Computer vision and computational neuroscience: a mathematical perspective

Olivier Faugeras

NeuroMathComp Laboratory - INRIA Sophia/UNSA LJAD

Okawa Prize symposium lecture
University of Tokyo
March 5th, 2015

Introduction

Geometric Computer Vision
- A bit of history
- Projective modeling
- Euclidean modeling

Variational Approaches
- General statements
- The framework
- Finding blood vessels in MRA images
- Image warping
- Shape and image means
- Shape priors in segmentation

Information representation in the brain
- Introduction
- The example of V1
- Toward a new theory

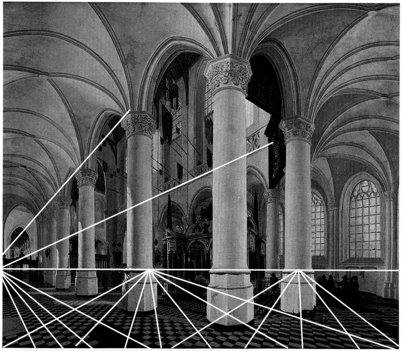
Conclusion

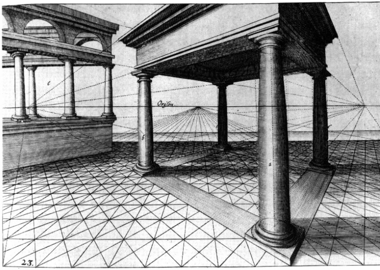
The Unreasonable Effectiveness of Mathematics

- Eugene Wigner (1960): "The Unreasonable Effectiveness of Mathematics in the Natural Sciences".
- David Marr (1980): Vision is information processing; computational theory of a process.

The invention of perspective

- What is the link between the paintings of a Gerard Houckgeest or a Vredeman de Vries and image processing, computer vision or virtual reality?





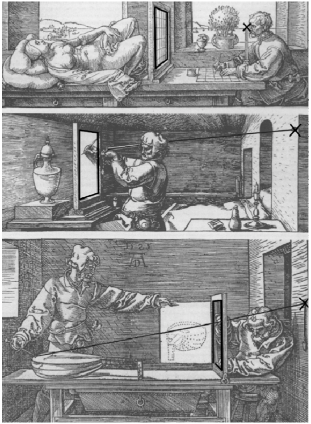
- Answer: the notion of perspective, a subfield of projective geometry that is the key to modeling systems of cameras.

The Erlangen program (1872)

- What is the link between Felix Klein and Hermann Weyl and image processing, computer vision or virtual reality?
- Answer: the idea of a quantity that is invariant to the action of some group of transformations.
- We will be interested here in the projective and Euclidean groups.

A camera is a projective engine

- The Renaissance approach:



A camera is a projective engine

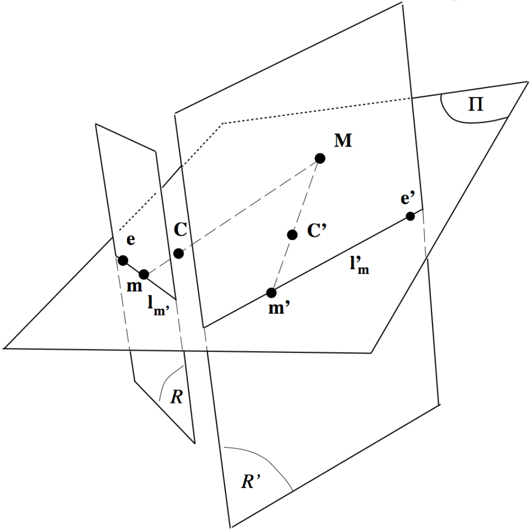
- The modern approach:

$$m = \mathcal{P} M,$$

where $\mathcal{P}$ is a 3×4 matrix.

Two projective cameras

- Important fact: epipolar geometry,



Two projective cameras

- represented by a 3×3 matrix $\mathbf{F}$ called the Fundamental matrix.

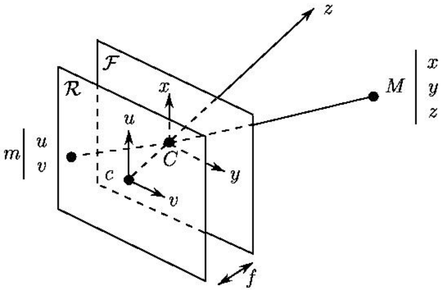
Two projective cameras

- represented by a 3×3 matrix $\mathbf{F}$ called the Fundamental matrix.
- It is estimated by point correspondences.

Two projective cameras

- represented by a 3×3 matrix $\mathbf{F}$ called the Fundamental matrix.
- It is estimated by point correspondences.
- The three-dimensional structure of the scene can then be recovered up to a projective transformation.

Back to the $\mathcal{P}$ matrix

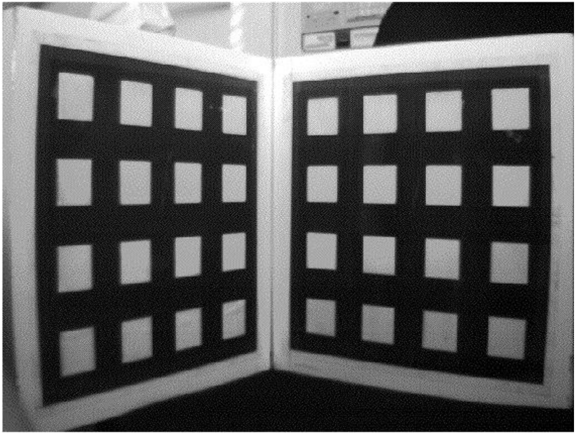


- The matrix $\mathcal{P}$ can be decomposed into intrinsic and extrinsic parameters

$$\mathcal{P} = \mathbf{A}[\mathbf{R}\ \mathbf{t}]$$

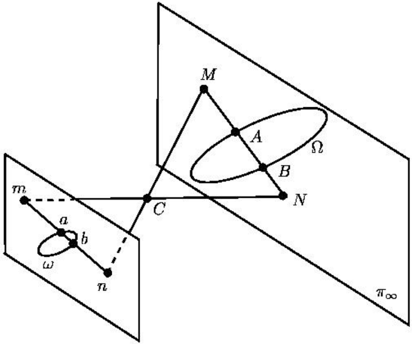
Back to the $\mathcal{P}$ matrix (continued)

- How do we recover these parameters from a few images (calibration)?
- The traditional approach uses a special calibration pattern



Back to the $\mathcal{P}$ matrix (continued)

- One can also use the umbilic or the absolute conic:



Faugeras-Luong-Maybank method (1995)

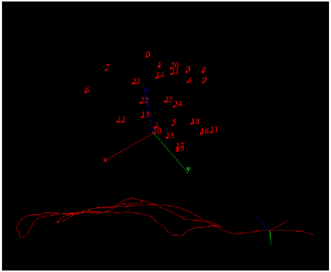
- Inspired from the work of Kruppa (1913)
- Given two views one uses the invariance of the umbilic w.r.t. the similitudes.
- Given the Fundamental matrix, one writes two equations, called the Kruppa equations, that connect the intrinsic parameters.
- In general three views are sufficient to calibrate.

Application to augmented reality (joint work with Rachid Deriche, Luc Robert, Quang-Tuan Luong, Imad Zoghlami) 1990-2000

- Goal: insert synthetic objects in a real image sequence

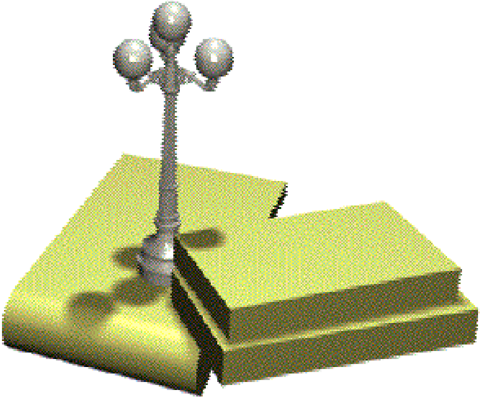
Application to augmented reality (continued)

- Through the analysis of the image sequence one computes the 3D motion of the camera.



Application to augmented reality (continued)

- Knowing this motion, one can model in 3D part of the scene and add a synthetic object

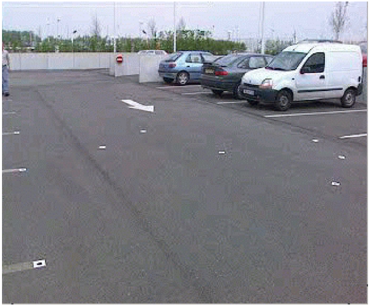


Application to augmented reality (continued)

- One can then insert this object in the initial sequence

Application to augmented reality (continued)

- Other examples (Courtesy RealViz, now AutoDesk)



Some of the goals and challenges of Computer Vision

- Finding interpretations of scenes: 3D structure, motion, shapes. . .
- Partial theories have been developed
- Many ad-hoc implementations have been proposed
- Proofs of correctness are scarce

Some of the goals and challenges of Computer Vision

- Integrating Computer Vision Modules within larger systems
- requires formal validation.

The concept of Energy

- We observe a series of images $I_1, I_2, \ldots, I_N$ of dimension n ($n = 2$ ou $n = 3$)
- We look for a shape, i.e. a model M (in effect a differential manifold of codimension 1 in a space of dimension n) that is a function of some parameters $\mathbf{p}$
- We define an energy E which measures the lack of adequation between the model and the observations, the images.

The concept of Energy (continued)

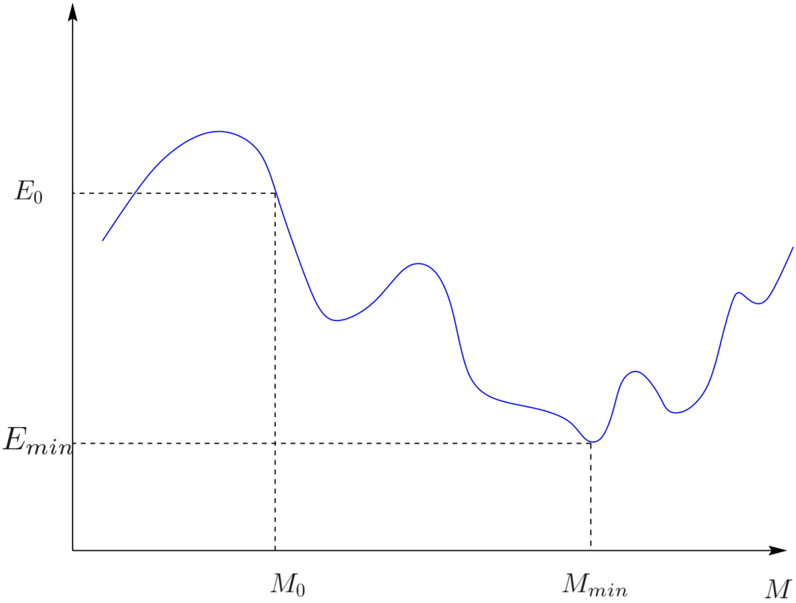
- We look for the model that minimizes the energy, hence which "explains" best the data.
- The principle is similar to the **Least Action Principle** in Physics: it is an extremality principle.

$$E(M, DM, \cdots) = \int L(I_1, \cdots, I_N, M, DM, \cdots, \mathbf{p})\, d\sigma$$

L is the analog of the Lagrangian
$M(\mathbf{p})$ is the model
$d\sigma$ is the area element of the manifold M.

- The methods used to solve this optimization problem are the methods of the calculus of variations.

Minimizing the Energy



Minimizing the Energy (continued)

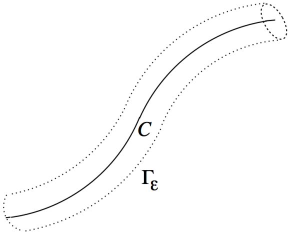
- We start with an initial model M_0 whose energy is E_0 and deform it until we obtain M_{min} with minimal energy.
- We define a family of models $M(\mathbf{p}, t)$ where $t \geq 0$ is an artificial time and

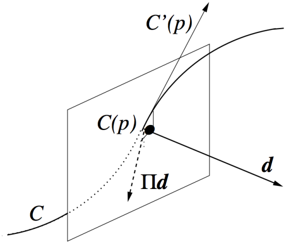
$$M(\mathbf{p}, 0) = M_0(\mathbf{p}),$$

- And solve a Partial Differential Equation (PDE).

Finding blood vessels in MRA images (joint work with Liana Lorigo and Eric Grimson, MIT)

- Goal: detect and characterize the shape and size of blood vessels in MRA images.
- Methodology: generalization of the previous approach to curves in 3D space through the idea of ε-level sets.

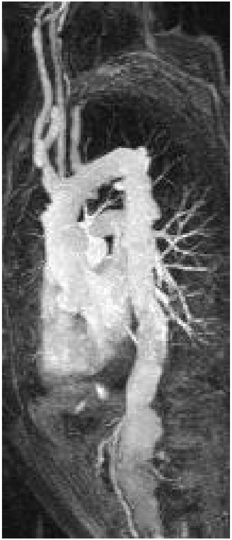


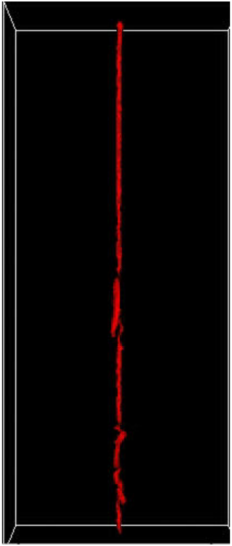


Finding blood vessels in MRA images (continued)

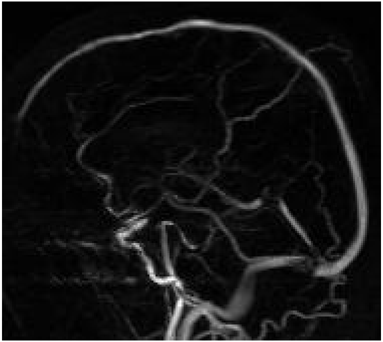
- The model M is a set of 3D curves, the medial axes of the blood vessels represented by their ε-level sets.
- The energy measures the coherence of the ε-level sets with the presence of a high image gradient.

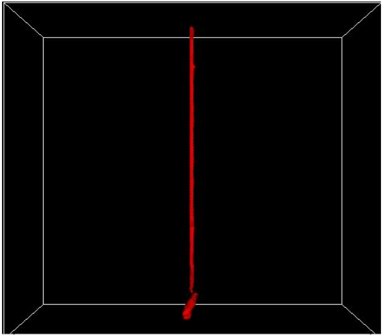
Example I: aorta data (courtesy Siemens)



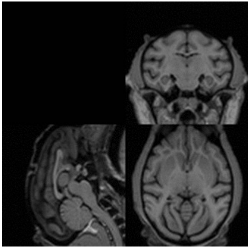
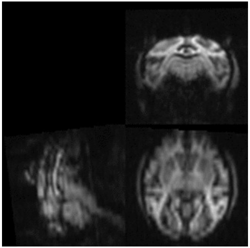
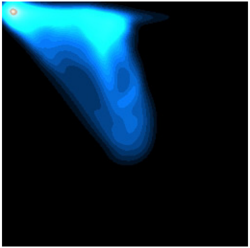
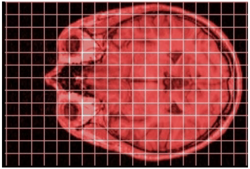
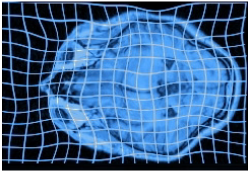
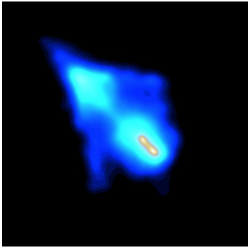


Example II: brain vessels data



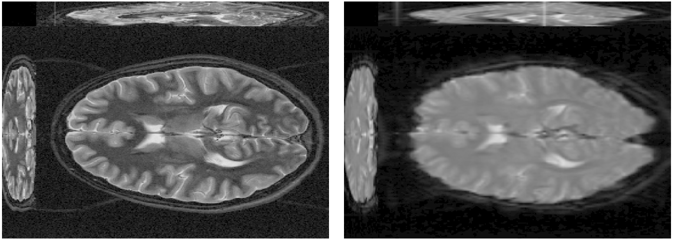


Non-rigid matching between images (joint work with Gerardo Hermosillo and Christophe Chefd'Hotel, INRIA)

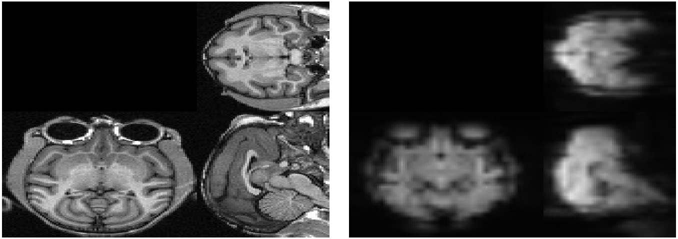







Applications in medical/cognitive imaging:

- Matching similar structures between several image modalities.



Human



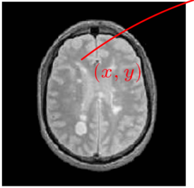
Monkey

Applications in computer vision:

- Matching similar structures
    - under varying lighting conditions.
    - with varying responses to similar lighting.

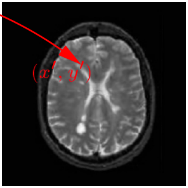
Problem formulation

$I_1 : \Omega \subset \mathbb{R}^n \to \mathbb{R}$

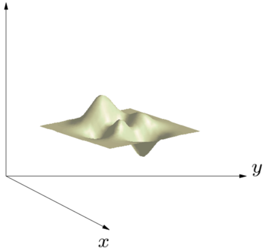


$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \Phi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$$

$I_2 : \Omega \subset \mathbb{R}^n \to \mathbb{R}$

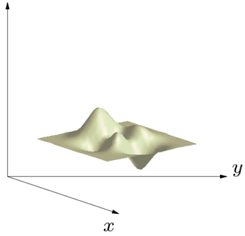


$I_1^\sigma = I_1 \star G_\sigma$



$$\Phi = \mathbf{Id} + \mathbf{h}$$

$I_2^\sigma = I_2 \star G_\sigma$



Problem formulation (continued)

We consider the minimization problem in a suitable functional space $\mathcal{F}$

$$\mathbf{h}^* = \arg\min_{\mathbf{h}\in\mathcal{F}} \mathcal{I}(\mathbf{h}) = \arg\min_{\mathbf{h}\in\mathcal{F}} \left(\mathcal{J}(\mathbf{h}) + \mathcal{R}(\mathbf{h})\right)$$

where:

- $\mathcal{J}(\mathbf{h})$ penalizes "statistical dissimilarity" between I_1^σ and $I_2^\sigma(\mathrm{Id} + \mathbf{h})$
- $\mathcal{R}(\mathbf{h})$ penalizes fast variations of $\mathbf{h}$.

Problem formulation (continued)

$$\mathbf{h}^* = \arg\min_{\mathbf{h}\in\mathcal{F}} \mathcal{I}(\mathbf{h}) = \arg\min_{\mathbf{h}\in\mathcal{F}} \left(\mathcal{J}(\mathbf{h}) + \mathcal{R}(\mathbf{h})\right)$$

We compute the Gradient $\nabla_H \mathcal{I}(\mathbf{h})$ of $\mathcal{I}(\mathbf{h})$, and consider the initial value problem:

$$\left\{ \begin{array}{l} \dfrac{d\mathbf{h}}{dt} = -\nabla_H \mathcal{I}(\mathbf{h}) = -\Big(\nabla_H \mathcal{J}(\mathbf{h}) + \nabla_H \mathcal{R}(\mathbf{h})\Big), \\ \mathbf{h}(0) = \mathbf{h}_0 \in H, \end{array} \right. ,$$

and prove that it is well-posed.

One then solves a PDE.

Image matching: experiments

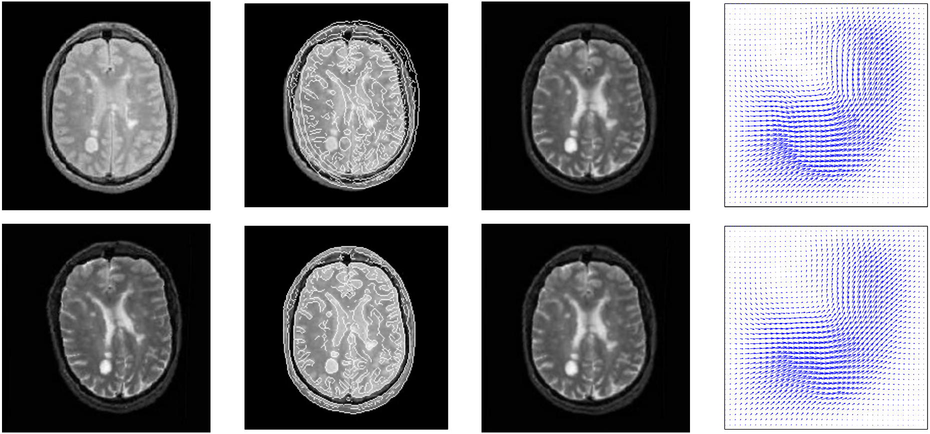


Image matching: experiments (continued)

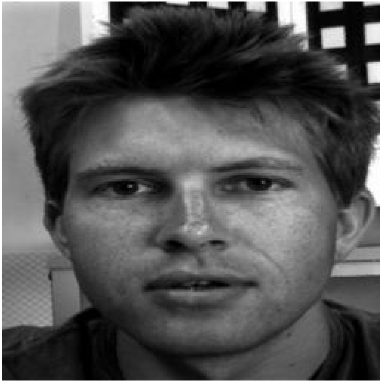


Image matching: experiments (continued)

Image matching: experiments (continued)

Application: computation of the empirical mean of N images (joint work with Guillaume Charpiat)

- Given N images $I_1, \ldots, I_N$,
- The mean image M is defined (implicitely) by the minimum of

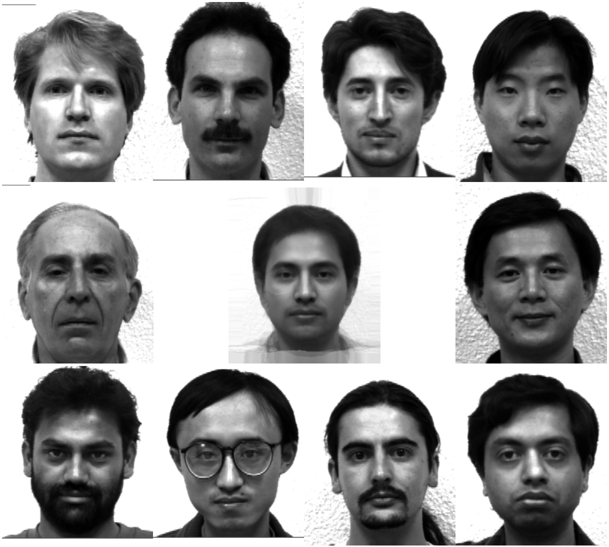
$$\mathcal{I}_M(\mathbf{h}_1, \cdots, \mathbf{h}_N) = \sum_{i,j=1,\, i<j}^{N} \mathcal{J}_{ij}(\mathbf{h}_i,\, \mathbf{h}_j) + \sum_{i=1}^{N} \mathcal{R}_i(\mathbf{h}_i)$$

- $\mathcal{J}_{ij}(\mathbf{h}_i,\, \mathbf{h}_j)$ is the statistical "dissimilarity" between $I_i(\mathrm{Id} + \mathbf{h}_i)$ and $I_j(\mathrm{Id} + \mathbf{h}_j)$.
- M is defined as

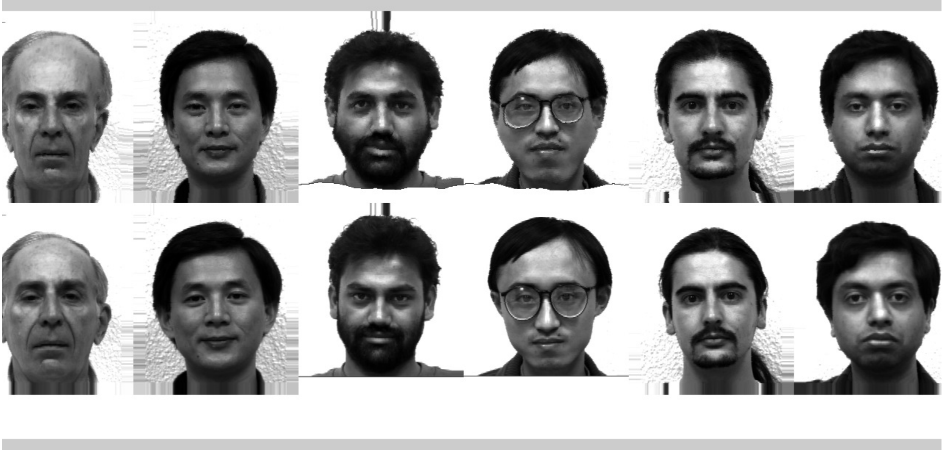
$$M = \frac{1}{N} \sum_{i=1}^{N} I_i(\mathrm{Id} + \mathbf{h}_i^*)$$

Results: the mean

Results: the warping applications



The warping applications: a summary

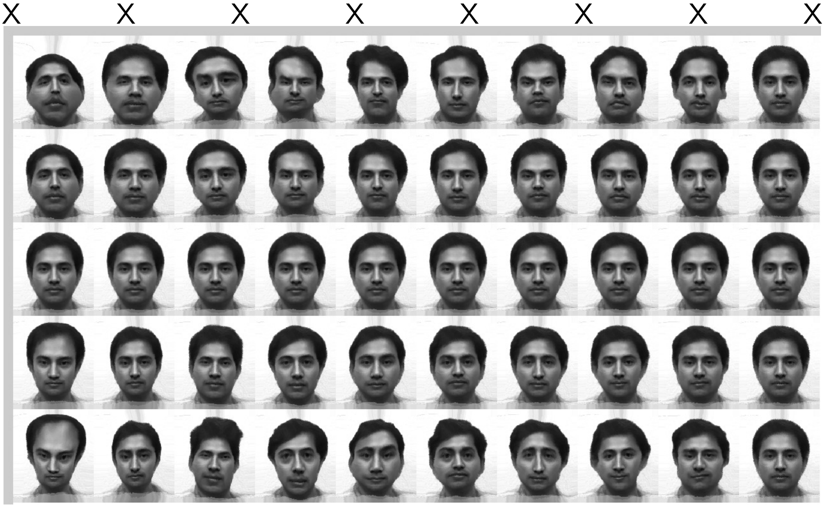


Application: computation of the empirical geometric covariance of N images

- Consider $\mathcal{I}_i(\mathbf{h})$, the dissimilarity between I_i and $M(\mathrm{Id} + \mathbf{h})$
- Compute the covariance matrix of $\nabla_H \mathcal{I}_i(\mathbf{h})$
- The principal modes are its eigenvectors $\mathbf{h}_i$
- To look at how the mean image changes geometrically with respect to the ith mode, solve:

$$\begin{cases} \dfrac{d\mathbf{h}}{dt} = -\mathbf{h}_i(t) \\ \mathbf{h}(0) = \mathbf{h}_i(0) \end{cases}$$

Encoding the geometry: the spatial modes of variation



Computing statistical 2D and 3D shape models

- We compute a statistical shape model over a training set of shapes:
    1. Corpus callosum.
    2. Vertebrae.

Computing statistical 2D and 3D shape models (continued)

- We compute the mean:
  1. Mean corpus callosum.
  2. Mean vertebra.
- and the modes:
  1. Modes for corpus callosum
  2. Modes for vertebrae.

Using statistical shape priors in segmentation (joint work with Michael Levinton and Eric Grimson, MIT)

- Active contours or snakes: to evolve the region boundaries
- Bayesian framework: use of statistical priors to constrain the snake evolution

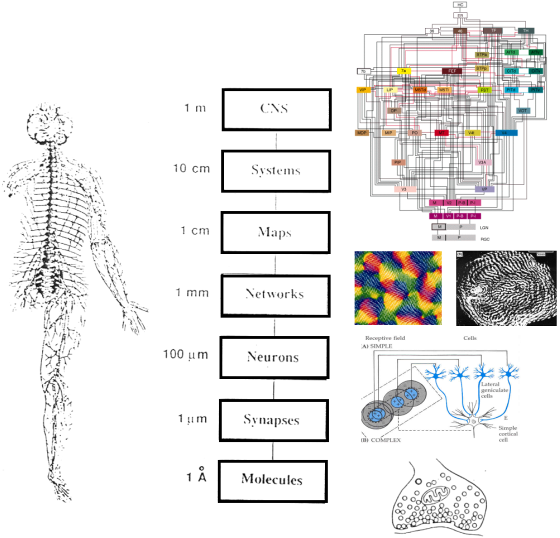
Results for 2D image segmentation

- One slice of a femur.
- Two differents corpus callosa: 1 and 2.

Results for 3D image segmentation

Example of the vertebra

Scales in the CNS

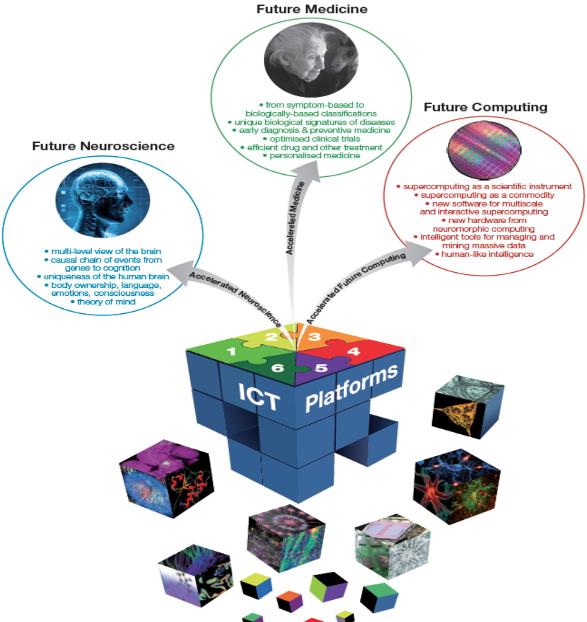


Mesoscopic and macroscopic models

Electroencephalography (EEG) and Magnetoencephalography (MEG)

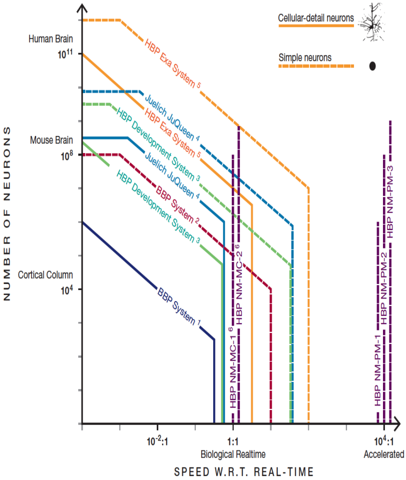
The Human Brain Project

- 10 year project (1 billion Euros)
- Goals:
  1. simulate a human brain
  2. equivalent for neuroscience to High Energy Physics equipment (e.g. CERN)
- The center is at the EPFL



Impact on High Performance Computing

- Exascale computation (10^{18} Flops/second).
- Neuromorphic acceleration
- Hierarchical memories (Petabytes).
- Several Terabits/second I/O



Assumptions

1. BBP System hosted by CADMOS is a BG/P (4 racks), detailed cellular Blue Brain Model (non-memory optimised)
2. BBP System hosted by CADMOS is a BG/P (4 racks), simplified network, numbers are extrapolation from Modha 2009
3. Planned HBP Development System tries to maintain memory:flop ratio of BG/P, numbers are extrapolated numbers based on system characteristics and memory optimisation
4. Jülich JuQueen, 28-rack BG/Q system, numbers are extrapolated numbers based on system characteristics; faster processors (as compared to BG/P) should allow better strong scaling
5. Planned HBP Exascale computer, will have large memory and strong scaling components
6. The HBP-NM systems are limited by the connections/s and for the above graph assumes firing rates of a few Hertz

Theoretical neuroscience

- Goal: better understand the fundamental mechanisms that govern cortical behaviours.
- Tools: mathematics and numerical simulations.
- An example in the field of neuroscience of "The Unreasonable Effectiveness of Mathematics in the Natural Sciences"(Eugene Wigner, 1960)

Motivations

- Represent the neuronal activity at different scales (**sparsity**)

Motivations

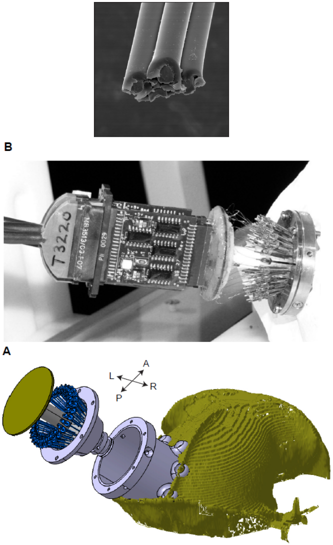
- Represent the neuronal activity at different scales (**sparsity**)
- Predict the occurence of new, **emerging**, neural phenomena

Motivations

- Represent the neuronal activity at different scales (**sparsity**)
- Predict the occurence of new, **emerging**, neural phenomena
- Understand the role of **randomness**

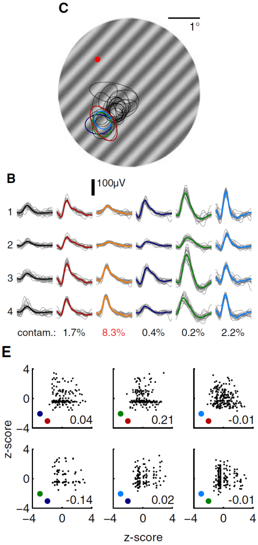
Neuronal activity in V1

- The recording of neuronal activity in V1



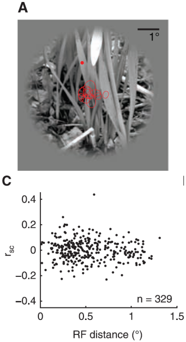
Neuronal activity in V1

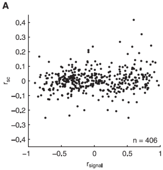
- shows that it is highly **decorrelated** for synthetic



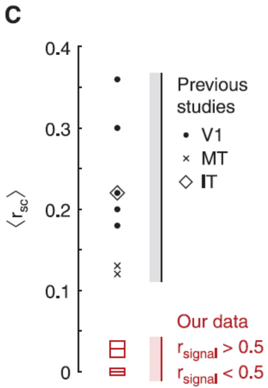
Neuronal activity in V1

- and natural images,





Neuronal activity in V1



- in contradiction with current belief.

A mathematical model of a cortical area (joint work with J. Baladron, D. Fasoli, and J. Touboul)

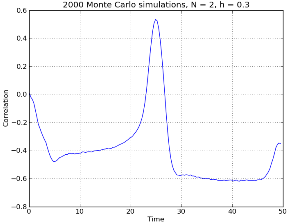
- The individual neurons model: Hodgkin-Huxley
- The type of synapses: chemical or electrical
- The synaptic weights: excitatory and et inhibitory.
- Many sources of noise: stochastic model.

Letting the number of neurons grow very large (to ∞!)

- The neurons in the model become independent: propagation of chaos.
- All neurons are described by the same stochastic process.
- Provides a sparse and accurate description of the network:
    - the millions of equations describing the network are summarized by less than 10!,
    - opening the door to very efficient numerical simulations.

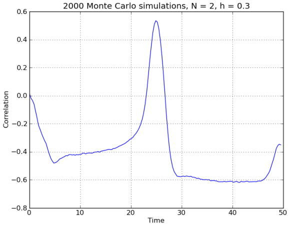
Numerical validation

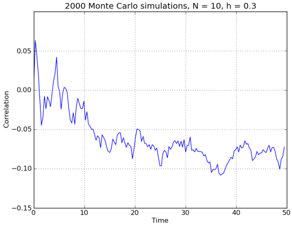
- The propagation of chaos effect appears for populations of neurons of relatively small sizes: in agreeement with biology.



Numerical validation

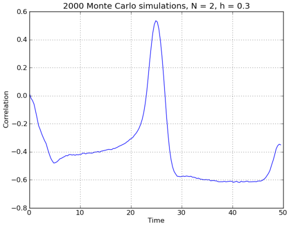
- The propagation of chaos effect appears for populations of neurons of relatively small sizes: in agreeement with biology.

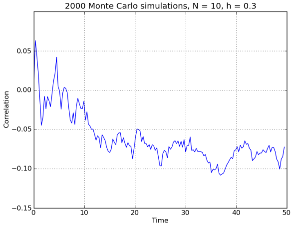


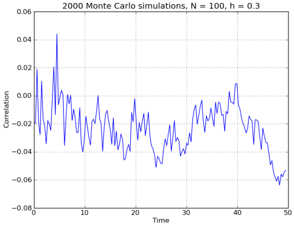


Numerical validation

- The propagation of chaos effect appears for populations of neurons of relatively small sizes: in agreeement with biology.

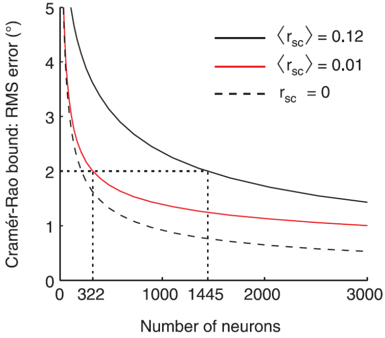






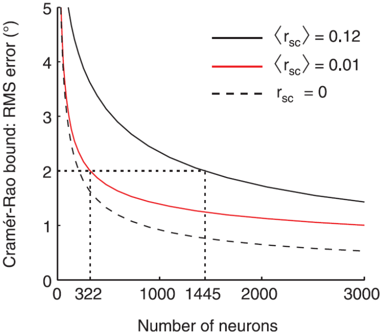
Connection with information theory

- The visual cortex behaves optimally from the viewpoint of information theory (signal decorrelation), and



Connection with information theory

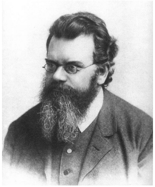
- The visual cortex behaves optimally from the viewpoint of information theory (signal decorrelation), and
- neurons appear to be coding **probability laws.**



Taming the complexity

- Develop a theory of statistical neuroscience

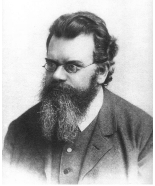
Taming the complexity



- Develop a theory of statistical neuroscience

- Ludwig Boltzmann : Inventor of statistical mechanics.

Taming the complexity



- Develop a theory of statistical neuroscience

- Ludwig Boltzmann : Inventor of statistical mechanics.
- Accounts for and explains how the properties of atoms (mass, electrical charge, structure . . .) determine the macroscopic properties of matter(viscosity, thermal conductivity, diffusion . . .)

Wigner's and Marr's statements

Wigner The effectiveness of mathematics in the sciences of the artificial has been clearly verified.

Wigner Their effectiveness for the sciences of the living is still a bit more unclear.

Marr That vision/cognition is a process independent of its substrate is still debatable.

Thank you to the Okawa foundation and to the members of the Selection Committee of the Okawa Prize.