Simulating the Emergence of Early Physical and Social Interactions : A Developmental Route through Low Level Visuomotor Learning

Raphael Braud, Ghiles Mostafaoui, Ali Karaouzene, Philippe Gaussier

Simulating the emergence of early physical and social interactions : A developmental route through low level visuomotor learning

Raphael Braud, Ghiles Mostafaoui, Ali Karaouzene, and Philippe Gaussier

Neurocybernetic team, ETIS
ENSEA, University of Cergy-Pontoise
95302 France
{raphael.braud,ghiles.mostafaoui,ali.karaouzene,gaussier}@ensea.fr
http://www-etis.ensea.fr/index.php/equipe-neurocybernetique.html

Abstract. In this paper, we propose a bio-inspired and developmental neural model allowing a robot, after learning its own dynamics during a babbling phase, to gain imitative and shape recognition abilities leading to early attempts for physical and social interactions. We use a motor controller based on oscillators. During the babbling step, the robot learn to associate its motor primitives (oscillators) to the visual optical flow induced by its own arm. It also statically learn to recognize its arm by selecting moving local view (feature points) in the visual field. We demonstrate in real indoor experiments that, using this same model, early physical (reaching objects) and social (immediate imitation) interactions can emerge through visual ambiguities induced by the external visual stimuli.

Keywords: visuomotor learning, developmental learning, neural networks, human robot interaction

1 Introduction

For future interactive robots, expected to cohabit with us in social environments, the ability to perceive, recognize and learn human actions remain a difficult but crucial question. These new artificial agents must be capable of detecting and predicting human movements to adapt their behaviors in social contexts. Consequently, it seems important to understand the human development process leading to early physical and social cognition in order to build bio-inspired robots permitting safe and intuitive human robot interactions.

A first rising question is how to perceive biological motion which is an important primitive of communication, learning and imitation in human-human interactions. The human ability to perceive biological motion (movements of living beings) is remarkably robust. We can consider that the widespread recognition of biological movements is based on specific characteristics but the exact nature of these features remains not clearly defined, scientists being divided between shape (ventral pathway in the brain) and kinematics (dorsal pathway in the brain) [1]. The roles of each pathway stay confused. A fairly complete neural model summarizing a possible integration of the two pathways for biological motion detection can be found in [2].

Additionally, this remarkable capacity to perceive biological motion seems to appear at early stages of infant development. In fact, psychological studies point out the neonates capacities to imitate simple facial expressions as demonstrated by the studies conducted by Meltzoff and Moore [3]. Considering the very basic visual perception abilities of the newborns we may question the reason of this early emergence (or presence) in human development, of a particular sensibility or competence for human motion perception. In [4], Meltzoff suggested in his "Like Me" theory that humans tends to recognize cross-modal equivalence between perceived actions and a self representation of their own movements. The author argued that this way of recognizing self in others can be a prime step for social cognition as it can be used to analyze, imitate and learn biological movements (others actions). Consequently, biological motion detection can be defined as a "resonance measurement " system that compares proprioception (perception of our own motor dynamics) and exteroception (perception of others movements). The evidence of the motor controllers influence on learning and perceiving motion was described by numerous other psychological studies.In [5], Viviani and Stucchi showed the coupling between motor and perceptual processes while perceiving doted points moving with trajectories respecting the *two third power low*. Recent studies point out a strong link between perceiving and executing movements [6] .This resonance between producing actions and perceiving others movements was also highlighted by the importance of synchrony during human social interactions. Developmental studies acknowledged synchrony as a prime requirement for interaction between a mother and her infant. An infant stops interacting with her mother when she stops synchronizing her movements [7]. These observations also imply the importance of a dynamical loop of treatment between motor production (proprioception) and visual perception.

Keeping in view the importance of motor resonance in social interaction, it has also been widely studied and used to improve human robot communications in particular through the notion of learning by imitation [8][9]. Numerous different works used motor babbling as a starting point to obtain imitative behaviors [10][11]. A possible bio-inspired approach is to rely on mirroring systems which constitute one of the main way to explain imitation behavior [12]. However, many of these works are based on internal models dedicated to specific behaviors. Furthermore, assessing whether there is imitation or not (goal directed Imitation vs simple immediate movement imitation), and consequently guessing what should be the underlying mechanism still a challenging question for developmental studies. Our approach will tend to examine very early mechanisms leading to imitation and reaching behaviors without any use of a specific pre-defined internal model. We will demonstrate that these capacities can possibly emerge through visual ambiguities as proposed in [13] or [14]. Additionally, we will question a possible use of a set of oscillators as motor primitives.

Inspired by the above state of the art, we will investigate in this paper the two main following questions : i) How can a robot gain, from a developmental learning, a cross-modal knowledge linking motor production and visual perception ? ii) How the robot can use this self-expertise to acquire an early social cognition : emergence of imitative capacities and interaction possibilities with the surrounding physical word (humans, objects etc.) ? The precise theoretical and experimental context of the presented work is defined in the next section.

2 Theoretical and Experimental Context

To answer the above theoretical questions, we wish to explore the recognition and the imitation of actions or gestures as a filter that could be built during early interaction (learning to recognize the motor dynamics of self in the perception of the others movements) and not as a pre-defined cascade of ad-hoc filters, leading to the building of the notion of self and others through actions. We defend the idea that intuitive interactions can be seen as an emergent function of sensori-motor dynamics.

To confirm our assumptions, we propose here to simulate, on a robotic platform, the behavior of infants aged approximatively from 0 to 3 months in the specific context of early simple gestures imitation and reaching trials triggered by a visual stimulus. Infants competences regarding the pre-cited context can be coarsely summarized by the following development schedule extracted from [15]:

- Pre-Natal: Grasp reflex on tactile feedback, Proprioceptive-motor mapping (Arm babbling)
- 1 month: Learning of saccade mapping (Moving Eyes and head to targets), Initial mapping of movements and vision (directed but unsuccessful hand movements), Initial goal directed reaching triggered by a visual stimulus without using visual feedbacks to mid-reach movement correction
- 3 months: Reach and miss (with contacts) triggered by visual stimulation, Initial learning of eye-hand mappings, Reaches are visually elicited but without continuous feedback (the gaze still focused on the target and not the hand)
- 4 months: Primitive hand-eye mapping : Successful visual goal directed reaching appears around 3-4 months after birth

We invite the reader to refer to [15] for a complete detailed and referenced development calendar.
We use a minimal setup including a Katana arm, a pan tilt camera, different objects (for reaching trials) and a human partner. Our objective here is to simulate the above behavioral development process by giving the robot the ability :

- to obtain a cross-modal visuomotor knowledge from a babbling step using a very coarse motor controller (oscillators) and low level visual features (optical flow)
- to imitate the human partner on the basis of visuomotor resonance
- to learn its arm shape and to focus its visual attention on it through statistical integrations of visual saccades during the babbling
- and finally to initiate an emerging reaching trial directed by an external visual stimuli (attractive objects) through visual ambiguities

We will experimentally show that all these early capabilities can possibly emerge from very low level visuomotor learning. The developed neural model will be detailed bellow after presenting, in the next section, the considered motor and visual primitives.

3 Motor and visual primitives

3.1 The motor controller

Recent studies suggested that the motor cortex responses during reaching contain a brief but strong oscillatory component, even if the movement itself is not oscillatory [16]. Inspired by these recent neurobiological findings, we will investigate in this study the notion of rhythmic patterns and motor control using oscillators. The other underlying reason behind this choice is to avoid the use of complex motor controllers implying a substantial refined proprioceptive knowledge which is not expected to be found during the early stages of development.

Our motor controller is illustrated in figure 1. Each articulation of the Katana arm is fed by a set of oscillators. Each oscillator is based on a simple model made of two neurons $N1$ and $N2$ [17]. The frequency of the oscillator depends on the three parameters $\alpha 1$, $\alpha 2$ and β :

$$N_1(n+1) = N_1(n) - \beta N_2(n) + \alpha 1 \ \textit{and} \ N_2(n+1) = N_1(n) + \beta N_2(n) + \alpha 2 \quad (1)$$

The control signal feeding each articulation is then obtained by a weighted sum of the different oscillators :

$$\theta_j(t) = \sum_{i=1}^{n} w_i^j . O_i(t). \quad (2)$$

$O_i(t)$ is the output signal of the oscillator i and w_i^j is a weight representing the contribution of the oscillator i to the control signal $\theta_j(t)$ of the articulation j.

3.2 Motion direction-selective neurons for low level visual features extraction

Neurobiological records of cells from V1 and MT brain areas showed that the V1 neurons and most of the MT neurons are sensitive to preferred motion directions, these neurons were called *component direction-selective neurons* by Movshon [18]. A smaller part (20%) of the MT neurons respond best to patterns directions, they are called *pattern direction-selective neurons*.

To simulate these motion selective neurons, we first estimate, for each pixel of the image, the velocity vectors induced by movements in the robot visual field. We used a hierarchical implementation of the classical Horn & Shunk optical flow algorithm [19] based on the works of Amiaz et al [20] . Using the extracted optical flow, we will now define the component direction-selective neurons. The firing of each of these neurons (A_i) is proportional to the angular distance between the visual stimulus (optical flow) and its preferred direction weighted by the motion intensity as : $A_i = \exp^{-(\frac{(\beta-\beta_i)^2}{2\tau_1^2})} \cdot (1 - \exp^{-(\frac{V^2}{2\tau_2^2})})$.
β is the direction of the computed optical flow,β_i is the preferred direction of the direction-selective neuron i, V is the motion intensity, τ_1 and τ_2 are the coefficients regulating the dynamic of the neuron activation respectively for the motion direction and the motion intensity. τ_1 and τ_2 are experimentally set to a value of 20 to optimize the neurons dynamic reacting to the observed range of motion intensities.

Further studies on selective-directional neurons showed that the reactivity range (around the preferred direction) of these neurons is about 40 to 60 degrees [21]. Consequently, we defined 6 different classes of selective neurons reacting for the given preferred motion direction : $0^0, 60^0, 120^0, 180^0, 240^0$ and 300^0. As we are using image coordinates, the Y axis is directed to the south (90^0). For high motion intensities, this type of neuron will respond with a value of 1 if the optical flow direction is equal to its preferred one, its firing will decrease gradually for lower motion intensities and optical flow directions far from the preferred one. Each velocity vector computed (for each pixel of the image) by the optical flow algorithm is then coded by 6 neurons sensitive to different motion directions . As a result, for an image of 640x480 pixels we will obtain 6x640x480 direction-selective neurons.

Pattern direction-selective neurons are then introduced to integrate the responses of the direction-selective neurons. For simplicity and to insure a real time interaction, we define only 6 pattern direction-selective neurons corresponding to the above preferred motion directions (see Figure 1). To obtain the response of a pattern direction-selective neuron sensitive to a motion direction i we integrate the activations of all the direction-selective neurons sensitive to the same motion direction i. A video illustrating the direction selective neurons responses to different movements can be found on our website [1].

4 Visuomotor Learning and low level imitation

In previous studies, we have proposed that low-level imitation (imitation of meaningless gestures) can be an emergent property of a simple perception-action homeostat based on perception ambiguity [13]. Based on this assumption, we

[1] http://www.etis.ensea.fr/~neurocyber/Videos/lowLevel_Reaching/video_directionSelectiveNeurons.avi

present in this section a neural network model (see figure 1 A) permitting to the robot, to learn perceiving its own motion and to imitate a human partner owing to visual ambiguities .

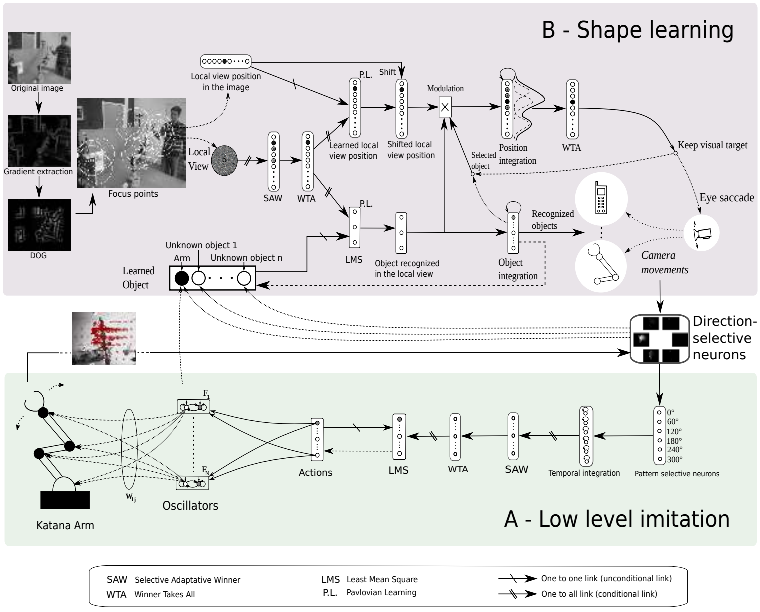

Fig. 1. Global Architecture, Part A : Neural model for motion learning and immediate imitation, Part B : Neural model for object and arm learning and recognition

Here, the motor controller is composed by 9 oscillators with 3 different frequencies and 3 different phase shifts. The control signal for each robot joint j is obtained by summing the different oscillators output modulated by the weights w_i^j (see equation 2). As a first trial, the weights w_i^j are randomly chosen, the robot starts to move according to this set of parameters. Unfortunately, even with this small number of oscillators, most of the obtained actions were difficult to analyze and not really biologically plausible (in particular because of the mechanical characteristics of the Katana arm). For simplicity sake, in this experiment we decided to settle the weights w_i^j to obtain only 3 different rhythmic actions : an horizontal motion (A1) and two diagonal ones (A2 and A3).

The model illustrated figure 1-A works in two phases :

– First, during a very simple babbling step the robot learn its own dynamics. The the robot starts moving by altering randomly the three actions A1,A2 and A3. A Selective Adaptive Winner (SAW) which is an ART-based neural networks, is fed by a time integration of the 6 pattern direction-selective neurons responses which react differently to the perceived robot's arm actions. Depending on the vigilance parameter of the SAW, if the new inputs are too different from the neurons encoding the previous ones, new encoding neurons are recruited. A Winner Takes All (WTA) is then used to select the relevant SAW neurons encoding at best the inputs from the pattern-directional neurons. These selected neurons represent the unconditional inputs of an LMS (Least Mean Square) network which learn to associate the so encoded visual stimuli (optical flow) with the motor actions represented by the sets of w_i^j parameters.

- After the learning phase, when a human starts moving in the visual field of the robot, the selective-direction neurons are activated accordingly to the different motion directions present in the visual stimuli. The integration made by the 6 pattern selective-direction neurons are then representative of the human motion visual pattern. As stated before, we suppose that the arm controller is an homeostatic system trying to maintain a coherence between the produced and the perceived actions, whether those perceived actions are performed by the robot itself or by a human. If the visual pattern induced by the human movements is close to one of the previously learnt movements, the LMS triggers the corresponding oscillators parameters w_i^j. Thus, the robot will start lunching the corresponding action and consequently imitate the human movement. A video of this experiment can be found on our website[2].

5 learning the arm shape during the babbling phases

As detailed section 2, we are aiming to simulate the emergence of a visual goal directed reaching (section 6). We will start, in this section, by explaining how to obtain an initial learning of eye-hand mappings using eye saccades. Our objective is to make the robot recognize its arm shape and to focus its visual attention on it. For doing that, we will use and define a neural model for object recognition inspired by the works in [22] and [23]. The general principal of this model is to learn local views of the objects in the basis of point of interest detection (focus points simulating eye saccades). As illustrated Fig. 1-B, the spatial gradient information is first extracted from the grayscale images. The resulted image gradient is then convolved by a DOG (Difference Of Gaussian) filter. The output of this process is a saliency map which highlight regions in the image having a local structure shaped as corners. Local maxima are then selected from this saliency map.

Local views collecting the pixel around each detected interest points (here on a radius of 20 pixels) are then extracted and filtered by a log polar transform in order to be robust to scale changing and rotational variations. The filtered local views feed the Selective Adaptive Winner (SAW), if the new inputs (local views) are too different from the previous ones, new encoding neurons are recruited. A Winer Take All (WTA) is then used to select the winning local views.

The model presented in (Fig. 1 B) can after be divided into two parts. The recognition of what is the object, and the localization of "where" is the object. A

[2] http://www.etis.ensea.fr/~neurocyber/Videos/lowLevel_Reaching/video_lowLevelImitation.avi

first LMS (Least Mean Square) algorithm is used for the "what" pathway to learn the local views associated to each object. The number of neurons in the LMS is then corresponding to the number of possible objects to learn. In the where pathway, two LMS are used to associate the object center position respectively on the x and y axis relative to the local views belonging to it. As presented figure 2, after the learning phase, each selected local view (point of interest) has its own prediction of the object center. If most of them predict the same position, the object is well recognized. In the opposite case, several positions of the object center will be predicted without a majority vote permitting to identify a winner (see figure 2). If an object is learned at a given position and detected in another one, the output of the LMS shift the learn position relative to the actual position allowing to predict the object position.

The previous model is then used to learn the robot its own arm shape without any added a priori knowledge. To do so, during the babbling phase, the robot will also start to detect feature points in the visual field using the general model for object recognition previously described. Additionally, the saliency maps resulting from the DOG filtering is modulated by the motion intensities (optical flow). Thus, if we assume that the robot will be able to perceive its moving arm during the whole babbling step, statistically, the detected feature points (or focus points) will mostly belong to the arm of the robot. The robot will consequently "statistically" learn the shape of its arm rather than other objects from the background.

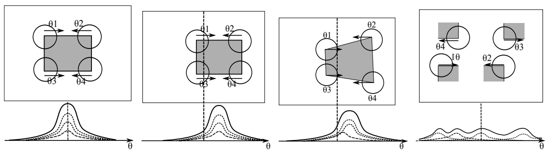

Fig. 2. Schematic example of object position estimations

6 Visual ambiguities and emergence of a visual directed reaching

After the learning phases (babbling), our robot is able to recognize its arm and to locate its visual focus of attention on it. Lets now consider, besides the robot arm, the presence of an added visual stimulus attracting the robot's visual attention. To simulate that, we will first simply and similarly use the previous shape learning neural model to make the robot learn a new object. The considered object is then "shacked" in front of the robot while its arm is not moving, consequently the learned local views will statistically belong to this new added object. As explained in the previous section, the model we used for object recognition can locate objects positions relative to the center of the image. Thus, if

an object (or the robots arm) is located in the images, we can use its predicted position to shift the pan-tilt camera to center the recognized object in the image using a neural field. If two recognized objects are seen (shared visual attention) by the robot, its pan-tilt camera move alternatively from an object to another, this process simulate human eye saccades.
As a consequence, if the robot can perceive its arm and focuses its visual attention on it, and if simultaneously a human shows a learned object somewhere in the visual field, the camera start moving alternatively from the arm to the object. An optical flow is then generated because of this shared visual attention. By using the neural model for imitating actions (see section VI and figure 1), this induced optical flow will be coded by the directional and pattern selective neurons leading the LMS to trigger the corresponding motor primitive which produces an oscillatory arm movement in the direction of the located object.
An experimental example is illustrated in figure 3. After the learning steps, the robot can perform and imitate 3 different movements as described section VI. It also learned to perceive and focus its attention on its arm and another external object (here, an Aibo robot). 3-A represent three snap shots from the experiment; 3-B illustrate the activities of the LMS neurons dedicated to object recognition (green for the arm and red for Aibo), the threshold deciding if an object is recognized or not is represented by the doted straight line; 3-C and D shows the X and Y neural fields activities (blue line) and the Pan and Tilt camera movements (doted green line); 3-E highlights the 6 pattern selective directional neurons firing; 3-F illustrate the activities of the LMS neurons dedicated to the 3 actions recognition (doted green, red and blue lines), the plain green, red and blue lines represent the lunched actions triggered when the LMS activity is over the threshold (straight black doted line); finally on 3-G and H we can see the evolution of the arm pliers real cartesian x and y positions deduced from the Katana physical model.
Lets now consider the whole scenario. First the robot perceive only its arm. The arm is recognized and located in the center of the image. The learned object (Aibo) is then presented by a human interactant on the left upper side of the image. Consequently, the arm and the object are simultaneously recognized and located (brown area of the figure 3 B). Because of this shared visual attention, the pan-tilt camera (controlled by the Neural Fields) start moving toward the recognized object (figure 3 C and D) inducing an optical flow in the visual field. This optical flow in then encoded by the directional and pattern directional selective neurons (figure 3 E). As the camera is moving from the right to the left upper side (toward Aibo), it generates an inverted optical flow directed to the right and down side (O^o and 60^o because of the inversion of the Y axis). The directional neurons trigger the corresponding learned oscillatory arm movement (action 2 on figure 3 F). Consequently, as proved in figure 3 G and H, the robots arm start moving toward the detected learnt object (Aibo) : left direction on the x axis and upward on the y axis.
This behavior simulate an emerging visual goal directed reaching trial induced

by visual ambiguities. A video of this final experiment can be find on our web site[3]

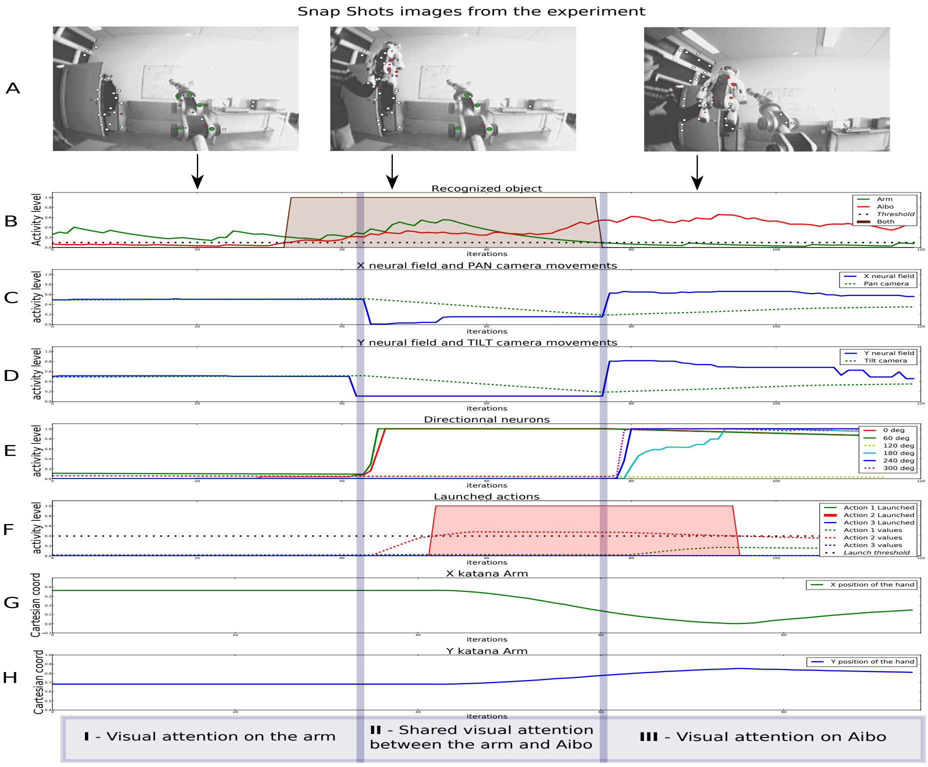

Fig. 3. Experimental results

7 Conclusion

We presented here a developmental approach for investigating the emergence of early physical and social interaction from a learning stage of visuomotor cross-modal knowledge by using a neural network model. We simulate here, on a robotic platform, the developmental behavior of infants aged approximatively from 0 to 3 months in the specific context of initial simple gestures imitation and early reaching trials triggered by an external visual stimulus. First, during a babbling step, the robot learn to associate its motor primitives to the optical flow induced by its own arm. in parallel, the robot statistically learn its arm shape by modulating, using motion intensities (optical flow), the feature points

3 http://www.etis.ensea.fr/~neurocyber/Videos/lowLevel_Reaching/video_LowLevelReaching.avi

detection and the local views learning . Similarly, if the arm stops moving, the robot can learn new shacked external objects in the visual field. After the learning phase, if a human start to move in the visual field, his movements induce visual ambiguities which make the robot start imitating the human as it will try maintaining the balance between the visual stimuli and the motor controller as learned during the babbling step. If the robot locate its arm and another learned object in the visual field, its camera start moving to center alternatively the arm and the detected object simulating ocular saccades. As for immediate imitation, the camera oscillations induce ambiguous optical flow making the robot initiating arm movements toward the located object. The efficiency of the proposed neural architecture was demonstrated by experiments in a real indoor and non constrained environment using a Katana arm and pan-tilt camera.

Despite the interesting obtained results, numerous outstanding questions remains. For example, how to the fill, from a developmental approach, the gap between this early coarse oscillatory motor control to a more refined one leading to more precise interactions and imitation games (object grasping) ? How to gain a better knowledge about spatial information and peripersonal space leading to social cognition? These harsh problematics among others related to physical and social development remain, obviously, opening questions. Nevertheless, our experimental approach demonstrates that early simple physical and social interactions can possibly be mediated by visual ambiguities through visuomotor learning rather than complex representations of "self" versus "Other" especially at an early stage of development.

Our short-term future works are aiming to use this model on an hydraulic robot in order to obtain a realistic force controlled arm leading to more "natural" movements. In order to maintain the interaction and to give the robot capabilities to learn new and more complex movements, we are also planing to introduce the notion of synchrony detection between the oscillators controlling the arm and the visual stimuli as proposed in our previous work on simple oscillatory movements [24]. Thus, a refined interaction can be obtained during immediate and differed imitation games. More precisely, the weights modulating the influence of each oscillator on the arm joints (set to a fixed value in this article) must be learnt during bidirectional imitation games.

References

1. Lange, J., Lappe, M. (2007). The role of spatial and temporal information in biological motion perception. Advances in Cognitive Psychology, 3(4), 419.
2. Giese, M. A., Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. Nature Reviews Neuroscience, 4(3), 179-192.
3. Meltzoff, A. N., Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. Science, 198(4312), 75-78.
4. Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. Developmental science, 10(1), 126-134.
5. Viviani, P., Stucchi, N. (1992). Biological movements look uniform: evidence of motor-perceptual interactions. Journal of Experimental Psychology: Human Perception and Performance, 18(3), 603.

6. Casile, A. and Giese, M.A. (2004) Possible influences of motor learning on perception of biological motion. J. Vis. 4, 221a.
7. Nadel, J., Carchon, I., Kervella, C., Marcelli, D., Rserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. Developmental science, 2(2), 164-173.
8. Breazeal, C., Scassellati, B. (2002). Robots that imitate humans. Trends in cognitive sciences, 6(11), 481-487.
9. Lopes, M., Santos-Victor, J. (2007). A developmental roadmap for learning by imitation in robots. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 37(2), 308-321.
10. Rao, R. P., Shon, A. P., Meltzoff, A. N. (2004). A Bayesian model of imitation in infants and robots. Imitation and social learning in robots, humans, and animals, 217-247.
11. Demiris, Y., Meltzoff, A. (2008). The robot in the crib: A developmental analysis of imitation skills in infants and robots. Infant and Child Development, 17(1), 43-53.
12. Oztop, E., Kawato, M., Arbib, M. (2006). Mirror neurons and imitation: A computationally guided review. Neural Networks, 19(3), 254-271.
13. Gaussier, P., Moga, S., Quoy, M., Banquet, J. P. (1998). From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. Applied Artificial Intelligence, 12(7-8), 701-727.
14. Nagai, Y., Kawai, Y., Asada, M. (2011, August). Emergence of mirror neuron system: Immature vision leads to self-other correspondence. In Development and Learning (ICDL), 2011 IEEE International Conference on (Vol. 2, pp. 1-6). IEEE.
15. Law, James and Shaw, Patricia and Earland, Kevin and Sheldon, Michael and Lee, Mark H, "A psychology based approach for longitudinal development in cognitive robotics", Frontiers in Neurorobotics, 2014, Volume 8, number 1
16. Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., Shenoy, K. V. (2012). Neural population dynamics during reaching. Nature.
17. Revel, A., Andry, P. (2009). Emergence of structured interactions: From a theoretical model to pragmatic robotics. Neural networks, 22(2), 116-125.
18. Movshon, J. A., Adelson, E. H., Gizzi, M. S., Newsome, W. T. (1985). The analysis of moving visual patterns. Pattern recognition mechanisms, 54, 117-151.
19. Horn, B. K., Schunck, B. G. (1981, November). Determining optical flow. In 1981 Technical Symposium East (pp. 319-331). International Society for Optics and Photonics.
20. T. Amiaz, E. Lubetzky, and N. Kiryati. Coarse to over-fine optical flow estimation. Pattern recognition, 40(9) :24962503, 2007.
21. K. Hol and S. Treue. Different populations of neurons contribute to the detection and discrimination of visual motion. Vision research, 41(6) :685689, 2001.
22. Gaussier, P., Joulain, C., Zrehen, S., Banquet, J. P., Revel, A. (1997). Visual navigation in an open environment without map. In Intelligent Robots and Systems, 1997. IROS'97., Proceedings of the 1997 IEEE/RSJ International Conference on (Vol. 2, pp. 545-550). IEEE.
23. Lepretre, S., Gaussier, P., Cocquerez, J. P. (2000). From navigation to active object recognition.
24. Hasnain, S. K., Mostafaoui, G., Gaussier, P. (2012). A synchrony-based perspective for partner selection and attentional mechanism in human-robot interaction. Paladyn, 3(3), 156-171.