

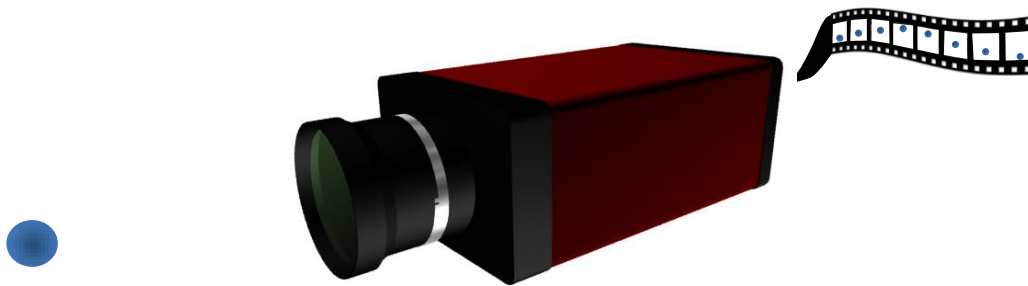
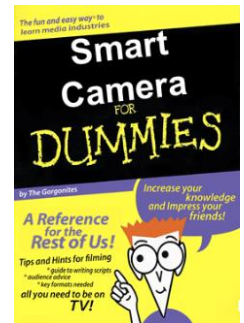
# FPGA-based embedded vision systems

François BERRY et al.

[francois.berry@uca.fr](mailto:francois.berry@uca.fr)

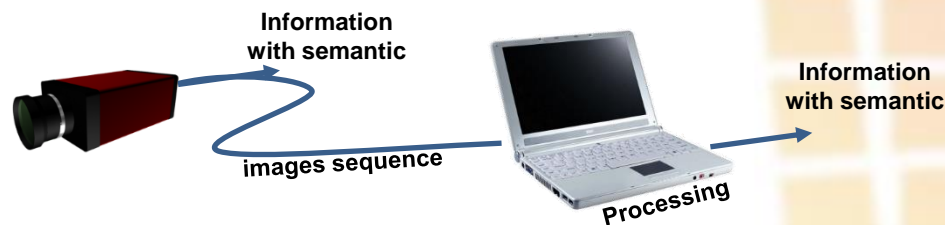


# Camera and « Smart camera »



Caméra = Device providing images or sequence, video,....

## Traditional machine vision system

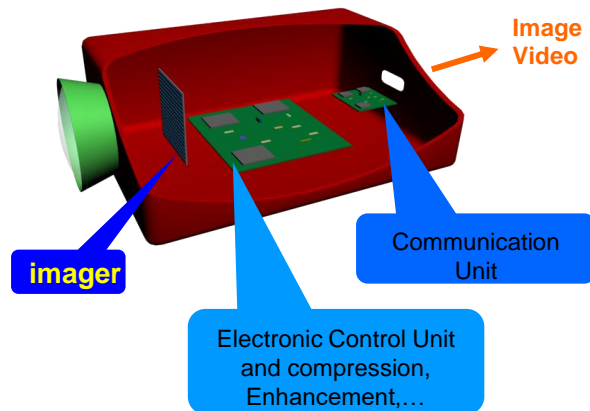


# Camera and « Smart camera »

## Traditional camera

- Imager
- Electronics
- Interfaces

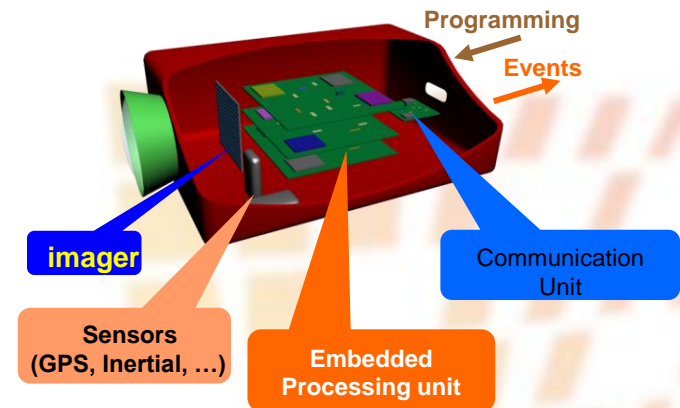
Traditional camera provides images and videos.



## Smart camera

- Imager (and sensors)
- Onboard computer
- Interfaces

Smart camera delivers **abstracted image data** and is also **configurable** and **programmable**

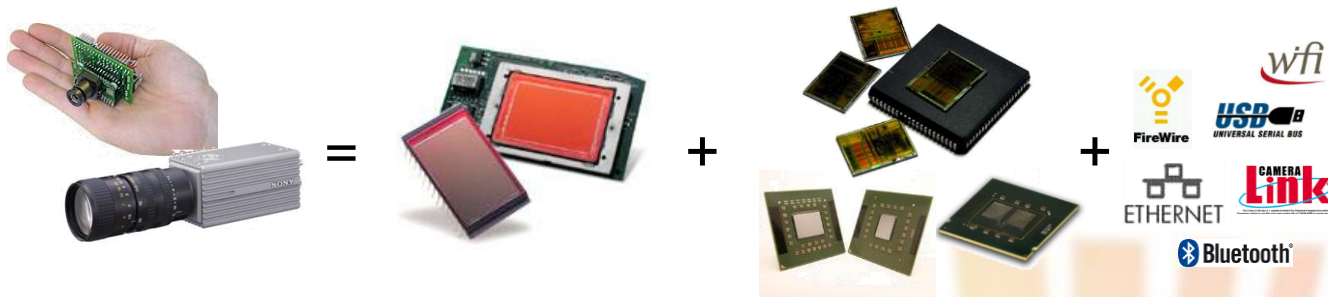


# Smart camera

## A definition

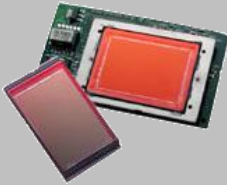
A **smart camera** is an embedded [machine vision](#),

- in addition to image capturing circuitry,
- includes a processor, which can extract information from images without need for an external processing unit,
- interfacing devices used to make results available to other devices.



$$\text{Smart Camera} = \sum \text{Capture} + \text{Processing} + \text{Communication}$$

# Anatomy of a smart cam



## Sensing devices:

Imager: CMOS, CCD,...

Proprioceptive devices: Inertial set,...

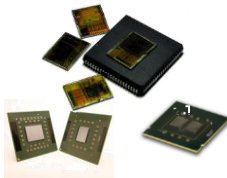


## Communication unit:

Wireless: Wifi, Bluetooth, ZigBee...

Wired: USB, FireWire, Ethernet

Gigabit, Camlink,...



## Processing Unit:

Micropocessor

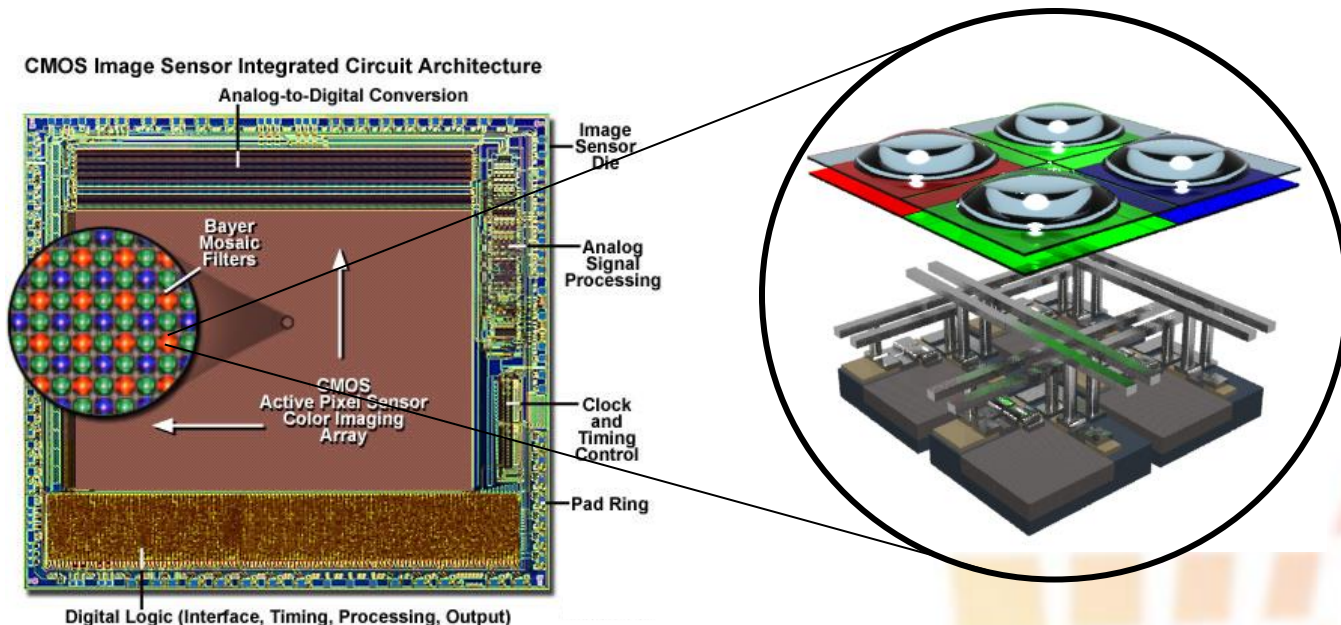
DSP

FPGA

ASIC, GPU,...

# CMOS Imager

The major advantage that **CMOS imagers** enjoy over their CCD counterparts is the ability to integrate a number of processing and control functions directly onto the sensor integrated circuit, which lie beyond the primary task of photon collection.



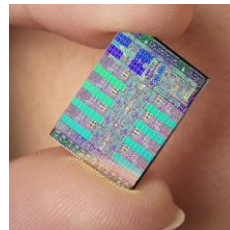
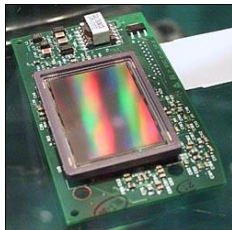
# Common CMOS features

CMOS sensors use the same technological process as modern microchips:

- Many foundries available worldwide
- Cost efficient
- Latest processes available down to 90/65 nm (Sony)

CMOS process enables integration of many additional features:

- Random pixel access, windowing, sub-sampling and binning
- Various pixel circuits from 3 transistors up to many 100 transistors per pixel
- Analog signal processing (e.g. CDS, programmable gain, noise filter)
- A/D conversion
- Logic (timing control, digital signal processing, etc.)

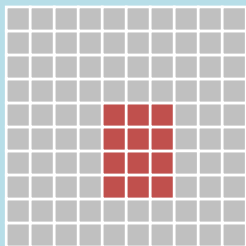


# Random pixel access

- Different scanning methods are available to reduce the number of pixels being read:
  - Allows for higher frame rate or lower pixel rate (reduction in noise)
  - Can reduce power consumption due to reduced data

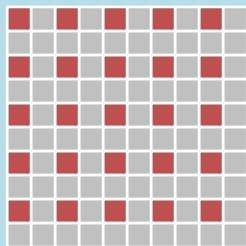
## Windowing

- Reading of one or multiple rectangular subwindows
- Used to achieve higher frame rates (e.g. AO, guiding)



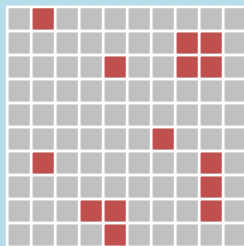
## Subsampling

- Skipping of certain pixels/rows when reading the array
- Used to obtain higher frame rates on full-field images



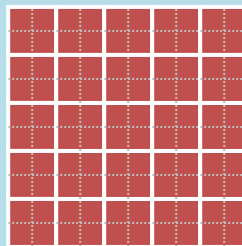
## Random Read

- Random access (read or reset) of certain pixels
- Selective reset of saturated pixels
- Fast reads of selected pixels



## Binning

- Combining several pixels into larger super pixels
- Used to achieve lower noise and higher frame rates



# Others sensing devices

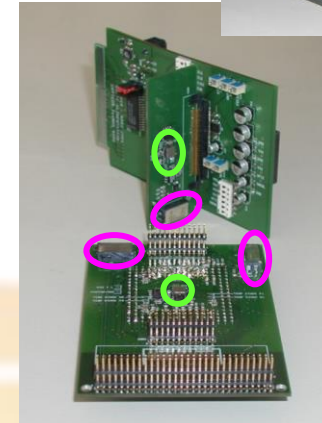
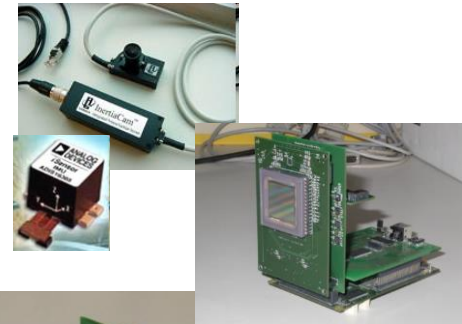
Imaging unit can be improved by others sensors such as:

- Inertial measurement unit:  
(3 linear accelerations and 3 rotational velocities)

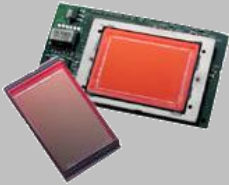
Used for:

- Image stabilization
- Inertial and Image fusion
- Inclinator (with gravity)
- ....

- GPS module  
provides reliable positioning, navigation,  
and timing services



# Anatomy of a smart cam



## Sensing devices:

Imager: CMOS, CCD,...

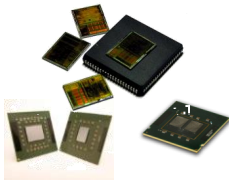
Proprioceptive devices: Inertial set,...



## Communication unit:

Wireless: Wifi, Bluetooth, ZigBee...

Wired: USB, FireWire, Ethernet Gigabit, Camlink,...



## Processing Unit:

Micropocessor

DSP

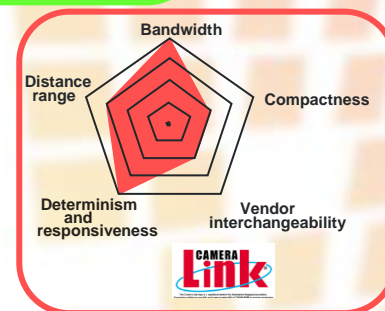
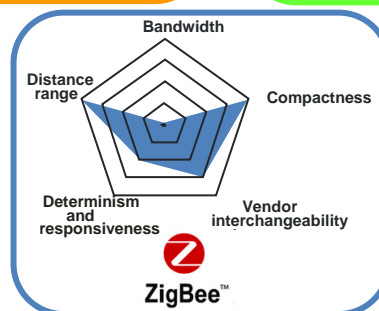
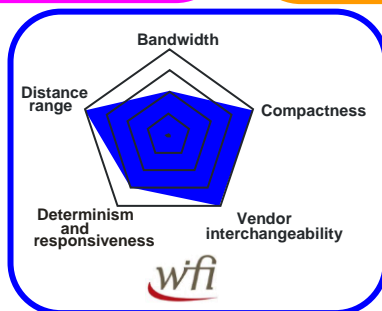
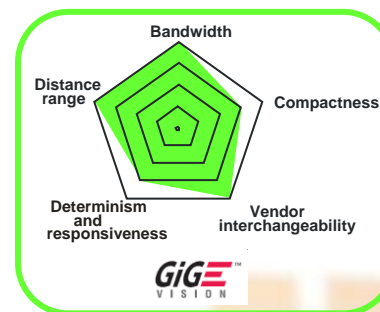
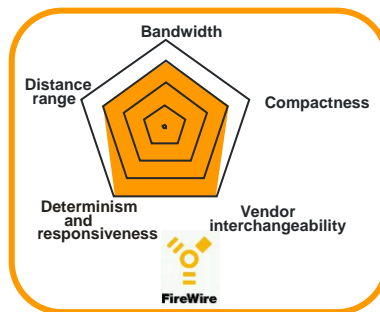
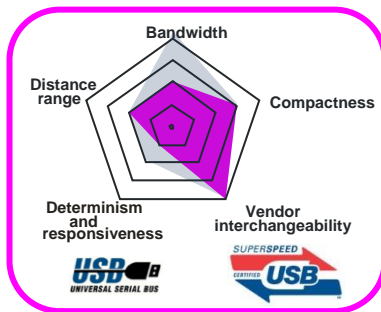
FPGA

ASIC, GPU,...

# Taxonomy of Communication units

Communication channel can be classified according to 5 factors (not exhaustive!):

- Bandwidth
- Distance range
- Compactness
- Determinism and Responsiveness
- Vendor interchangeability



# Anatomy of a smart cam



## Sensing devices:

Imager: CMOS, CCD,...

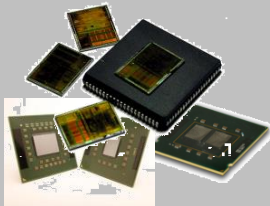
Proprioceptive devices: Inertial set,...



## Communication unit:

Wireless: Wifi, Bluetooth, ZigBee...

Wired: USB, FireWire, Ethernet Gigabit, Camlink,...



## Processing Unit:

Micropocessor

DSP

FPGA

ASIC, GPU,...

# Processing Unit

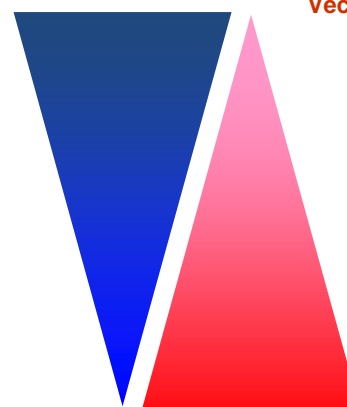
The objective of Vision system is to measure different features of a specific object.

As every image contains a lot of unnecessary information,  
**so data reduction is key for efficiency.**

Mostly image processing tasks can be divided into 3 levels:

- **Parallel data reduction**, where the operations depends on the closest surroundings (e. g. noise reduction, edge detection).
- **Serial data reduction**, where data from the entire image is needed for the processing task (e. g. count the number of elements).
- **Classification**, that try to inform what we see.

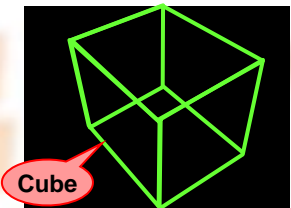
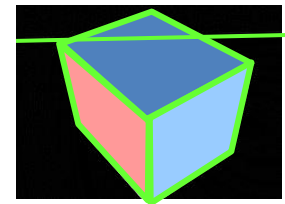
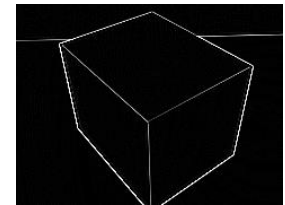
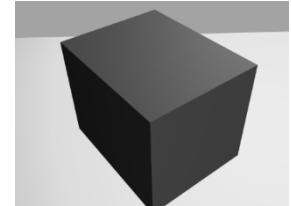
Amount of data



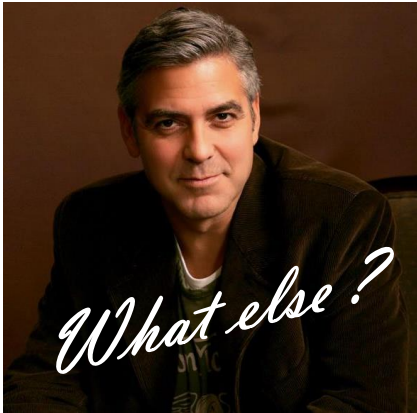
Algorithm  
Complexity

Image processing  
Vector Processor, SIMD,  
FPGA  
50 GOPS

Application  
processing  
CPU 10MOPS



# Why FPGA-based Smart -cams ?



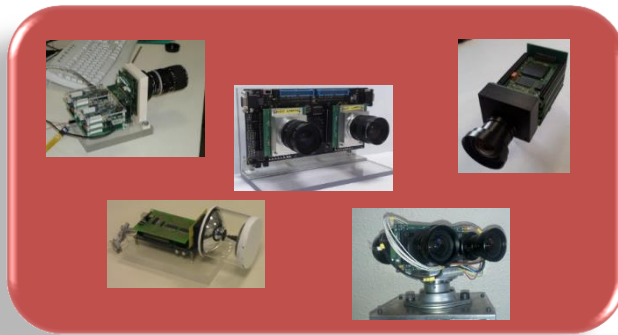
- General purpose processors (GPPs) (single/multi-core)
- DSPs
- Parallel architectures built on the former
- GPUs
- ASICs

# Why FPGA-based Smart -cams ?

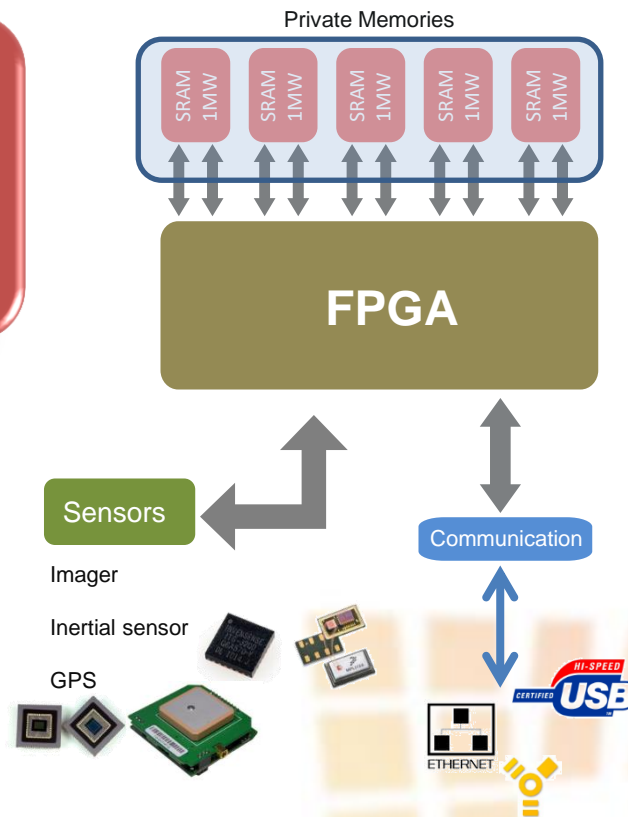
## The **hardware reconfiguration facilities** offered by FPGAs

- brings **flexibility** (from full-fledged multi-softcore architectures down to low-level HW interfacing tasks)
- naturally supports **co-design**-based approaches
- is suitable for **tightly embedded** systems (with strong operational constraints)
- offers vast opportunities for **design-space exploration**

# Synoptic of our smart cameras

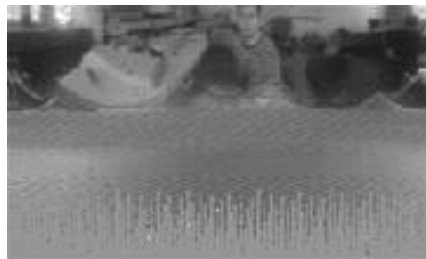
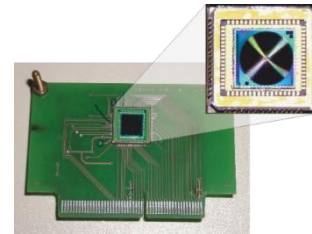
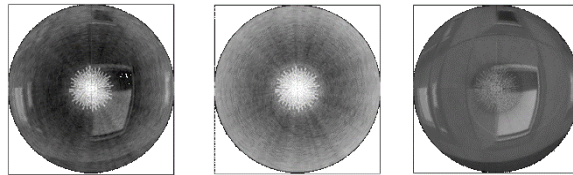
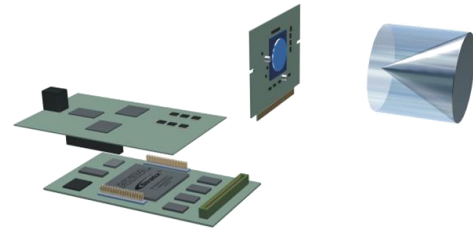
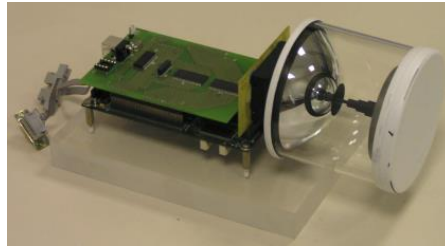


Ours Smart cameras are  
FPGA-based systems  
and modular.



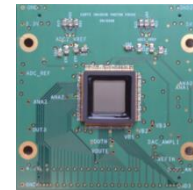
# Few FPGA-based smart camera

OmniMOS



# Few FPGA-based smart camera

BiSeeMOS



a) Left Frame



b) LoG Frame



c) Depth Frame

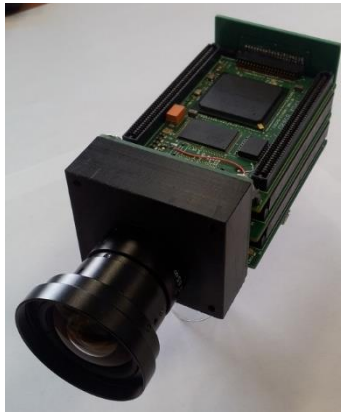


d) Depth Frame with a Threshold of 64/120

Disparity map 1024x1024 @ 200 fr/s

# Few FPGA-based smart camera

DREAM-CAM



Imaging device:  
APTINA  
CMOS 1024x1024  
Couleur/NB

E2V  
CMOS 2048x2048  
Couleur/NB

Sensor Board:  
GPS+ Compass + Inertial

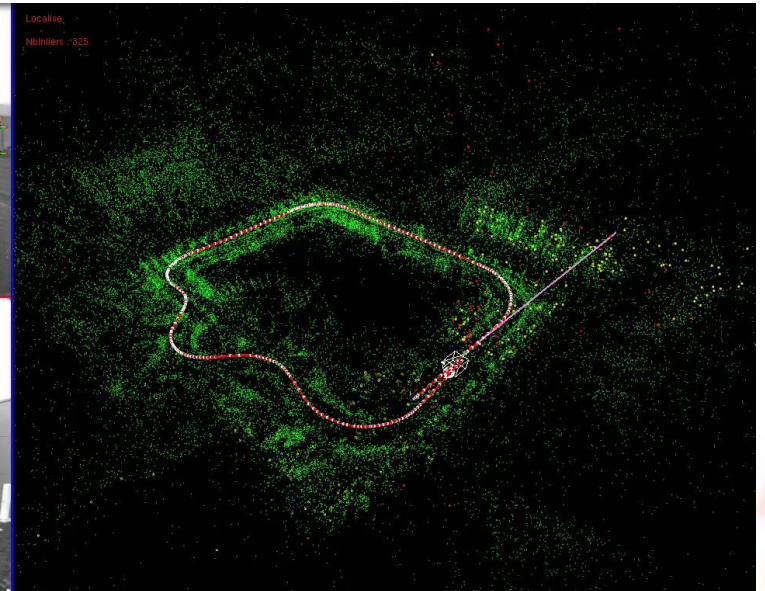
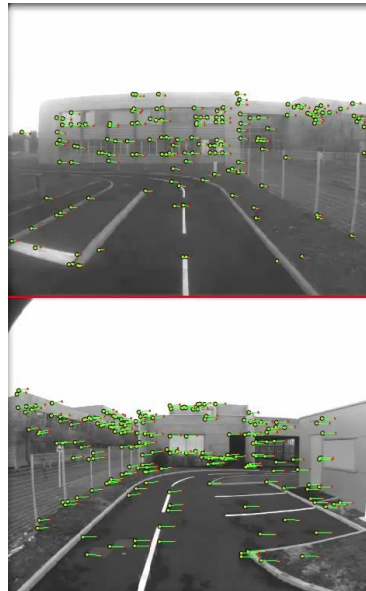
Memory Board:  
SRAM  
6 x 1MB @ 100MHz

Main Board  
Cylcone III EP3C120

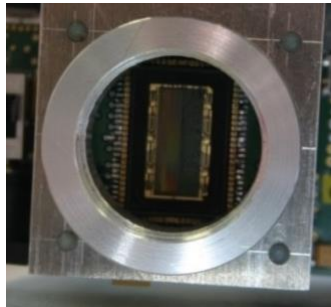
Power Supply  
Provides 10 voltages  
(1,2v to 5v) @5 A  
+ JTAG Programmer

Communication Board:  
USB2,0, Giga-ethernet

# Application: Navigation



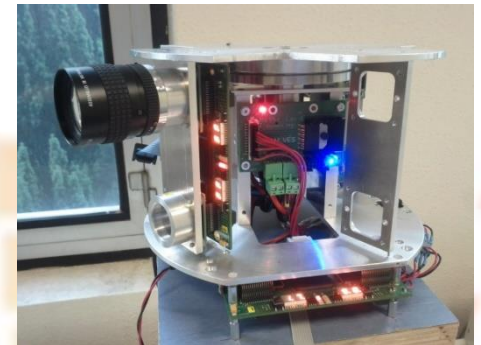
# Grey-level Scanning platform: PANORAMOS



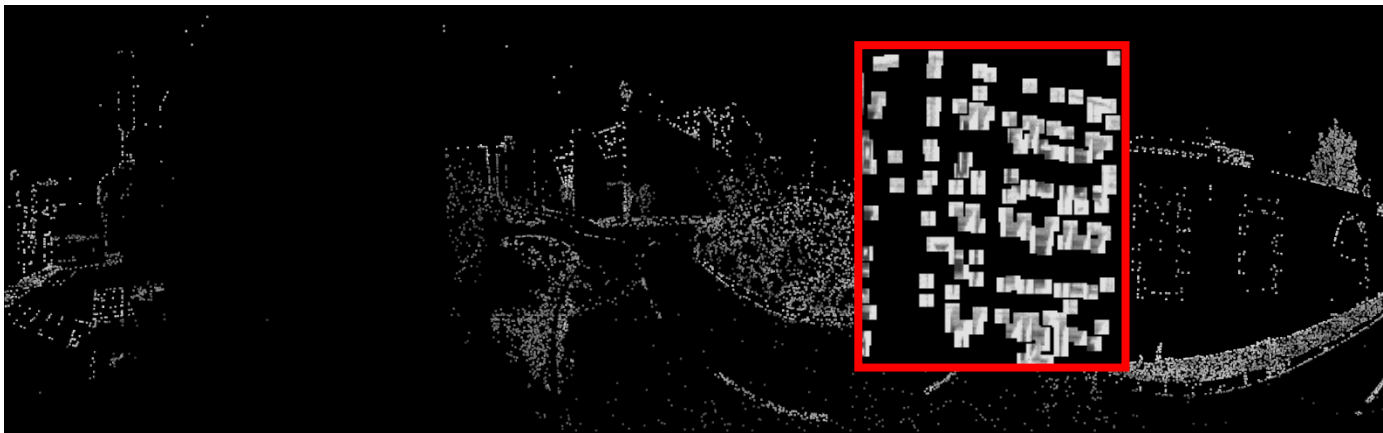
- Allows to acquire Grey-level image
- Extracts visual features on the flow
- Give a description by coordinates and Grey-level template
- Synchronization between mechanics and sensing

## RESULTS:

- Innovative mechanical structure with hanging rotating part
- Full matching between mechanics and sensing



# Grey-level Scanning platform: PANORAMOS



# Processing



Image processing  
Computer vision

# Deep learning approach

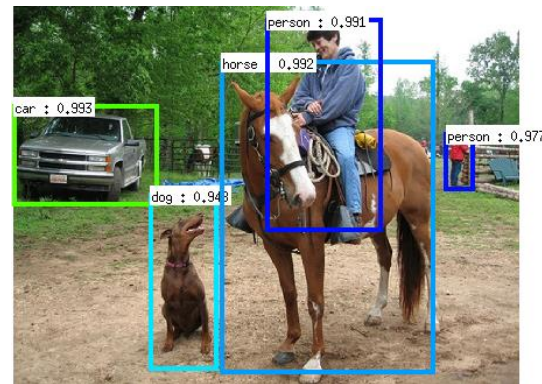
- Invention associated to LeCun in the late 90s [1]
- The state of the art in image classification[2], detection[3] ...



“ five “



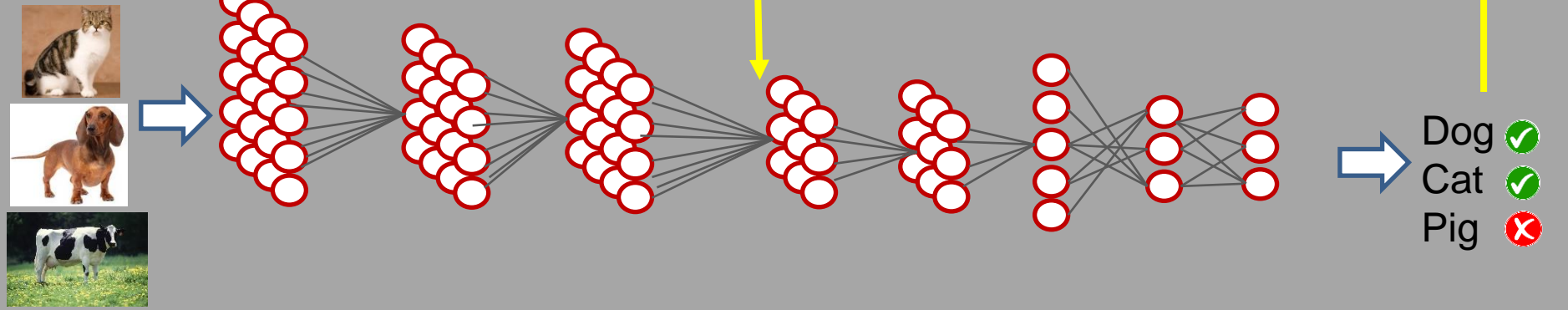
x1M  
More data and  
computations



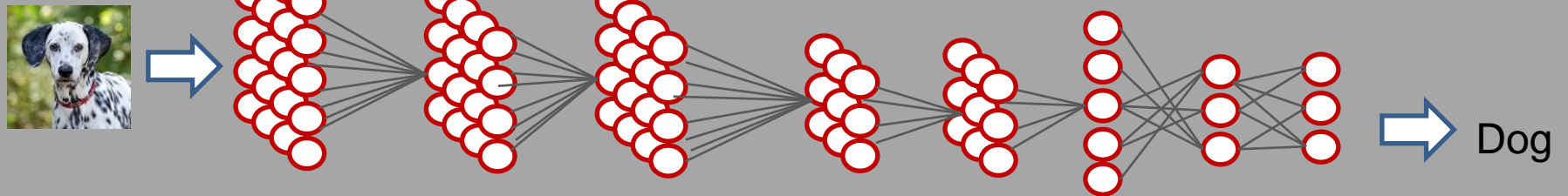
- 
- [1] LeCun et al. Gradient Based Learning Applied to Document Recognition - IEEE 1998  
[2] Krizhevsky et al. ImageNet Classification with Deep Convolutional Neural Networks - NIPS'12  
[3] Girshick et al. - Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation - CVPR 2014

# Deep learning approach

## Training

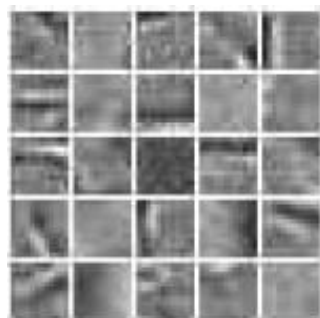


## Deploy

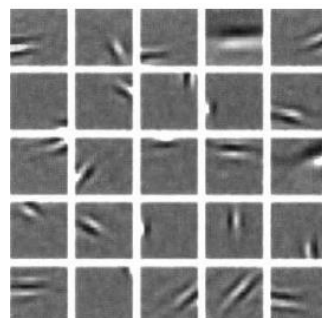


# What features are learned?

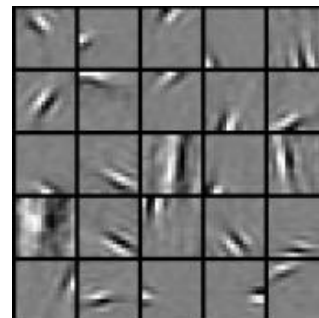
- Applied to image patches, well-known result:



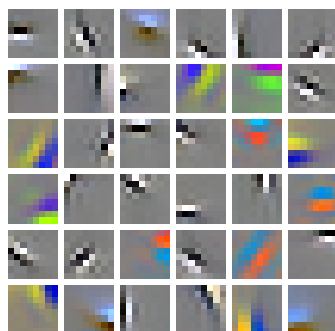
Sparse auto-encoder  
[Ranzato et al., 2007]



Sparse coding  
[Olshausen & Field, 1996]



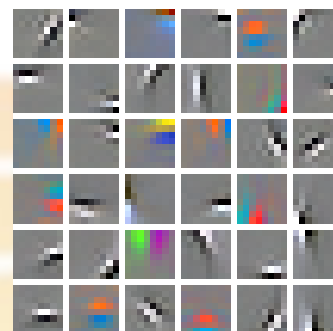
Sparse RBM  
[Lee et al., 2007]



Sparse auto-encoder

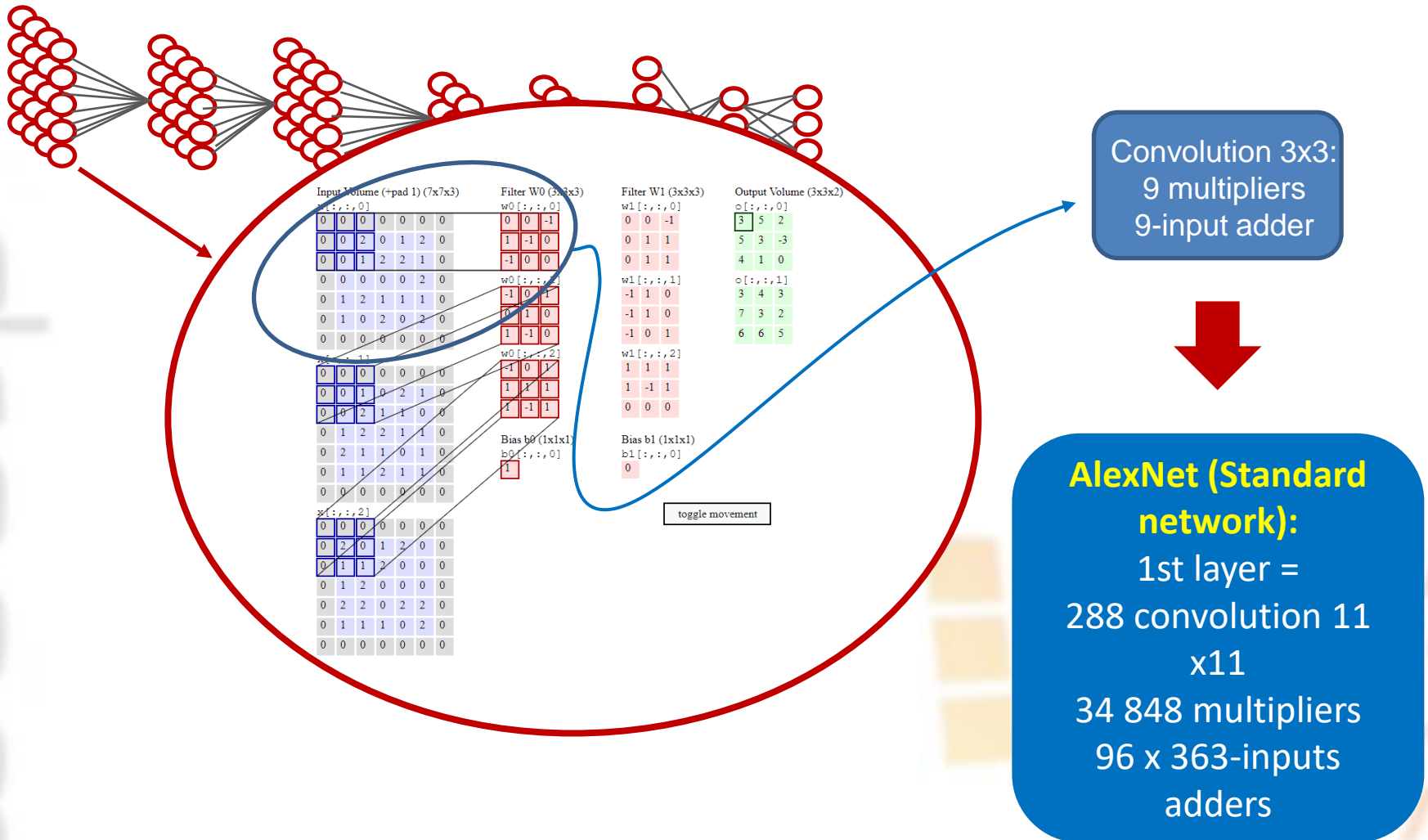


K-means



Sparse RBM

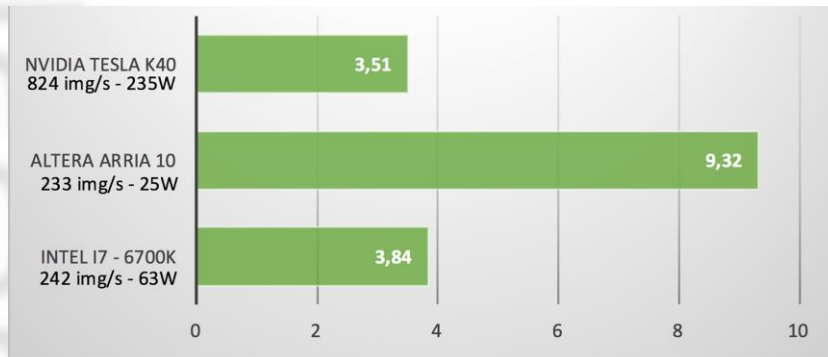
# Deep learning vs Computation....



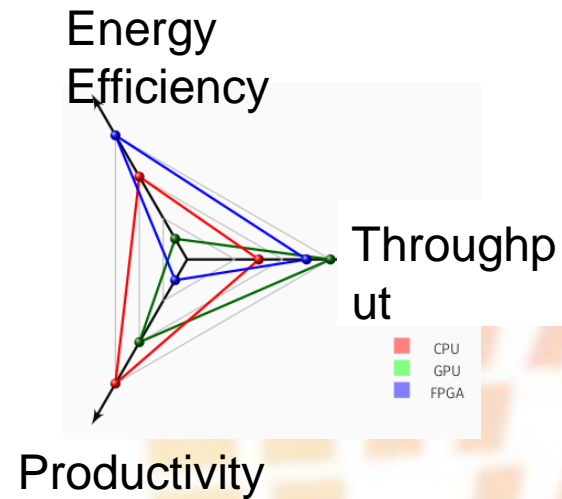
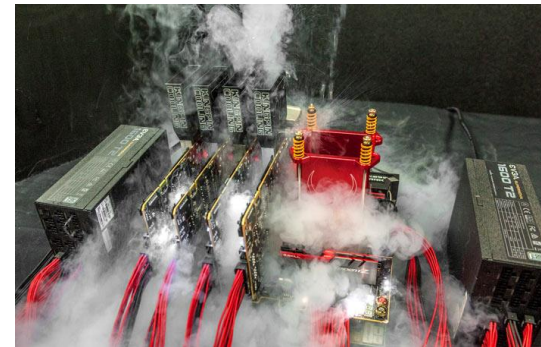
# Hardware Accelerators for DL

## FPGAs for DL

- DSP Blocks (variable precision)
- *In situ* BRAM/SRAM
- Reconfigurability vs ASICs
- **Logic fabric (we'll talk about that later)**



Energy efficiency of AlexNet  
inference



# Approximate Computing for Deep Learning



**Twist plot\*** : DL tolerate approximate computing (AC)

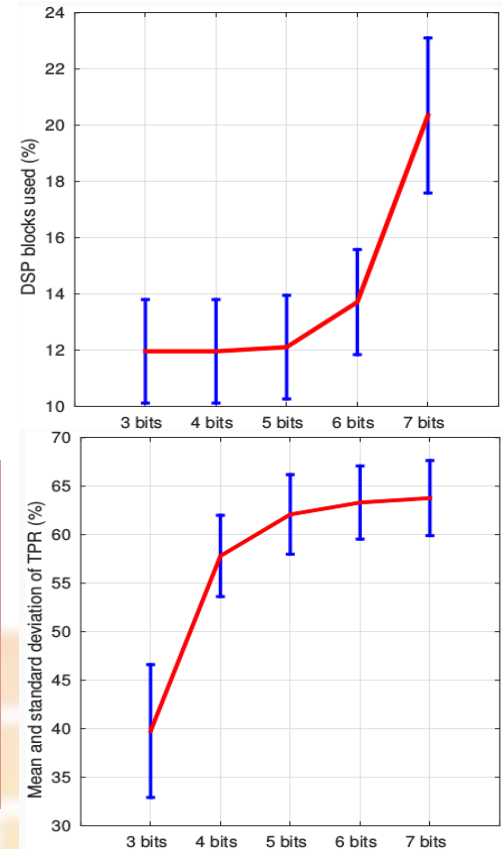


## Approximate arithmetic:

- Fixed point
- Dynamic Fixed point
- Binary and pseudo-binary Nets

## Reduce operations:

- Prune neurons
- Low rank approximation of weights



\* Well, no that much ...

# Thank You



[francois.berry@uca.fr](mailto:francois.berry@uca.fr)