Project: building a software for online spike sorting

Résumé en français :

      La rétine humaine ou animale s'étudie "in vivo" en mesurant à l'aide d'un tableau d'électrodes les potentiel d'actions générés par une population de neurones qui réagit à un stimulus donné. Sur chaque électrode est mesuré un voltage qui reflète le mélange des réponses de plusieurs neurones. C'est un défi de pouvoir identifier puis trier les événements correspondants aux réponses de chaque neurone. Nous venons de publier un algorithme réalisant ce tri. Il est proposé ici de reprendre cet algorithme et d'en faire un logiciel performant et utilisable par des biologistes. L'algorithme devra être parallélisé sur les processeurs graphiques de l'ordinateur qui exécute le calcul, pour obtenir un temps d'exécution compatible avec une séance d'expérimentation biologique.
Un contrat de six mois de conception et développement logiciel est proposé ici pour mener à bien ce projet.

Problem position:

The retina is a thin sheet of neural tissue in the back of the eye where vision begins. The optics of the eye focus light down onto the plane of the retina, where it is absorbed by a layer of photoreceptor cells and converted into a neural signal. This signal flows through a layer of bipolar neurons to the final layer of ganglion cells, which send brief electrical pulses, known as spikes, down the optic nerve to the central brain. In addition, classes of interneurons called horizontal cells and amacrine cells spread signals laterally within the plane of the retina.

Our purpose is to understand the processing that transform the light stimulation into a complex sequence of spikes sent by the ganglion cells to the brain. We have developed a new technique to record a large population of ganglion cells from the retina. This technique combines the use of dense and large Micro-Electrode Arrays (MEA, 252 electrodes), and a new technique to efficiently sort all the spikes out of these simultaneous extracellular recordings. Upgrading the classical MEA technique to a new level, it allows simultaneous recording of almost all the ganglion cells in a patch of the retina containing 200+ cells, providing an essentially complete picture about the information the brain can read about this region (Marre et al, 2012). Note that this technique is not only novel because of the number of neurons we recorded is high, but also because the *fraction* of neurons we record is very high, a feat that is currently out of reach in the cortex.

The data flow generated by these data is huge: a typical afternoon of experiment will generate 50 Gb. So far, our algorithm has been running offline, overnight, using a computing cluster, and was not optimized at all. However, the algorithm is entirely parallelized.

Proposed work:

The purpose of this internship is to transform this code into a software that can be used daily by biologists. This would include speeding up the current algorithm by optimizing the current code, as well as adapting the code so that it can be run on GPUs. The algorithmic part of the existing code is in C/C++ and Matlab, and the interactive part is in Matlab. A user-friendly graphical interface would also be useful to visualize the results, using a standard GUI framework (e.g. QT, or equivalent libraries). From our estimation, we think that it is possible to reach a speed of execution of the algorithm that enables sorting the data "online", i.e. during the time course of the experiment. This would open the possibility to perform new kinds of experiment, where the stimulus is modified according to the responses of ganglion cells.
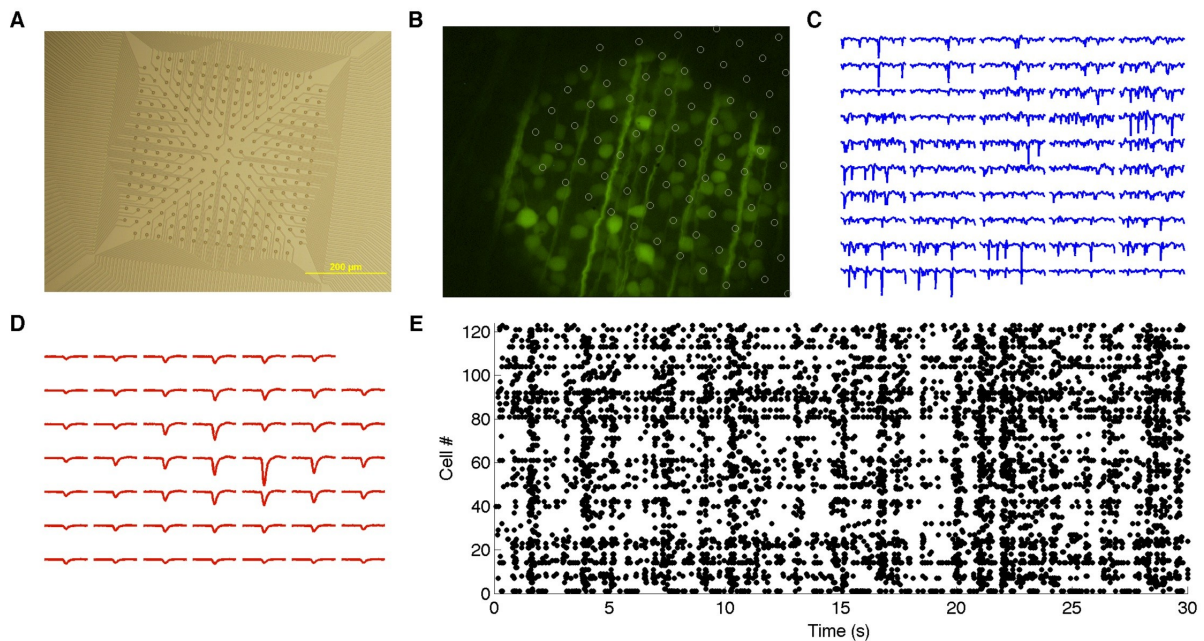
Illustration 1: A: an example of a dense multi-electrode array. B: Example of a retinal patch over the multielectrode array. The ganglion cells are stained with a dye and shown here in green. The circles represent the position of the electrodes C: Example of raw data recorded from a retina pressed against the array. Each trace represents the voltage measured on one electrode. Here we represent traces from 50 electrodes. D: example of template extracted by the algorithm. E: spiking activity of 123 cells recorded simultaneously. Each point represents a spike, and each line is a different cell.

Required profile:

Competences in algorithmic software development and in GPU implementation are mandatory. Basic methods of data processing (signal filtering, data clustering) must be known. No competence in neuroscience or biology is required.

Practical information:

The internship is for six months and is going to take place at http://www.institut-vision.org 17 rue Moreau, 75012 Paris, under the supervision of Olivier Marre. The position is open, the work may start one month after the selection of the candidate (two months for an international profile, depending on administrative delays) for a salary of 2.3 to 2.5K€ / month (charge free) depending on the diploma.

Reference:

Marre, O., Amodei, D., Deshmukh, N., Sadeghi, K., Soo, F., Holy, T.E., Berry II, M.J. Mapping a Complete Neural Population in the Retina. The Journal of Neuroscience, 24 October 2012, 32(43):14859-14873

Contact :

Thierry.Vieville@inria.fr
Olivier.Marre@inserm.fr