Modeling the Human Sense of Touch and Agency in Multi-Modal Networks of Spiking Neurons

Alexandre Pitti*, Hassan Alirezaei+, Yasuo Kuniyoshi**
*ERATO Synergistic Intelligence Project, JST, Japan, alex@jeap.org
+Lab. for Intelligent Systems and Informatics, Dept. of Mechano-Informatics, Grad. School of Information Science and Technology, Univ. of Tokyo, Tokyo, Japan.

Summary

We present a biologically plausible framework to model the human sense of touch and of agency based on the cross-modal integration of spiking neurons. We investigate how networks of spiking neurons represent actions, such as “grasping” an object, by simulating the visuo-tactile integration done in the mirror neuron system (MNS, see Figure 1). The mechanism responsible for synchronizing the neural pairs is the one of spike timing-dependent plasticity (STDP, see Figure 2). It contributes to the visuo-tactile integration to represent one's own action (Figure 4) but also to simulate those of others by re-activating the missing modality (e.g. tactile information) based on cross-modal association (Figure 5). Moreover, the synchrony developed in the sensorimotor maps permits also to detect the signals' coherence and contingency necessary for self-perception (Figure 3). Our results suggest shared representations from low-level sensorimotor processes across multiple modalities permit to represent one's own action and body image but also to simulate those of others.

Motivation

The body image, which means the neural representation of our body, emerges from the dynamical integration of multitudinous signals across different senses. Understanding its neural mechanisms could permit to make robots aware of their own body in order to perceive real-world objects. To this end, we propose to model the MNS which is hypothesized to represent one's own actions. The MNS responds to either visual or somatopic inputs when the person executes one specific action or when she observes it performed by someone else. This exhibits not only how visuo-tactile integration contributes to represent our own action but also how we simulate those of others by using our own representation.

Results

In our experiments, visuo-tactile networks of spiking neurons learn the cross-modal associations of self-observed arm’s motion performed by the experimenter during a grasping sequence using a tactile sensor sheet placed at the object surface (Figure 1) where pixels value and pressure sensors are translated into on/off spikes. At the time-to-contact, the most congruent sensorimotor patterns strengthen their links whereas the incongruent ones weaken theirs. This leads to a first stage of action-oriented sensory fusion exhibiting the attributes of an associative memory (Figure 4). It follows that observing someone else grasping the object (no tactile modality) activates the vision neurons that trigger some others in the somatosensory map and reconstructing the missing tactile modality, the property of MNS (Figure 5).

**Fig. 1:** The network (bottom-right corner) experiences the co-occurrent visuo-tactile perception during grasping (upper-left corner) by receiving the filtered visual information from the camera (bottom-left corner) and from the pressure sensitive device (upper-right corner).

**Fig. 2:** Mechanism of STDP with $A_+ = -A^- = +1$ and $\sigma^- = \sigma^+ = 20\text{ms}$. Each time a post-synaptic neuron fires, its synaptic weights are decreased by $A^-$ (LTD), whereas its synaptic weight is incremented by $A^+$ (LTP) when the synapse receives an action potential.

**Fig. 3:** Contingency detection. The sense of agency arises from the matching between afferent (proprioception) and efferent information (motor prediction). In our model, it corresponds to the synchronization between the incoming input stimulus #2 and the pre-synaptic neuron #1 that activates in the same time the neuron #2. To fire contingently with the input, the pre-synaptic neuron #1 must trigger in advance.

**Fig. 4:** Neural dynamics of the visuo-tactile maps during physical interactions. In red (resp. in cyan) the synaptic activation from the neurons of the vision map (resp. the somato map). The retina anticipates before the time-to-contact the perceptual stimulus in the somatosensory map.

**Fig. 5:** Activity of the neural maps in response to action of grasping observed (no tactile information). The neuron-to-neuron synaptic links in red (resp. cyan) correspond to neural activation having for pre-synaptic neurons those belonging to the vision map (resp. somatosensory map). At the time-to-contact and during handling, the visual map activates the same somatosensory activity as during enaction.