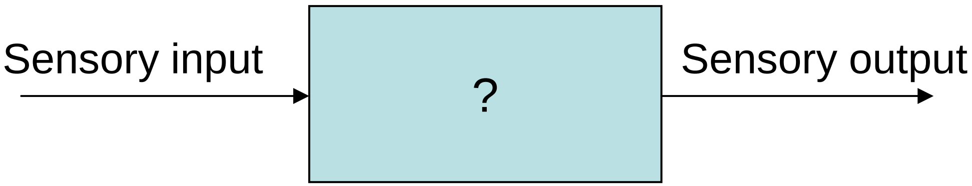
Systems Neuroscience
& Neurotechnology Unit

Kernel Machines
Support Vector Machine Regression

Prof. Dr. Dr. Daniel J. Strauss

Initial situation

- Research, with regard to multiple scales of the structure of the human brain is meanwhile very advanced and largely investigated.
- But up to now, the relationship between signals at one scale and those at others is poorly understood.
- One of the most fascinating questions in neuroscience is the relationship between neural activity of individual neurons and their behavior or interaction within the neural system.
- The complete progress of understanding the functioning of the human brain can approximately be limited to the fact, that sensory input and output operations are playing the key role



- But little is known about the interactions and information process between this operations.

Objectives

- Mapping physiological features with in individual brain areas represents a powerful tool to understand how brain function is organized and how it takes place.

→ Support Vektor Maschinen (SVMs)
with compactly supported RBFs

- 64-channel electrophysiological recordings by placing a flexible multielectrode array on the auditory cortex of a adult Wistar rat.

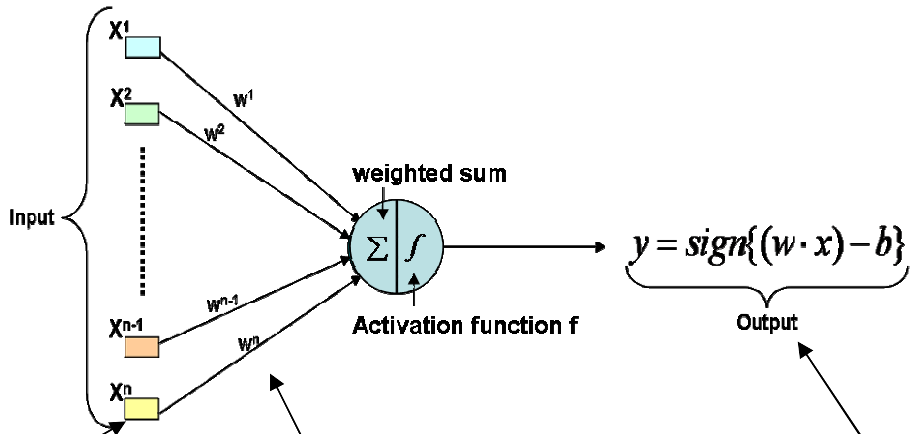
- Basically, when plotting or imaging recorded data the result will be a kind of a "snap-shot"

Intention
→„Dynamical map“
→ Each point is given a value
→ Transform the coarse grid into a **continuum**

Neural Networks - Forerunner

- In 1957 Frank Rosenblatt invented the first algorithmically described neural network (i.e. first model of a learning machine)

  →Rosenblatts perceptron
- This model is deemed to be the first model for learning with a *“teacher”.*
- *Based on a “Binary threshold unit”*



X_i represents the input to the neuron

Each neuron is multiplied by its weight w_i

y represents the output

Support Vector Regression (SVR)

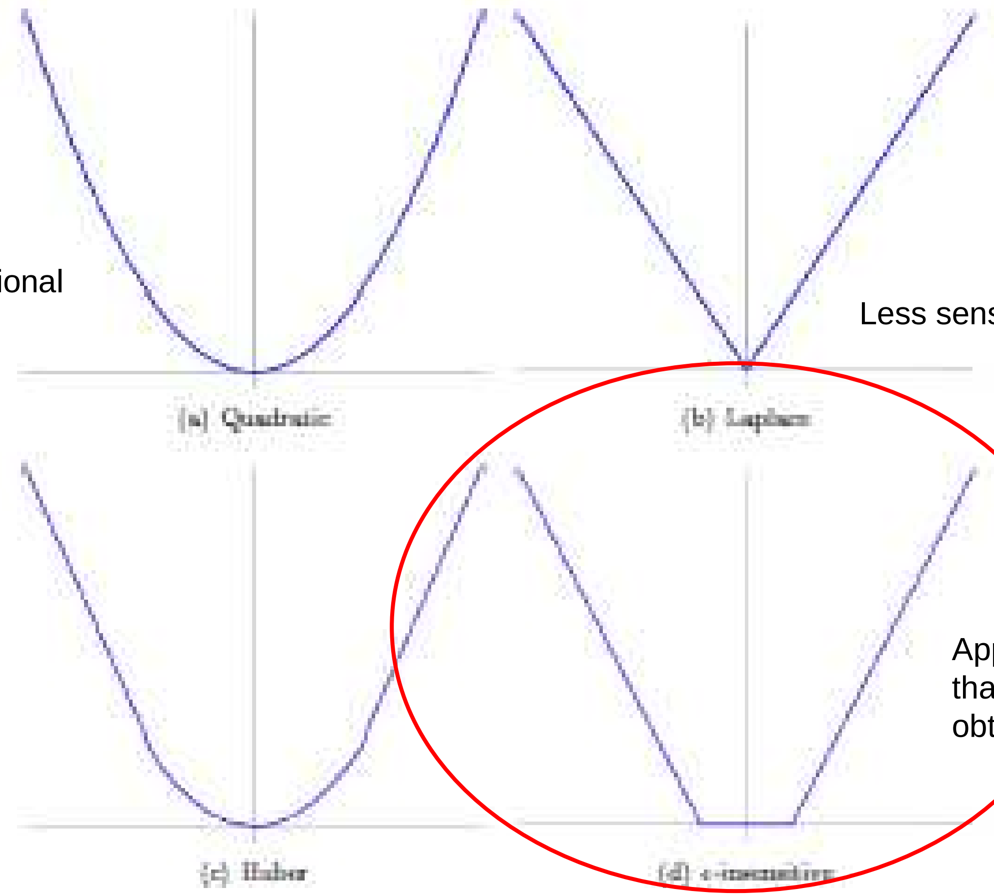
- Basically SVMs were developed to solve the classification problem
- But recently they have been extended to the domain of regression problems (Vapnik 1997)
- Brief: SVC leads onto mapping the input into a higher dimensional feature space by a suitable choice of kernel function

Support vector regression

- SVR can be applied to regressin problems by the introduction of an loss function

corresponds to the conventional
least sqaures error criterion

Less sensitive to outliers than the quadratic



Robust loss function that
has optimal properties

Approximation to Huber's loss function
that enables a sparse set of SVs to be
obtained

Linear Regression

Initial situation

Consider the problem of approximating the set of data

$$\mathcal{D} = \left\{(x^1, y^1), \ldots, (x^l, y^l)\right\}, \quad x \in \mathbb{R}^n, y \in \mathbb{R},$$

The optimal regression function is given by the minimum of the functional

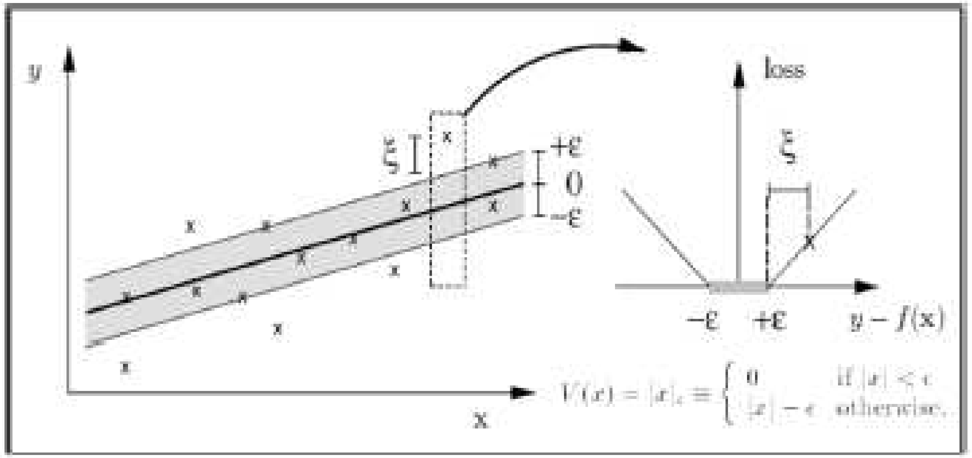
$$\lambda \sum_{i=1}^{M} V_\varepsilon(y_i - f_i) + \frac{1}{2} \|h\|^2_{\mathcal{H}_K}, \qquad V_\varepsilon(x) = \max\{0, |x| - \varepsilon\}$$

denotes Vapnik's e-insensitive loss function, where $V(x)$ s some error cost function that is used to measure the interpolation error.

The first term is enforcing closeness to the data, the second smoothness and λ controls the tradeoff between two terms

- By the Representer Theorem, the minimizer of $\lambda \sum_{i=1}^{M} V_\varepsilon(y_i - f_i) + \frac{1}{2} \|h\|^2_{\mathcal{H}_K}$, can be written

$$f(\boldsymbol{\nu}) = \sum_{j=1}^{M} c_j K(\boldsymbol{\nu}, \boldsymbol{\nu}_j)$$



Non-Linear regression

- In the same manner as the Support vector classification approach, a non linear mapping is used to map the data into a high dimensional feature space where linear regression is performed.

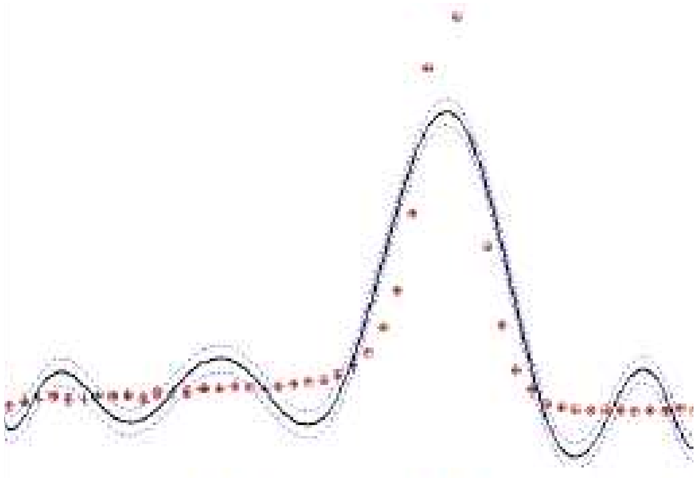
Example of a non-linear regression problem

Three methods of controlling the regression model

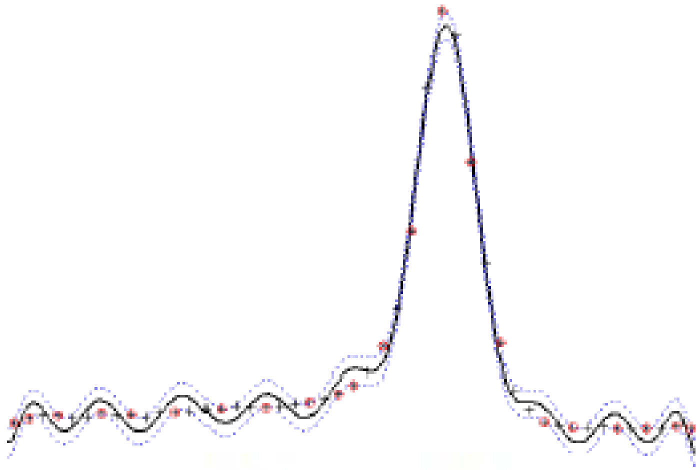
→ Loss function

→ kernel

→ Capacity control



$(\epsilon = 0.05, \sigma = 1.0, C = \infty)$



$(\epsilon = 0.05, \sigma = 0.3, C = \infty)$

Solution for a gaussian RBF kernel.

→Whats the difference between both???

→ It can be seen that the kernel is to wide to accurately model the data

CSRBFs

The RBF we used up to know is the Gaussian kernel

$$k(x, x') = exp(-\|x - x'\|^2 / \sigma^2)$$

where $\sigma \in \mathbb{R}$ is a tuning parameter (or bandwidth) of the model

But most RBF kernels are „globally" supported
→ Can be ill conditioned for large datasets

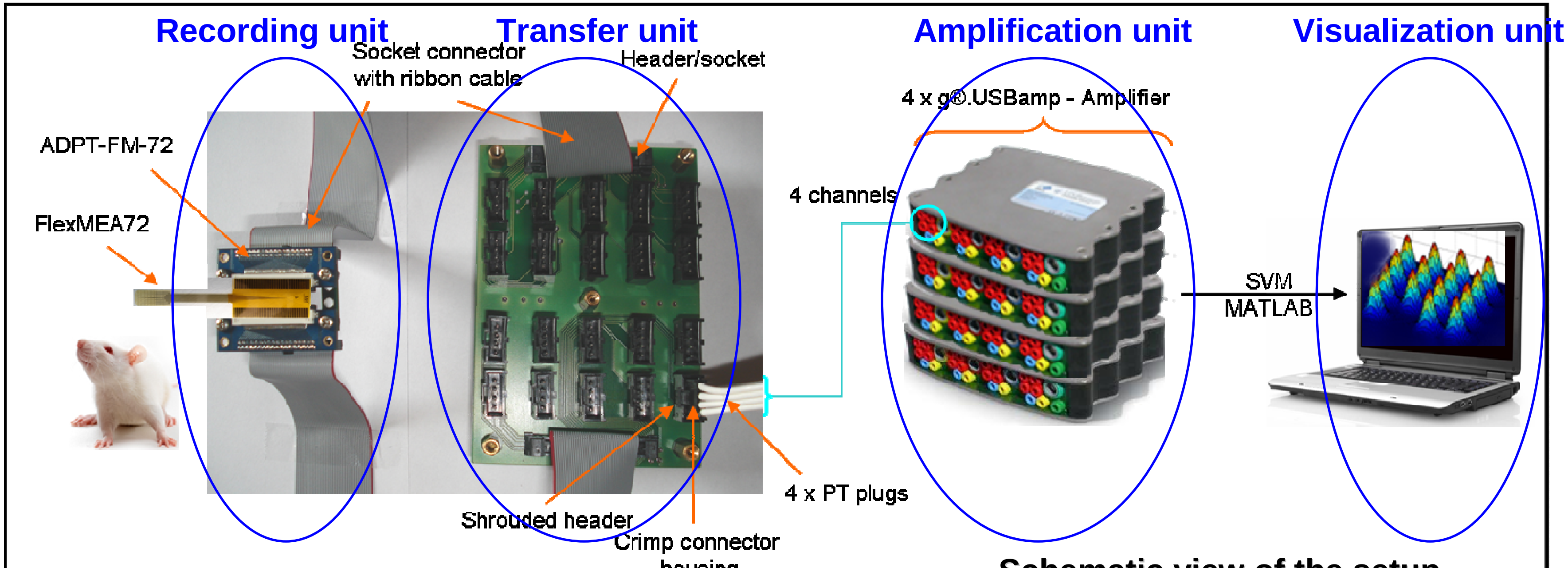
What to do if having large datasets?????

→ **RBFs wich are „local" with compact support.**

Compactly Supported Radial Basis Functions for the Multivariate Approximation of Neural Field Potentials in MEA Measurements

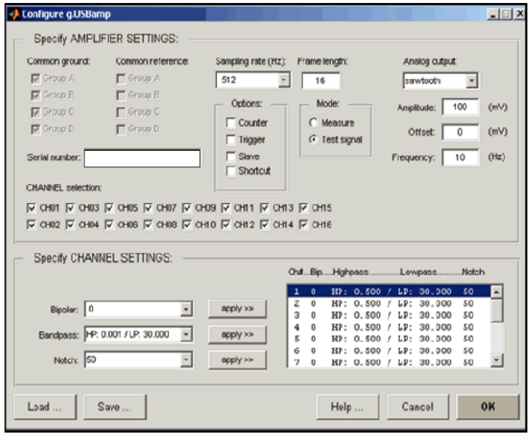
Materials and Methods

- To carry out an electrophysiological study, with respect to animals, no ubiquitous experimental setup is available.
- Intention
  - Place a flexible multielectrode array (FlexMEA72) onto the auditory cortex of a rat and performe electrophysiological recordings.
  - Provide the data to the support vector machine (SVM) for generating a „dynamical map“ of the cortex-area.

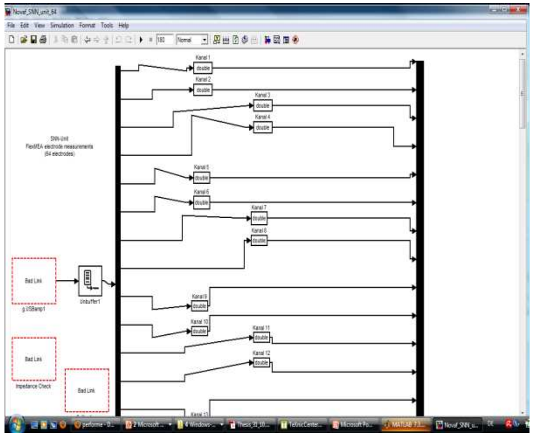


Schematic view of the setup

Hardware and Software

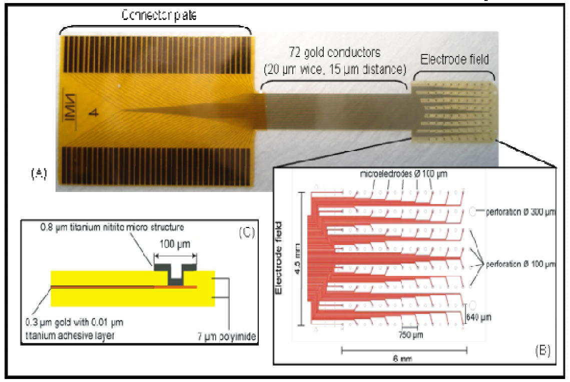


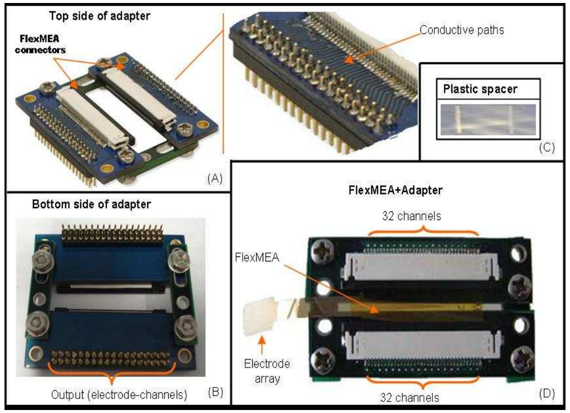
MATLAB and Simulink

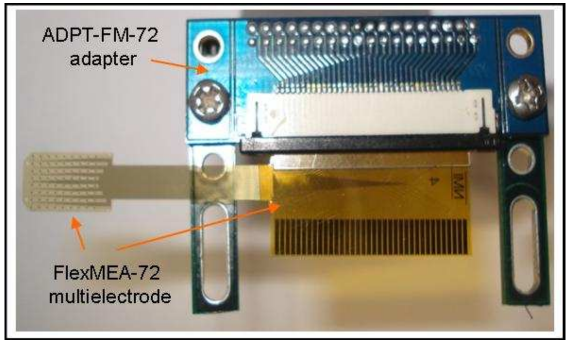


g.USBamp amplifier

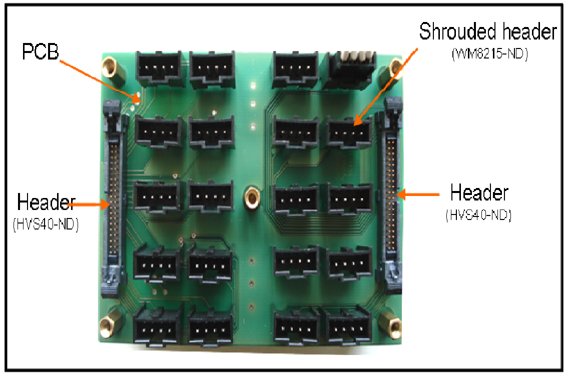
FlexMEA72 and ADPT-adapter

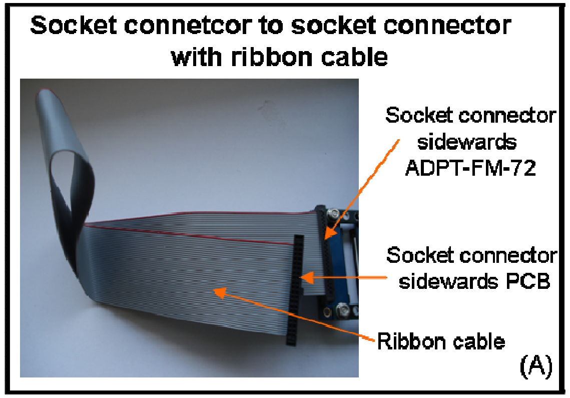


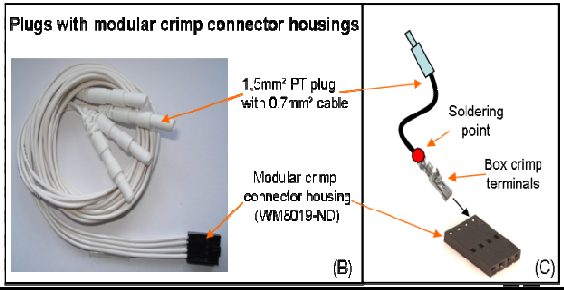




PCB and accessoiries

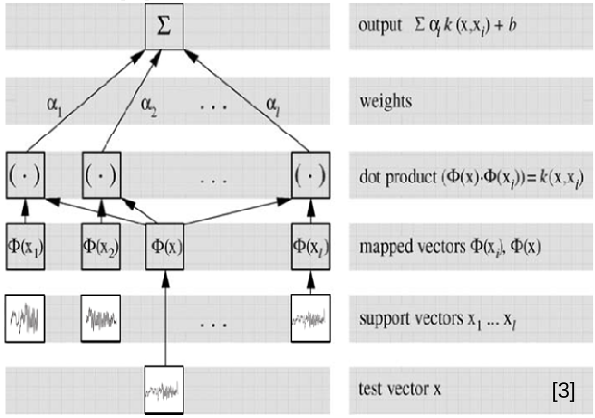






Methods

Support vector machine



- „Teach“ the machine (training set)
- Construct hyperplane (Support vectors (SVs))
- Map input vectors into the feature space („*Kernel-Trick*“)
- Compute the dot products under the map (*“Gaussian-kernel”*)
- By adding the dot products, plus the constant term b the machine computes the final prediction output

Data aquisition and analysis

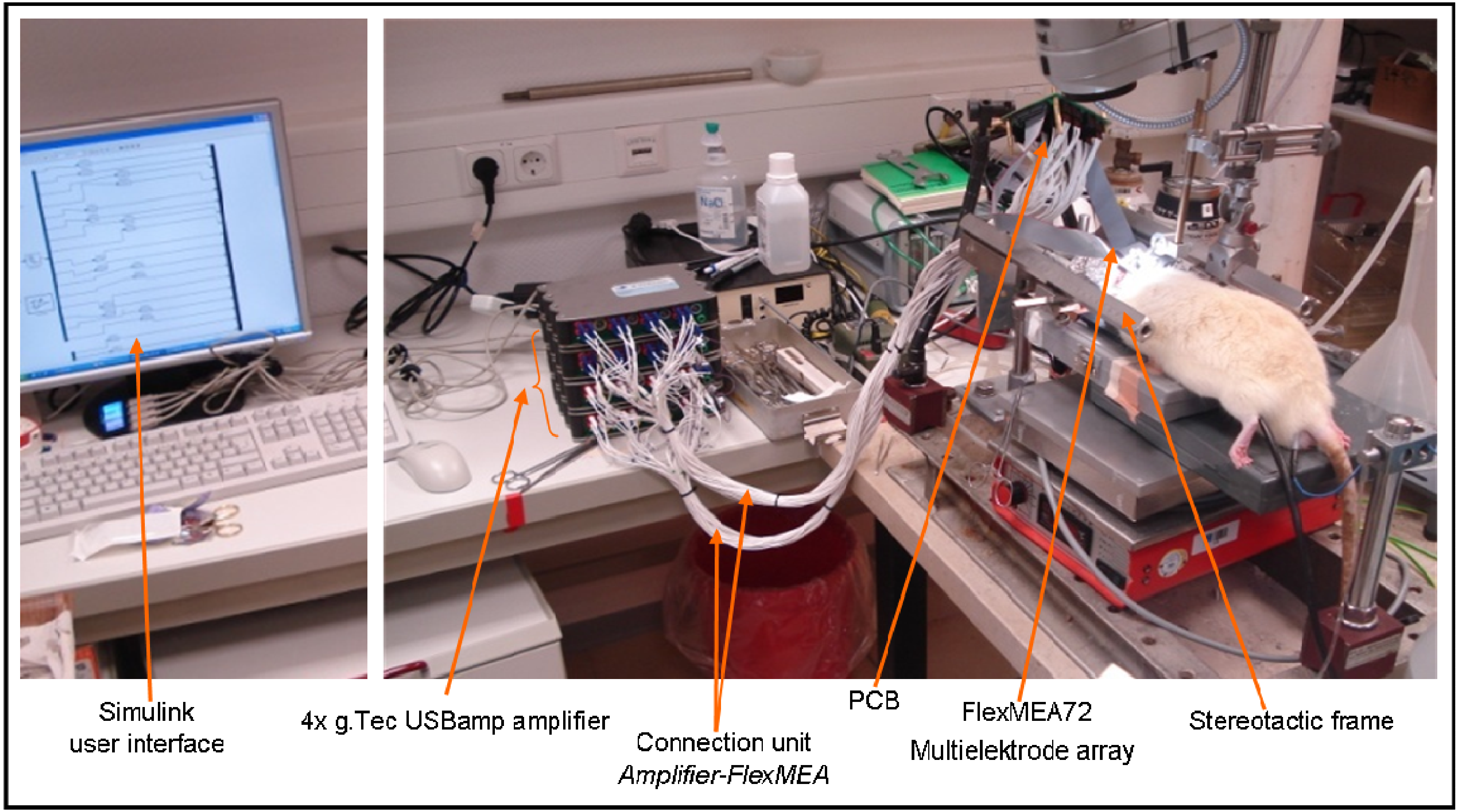
- For recording settings Simulink was used.
- Sampling frequency 19,2 kHz. (Important to fulfill the Nyquist-theorem f(s) > 2 f(c))
- Recording of neural field potentials (NFPs)
- NFPs contain local field potential (LFPs) and multi unit activities (MUA)
- Extract LFPs (30-200 Hz)
  - Low pass filtering at 300 Hz
- Since the collected amount of data prevented the usage of MATLAB - for analysis and SVM regression estimation - we also decided to decimate, and thereby reducing the length of the available signals.

Animal preperation

- Experimental procedures were approved by the Saarland University Ethics Committee and in accordance with section 5 TierSchG.
- Anesthesia was given by inhalation of the anesthetic agent isoflurane.
- The rat was placed in a stereotaxic frame, the head stabilized and the corneas were covered with ophthalmic ointment.
- A heating pad maintained a body temperature between 37-38℃.
- The skull was surgically exposed and a craniotomy was perforemd.

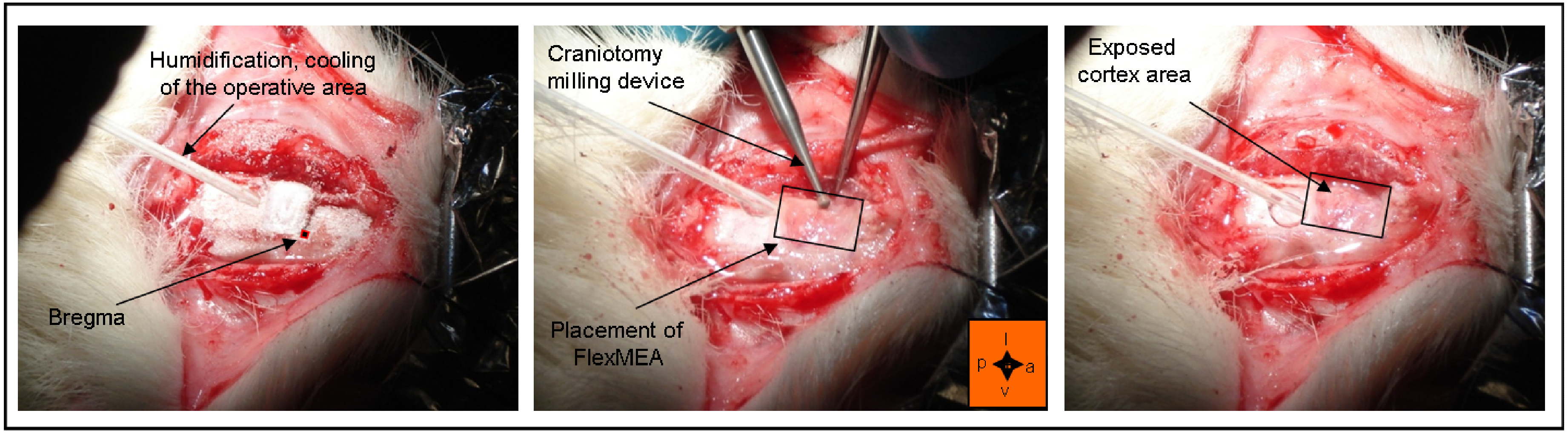
Results

Experimental setup

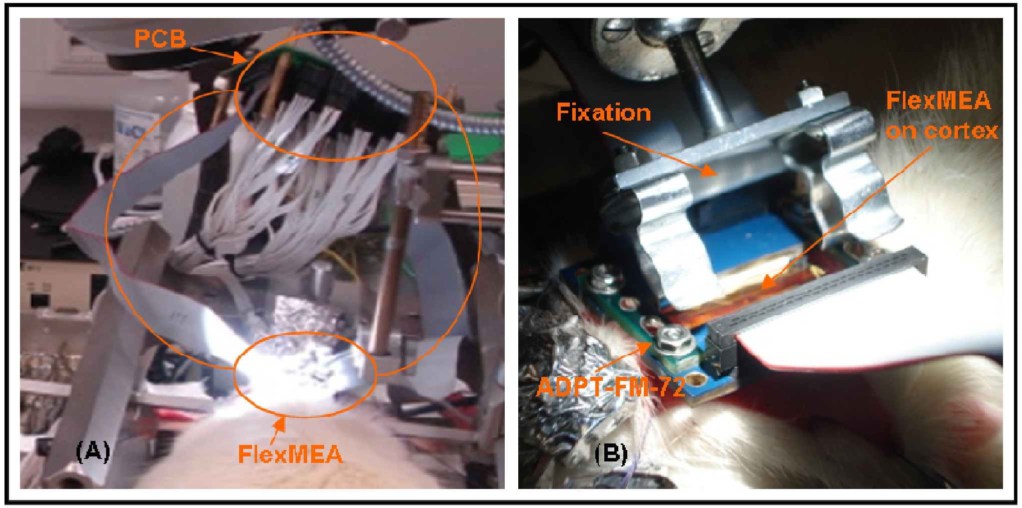


Results

Animal preperation



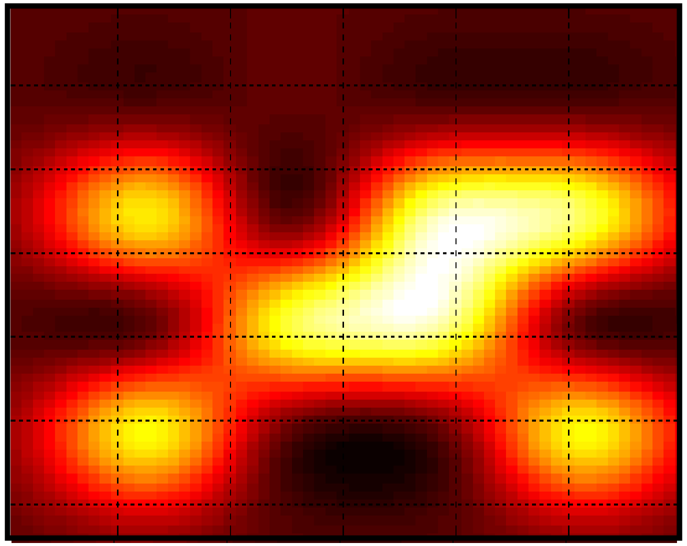
- Skull was surgically exposed in a sterile procedure
- The *Bregma* was identified, serving as a reference point for stereotactic surgery.
- A drilling a milling device was used to „remove" the skull (0.4 mm lateral and 0.2 mm anterior to posterior, relative to Bregma).
- The „Recording Unit" was then placed on the cortex with the bottom side pointing upwards and the FlexMEA array lying underneath.

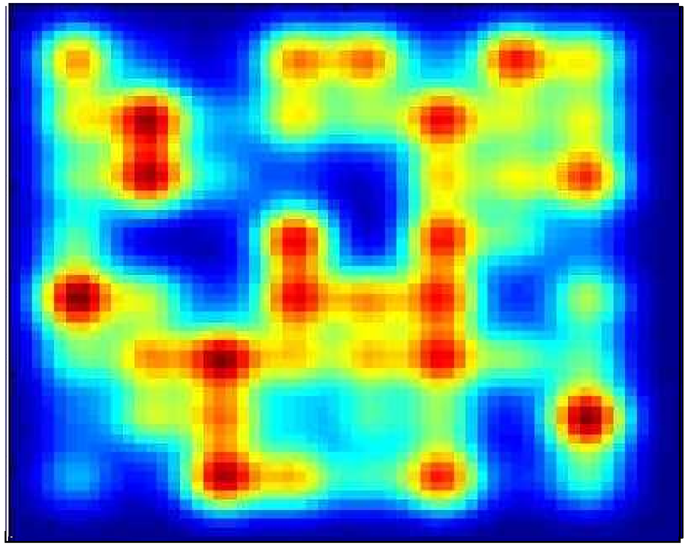


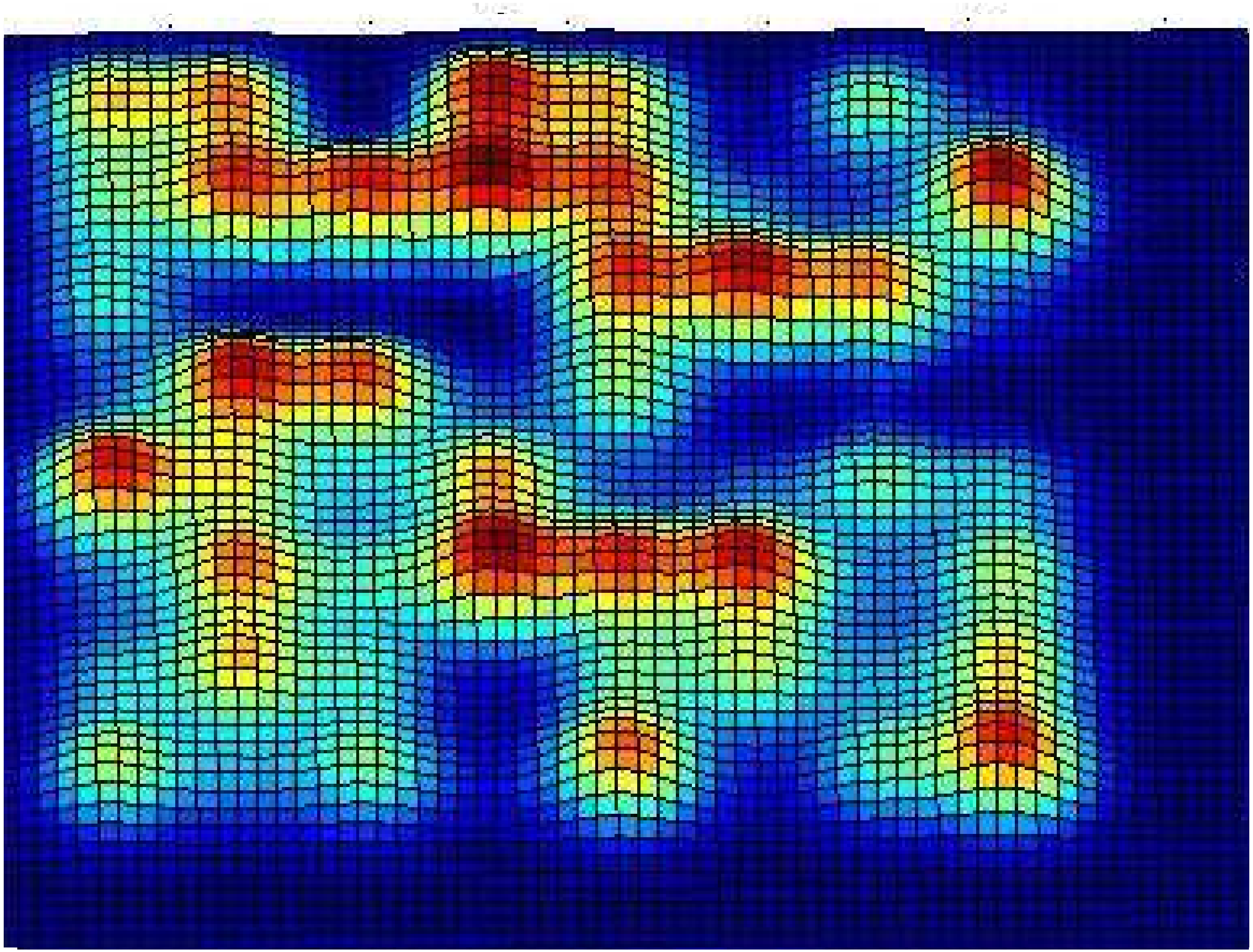
Results

SVM

- First, the recorded data was preprocessed (filtering, decimating)
- Afterwards the available data was provided to the SVM for regression estimation
- Generating a „dynamical map“ of the data recorded from the surface, picturing the possible "ongoing-actions“.

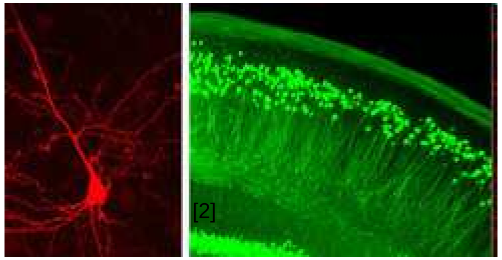






Summary and Outlook

- This work represents an attempt to understand cortical functioning in terms of a computational model.
- Electrophysiological data was recorded by placing a flexible multielectrode array (FlexMEA72) on the cortex of a adult Wistar rat.
- In this way our intention was to generate a "dynamical map" of the underlying electrophysiological data belonging to the surface, picturing the possible "ongoing-actions".
- Basically a great deal of work and efforts still requires to be undertaken in understanding the behavioral role of the "10 billion" neurons in the mammalian brain.
- Nevertheless, our results have shown that it is feasible in teaching a machine (SVM), using electrophysiological data, for constructing a "functional map" in a dynamic manner.
- Moreover, in addition to computer modeling techniques, optical recording methods such as voltage sensitive dye imaging (VSDI) can be used for measuring brain activity
- Thus, in future work both techniques could be compared with each other and mutually complete themselves.



References

- Burgers C.J.C.: *A tutorial on Support vector machines for pattern recognition, Springer*, Data mining and discovery, 2, 121-167 (1998)
- Eder M.: *Visualizing the electrical activity of neuronal networks – a promising method for basic psychiatric research,* Max-Planck-Institut für Psychiatrie, München (2011)
- Schölkopf B. & Smola A.J.: *Learning with kernels: Support vector machines, Regularization, Optimization and Beyond (Adaptive Computation and Machine Learning),* MIT Press, Cambridge, MA (2002)
- Vapnik V.N.: *The nature of statistical learning theory*, Springer Verlag, New York, NY, USA (1995)
- Molina-Luna K. et al.: *Cortical stimulation mapping using epidurally implanted thin-film microelectrode arrays, Elsevier,* Journal of Neuroscience Methods 161, 118–125 (2007)
- www.plexon.com
- www.multichannelsystems.com

Thank you for
your attention!