

embedded **VISION** SUMMIT 2018

What is Neuromorphic Event-based Computer Vision? Sensors, Theory and Applications

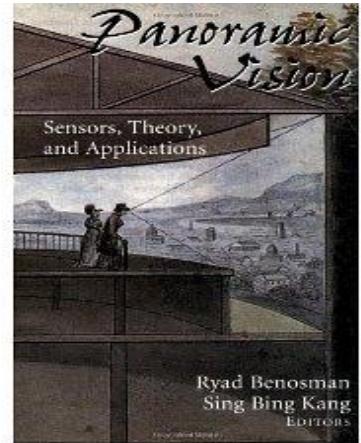
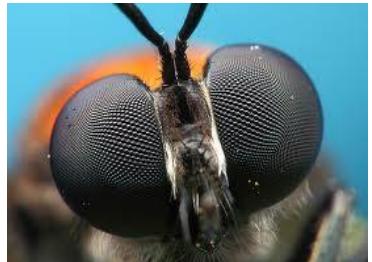


 UPMC
University of Pittsburgh
Medical Center

 Carnegie
Mellon
University  UPMC SOITONNE UNIVERSITÉS

R.B. Benosman
05/23/18

Omnidirectional Vision



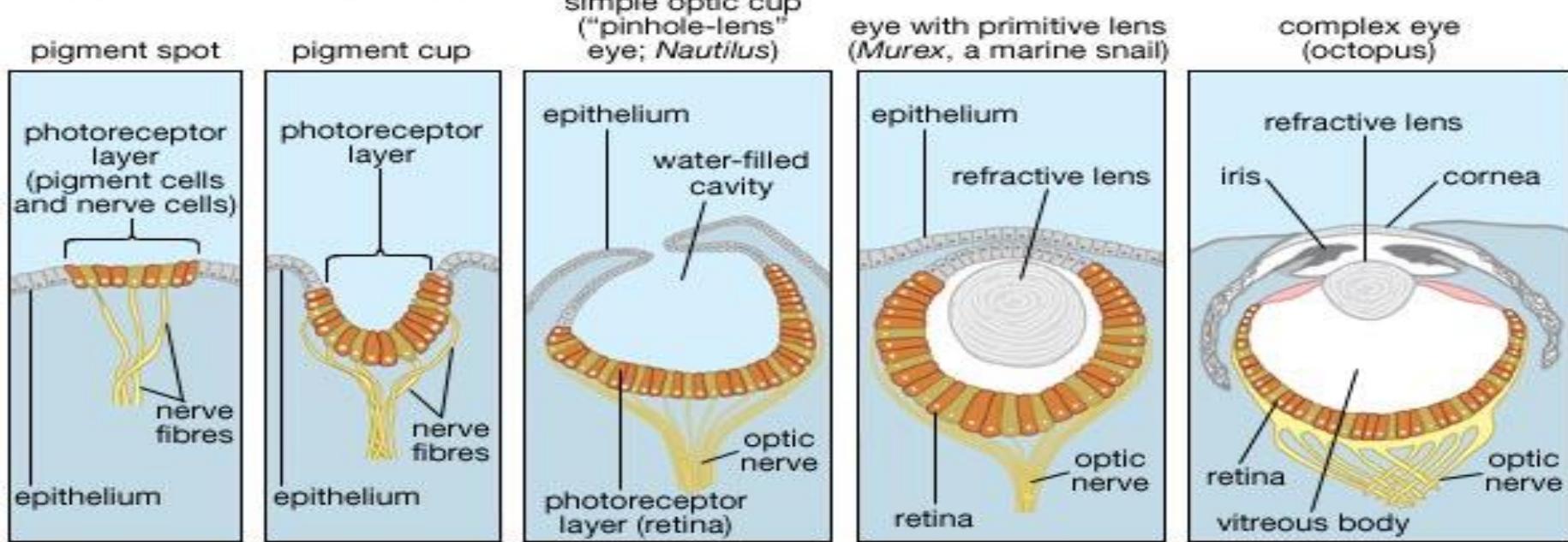
Format Kindle



Technologie et Capteurs pour la Vision 360°

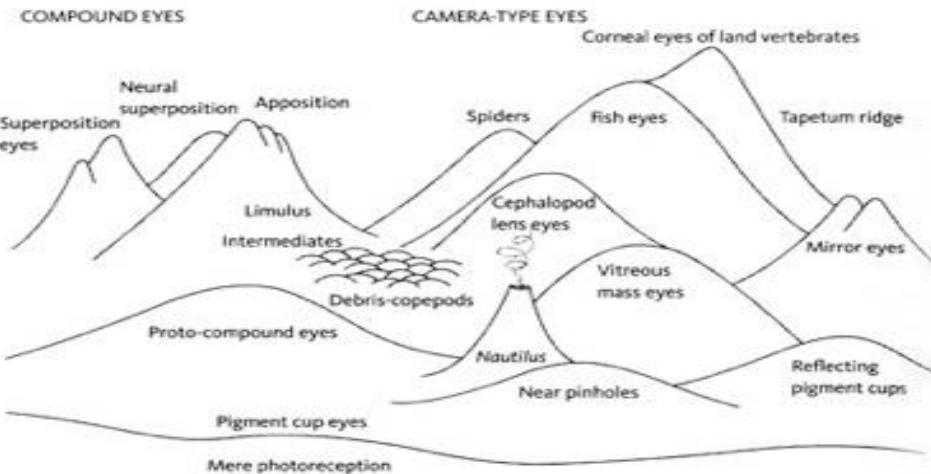
Origins of Eyes

Stages of eye complexity in mollusks



© 2005 Encyclopædia Britannica, Inc.

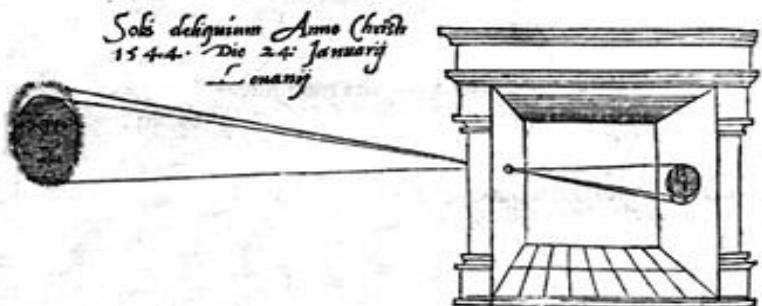
Landscape of possible eye forms (M. Land)



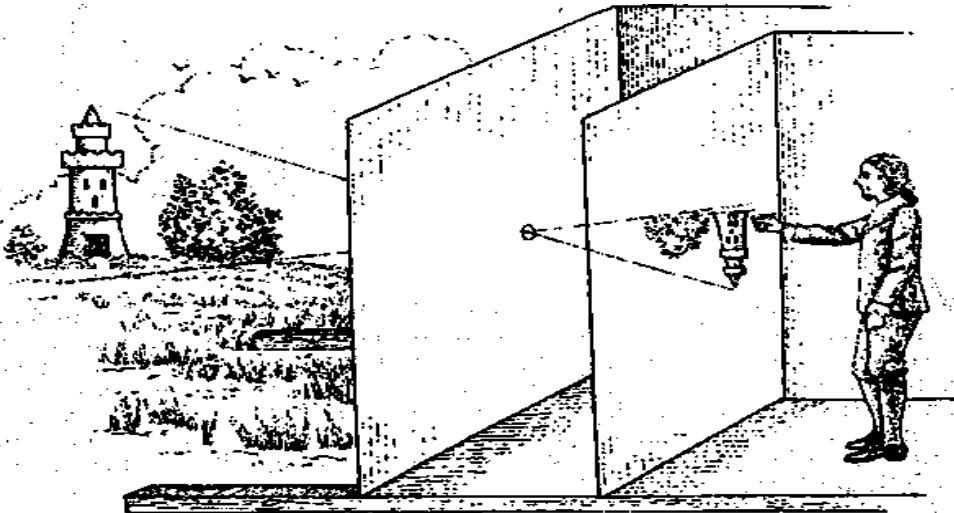
“the eye has evolved independently between 40 and 60 times around the animal kingdom, “It seems that life, at least as we know it on this planet, is almost indecently eager to evolve eyes. There are only so many ways to make an eye, and life as we know it may well have found them all

Origins of Imaging

illum in tabula per radios Solis, quām in cōelo contin-
git: hoc est, si in cōelo superior pars deliquiū patiatur, in
radiis apparebit inferior deficere, vt ratio exigit optica.

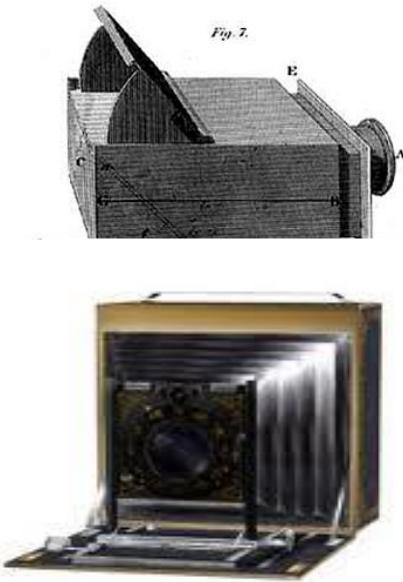


Sic nos exactē Anno .1544. Louanii eclipsim Solis
obseruauimus, inuenimusq; deficere paulo plus q; dex-



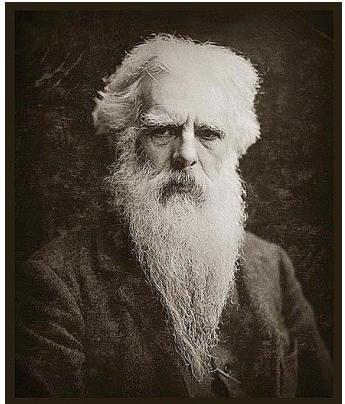
- Invention of the camera obscura in 1544 (L. Da Vinci)
- The mother of all cameras
- A more realistic and fast depiction of reality

Origins of Imaging

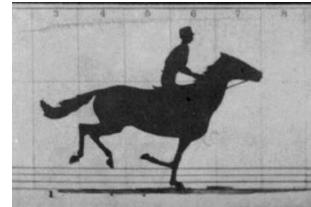
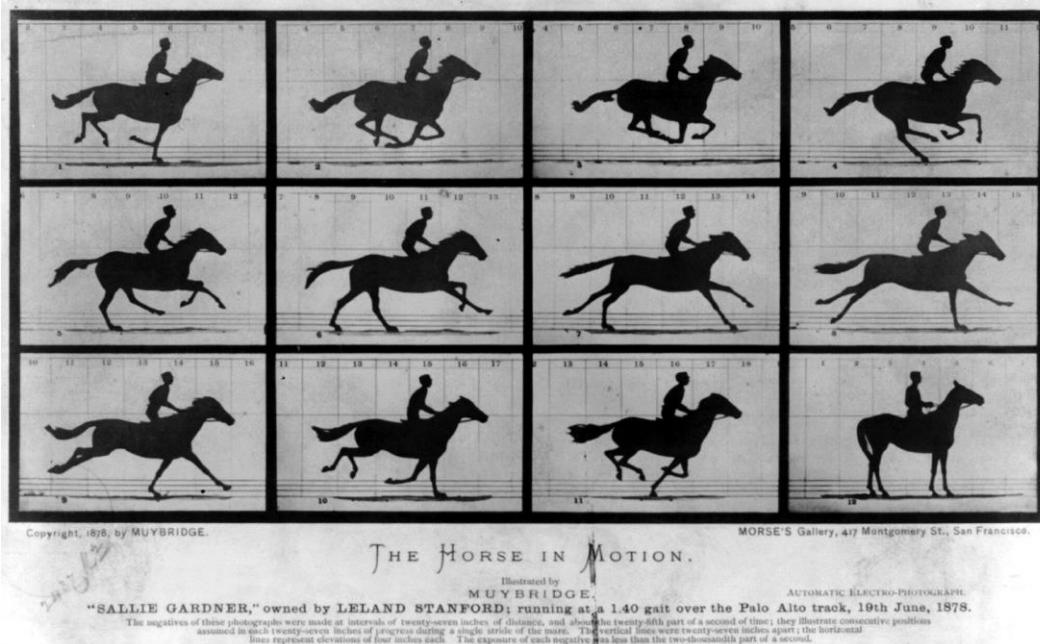


- Increasing painters profits: painting faster
- Evolution from portable models for travellers to current digital cameras
- Evolving from canvas, to paper, to glass, to celluloid, to pixels

Motion Picture: origins of video



Eadweard Muybridge
(1830-1904)



- Early work in motion-picture projection
- known for his pioneering work on animal locomotion in 1877 and 1878, which used multiple cameras to capture motion in stop-motion photographs

Why are images bad?

- When dealing with change or motion the universally accepted paradigm of visual acquisition becomes **FUNDAMENTALLY FLAWED!**
- **FIXED frame rate for all pixels is always WRONG**
- **NO relation** between frame rate and **scene dynamics**

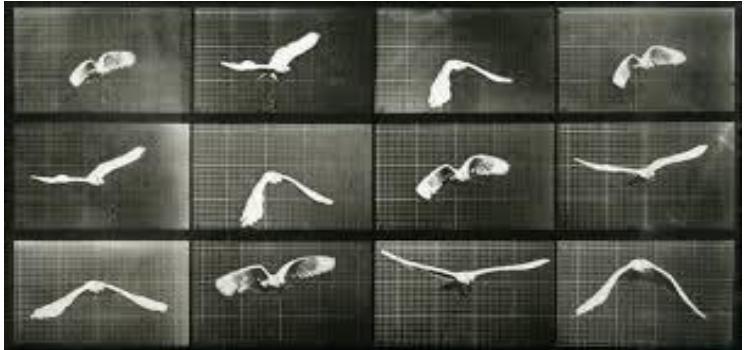


Under-sampling

- Motion blur
- Displacement between frames

Over-sampling

Why are images bad?



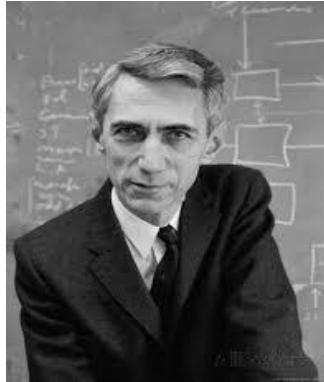
Good

1. Small pixels
2. Large fill factor, few wasted photons
3. Compatible with 50 years of machine vision theory

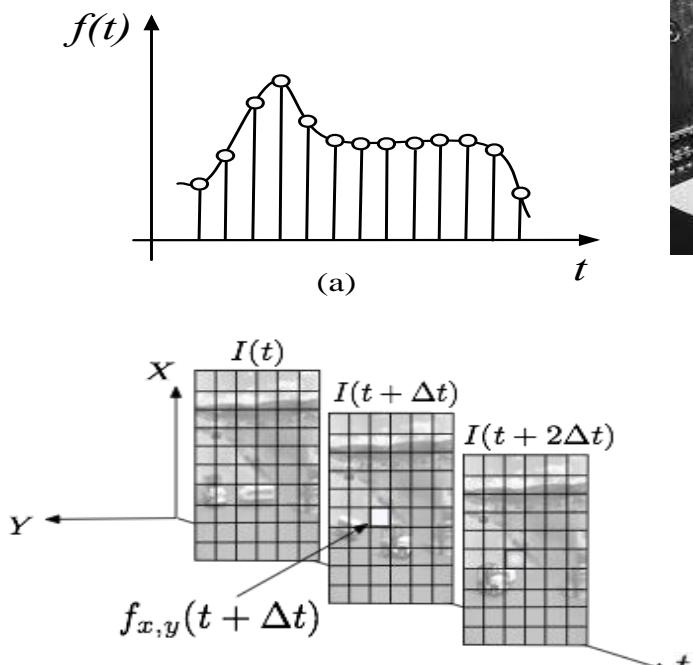
Bad

1. Pixels produce redundant output
2. Uniform integration time limits dynamic range
3. All pixels must be sampled at the same time rate sufficient for dynamics problems

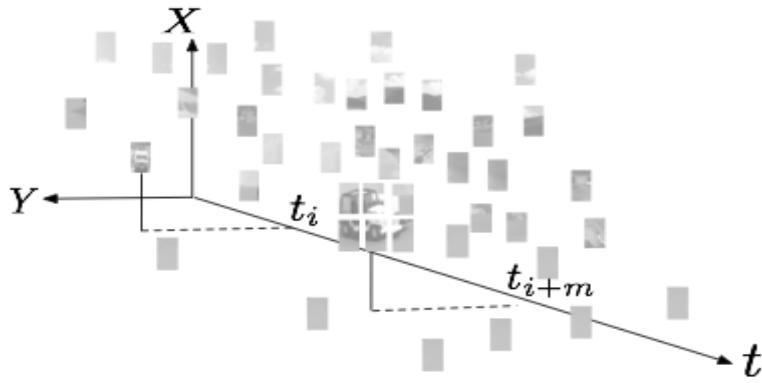
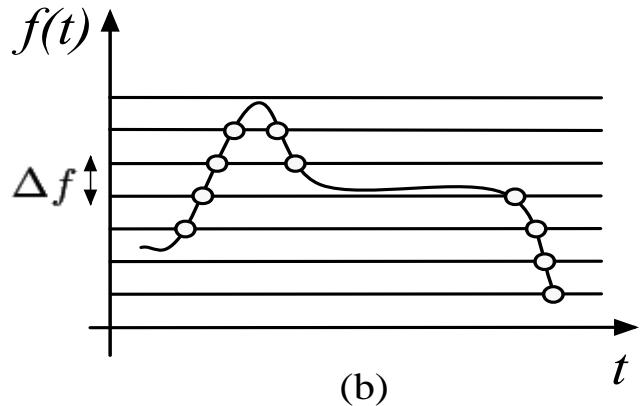
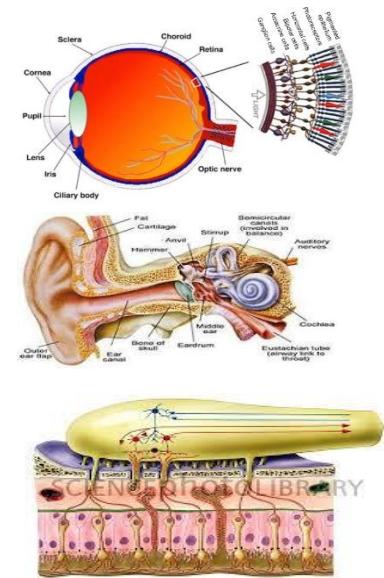
Beyond the Shannon-Nyquist Acquisition principle



Claude
Shannon
(1916-2001)



Neural acquisition



- Amplitude sampling
- Information is sent when it happens
- When nothing happens, nothing is sent or processed
- Sparse information coding
- Time is the most valuable information

Castle Metaphor: digital acquisition



+



+



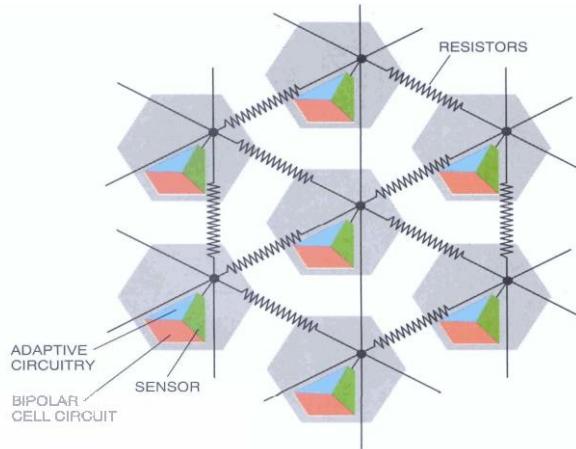
- Ensuring the safety of the castle
- Each sentinel has to wait for a drum beat to send information

The Silicon Retina



Misha
Mahowald
1963-1996

Carver
Mead

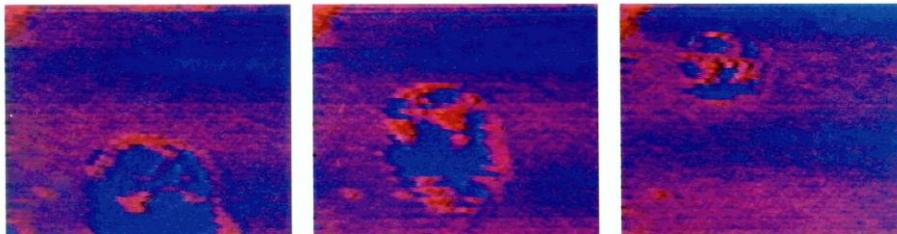


SCIENTIFIC AMERICAN

*Exploring the genetic heritage of racehorses.
Can anyone explain high-temperature superconductivity?
The impact of Kuwait's burning oil wells.*

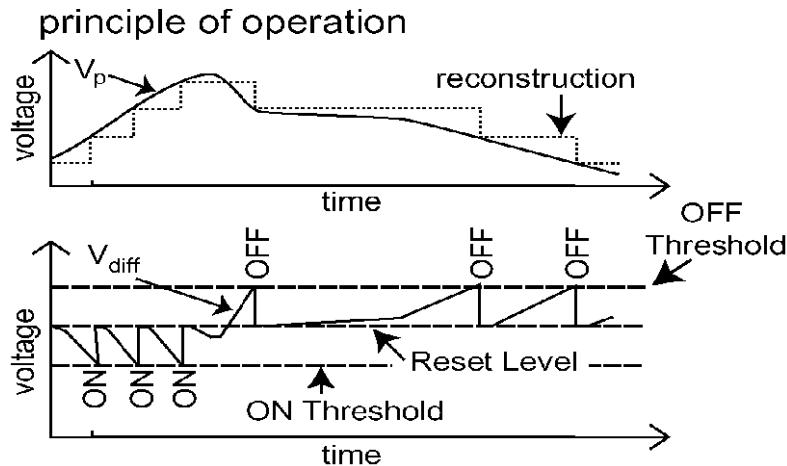
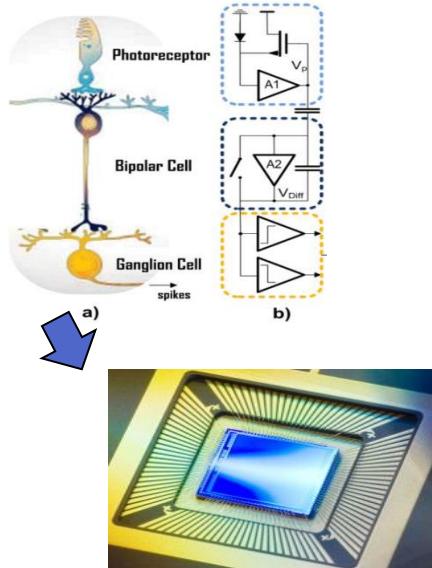


Silicon sees a cat. This retina on a chip mimics the functions of cells in the human eye.

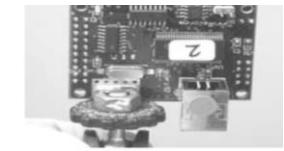


SOCCEBALL in motion shows how the delayed response of the horizontal cell network affects the retina's perception. The ball leaves behind a trail of excitation; bright where the dark spots have just passed; dark where bright parts have been.

Dynamic Vision Sensor



) Vision sensor USE



interface

P. Lichtsteiner, C. Posch, T. Delbrück,

- Each event represents a quantized change in log intensity (brightness or relative intensity change or “temporal contrast”)
- Artificial retinas **do not provide images**,, Data driven, energy efficient

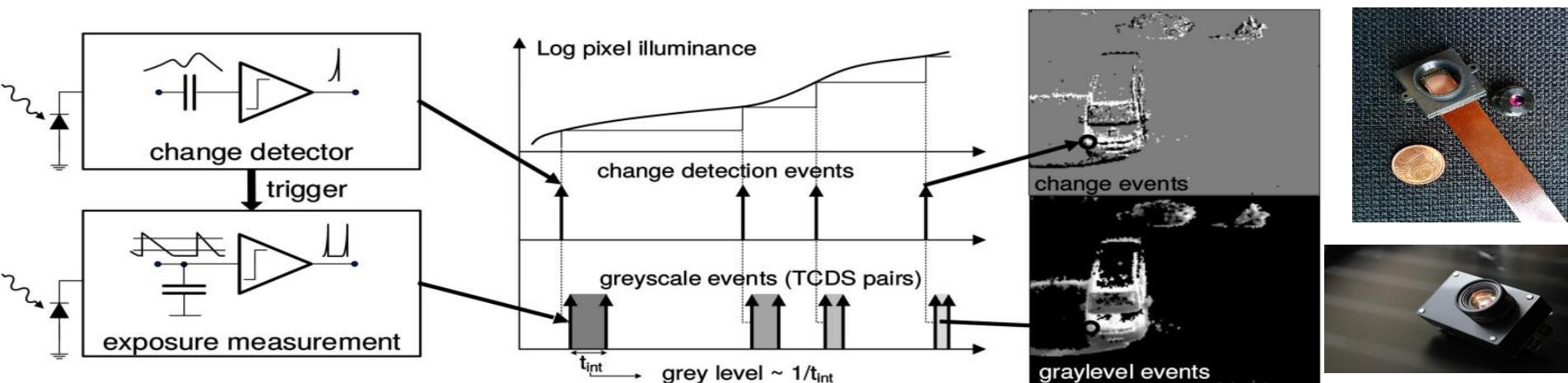
DAVIS(Dynamic and Active-pixel Vision Sensor)



- **Hybrid solution:** Frame+Events
- Low temporal resolution,
- Need of a **high bandwidth**
- Grey levels (absolute measurements) : low dynamic range (50db)

C. Brandli, M. Yang, S.-C. Liu, V. Villeneuva, and T. Delbruck,

ATIS (Asynchronous Time-based Image)



- Each event triggers an **intensity measurement encoding in time**
- **143 dB** Dynamic Range Frame-Free PWM
- Image Sensor With Lossless Pixel-Level Video
- Compression and Time-Domain CDS

C.Posch, Daniel
Matolin, and
Rainer
Wohlgenannt

Why Event Based sensors?

ATIS
vs.
Conventional Camera



Sample Data

Conventional Camera

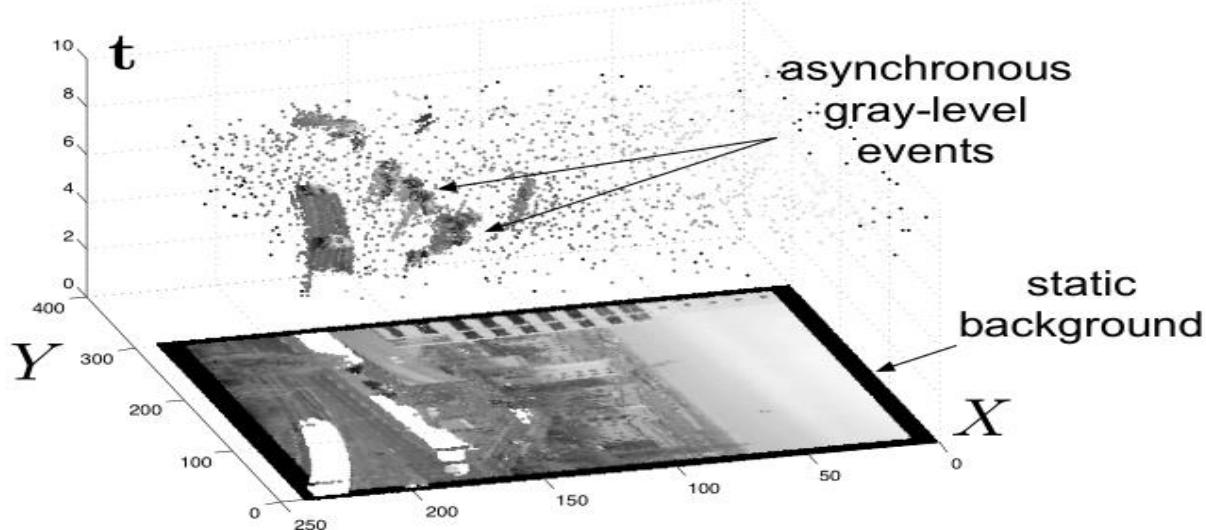


Event based Camera

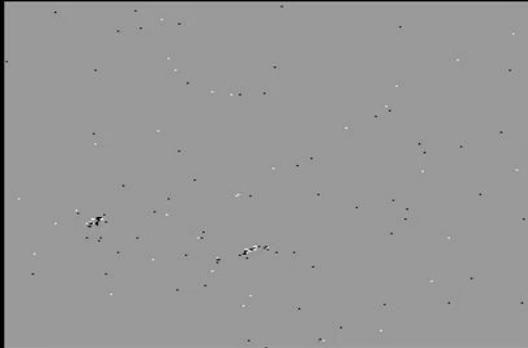


- Data driven: **only moving edges** produce data
- Temporal edges, precisely timed

Data Space of Events



Asynchronous Time-based Image Sensor



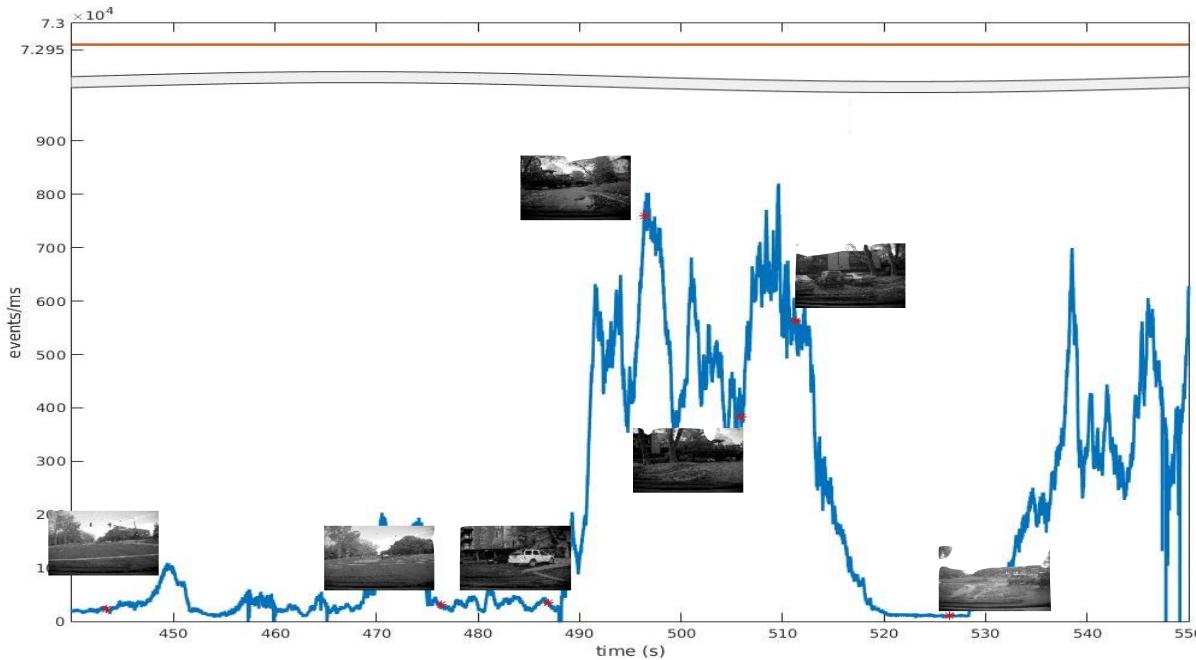
APS

ATIS

Asynchronous Time-based Image Sensor

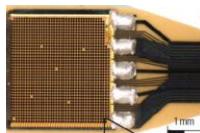
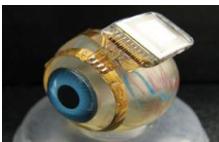
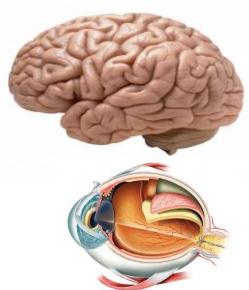


Low data bandwidth



- Comparing amount of data generated by a conventional camera running at 30Hz vs an event based camera (1MHz)
- 1M times faster for a lower data load (.1-20% less)

Neuromorphic engineering



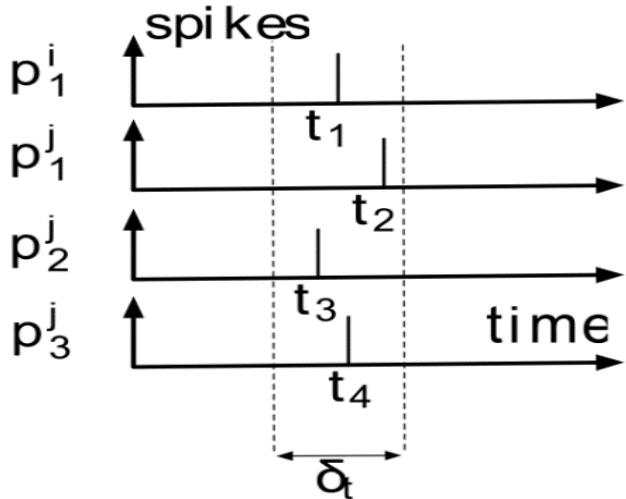
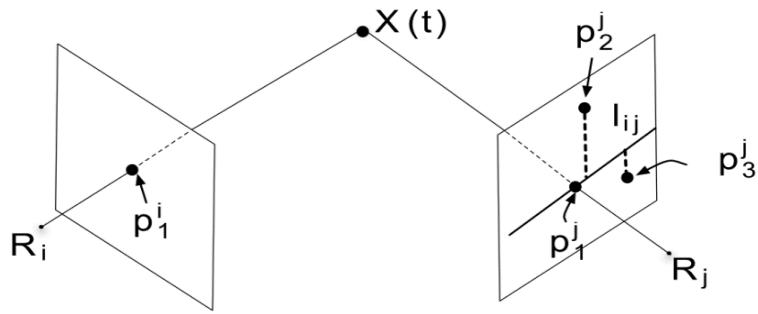
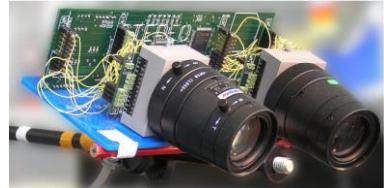
Physiology

- Models
- Hardware

- Prosthetics
- Robotics
- Computation

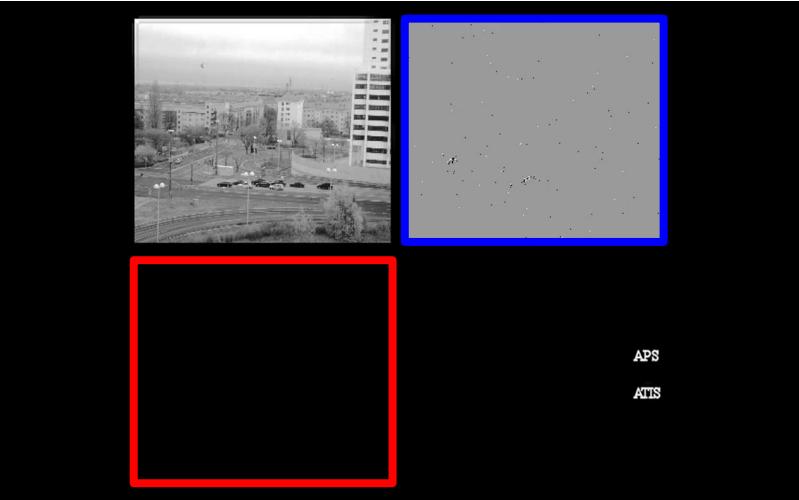
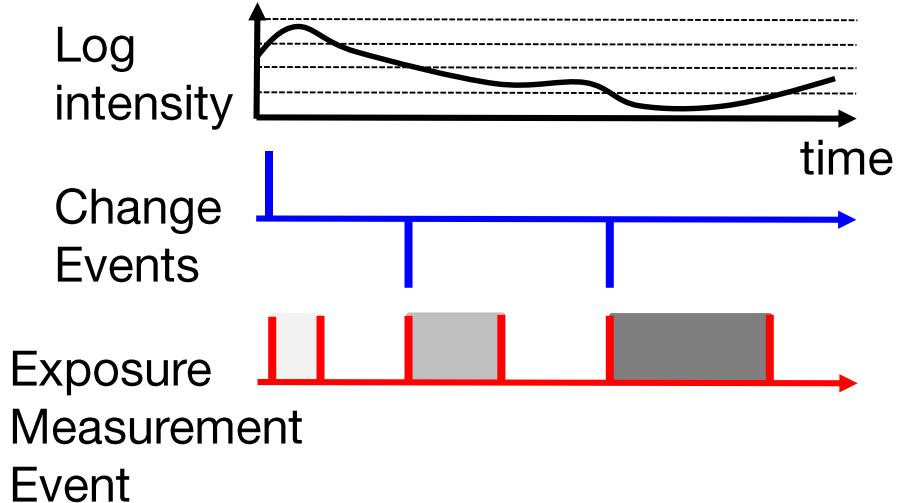
- Makes of **machine vision** a science!
- Develop new **bidirectional methodology** to understand the brain
- Merging **Computational and Biological Vision**

Applications: Stereovision



- Matching binocular events only using the time of arrival
- Two events arriving at the same time and fulfilling geometric constraints are matched

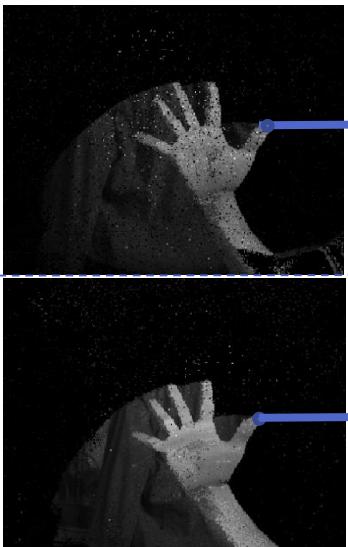
Applications: Stereovision



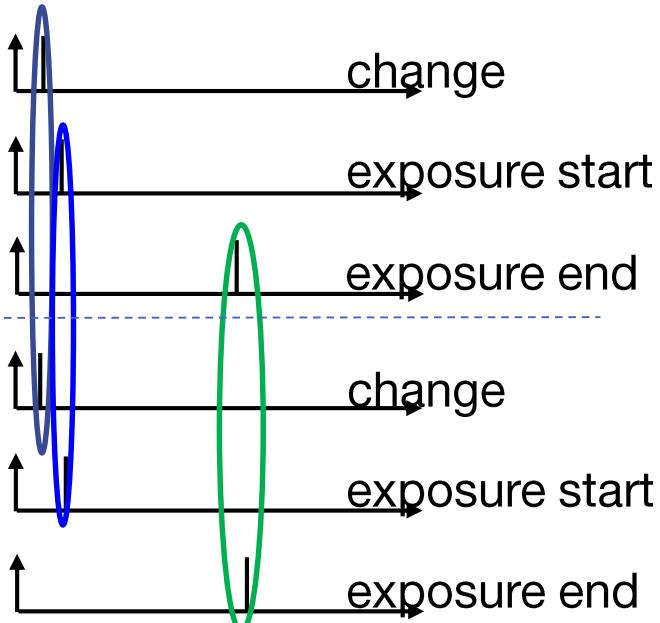
- Adding luminance constraints
- Luminance encoded in time implies 3 coincidences detection

Applications: Stereovision

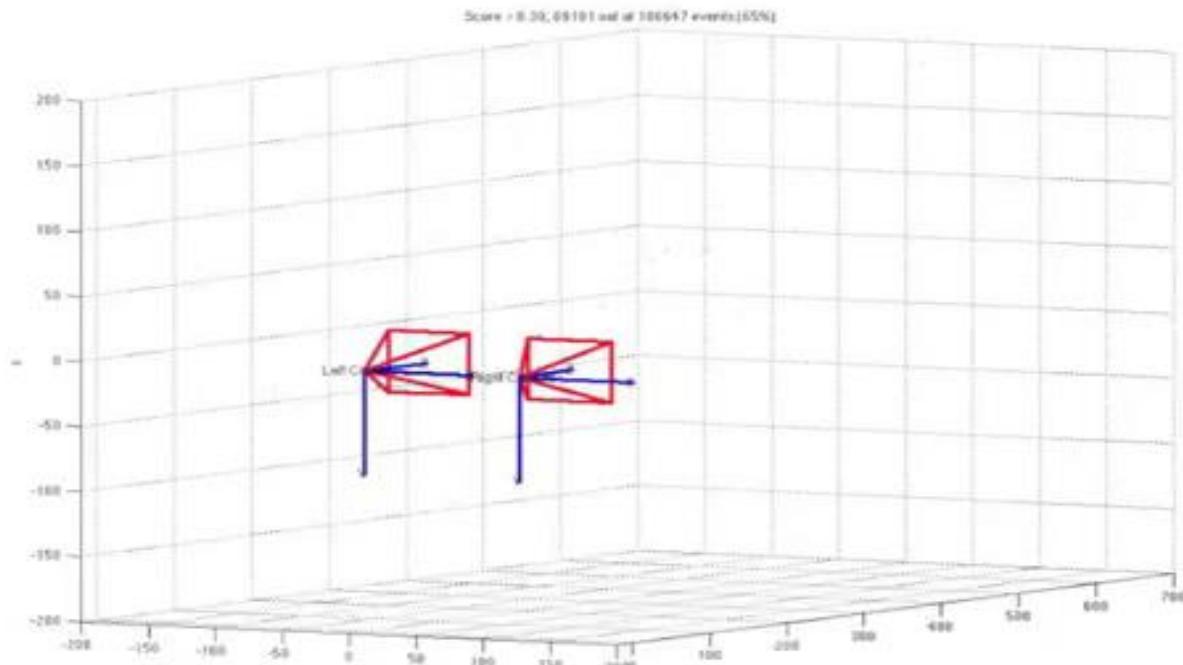
Time encoded gray-level images

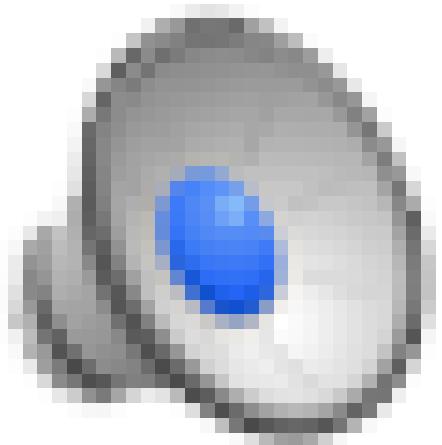


3 event coincidences

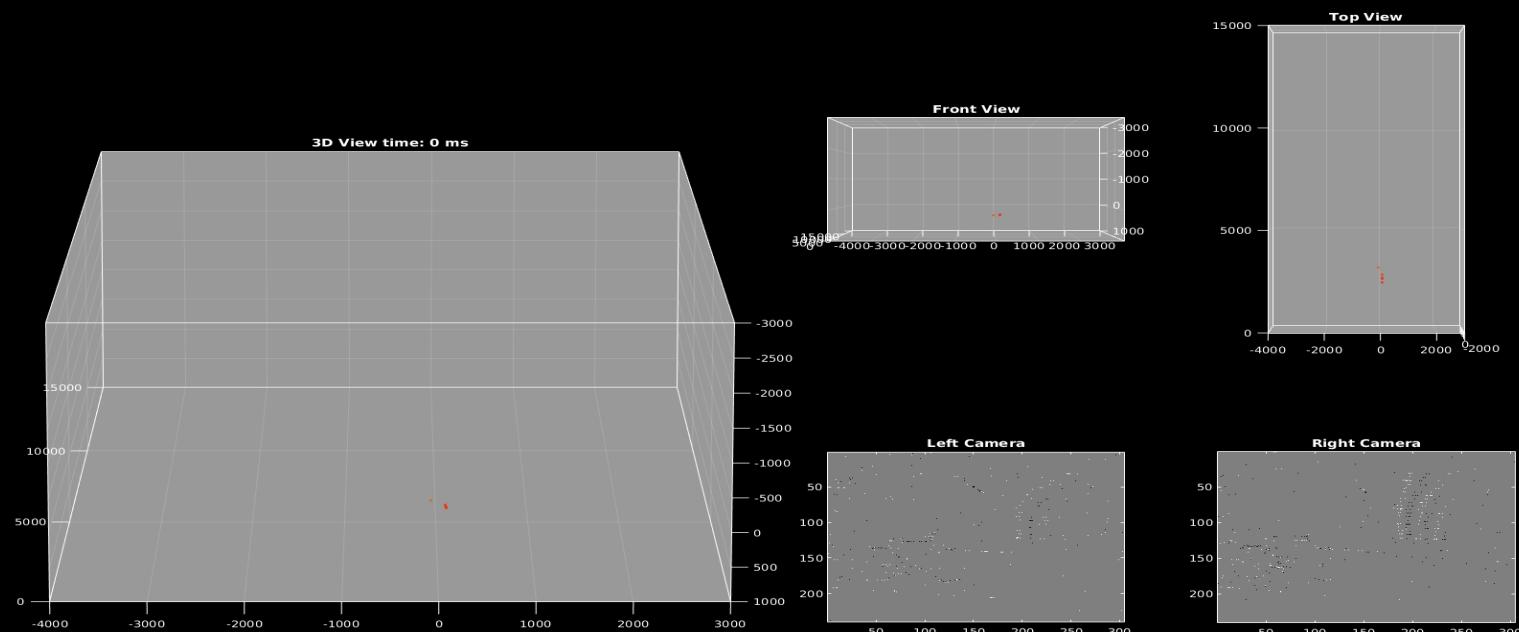


E_l = how well events match according to their luminance

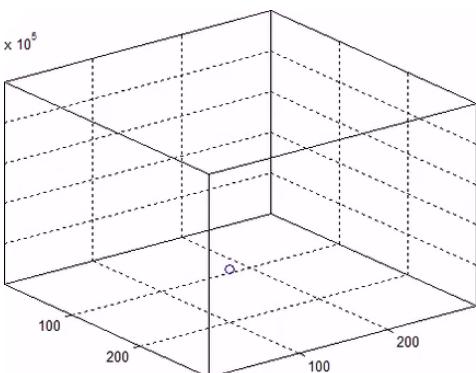
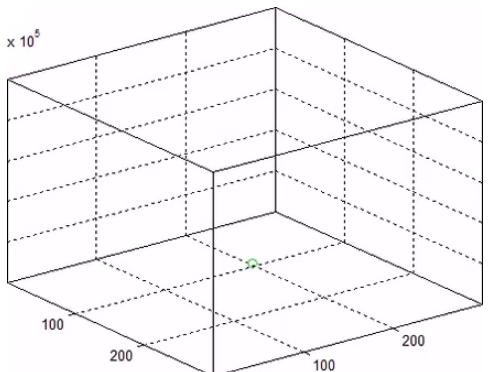
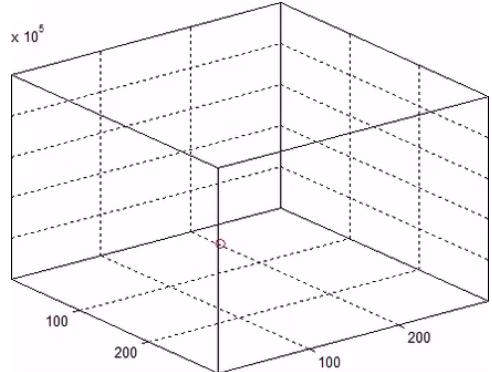
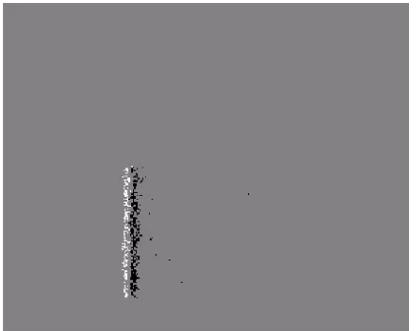
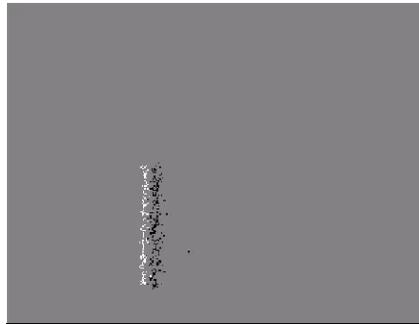




Reconstruction Examples



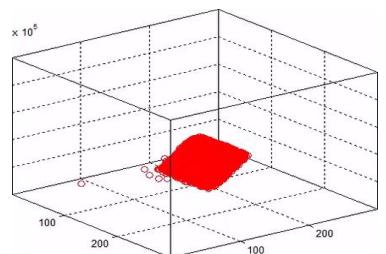
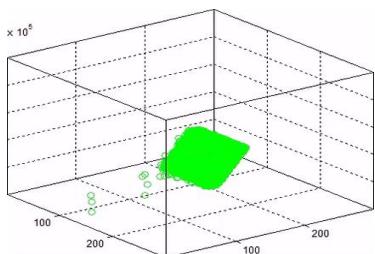
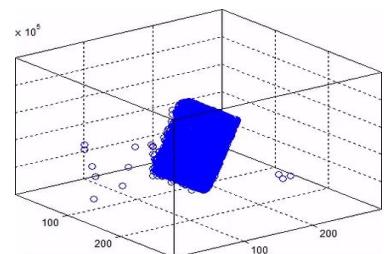
Motion estimation: optical flow



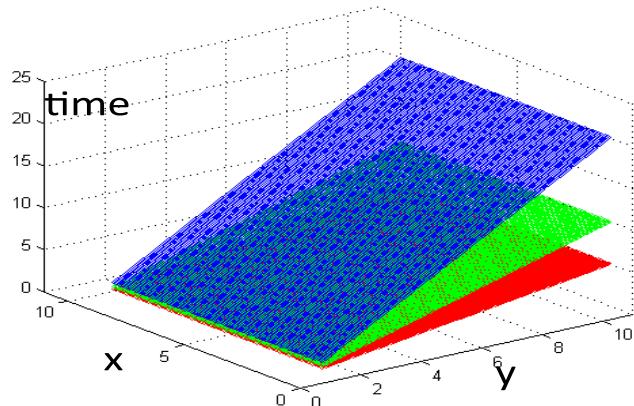
Motion estimation: optical flow



(a)



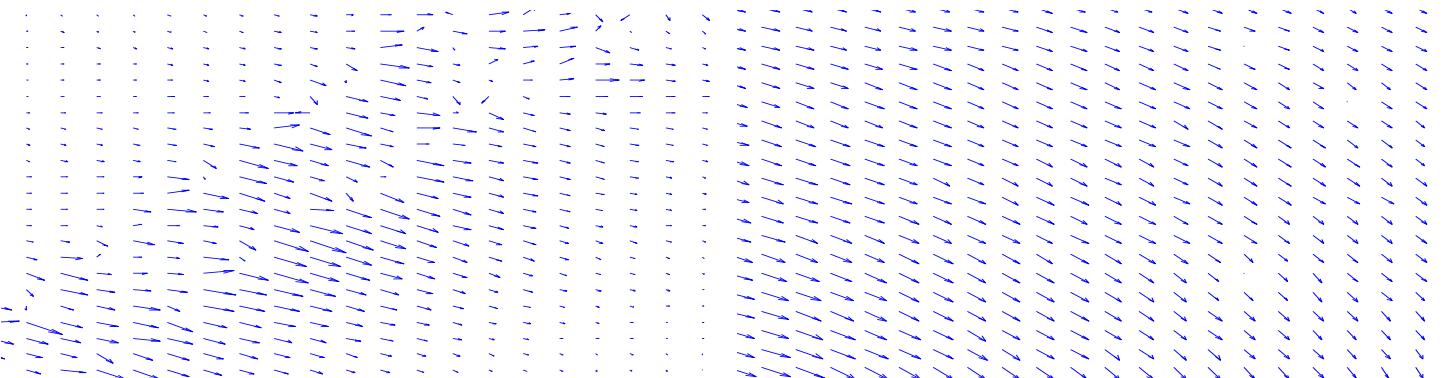
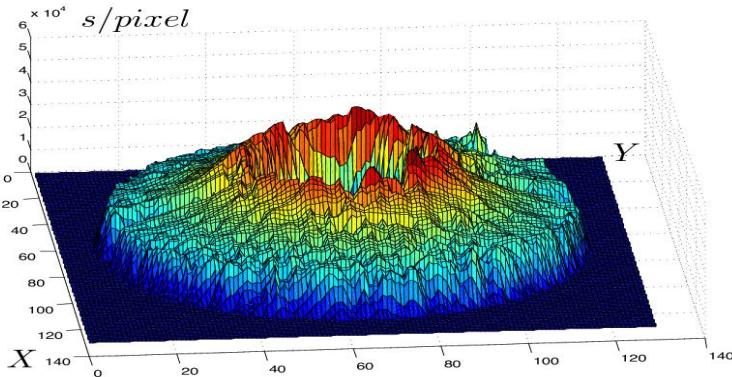
(b)



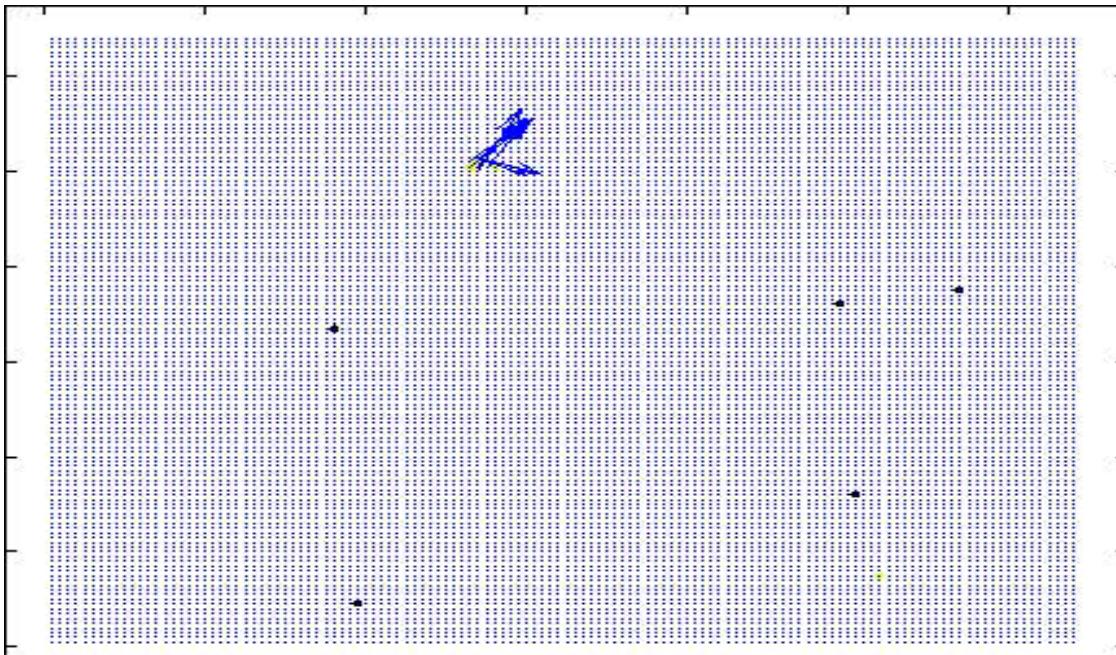
(c)

- High temporal resolution allows to generate smooth space-time surface
- The slope of the local surface contains the orientation and amplitude of the optical flow

Motion estimation: optical flow

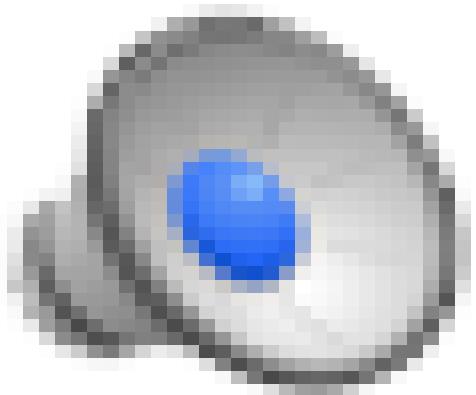


Motion estimation: optical flow



Single Pixel Tracking + Optical Flow

embedded
VISION
SUMMIT
2018



Carnegie
Mellon
University



Copyright © 2018 UPMC

Aperture robust optical flow



corrected visual flow



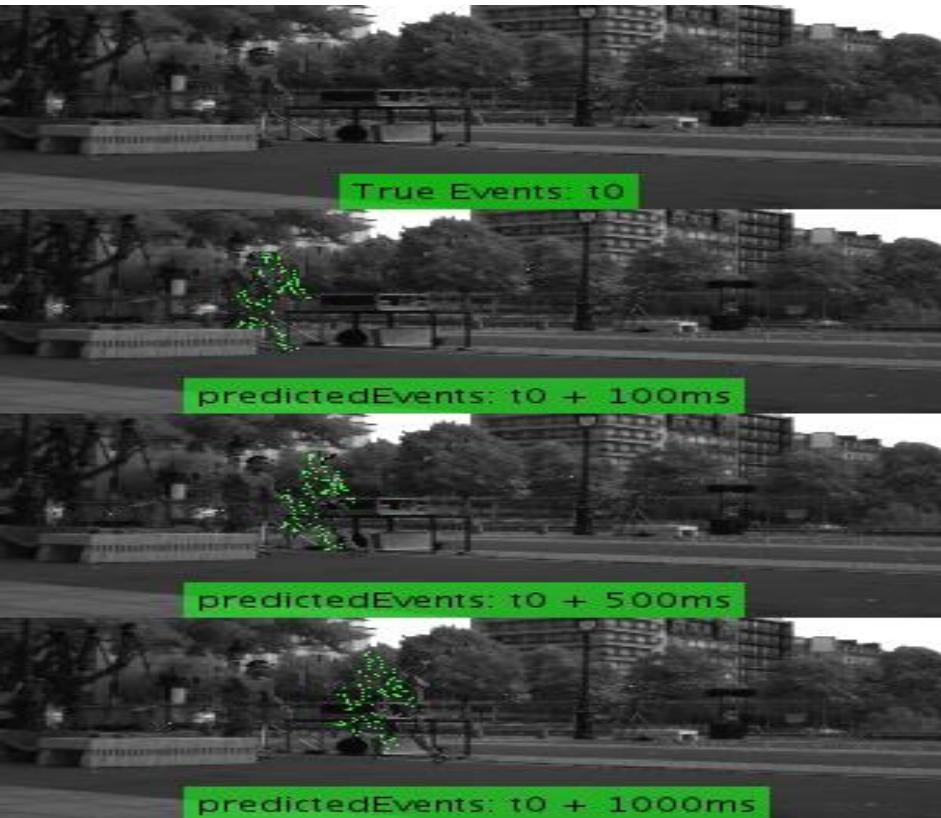
UPMC
University of Pittsburgh
Medical Center

Carnegie
Mellon
University

UPMC
Sorbonne UNIVERSITÉS

Copyright © 2018 UPMC

Prediction of future events



Tracking, Iterative Closest Point

embedded
VISION
SUMMIT
2018



UPMC
University of Pittsburgh
Medical Center

**Carnegie
Mellon
University**

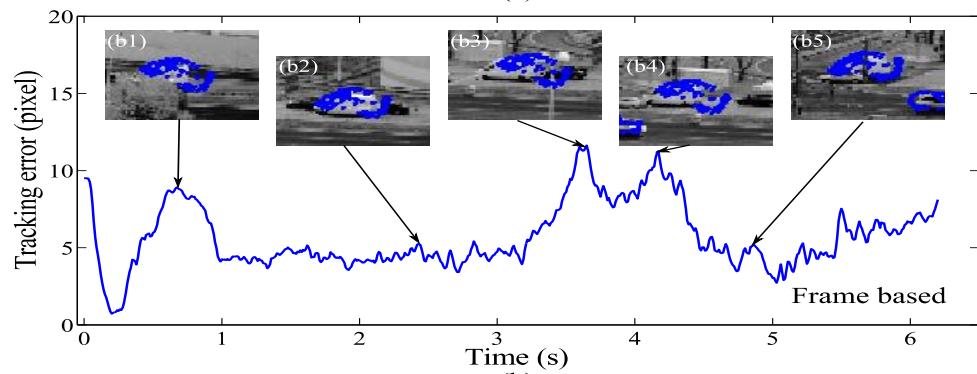
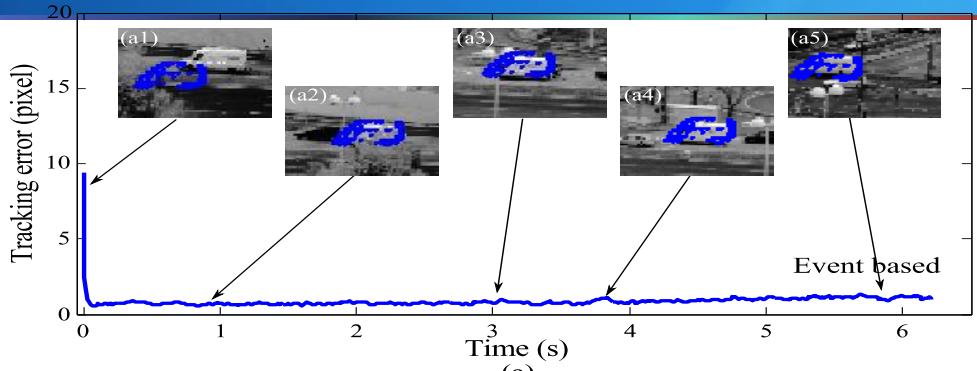
UPMC
Sorbonne UNIVERSITÉS

Copyright © 2018 UPMC

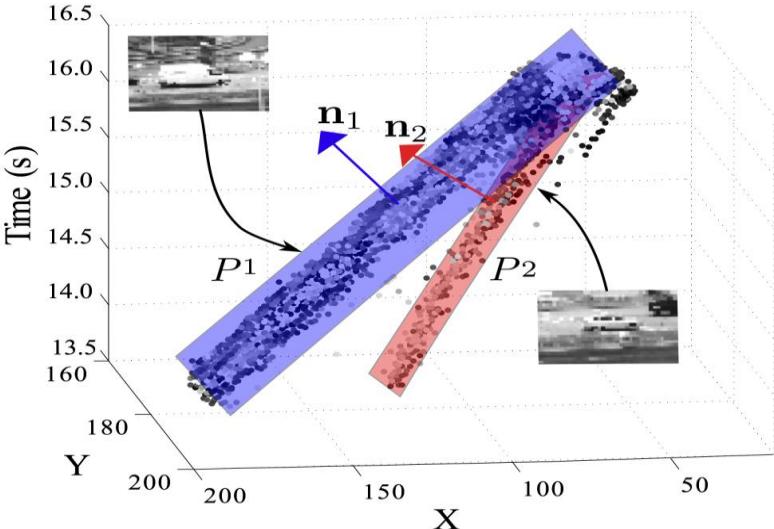
Tracking real-time outdoor scenes



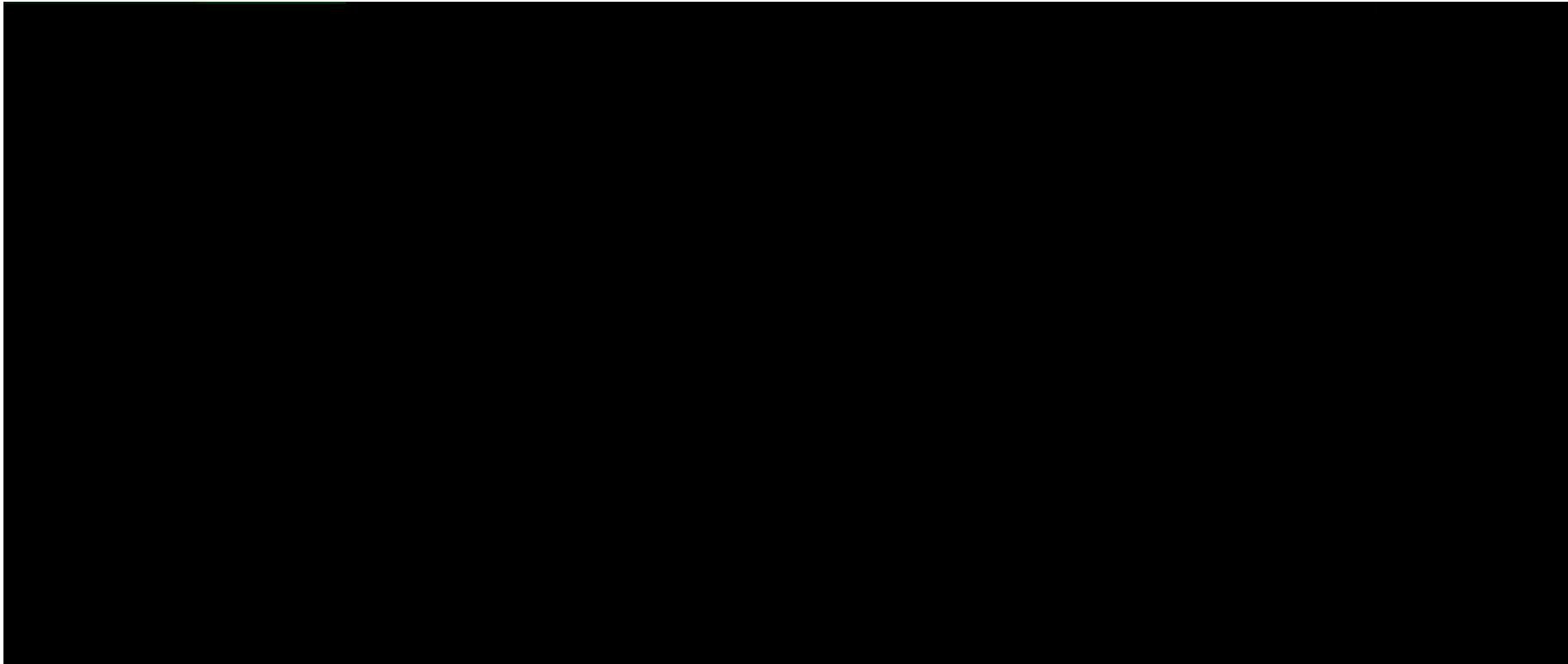
Outdoor vehicle tracking



- Space-time surfaces spanned by moving objects
- Space-time consistency



Event-based tracking with a moving camera



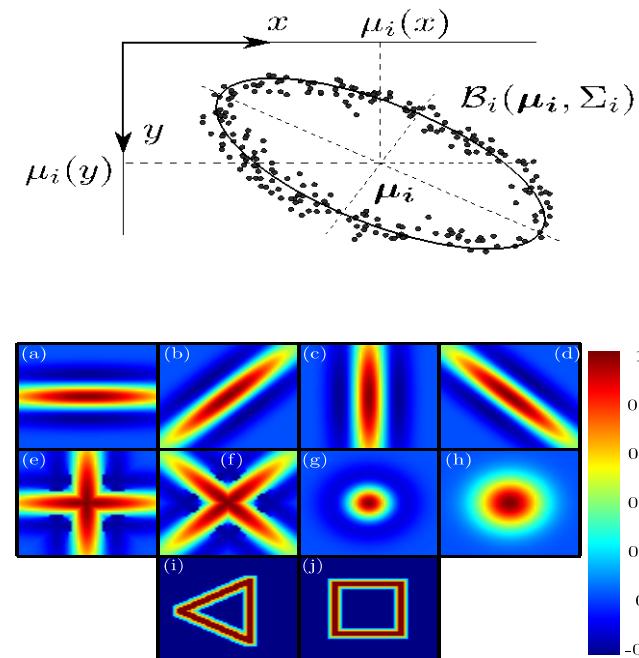
UPMC
University of Pittsburgh
Medical Center

Carnegie
Mellon
University

UPMC
SORBONNE UNIVERSITÉS

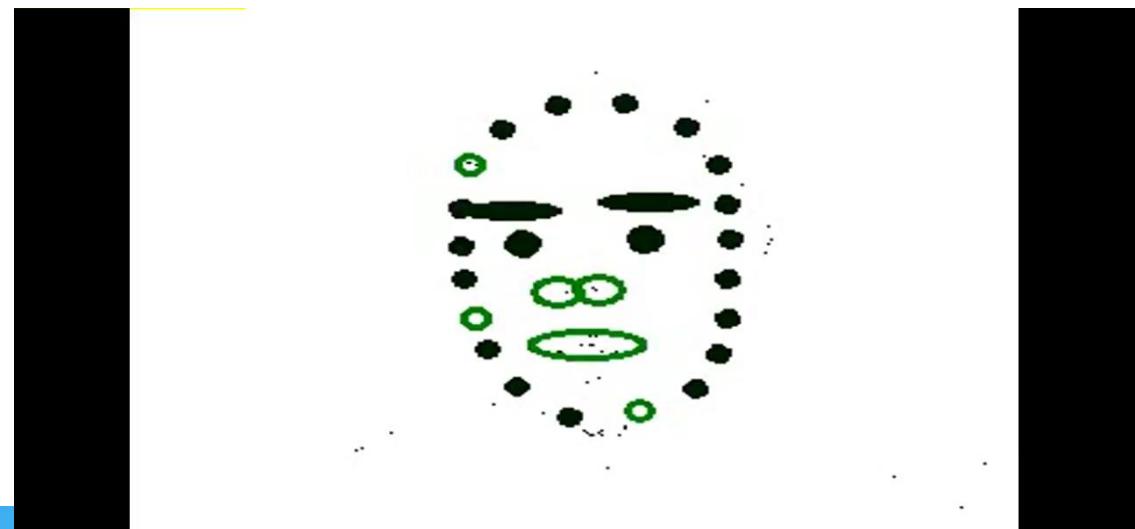
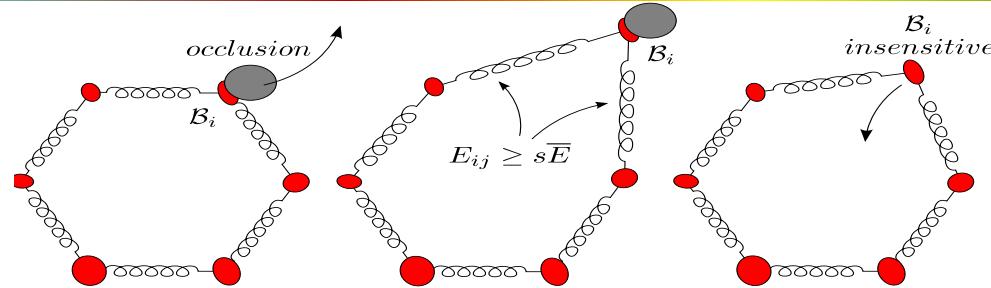
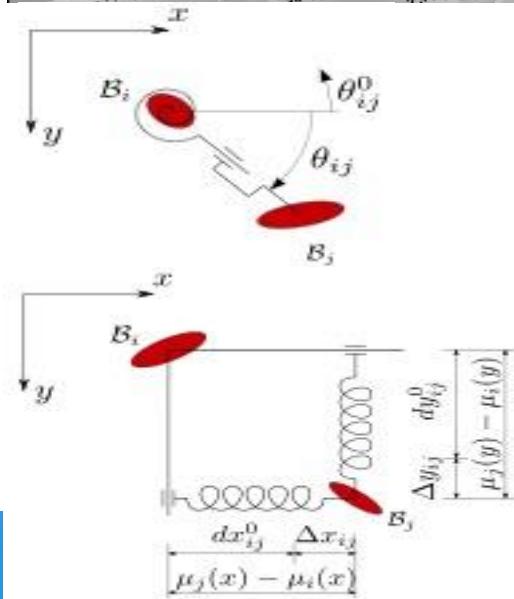
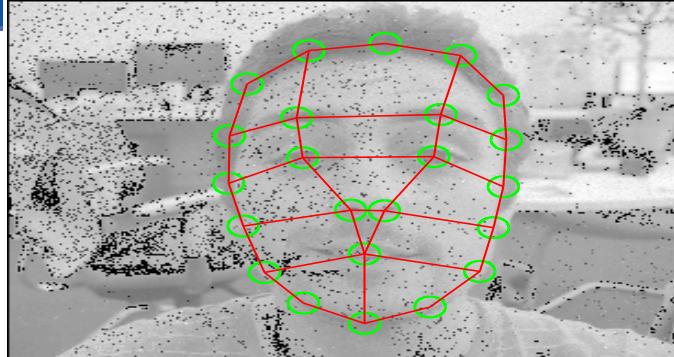
Copyright © 2018 UPMC

Multi-kernel High Speed Visual Features Tracking

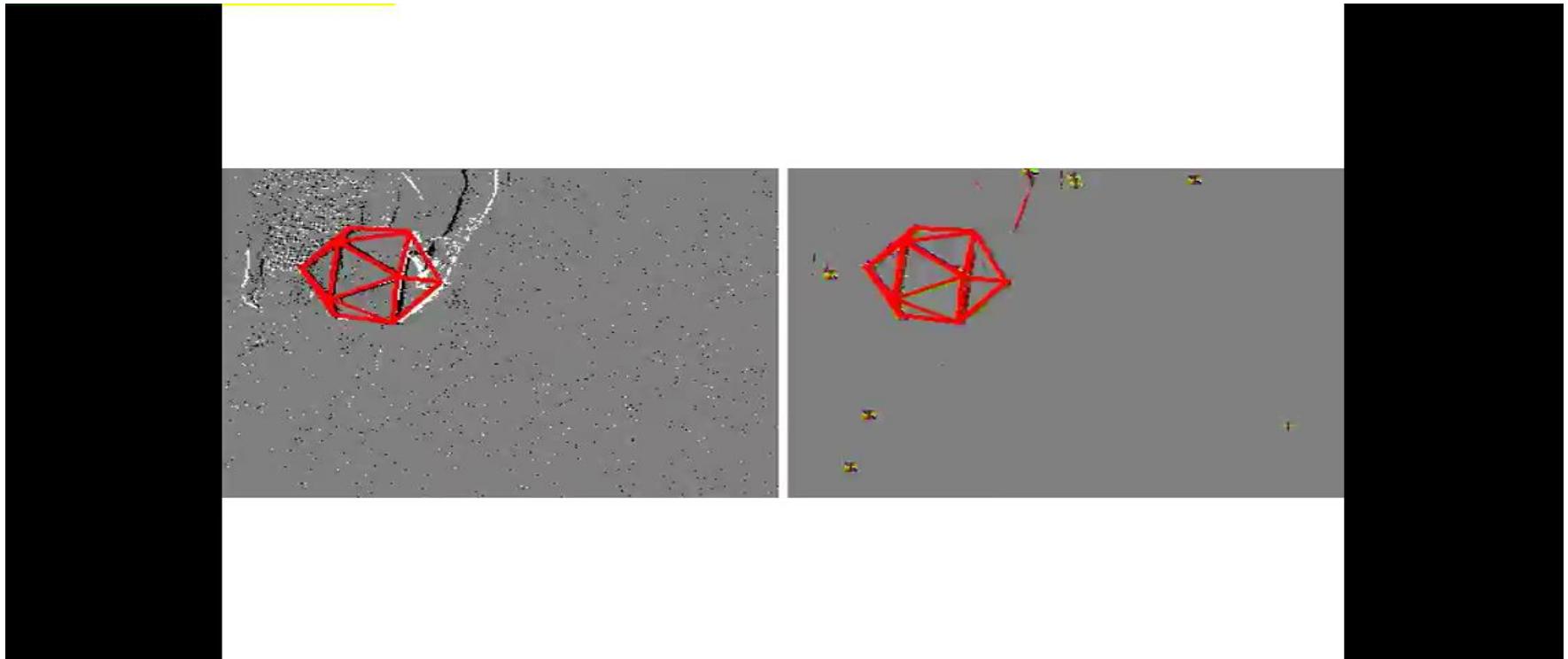


Asynchronous Event-based Trackers

Event-based Face Tracker

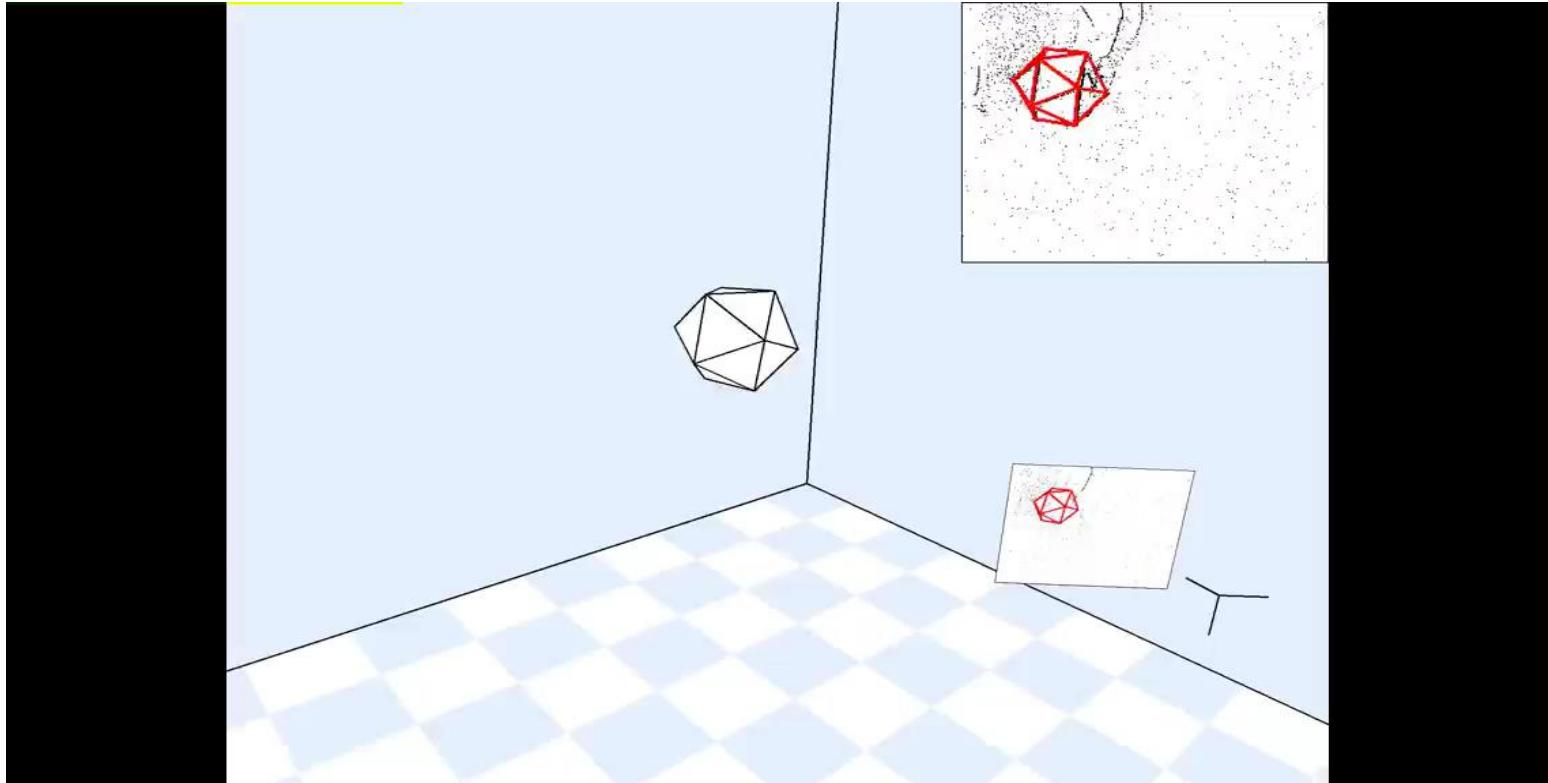


Model Based Pose Estimation



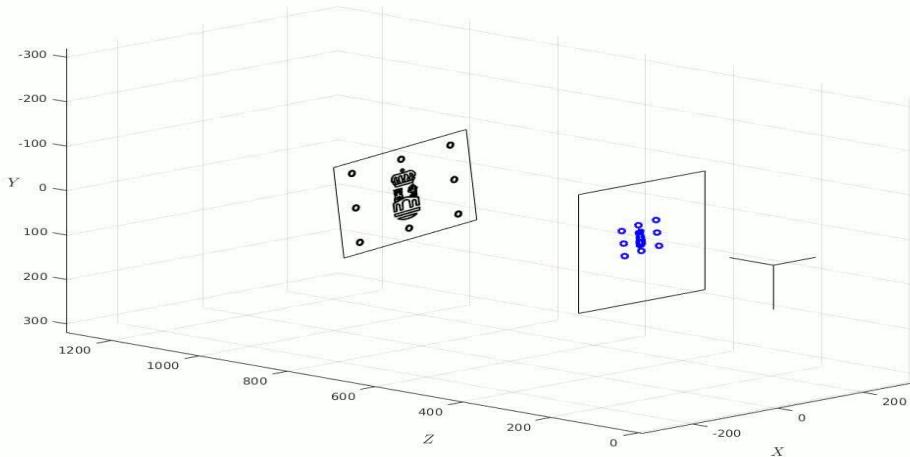
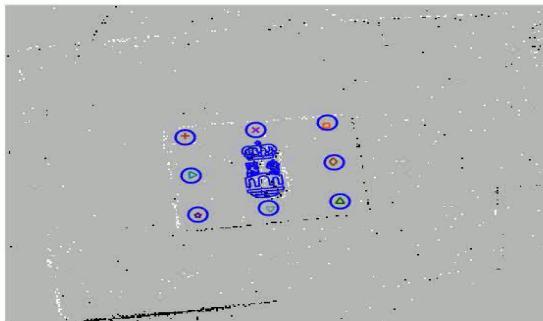
- Estimation of the pose of a camera given a known shape

Model Based Pose Estimation



Perspective-n-Point

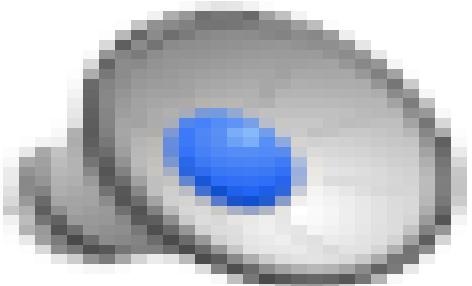
Event-Based Solution to the PnP Problem



- Estimation of the pose of a calibrated camera given a set of n 3D points in the world and their corresponding 2D projections in the image

Dense feature-less, Event Based Visual Odometry

embedded
VISION
SUMMIT
2018



UPMC
University of Pittsburgh
Medical Center

Carnegie
Mellon
University

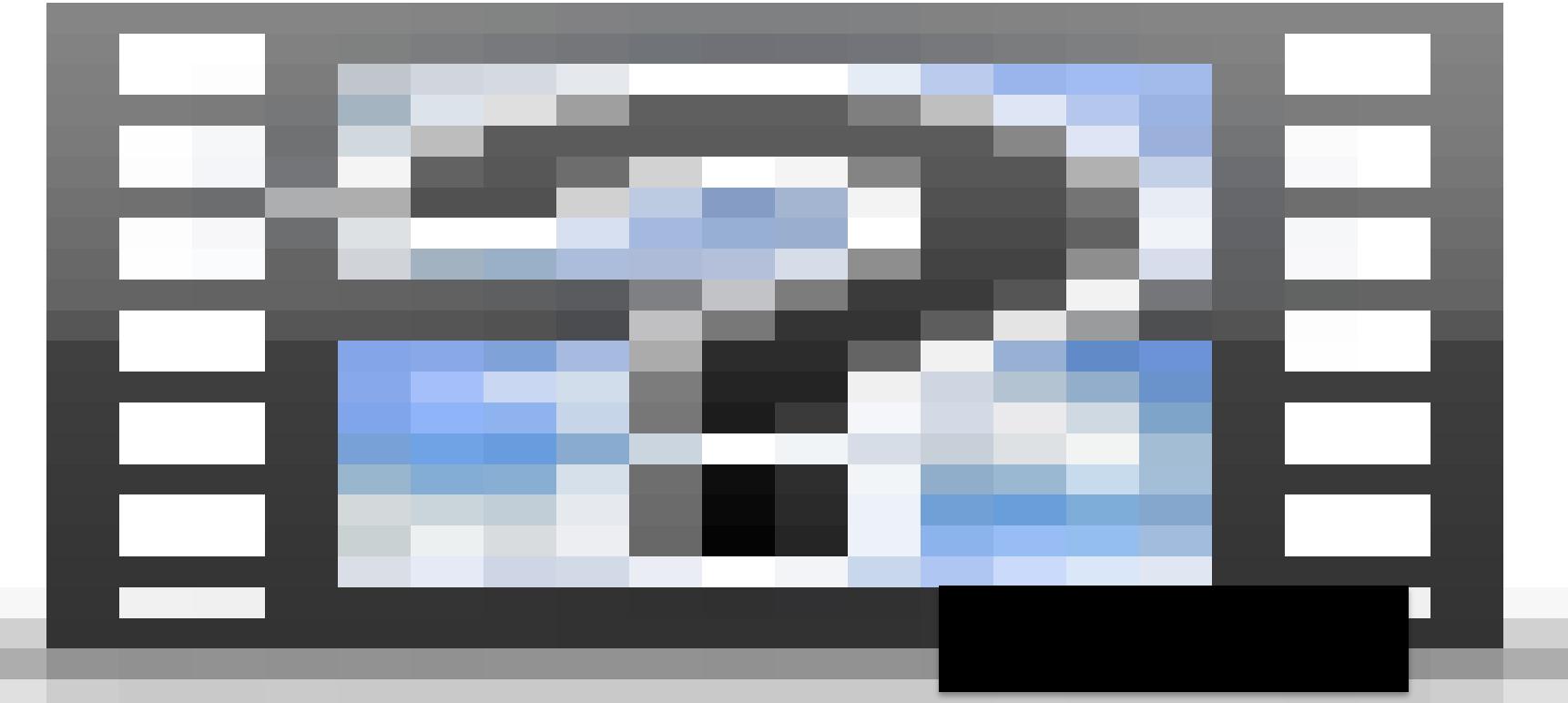
UPMC
Sorbonne Universités

Copyright © 2018 UPMC

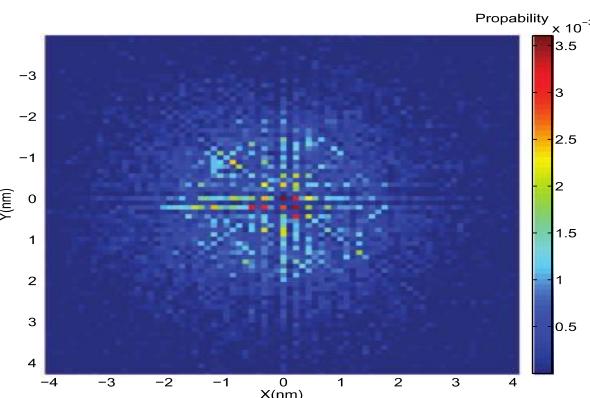
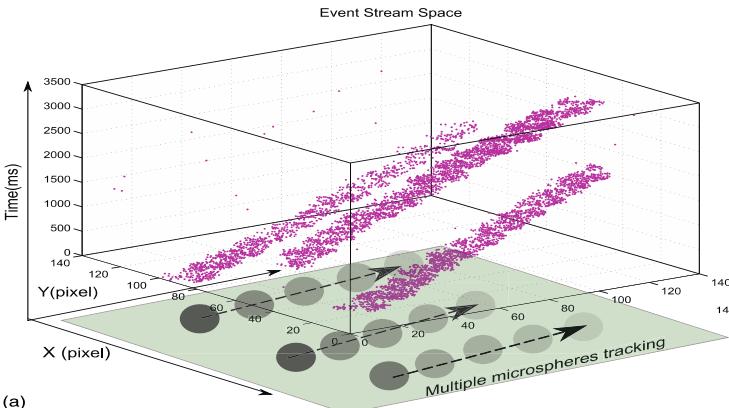
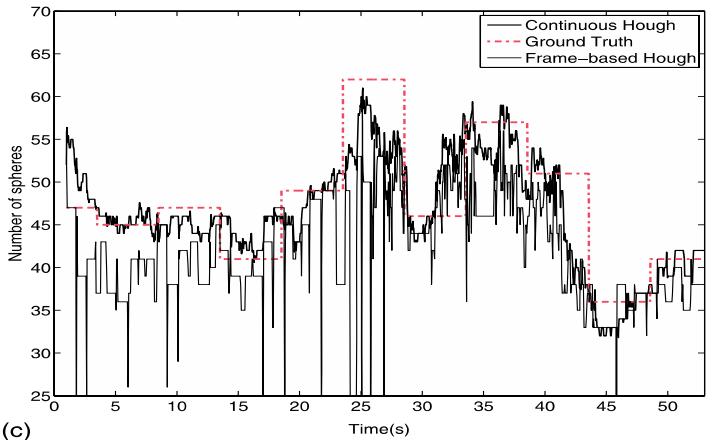
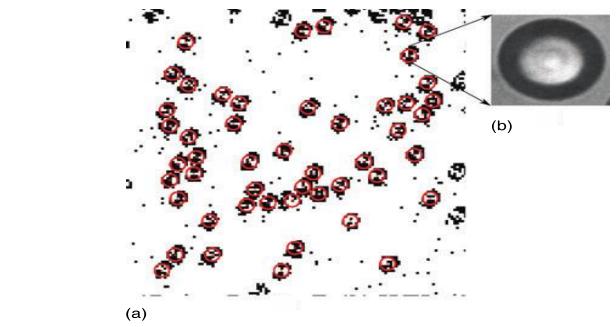
Dense feature-less, Event Based Visual Odometry



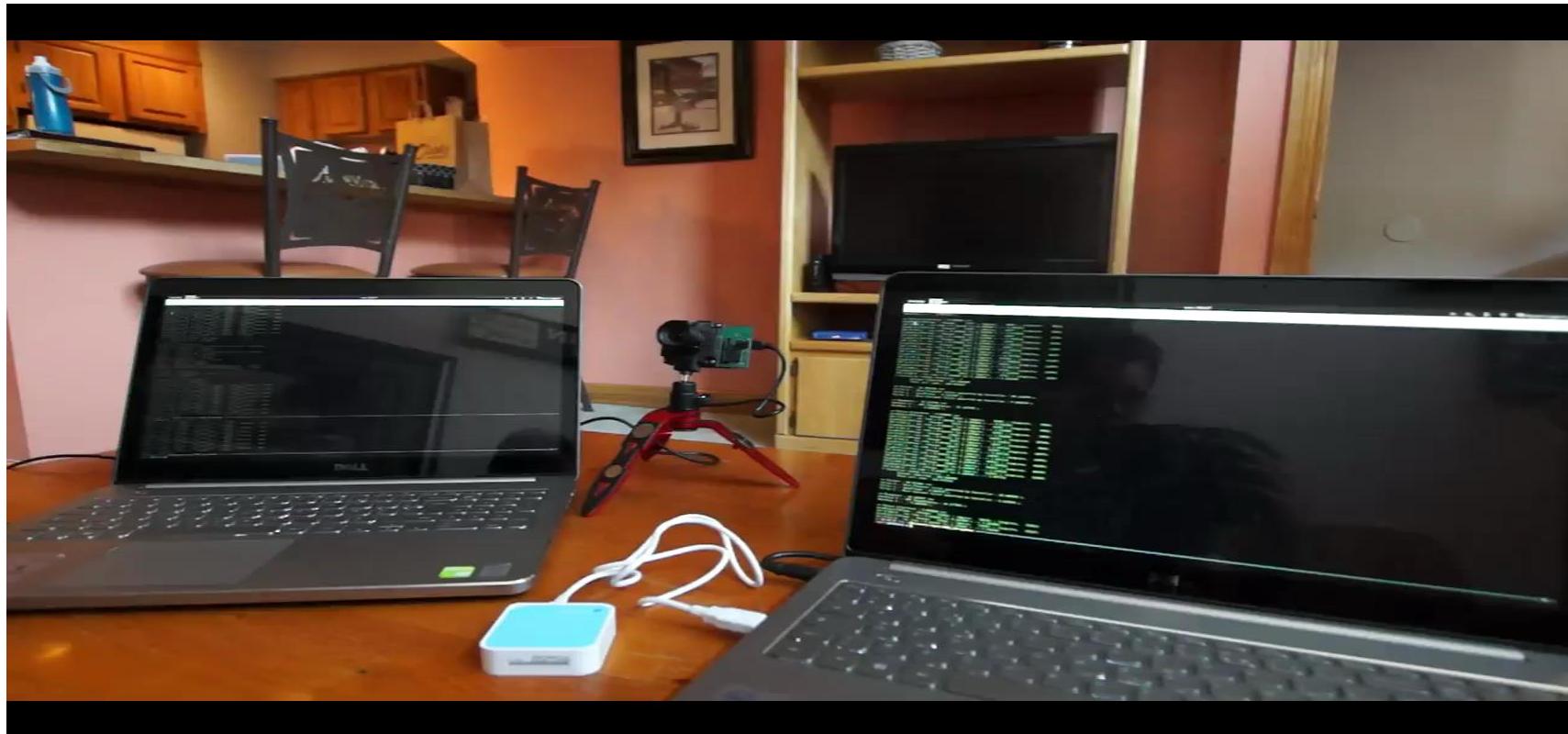
Point Based Event Driven SLAM



Micro-particule tracking



Extra low bandwidth low power streaming



What is NOT event based Machine Vision



(a) Event Data

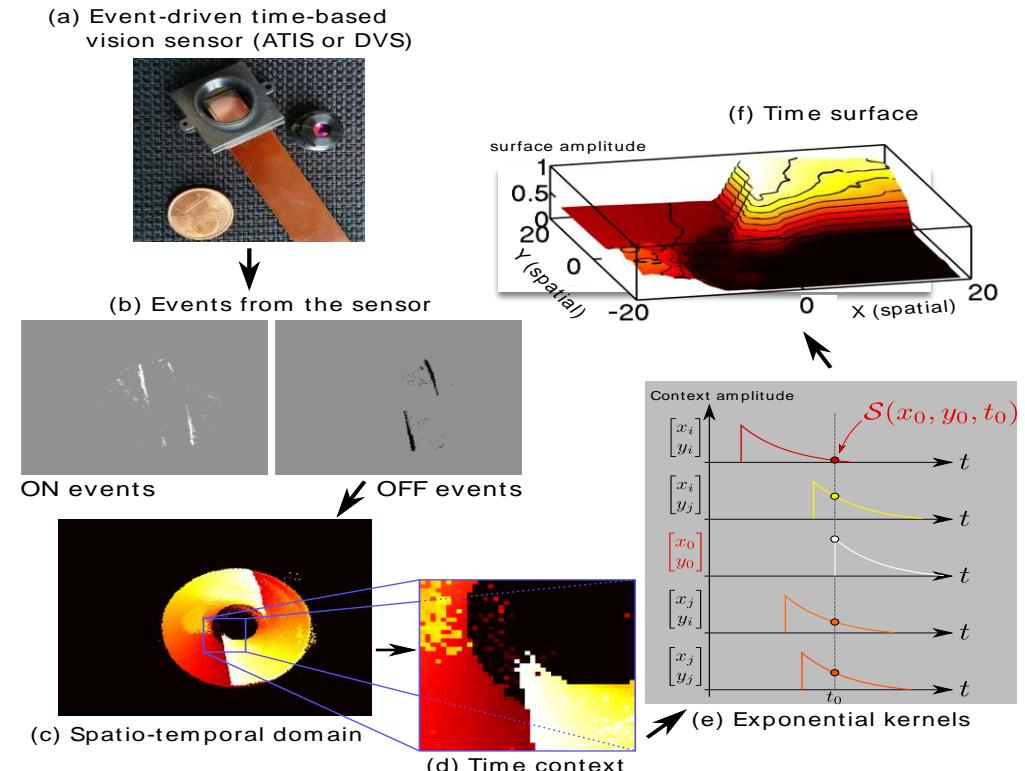


(b) Image Reconstruction

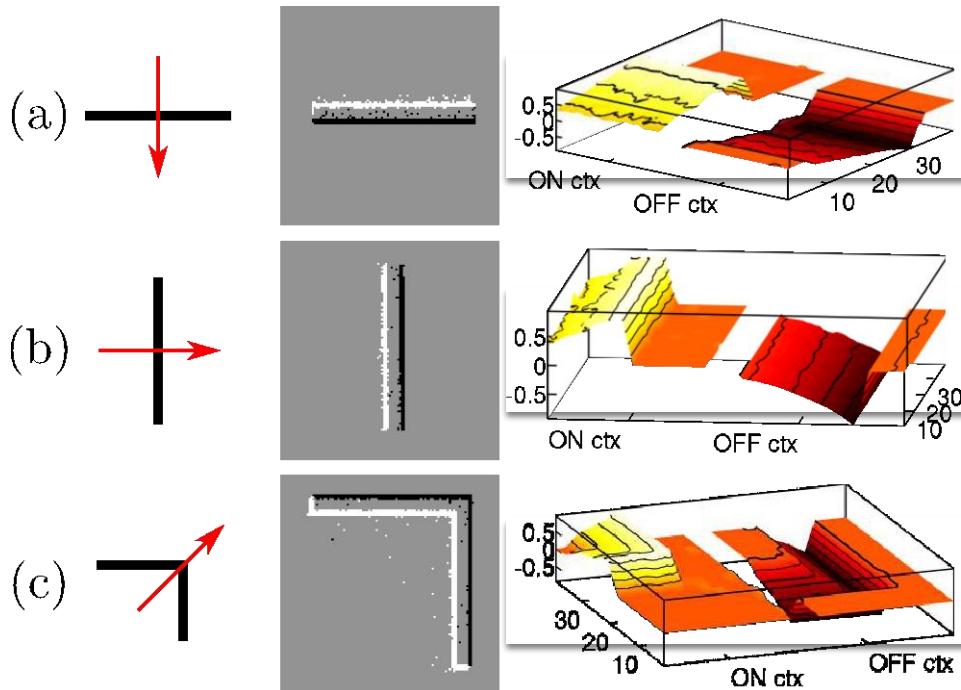
Do not generate Images from Events

- Not event based,
- Useless approach, GPU use, to generate 100-200Hz bad quality images
- Fake SLAM, using binary images....

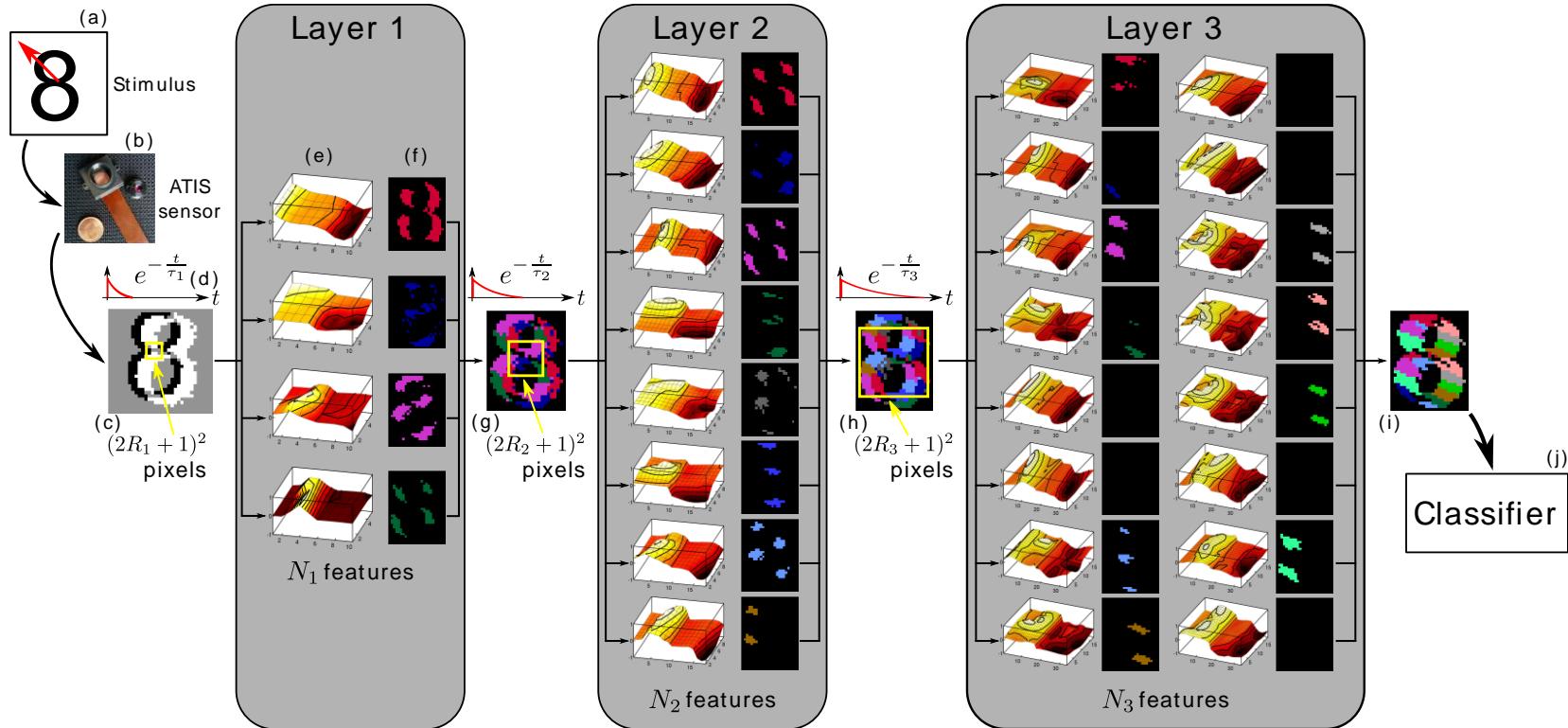
Dynamic Machine Learning



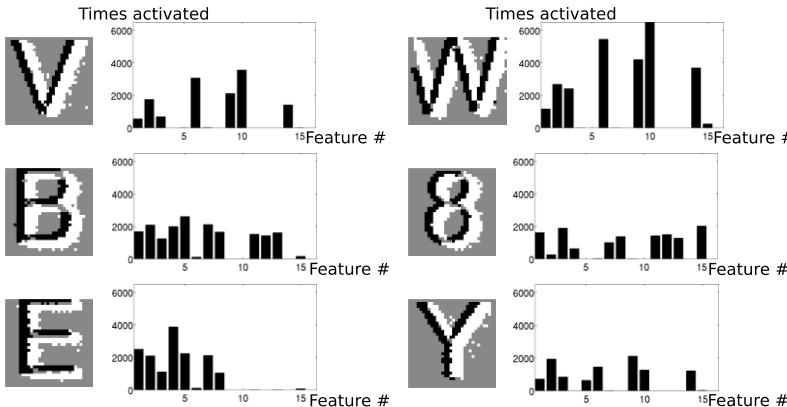
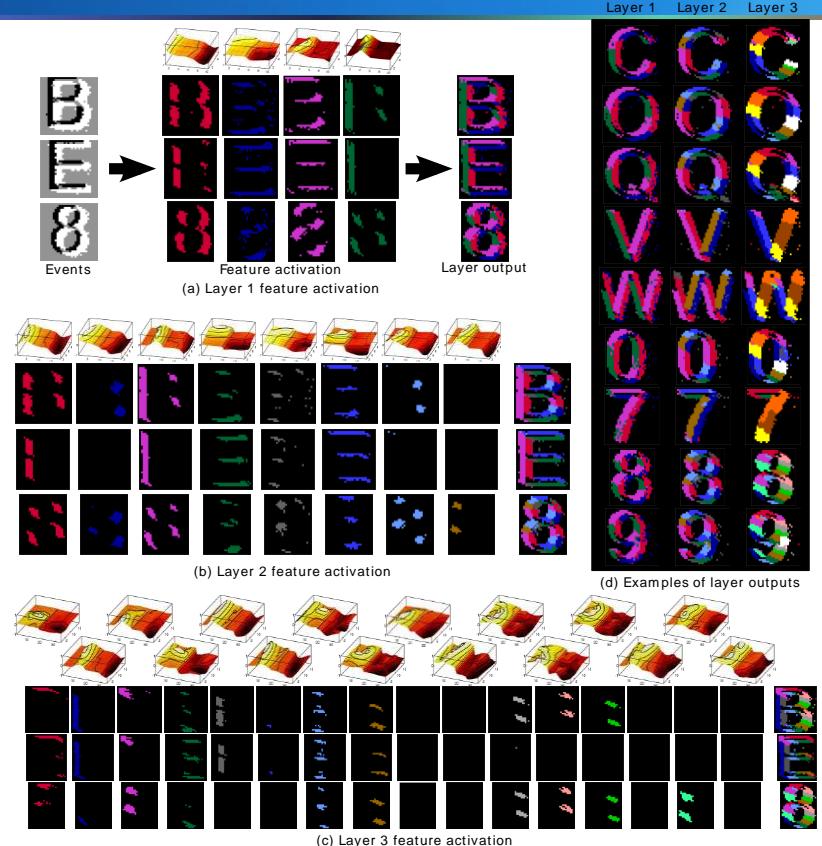
Event-based Time-Surfaces



HOTS: A Hierarchy Of event-based Time-Surfaces



Letters & Digits



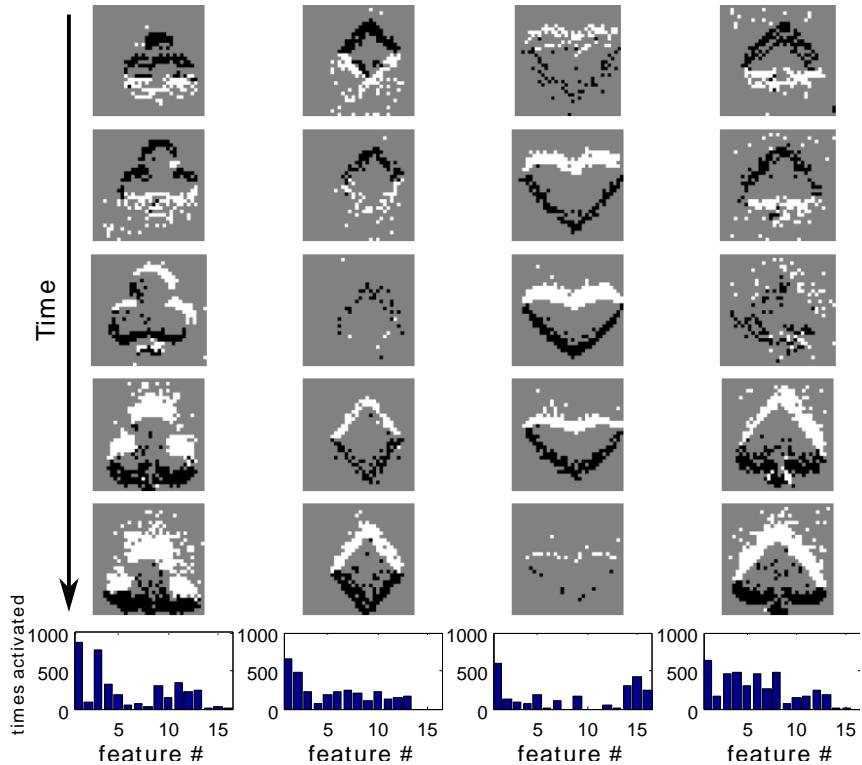
Playing cards experiment

High Speed Poker Card Symbols Recognition using Spiking ConvNets

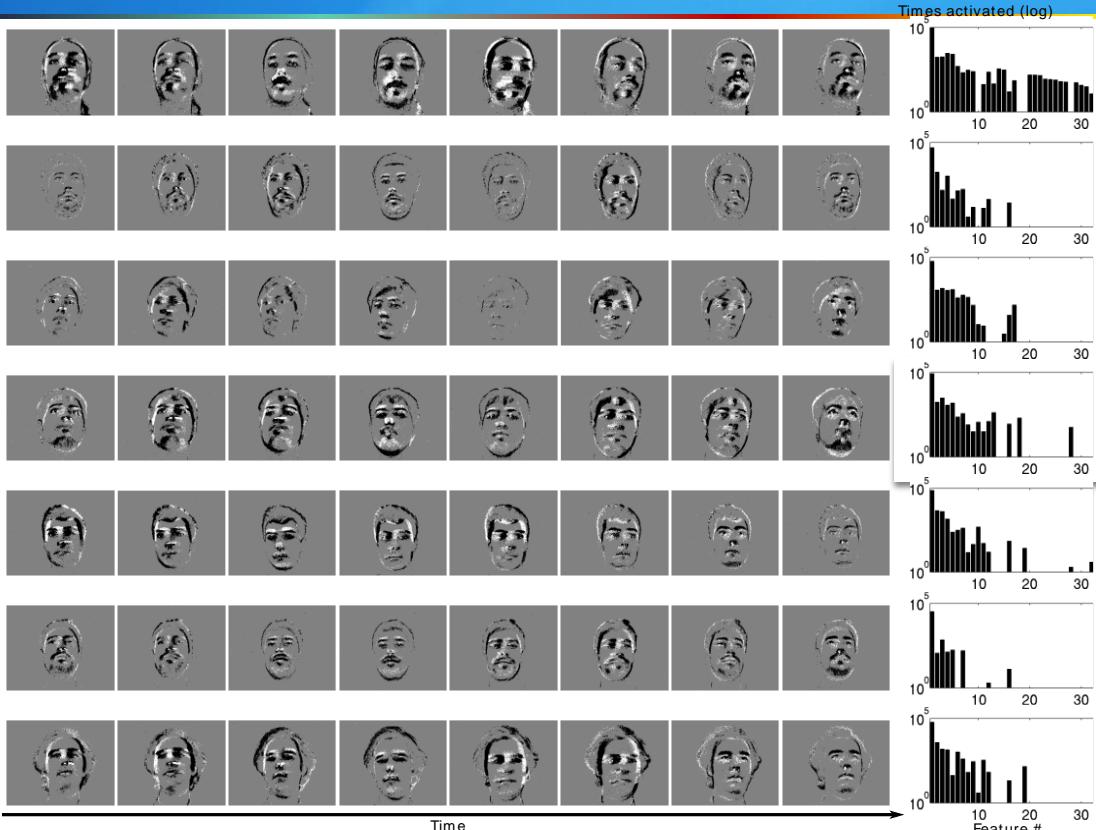
From IEEE TPAMI paper "Mapping from Frame-Driven to Frame-Free Event-Driven Vision Systems by Low-Rate Rate-Coding and Coincidence Processing. Application to Feed-Forward ConvNets"



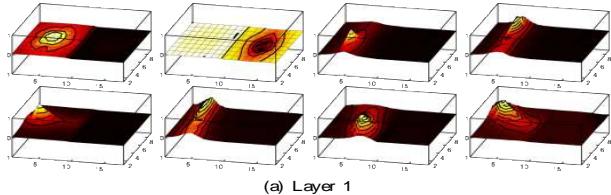
Playing cards experiment



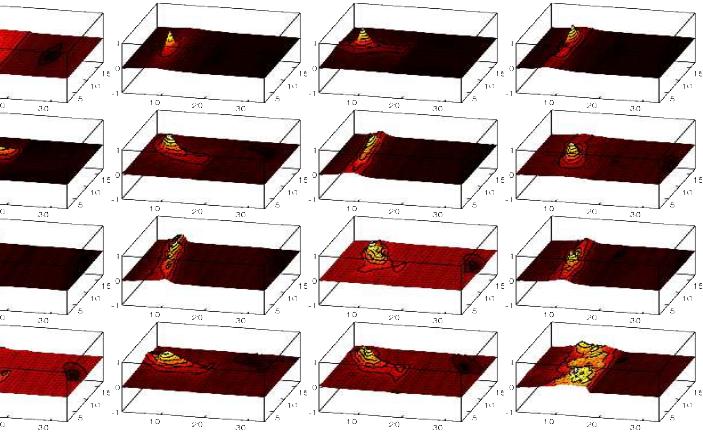
Dynamic Faces



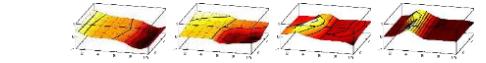
Dynamic Features



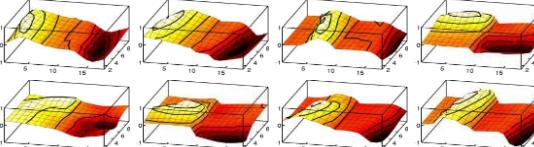
(a) Layer 1



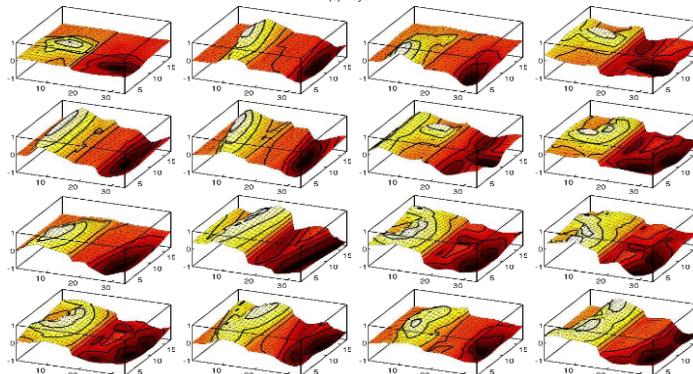
Faces



(a) Layer 1



(b) Layer 2



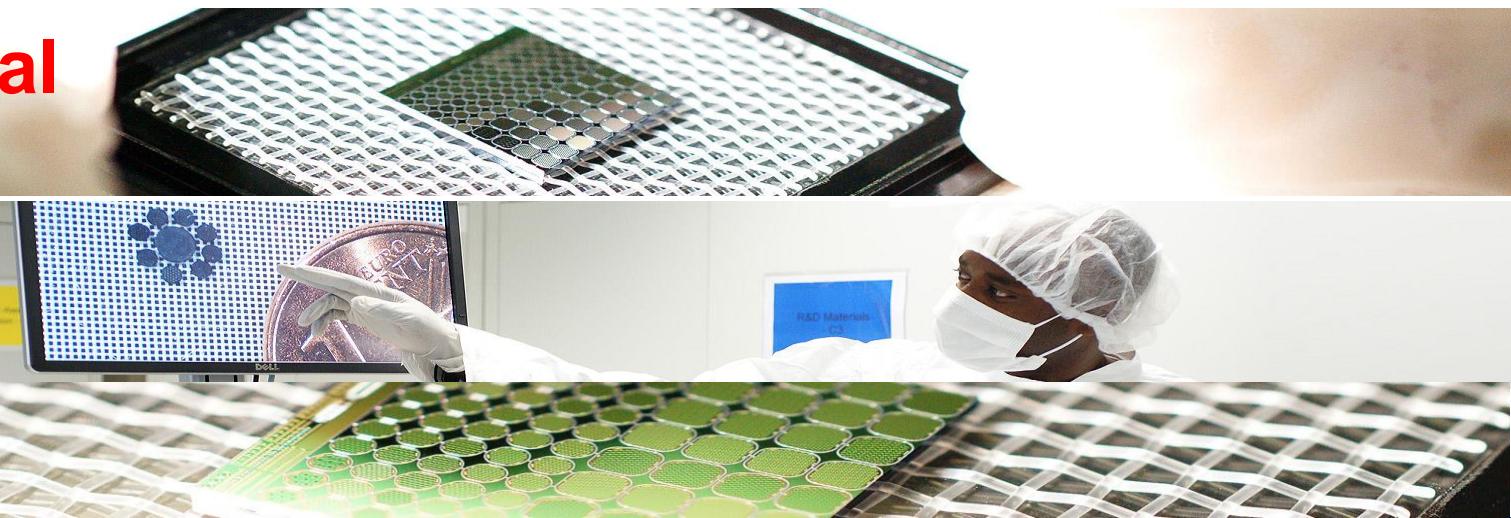
Letters and digits

Retina implants

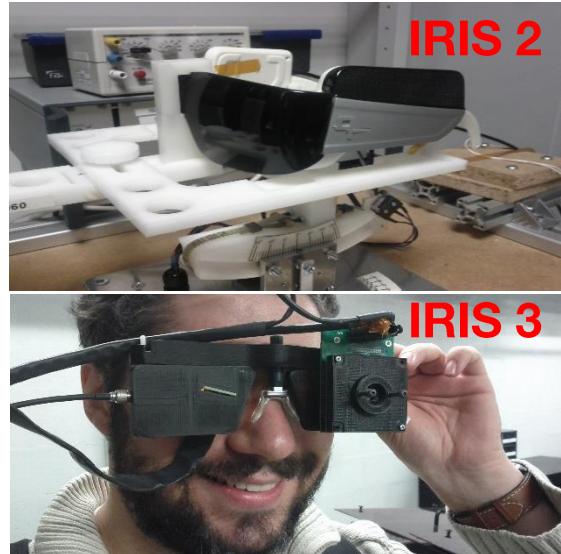
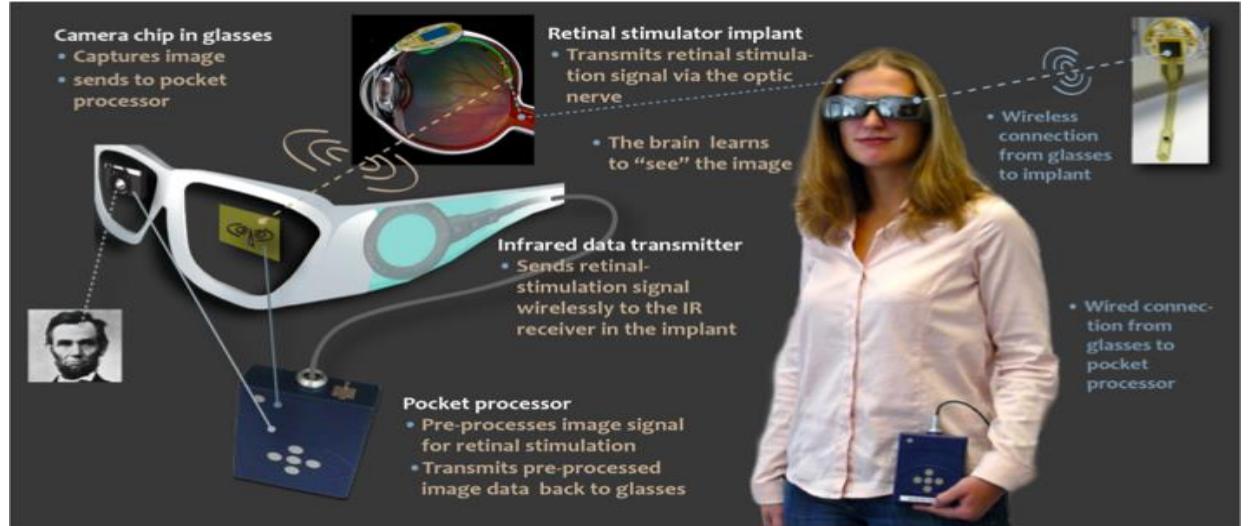
Epi-retinal



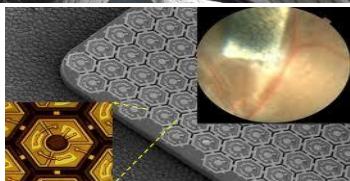
Sub-retinal



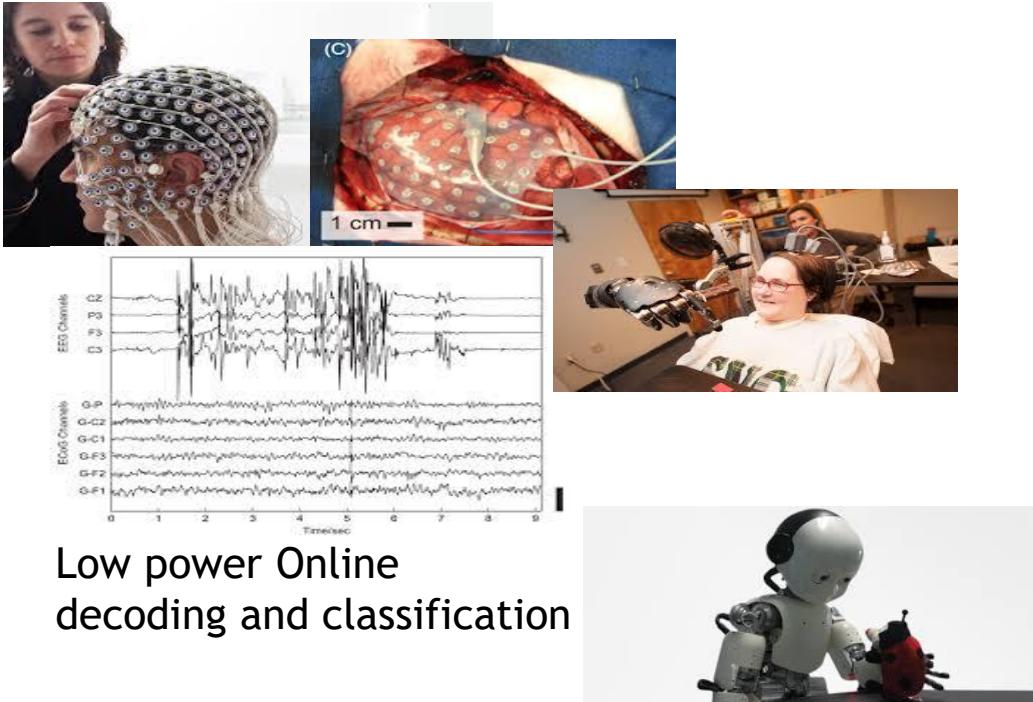
Asynchronous Display: retina prosthetics



- Development of retina stimulation goggles
- Asynchronous retina stimulation: prosthetics and optogenetics
- Development of stimulation optogénétique (GenSight, FMI),
- New generation of display for gaming and wearable devices



& much more....



Low power Online
decoding and classification

Robotics



Decision making: game theory
stock Market



Autonomous driving



Always on sensing



Conclusions

- A paradigm shift in machine vision
- Operate on time rather than luminance information
- Several possible sensors
- Adapted to IOT and low power computation
- Low data bandwidth
- Outperforms conventional image based acquisition

