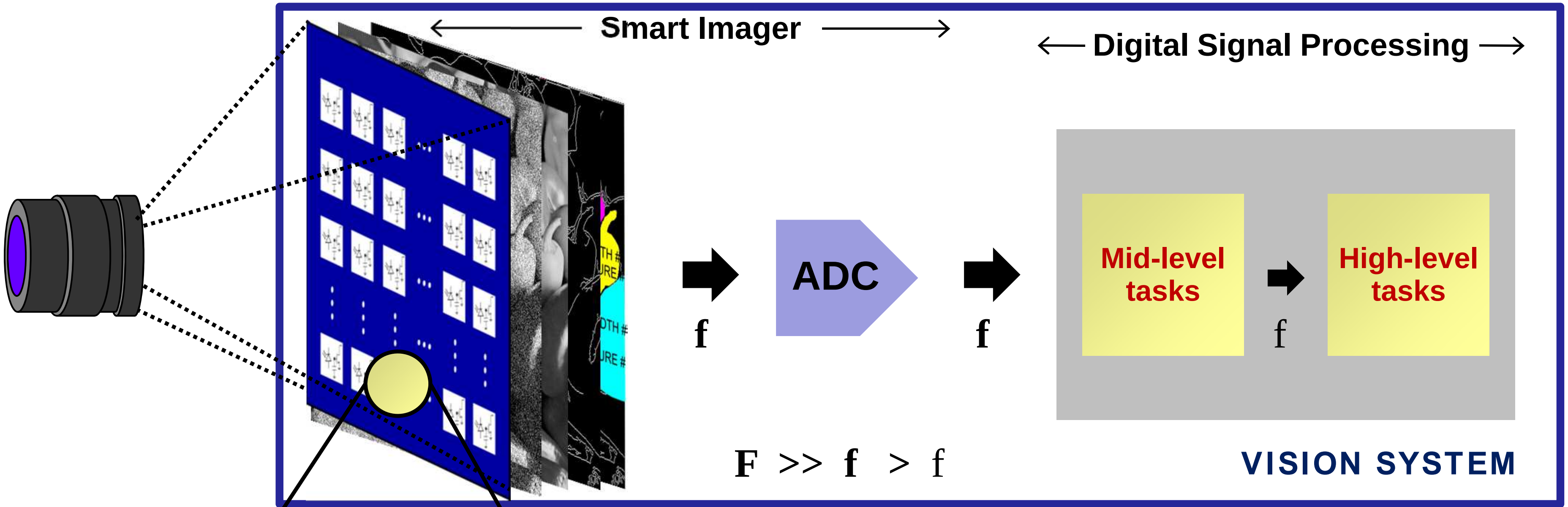
$$p_{ij_y} = p_{i-1,j-1} + 2p_{i-1,j} + p_{i-1,j+1} - p_{i+1,j-1} - 2p_{i+1,j} - p_{i+1,j+1}$$

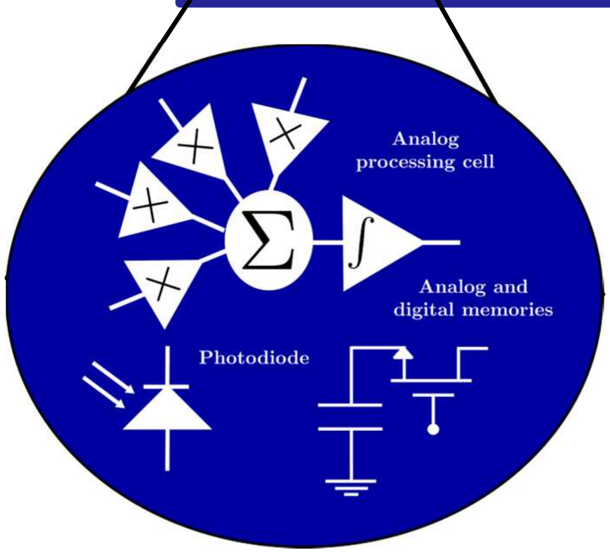
$$p'_{ij} = \sqrt{p^2_{ij_x} + p^2_{ij_y}}$$

- **Potential parallel operation**
- **Moderate accuracy required**

Our approach for smart imaging

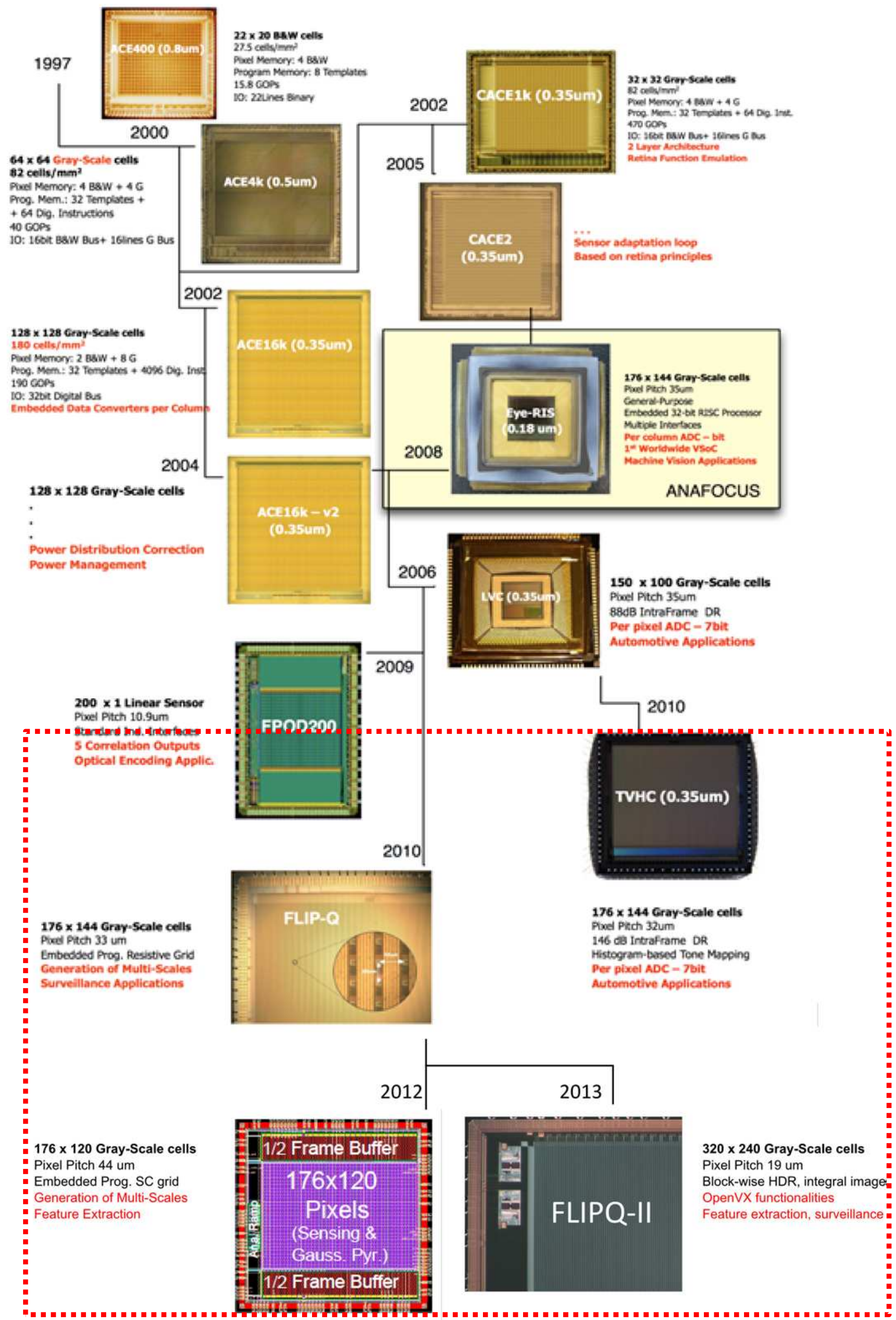
FOCAL-PLANE MASSIVELY PARALLEL EARLY VISION





- Processing based on a focal-plane array of **elementary mixed-signal sensing-processing cells**
- Single-Instruction Multiple-Data (SIMD) **massively parallel operation**
- **Distributed memory**
- **Bio-inspired model**: progressive reduction of information

Smart CIS based on FPP @ IMSE



Major achievements

- Concept demonstration
- Programmable embedded functionalities
- Image-to-Decision chain at >1,000fps using 60nW per pixel (industrial chip)
- Spatial Gaussian filtering @20nJ/filter
- Content-aware HDR acquisition with >145dB intra-frame DR

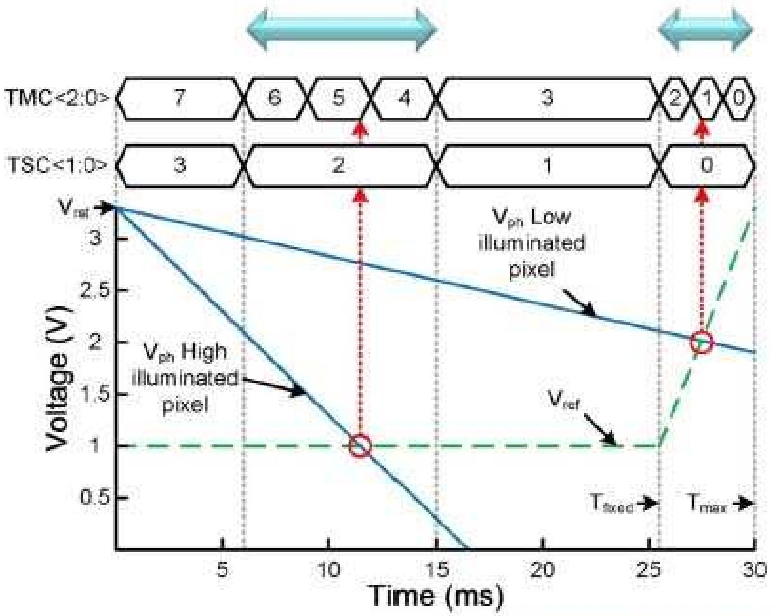
Major drawbacks

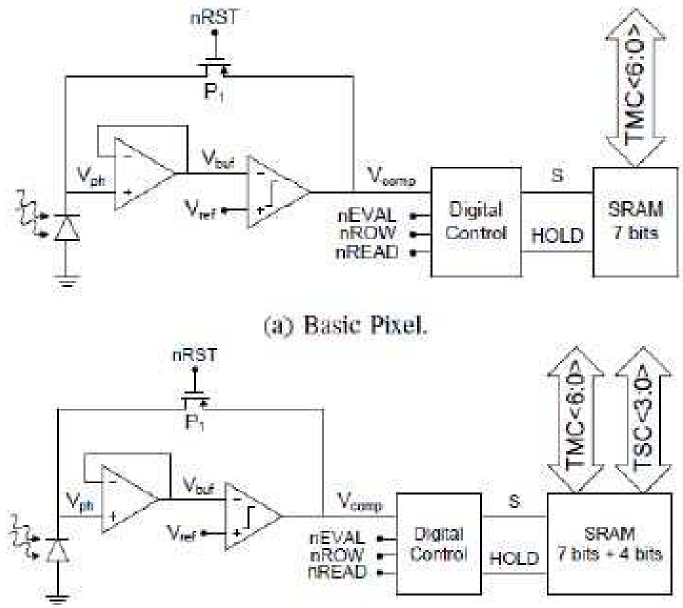
- Reduced fill factor
- Large pixel pitch
  - → **Limited sensitivity**
  - → **Small image size**
  - → **Spatial aliasing**

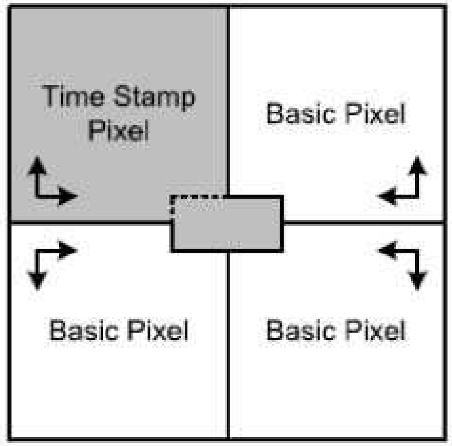
Major challenges

- Implementation of in-pixel embedded functionalities at minimum area cost
- Increase hardware-software integration

In-pixel HDR Tone-mapping Compression

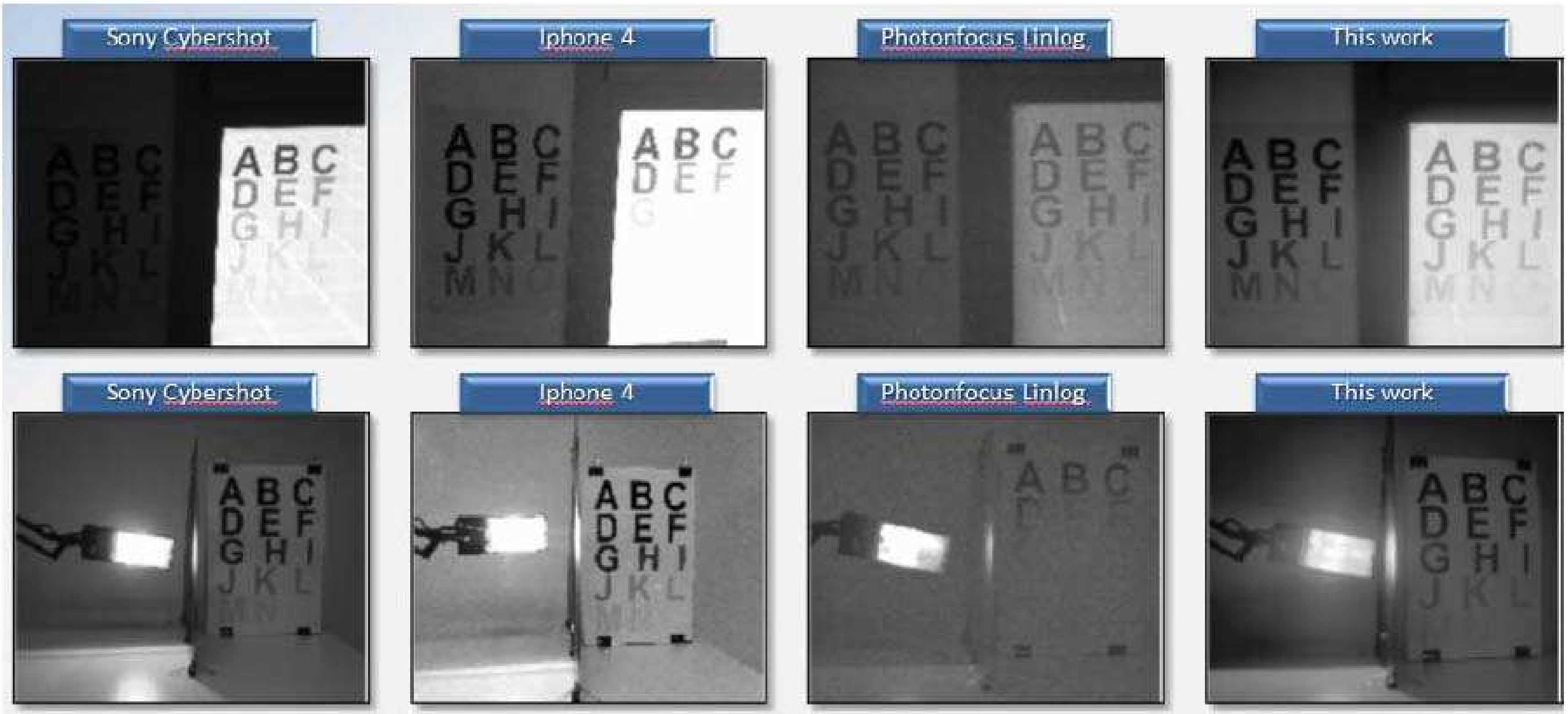






[S. Vargas et al, 2010]

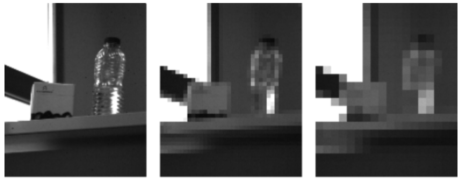
145dB intra-frame



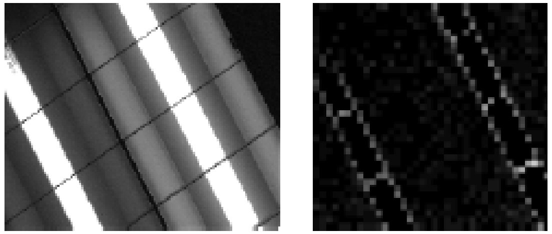
FLIP-Q: Focal-plane array for low-power image processing

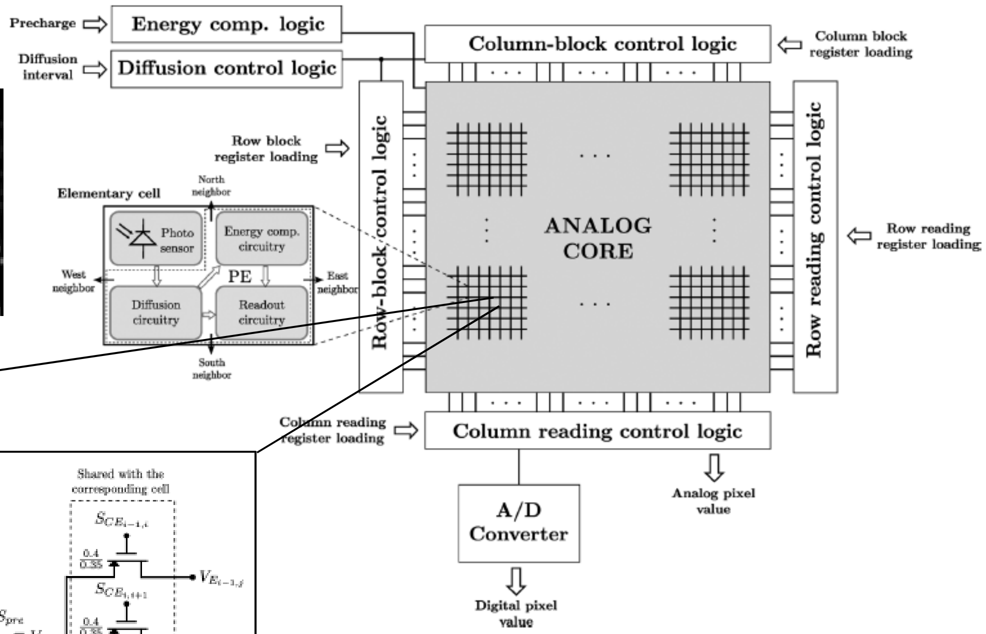
[Fernández-Berni et al, IJSSC 2011]

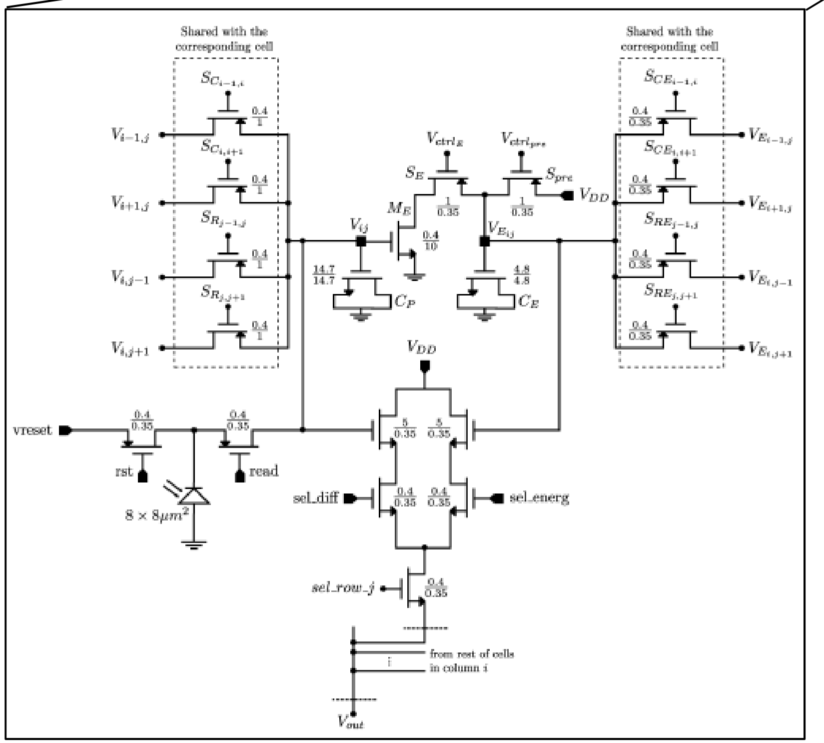
Multiresolution image representation

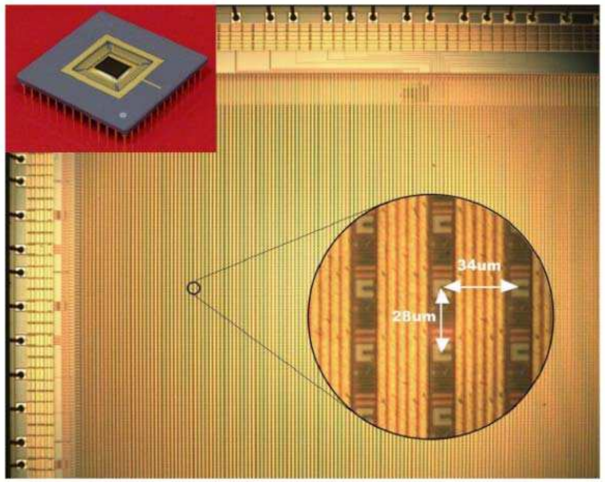


Energy-based image representation





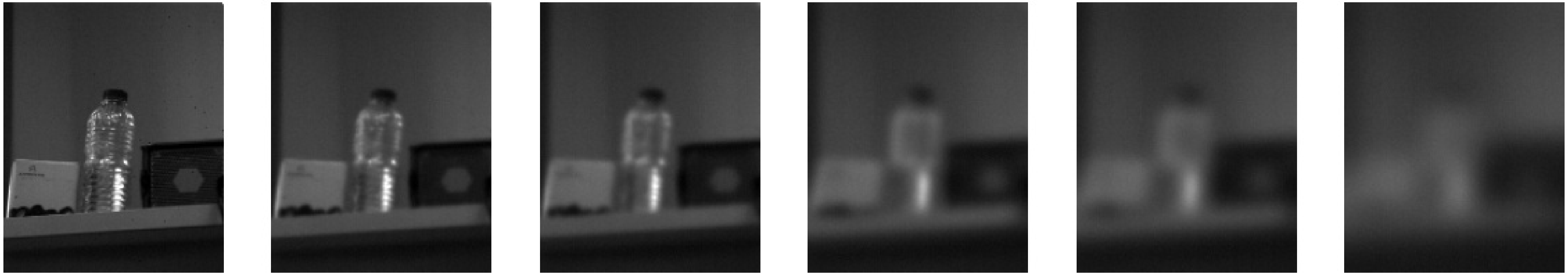




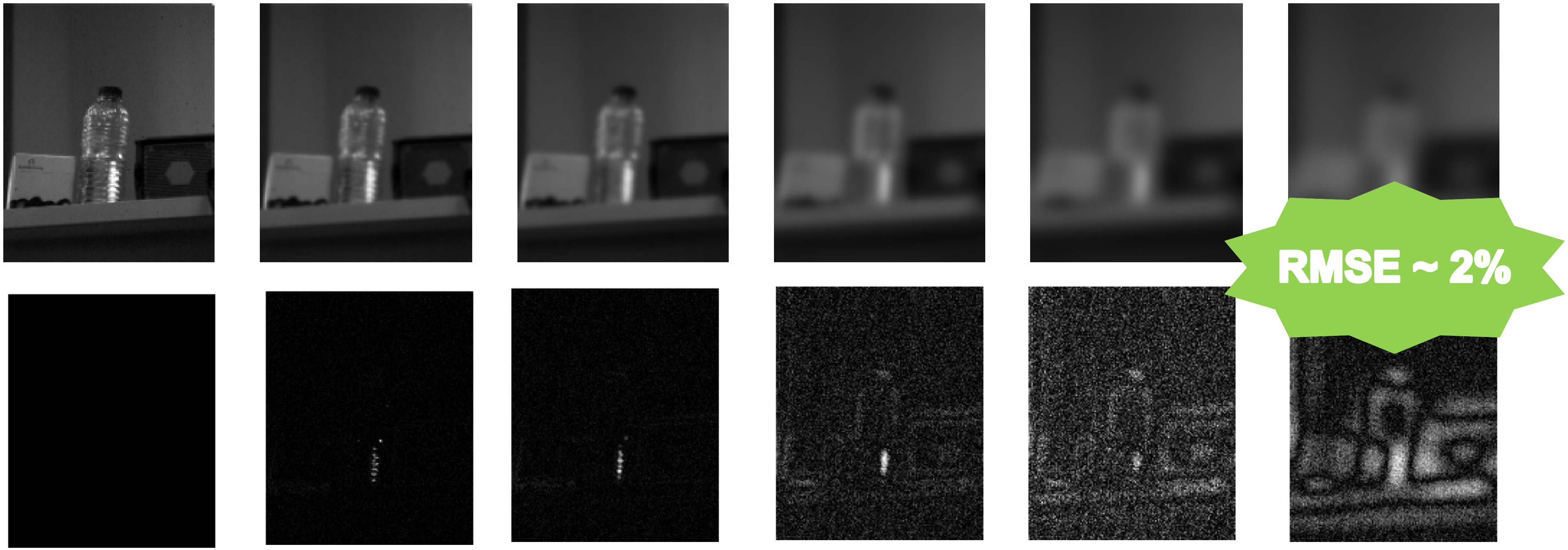
Technology	0.35μm CMOS 2P4M
Vendor (Process)	Austria Microsystems (C35OPTO)
Die size (with pads)	7280.8μm × 5780.8μm
Cell size	34.07μm × 29.13μm
Fill factor	6.45%
Resolution	QCIF: 176 × 144px
Photodiode type	n-well/p-substrate
Power supply	3.3V
Signal range	[1.5V, 2.5V]
FPN	0.72%
PRNU (50% signal range)	2.42%
Sensitivity	0.15V/(lux·s)
Measured power consumption (worst case)	5.6mW @ 30fps 22 × 18px
Predicted power consumption (worst case)	17.6mW @ 30fps 176 × 144px
ADC throughput	0.11MSa/s (9μs/Sa)
Internal clock freq. range	0.5-150MHz

[Fernández-Berni et al, IJSSC 2011]

Scale space generation on/off chip?

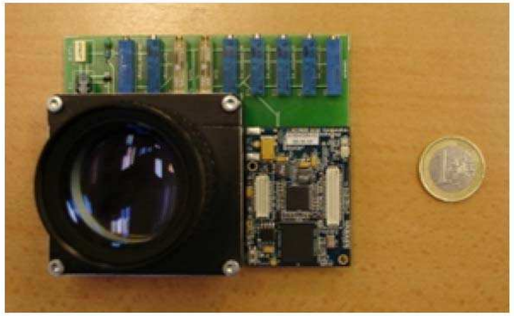


Scale space generation off/off chip?



FLIP-Q, application scenario: forest fire detection

[Fernández-Berni et al, Int. Journal of Wildland Fire 2012]



Wi-FLIP: Integration of FLIP-Q with a commercial WSN node (Imote2)

Real tests: prescribed burning of an area of vegetation

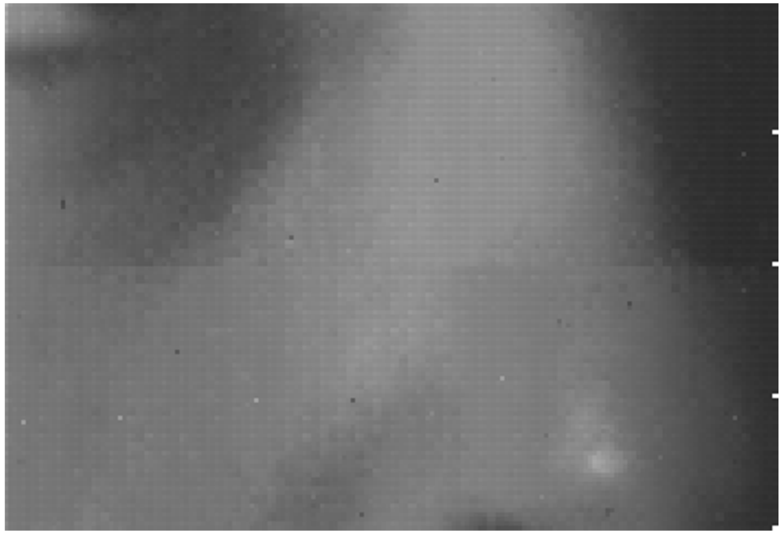
Image sent via radio to a base station when the fire was detected

Setting of algorithm parameters by using smoke generators in a public park

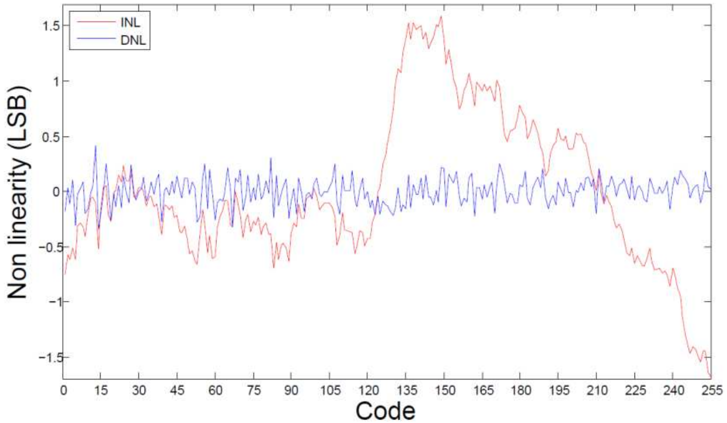
Hierarchical vision processing architecture

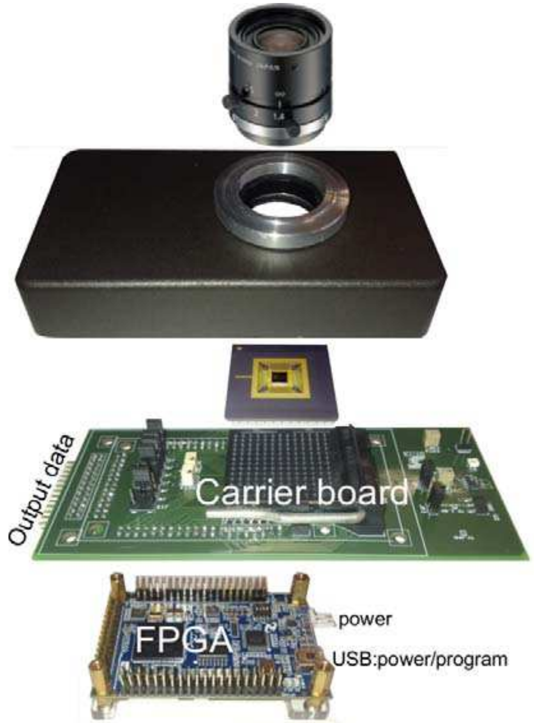
[M. Suárez et al, IEEE JETCAS 2012]

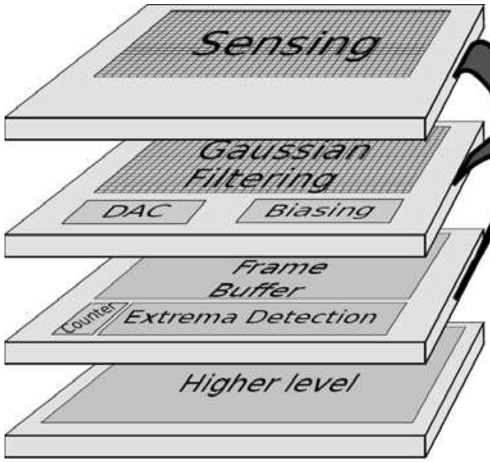
Image capture and step-by-step Gaussian filtering

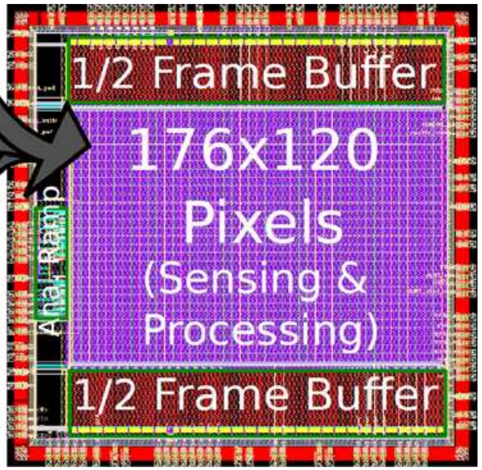


Fully-parallel conversion with ramp ADC

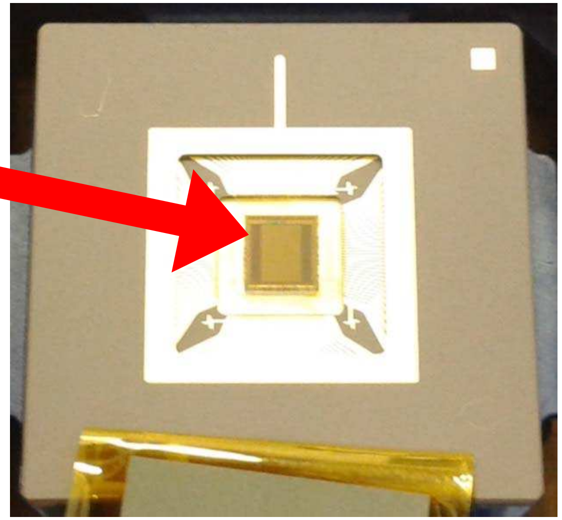






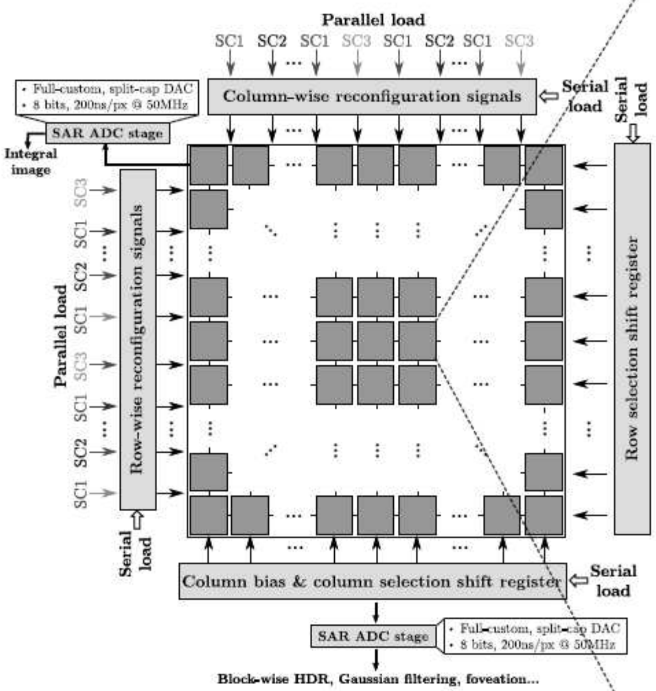


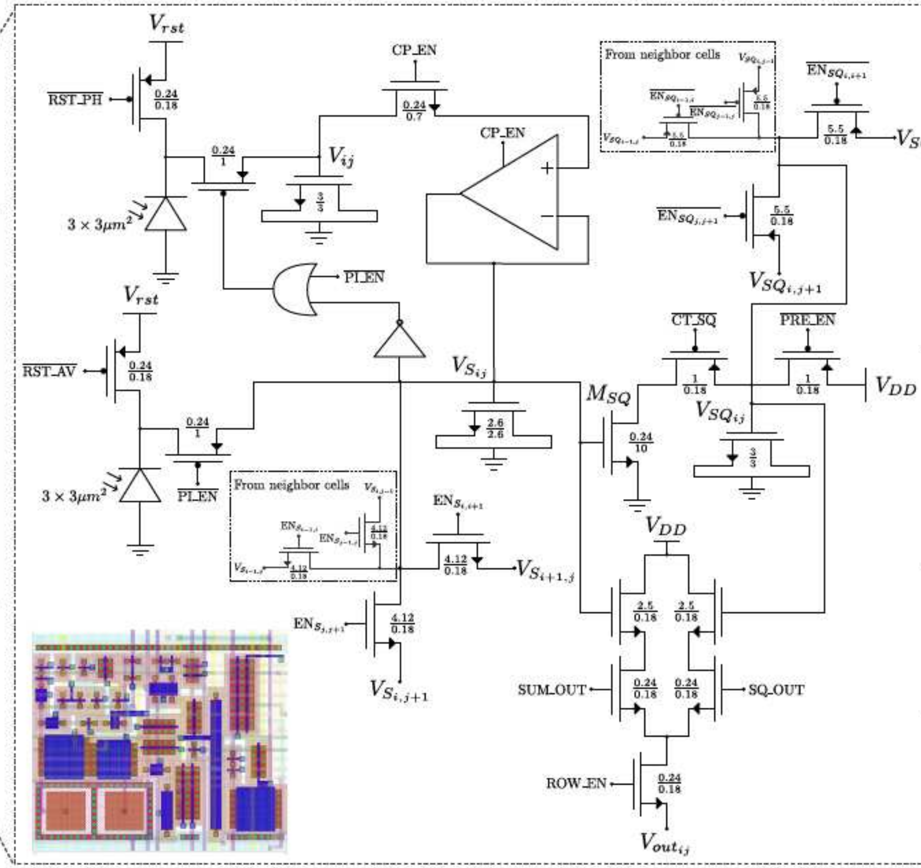
Technology	CMOS 180nm
No. pixels	176×120
No. PE's	88×60
Chip area	5×5 mm^2
Cell area	44×44 mm^2
Power	75nW/px @30fps
Min. filter	σ = 0.48



FLIPQ-II: QVGA smart image sensor

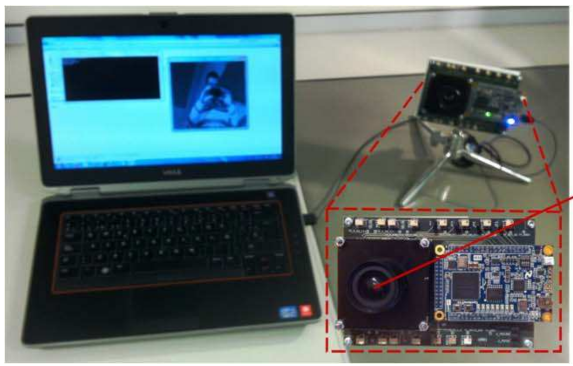
[Fernández-Berni et al, Int J. of Circuit Theory and Appl. 2014]

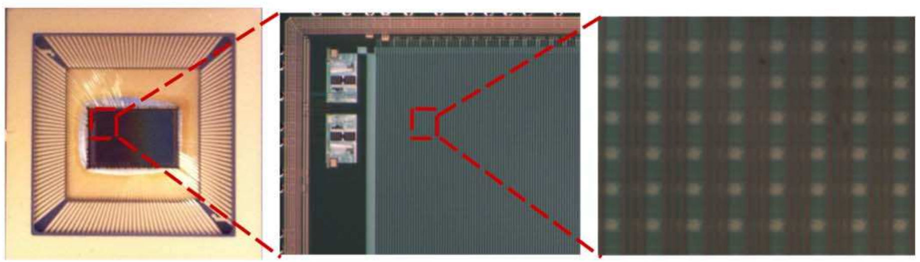


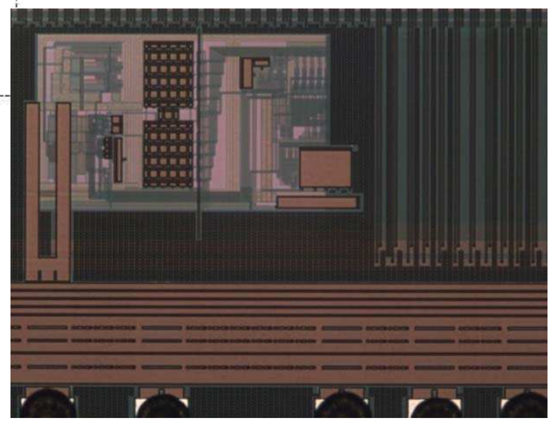


- Scale space
- Multi-resolution image representation
- Energy-based representation
- Block-wise HDR
- Integral image

Technology	Std 0.18μm 1.8V 1P6M CMOS process
Die size (with pads)	7.5mm × 5mm
Pixel size	19.59μm × 17μm
Fill factor	5.4%
Photodiode type	n-well/p-substrate
Power supply	3.3 (pads), 1.8V (core)
DSNU	1.7%
PRNU (50% signal range)	3.5%
ADC throughput	5MSa/s (200ns/Sa)
Power consumption@30fps	
Image capture	42.6mW
Block-wise HDR	42.6mW
Integral imaging	55.2mW
Gaussian filtering	43.1mW
Foveation	42.6mW



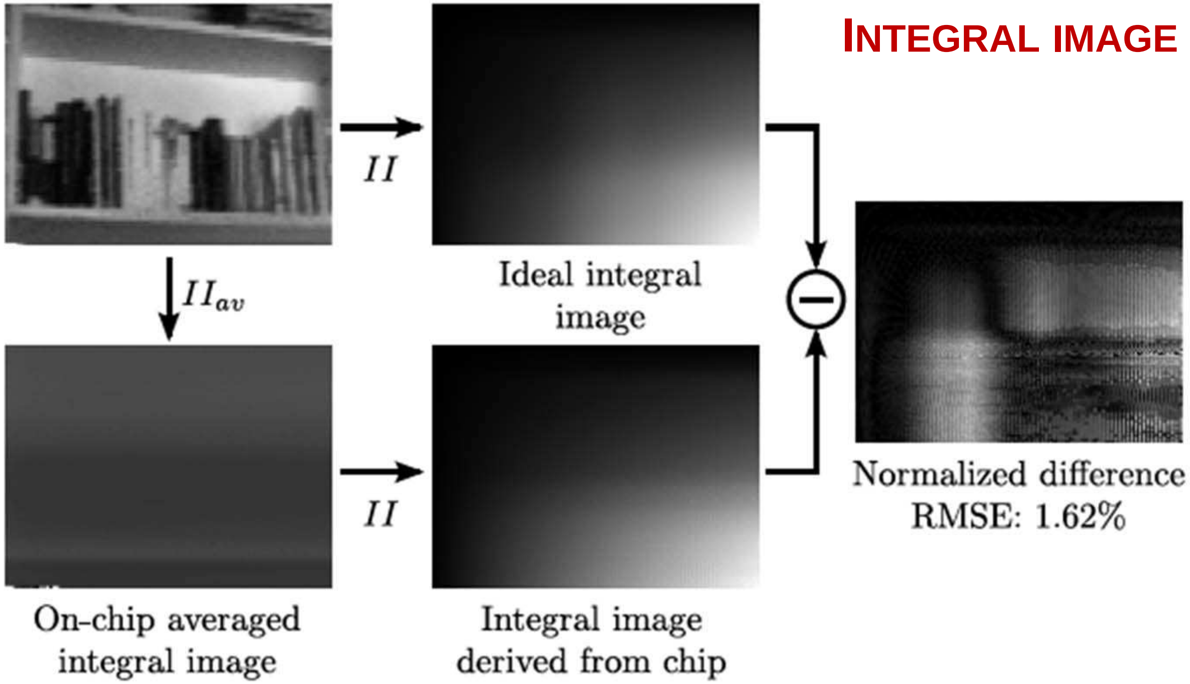




Full-custom SAR ADC @ 5Msa/s

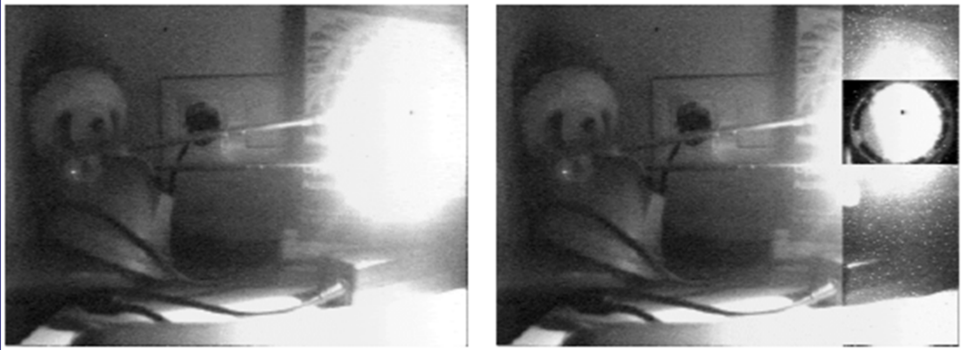
FLIPQ-II: experimental results

INTEGRAL IMAGE



GAUSSIAN FILTERING

BLOCK-WISE HIGH DYNAMIC RANGE



Global integration time control

Block-wise integration time control

FOVEATION AND INVERSE FOVEATION

And what about the industry?

CURRENT ART

- Very significant progress on CMOS Imagers
- Multi-Mega Pixel sensors with pixel pitch in the 1μm range
- Ultra-high speed sensors with data rates in the 10Gpixels/s range
- Very large embedding of image system functions on-chip: ADCs, image correction, thermal and energy management, etc.
- Progressive incorporation of smartness at sensor chip level, but still in infancy phase.

FORECAST

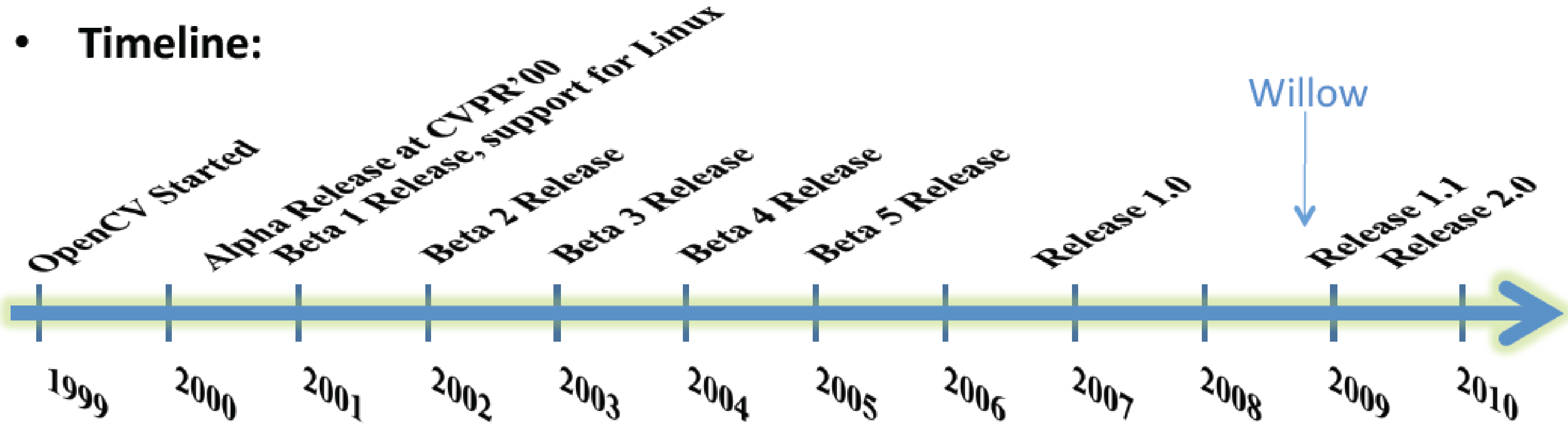
- More pixels, smaller pixel pitch, higher speed, etc.
- BASICALLY A DOMAIN FOR LARGE COMPANIES FOCUSED ON CONSUMER APPLICATIONS !!

OpenCV (Open Source Computer Vision)

- Open source computer vision and machine learning software library.
- OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in commercial products.
- BSD-licensed product, easy for businesses to utilize and modify the code.
- The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms.
- Employed by well-established companies like Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota as well as startups such as Applied Minds, VideoSurf, and Zeitera,
- It has C/C++ (aiming real-time operation), Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS. Primary vision package for ROS (Robot Operating System)
- 47 thousand people of user community and estimated number of downloads exceeding 7 million growing by nearly 200K/month (lots from Asia). De-facto standard.

OpenCV (Open Source Computer Vision)

- **Timeline:**



- **Since 2012, the library is developed by a company called Itseez located in Russia (four members of OpenCV dev. team who have been working on the library for more than 10 years are now with Itseez).**

- **November 2013: Khronos Releases OpenVX 1.0 Specification for Computer Vision Acceleration**

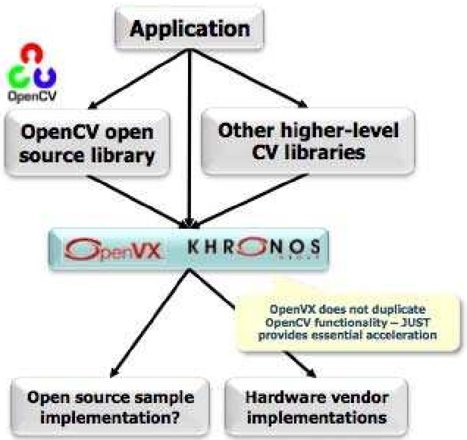
- **According to Itseez: "We work on OpenCV that has become a de-facto standard in the area of computer vision both for research and product prototyping. Now we are pushing the industry ahead by working on the OpenVX -- a computer vision hardware abstraction layer that will be standardized across many hardware vendors by the Khronos group."**

OpenVX: industrial standard

OpenVX

- **Vision Hardware Acceleration Layer**
  - Enables hardware vendors to implement accelerated imaging and vision algorithms
  - For use by high-level libraries or apps
- **Focus on enabling real-time vision**
  - On mobile and embedded systems
- **Diversity of efficient implementations**
  - From programmable processors, through GPUS to dedicated hardware pipelines

Dedicated hardware can help make vision processing performant and low-power enough for pervasive 'always-on' use



OpenVX and OpenCV are Complementary

	OpenCV	OpenVX
Governance	Open Source Community Driven No formal specification	Formal specification and conformance tests Implemented by hardware vendors
Scope	Very wide 1000s of functions of imaging and vision Multiple camera APIs/interfaces	Tight focus on hardware accelerated functions for mobile vision Use external camera API
Conformance	No Conformance testing Every vendor implements different subset	Full conformance test suite / process Reliable acceleration platform
Use Case	Rapid prototyping	Production deployment
Efficiency	Memory-based architecture Each operation reads and writes memory	Graph-based execution Optimizable computation, data transfer
Portability	APIs can vary depending on processor	Hardware abstracted for portability

OpenVX: industrial standard

OpenVX 1.0 Function Overview: ***low- and mid-level vision tasks that "lend themselves to an appreciate level of hardware-based optimization"***

Core data structures
- Images and Image Pyramids
- Processing Graphs, Kernels, Parameters

Image Processing
- Arithmetic, Logical, and statistical operations
- Multichannel Color and BitDepth Extraction and Conversion
- 2D Filtering and Morphological operations
- Image Resizing and Warping

Core Computer Vision
- Pyramid computation
- Integral Image computation

Feature Extraction and Tracking
- Histogram Computation and Equalization
- Canny Edge Detection
- Harris and FAST Corner detection
- Sparse Optical Flow

Conclusions… specifications?

- OpenVX is a crucial framework for us since it defines specific functionalities demanding vision hardware optimization
- If finally adopted by the industry, OpenVX will open the market to non-conventional vision hardware since software and system engineers will be able to reuse the same source code for different hardware solutions
- Supported by the frame of reference set up by OpenCV and OpenVX, we require specifications at application level in order to further exploit the functionalities provided by our chips
- In particular, we would like to map those high-level specifications in a HW/SW co-design framework enabling parameter tuning for the different system components at different abstraction levels
- Tight HW/SW integration based on well-established standards can lead to unprecedented performance figures for vision systems


INSTITUTO DE
MICROELECTRÓNICA
DE SEVILLA

Vision group at the Institute of Microelectronics of Seville

Spanish National Research Council-University of Seville

PI: Ángel Rodríguez-Vázquez

14 Doctors

- 2 Full Professors
- 2 Tenured Scientists
- 4 Associate Professors
- 6 Postdoc researchers

3 Doctoral Students

Vision group at the Institute of Microelectronics of Seville

Spanish National Research Council-University of Seville

Other research activities:

- Application of **Single-Photon Avalanche Photodiodes** to 3D Imaging and Medical Imaging
- **Compressed sensing** for exploitation of visual information sparsity and redundancy

Questions?

THANK YOU VERY MUCH FOR YOUR ATTENTION

berni@imse-cnm.csic.es