

A Synchrony-Based Perspective for Partner Selection and Attentional Mechanism in Human-Robot Interaction

Syed Khursheed Hasnain^{1*},
 Ghiles Mostafaoui^{1†},
 Philippe Gaussier^{1‡}

¹ Neurocybernetic team,
 ETIS, ENSEA,
 University of Cergy-Pontoise,
 95302 France.

Received 14-12-2012

Accepted 10-04-2013

Abstract

Future robots must co-exist and directly interact with human beings. Designing these agents imply solving hard problems linked to human-robot interaction tasks. For instance, how a robot can choose an interacting partner among various agents and how a robot locates regions of interest in its visual field. Studies of neurobiology and psychology collectively named synchrony as an indispensable parameter for social interaction. We assumed that Human-Robot interaction could be initiated by synchrony detection. In this paper, we present a developmental approach for analyzing unintentional synchronization in human-robot interaction. Using our neural network model, the robot learns from a babbling step its inner dynamics by associating its own motor activities (oscillators) with the visual stimulus induced by its own motion. After learning the robot is capable of choosing an interacting agent and of localizing the spatial position of its preferred partner by synchrony detection.

Keywords

Human Robot Interaction · Synchrony · Focus of Attention · Partner Selection · Dynamical Systems · neural networks

1. Introduction

Traditionally, robots are designed for a specific set of tasks in deterministic and highly constrained environment. The majority of these robots are used in the industry where the accuracy and speed are the priority. With technology and artificial intelligence advancements, this field is now looking to further challenges in non-constrained human social environments implying more flexible control systems making it possible to build adaptive agents able to learn dynamically and to achieve new tasks [1]. These future robots may co-exist with humans as a part of our social life and are expected to behave as companions by sharing the same working place in offices, factories and homes [2].

As the robots begin participating in human social environments, agency and sociality become very important [3][4]. Indeed, designing robots that dynamically interact with humans implies solving tremendous harsh questions. Among these issues, we will discuss two questions. First, how can robots select an interacting partner among many interactants? Second, how can they focus their attention on specific regions of interest? In other words, how can the robots be able to discriminate the relevant visual stimulus?

To tackle these two problematics, we will study the notion of "synchrony", and more precisely, "unintentional synchrony" which was suggested by both psychological studies of dyadic interactions and neurobiological data on motor coordination as an important parameter for human-human interaction.

In this paper, we study, in a developmental perspective, the unconscious or the unintentional synchronization during human-robot interaction. The presented neural network architectures allow the robot to first learn his own movements (babbling step) by associating its sensori-motor (proprioception) information with the induced visual stimuli (optical flow), second, be capable of automatically selecting and locating (focus of attention) a partner among many interactants using synchrony detection. In other words, we use immediate synchronous imitation (adaptation of other's synchronous behavior) as a communication tool. The robot imitates the other agent if it detects synchrony between its internal dynamics and the interactant's movements.

The paper is organized as follows: After a global overview on synchrony in dynamical systems in section 2, the experimental setup and methods are described in section 3. In section 4, a first simple architecture for human-robot interactions is presented. The architecture for partner selection on the basis of synchrony detection is explained in section 5. In section 6, the model of attentional mechanism or Focus of Attention (FOA) is detailed. Finally, before concluding, the experimental results are shown in section 7.

2. Synchrony in dynamical systems

Synchronization can be defined as an adjustment of frequency of oscillating objects due to coupling (energy) between them [5]. It is a non linear phenomena and it is common in physical and biological systems where two or more oscillating systems interact with each other and start to move together by adjusting their frequency [6][7]. The earliest known scientific discussion of synchronization started in 1657 when the famous Dutch physicist Christiaan Huygens observed and described the synchronization phenomenon. He discovered that two pendulum clocks mounted on a common base (beam) synchronize and move at the same frequency in the opposite direction. He did

*E-mail: syed-khursheed.hasnain@ensea.fr

†E-mail: ghiles.mostafaoui@ensea.fr

‡E-mail: gaussier@ensea.fr

not only give the exact description but also the explanation of mutual synchronization: clocks were synchronized (anti phase) because of the coupling through the beam (imperceptible vibration of the beam) [5][8]. Blekhman examined experimentally a similar system and found two possible stable synchronization states (in-phase and anti-phase)[9]. J. Pantaleone studied a variant of Christiaan Huygens's original system, he used two pendulum metronomes (with almost same frequency) and deflected the two pendulum bobs in the opposite direction, after few seconds of asynchronous movements the system evolved to steady state in-phase synchronization. Pantaleone observed some very interesting properties of synchronization during his experiments; 1) The anti-phase synchronization state can be made possible in metronomes system either by enhancing the damping (weak coupling) associated with the base motion or by going to very large oscillation frequencies. 2) If the natural frequencies of the two pendulums differ by more than a few percent, synchronization will not occur. 3) If the natural frequencies of two metronomes are different significantly but within an acceptable range, it leads to a synchronized state but with constant time lag between two oscillators [6]. As we will see in the next sections, most of these observations made by Pantaleone are also verified in our experiments.

Synchronizations are ubiquitous in nature. From a biological view, synchronization in coupled oscillators can be seen through out the natural world and specially conspicuous in living things. A good example of synchronization in nature is the flashing of fireflies. Thousands of coupled oscillator can be seen in the form of fireflies (paradigm of pulse coupled oscillators). At night, these insects gather at some place and flash in synchrony. Each insect has its own rhythm, they interact each other only when one sees the sudden flash of another and shift its rhythm [10]. Cricket chirping is another example of pulse coupled oscillators [11]. Many examples of the coupled oscillator can be found in biological systems: in the eighteen century, Jean-Jacques Dortous de Mairan discovered the circadian rhythm by observing the day and night oscillation of haricot bean leaves [12]; Birds in flocks synchronize take-off and landing [13]; Male and female mosquitoes synchronize wing beats [14]; pacemaker cells beat in the heart together [15] etc.

2.1. Synchrony in social interactions

Synchrony in dynamic systems, such as social systems, is a reciprocal adaptation of behaviors (temporal structures) between interactants [16]. During human communication some nonverbal languages like gestures, facial expressions and nodding are also involved.

As stated by L. Glass: "Complex bodily rhythms are ubiquitous in living organisms" [17]. But the arising query is how these synchronized behaviors or rhythms interact in human social communications [18]. For instance, singing in unison is a highly synchronized form of social interaction. Viktor et al. examined and unveiled group dynamics underlying temporally coordinated actions (choir singing). The authors revealed that phase synchronization in heart rate variability (HRV) and respiration was much higher during unison singing. They concluded that respiratory and cardiac coupling patterns render the physiological foundation for interpersonal temporally coordinated actions [19].

Fred Cummins augmented this discussion with an instance of aperiodic synchronization of complex action. He found aperiodic synchronization of complex action in his experimental task of synchronous speaking [20].

Several researches in psychology took into account the concept of synchrony in early human social interactions, they studied the temporal coordination of nonverbal behaviors as body movements, vocalizations, gaze and many others [21]. According to psychologists, when two hu-

mans interact or communicate with each other, they do not only use speech to convey the content of a message but also employ a large variety of non-verbal behaviors [22], for example: hand movements, pauses during discussion, facial expressions that show their attitude and their level of attention towards the partner etc. An important parameter of non-verbal communications is the temporal correlation or synchrony between the behavioral states of the interactants [23]. An interesting aspect of these synchronized behaviors during human interactions is its unintentional nature.

Moreover, developmental psychology also considered synchrony as an essential parameter for interactions between mothers and their children. In fact, if the mother loses synchrony, infant struggles to sustain the interaction [24]. Infants synchronize their legs motion with adult speech [25]. In addition, synchrony detection mechanism in young infants plays a pervasive role in learning and cognitive development [26] (word learning [27], object interaction skills [28], self-awareness and control [29], learning related to self [30] etc.)

From a neurobiological point of view, neuro imaging techniques enable us to observe synchrony from local scale (brain's local field potential and between distant brain regions) to inter-individual scale (in a social setting) [31]. Several studies used fMRI and EEG to record the brain activities during social interaction. Hasson et al. used fMRI neuro imaging to scan the brain activities of 5 participants (isolated) during watching a popular movie, strong functional and anatomical similarity was found in individuals who were immersed in the same natural settings [32]. Stephens et al. scanned the brain of a speaker and a listener. The authors revealed spatial and temporal correlation between the two brain signals during the speakers monologue. [33].

Recent work of Dumas et al. [34] using hyperscanning has revealed the emergence of inter-brain synchronization across multiple frequency bands during social interaction. The authors selected 11 dyads and recorded the brain activities during social interaction, more precisely, during spontaneous exchanges (between two participants) of intransitive bi-manual movements. Examining the phase synchronization between the two brains, it was revealed that these synchronous exchanges exhibit the emergence of an interindividual brain-web (linked to the sensorimotor information) across several frequency bands. Symmetrical patterns were found in low frequency bands (may be due to coordinated dynamics of hand movements) and asymmetric in higher frequency.

Interpersonal motor coordination studies also point out this fact of very low level mechanism of unintentional synchronization or communication. For instance, when people walk together, they synchronize unconsciously their foot steps by steadily regulating their step-size or frequency [35]. Neda et al. investigated the dynamics of rhythmic applause and the development of synchronized clapping, they observed that after a few seconds of random clapping people synchronizes gradually [36]. Moreover, in [35] Issartel et al. analyzed interpersonal motor coordination of participants instructed to not synchronize with each other. Interestingly, subjects could not abstain unintentional coordination. Consequently, we can deduce that immediate unconscious motor coordination could not be avoided when the subjects share the visual information.

Recently, Varlet et al. investigated social motor coordination of patients suffering from schizophrenia. The participants are advised to oscillate hand-held pendulums from the wrist. The results demonstrated that patients intentional motor coordination was altered while their unintentional motor-coordination was retained. This study concludes that unintentional motor-coordination preserves even for subjects affected by social interaction disorder [37].

2.2. Synchrony in Human-Robot Interaction

Taking into account the importance of synchrony in human-human interactions, numerous works used synchrony in the field of Human-Robot interaction. Andry et al. highlighted the importance of a learning rule associated with synchrony prediction, they presented a biologically inspired architecture proposing rhythm detection as an internal reward for learning [38]. Prepin et al. proposed an architecture to detect the level of synchrony between a robot ADRIANA (Adaptable Robotic for Interaction ANALysis) and a human. It was employed as a reinforcement signal for learning [39]. Blanchard et al. improve the reactivity of two robots by a velocity detection system capable of synchronizing the movements of the agents [40]. Qiming Shen et al. studied motor interference and motor coordination in human-humanoid interactions for different types of visual stimuli (robot, pendulum and moving dot). The authors concluded that participants tended to synchronize with agents having a better appearance, which means that a robot perceived as close as possible to a social entity may facilitate human-robot interaction [41]. In the same line, Marin et al. showed motor resonance between humans and artificial agents (robots) could enhance and optimize the social competence of HRLs [42]. Michalowski et al. developed a dancing robot to analyse the properties and significance of synchronized movement in general social interaction [43]. Ikegami and Iizuka [44] used the genetic algorithm technique and showed that coupling and turn-taking between two agents are sensitive to the dynamics of interaction. Crick et al. programmed a robot for drumming (with human drummers) by integrating multiple sensors input (oscillators). They showed that precise synchronization between humans and robots can be achieved by fusing multiple sensors input although the incoming data is imperfect [45]. Inspired by the infants development, Rolf et al proposed a model of bottom-up visual attention guided by audio-visual synchrony [46]. Moreover, Hafner and Kaplan presented the idea of interpersonal maps. These maps are the geometrical representation based on one's own behavior and the others. Using these maps, different types of interactions (for instance imitation) can be detected [47].

In the line of this state of the art, we assumed that unintentional synchronization could play a pervasive role for initiating human-robot interactions.

3. Materials and methods

A minimal experimental setup is used to avoid complexity (Figure 1) and focuses on the one problem (real size application is the focus of ongoing work and introduce a lot of other issues that we will discuss in the conclusion). The experimental setup includes a basic automaton, a Nao humanoid robot and a human partner. Practically, Nao robot has the capability of moving with multiple degrees of freedom but we used one dimensional arm movement only (up-down, one degree of freedom). The basic automaton (one degree of freedom) has the ability to oscillate at different frequencies. Instead of Nao's camera (frame rate limited to 10 Hz through an ethernet connection) an external camera is used to allow our architecture to work on 30 Hz.

To analyze synchrony, we need to investigate the dynamics of interaction between two signals. To do so, we use the Phase Locking value (PLV) which is a practical method presented by Lachaux et al. for detecting EEG synchrony in a band of frequencies [48]. The PLV for two signals is defined by

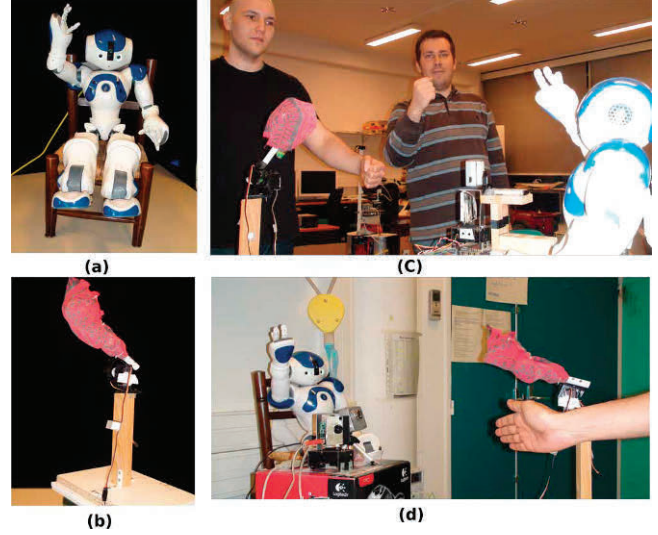


Figure 1. Experimental setup. (a) Nao (b) a basic automaton (1 degree of freedom) (c) & (d) Overall setup for human-robot interaction.

$$PLV_{x,y} = \frac{1}{N} \left| \sum_{t=1}^N \exp(i(\phi_x - \phi_y)) \right| \quad (1)$$

Where N is the number of samples and $\phi_x - \phi_y$ is the phase difference between two signals. The PLV value is close to 1 for synchronized signals and approaches 0 otherwise.

Videos of our experiments can be found on: <http://www.etis.ensea.fr/neurocyber/Videos/synchro/>

4. Human Robot Interaction using optical flow

As a starting point for human-robot interactions, a dynamical interaction model was developed to synchronize two agents influencing each other. Specifically, the proposed architecture permits the Nao robot to adopt the frequency and phase of its partner of interaction. A classical optical flow algorithm is used to estimate the velocity vectors of the perceived motions (in the robots visual field) [49]. These estimated vectors act as visual stimuli and inputs to the proposed architecture.

4.1. The oscillator Model

As shown in Figure 2(a) (dotted box), in our architecture, an oscillator module [50] controls the Nao's arm movements. It can also be seen as the signal representing the Nao's internal dynamics. It consists two neurons $N1$ and $N2$ inhibiting each other proportionally to the variable β . The oscillating frequency is a function of the variables $\alpha1$, $\alpha2$ and β :

$$N_1(n+1) = N_1(n) - \beta N_2(n) + \alpha1 \quad (2)$$

$$N_2(n+1) = N_1(n) + \beta N_2(n) + \alpha2 \quad (3)$$

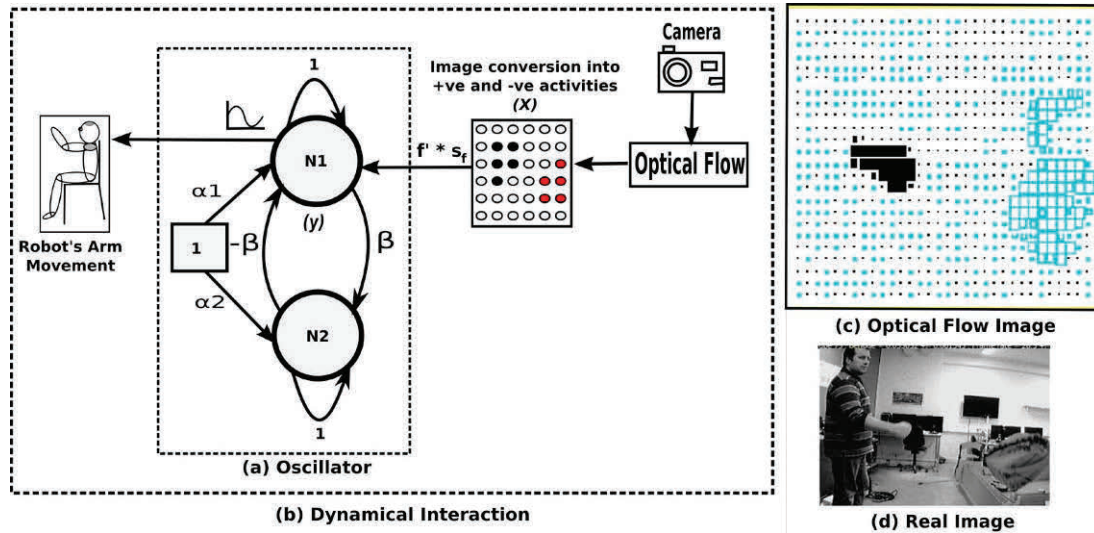


Figure 2. (a) Oscillator model (b) Dynamical Interaction model (c) Positive and negative activities deduced from optical flow: Upward movements (positive) are shown as fulfilled black color pixels, downward movements (negative) are shown by unfilled pixels. (d) A real image captured by the camera

In addition, reservoir of oscillators (echo state network) could be used to work with a larger range of frequencies.

4.2. Dynamical Interaction Model

As shows in Figure 2(b), the oscillator is connected to Nao's arm and oscillates at its own frequency and amplitude. Motion in the visual field of Nao is estimated by an optical flow algorithm, velocity vectors are then converted into positive and negative activities (Figure 2(b)). If the perceived movements are in the upward direction, the oscillator gets the positive activity and its amplitude increases on the positive side depending on the induced quantity of energy (motion). On the contrary, if the negative activity is perceived, the amplitude goes down. Figure 2(c and d) is a snapshot taken during the experiment illustrating positive and negative activities in the visual field deduced using the optical flow velocity vectors. There are two moving objects in the field of view of Nao. One moves upward and induces positive activities (shown by filled black color pixels) whereas the other moves downward and induces negative activities (unfilled pixels). These positive and negative activities can be learnt by the robot and modify the oscillator accordingly. When an agent interacts with a motion frequency close to Nao's frequency, Nao's oscillator can be modified (frequency and phase) within certain limits. Otherwise, it continues to its default frequency. The mathematical equation of the oscillator can be rephrased as:

$$N_1(n+1) = N_1(n) - \beta N_2(n) + \alpha_1 + f' \quad (4)$$

Where f' is the induced energy computed by an integrating, over time, all the active pixels in the image. This function can be defined as a time representation of the quantity of energy produced by the positive and negative activities in the image. f' may be negative or positive. It's worth noticing that this direct feeding of the motor controller by (f') reflects the influence of the visual stimuli on Nao's actions which makes the robot changing its behaviors (regarding to the motion in the visual field) in an unintentional manner.

It's important to note that the influence of the visual stimuli f' , is weighted by a coupling factor (or coupling scaling factor) S_f (see Figure 2(b)). Higher values of S_f induce a bigger influence of the visual stimuli on the robot motor controller. Consequently, the robot will be able to synchronize with a larger range of frequencies.

Figure 3 and Figure 4 illustrate the results of our experiments for two cases: first for Human-Nao interaction and second for Automaton-Nao interaction. Figure 3(a) details the motion signals of both Nao's arm and the human (arm or hand) while entraining each other. First, movements of both agents are not synchronized. Consequently, PLV (measure of synchrony) shows lowest value (see Figure 3(b)). As shown in Figure 3(a) and 3(b), during the interaction between Nao and the human both are synchronizing successively as time passes. The increasing tendency in PLV values reflects the emerging synchrony between the two agents. Figures 3(a) and 3(b) also illustrate that, after a certain time, both interactants are fully synchronized and the corresponding PLV values are at its maximum.

Lissajous curves show the dynamics of the two signals. If the signals have the same frequencies, their Lissajous curve will be a straight line with an angle of 45 degree with the horizontal axis. If the signals have the same frequencies but with a small phase shift, the lissajous curve take the form of an ellipse. Gradually, increase in the phase shift makes the ellipse wider, with a phase shift of 90 degree, the curve becomes a circle. In our case, if the two signals are identical and synchronized with each other with small phase shift then the Lissajous curves between the signals of robot and the interactant should be an ellipse. Figure 3(c) shows Lissajous curve between $N(t)$ (Nao's oscillation) and $H(t)$ (human's movements). The elliptic shape of the curve indicates that both signals are almost identical. For this experiment, Nao's standard frequency was 0.428 Hz and human oscillations were between 0.4615 Hz (7.8% higher than Nao's frequency) to 0.476 Hz (11% higher) and the scaling multiplying factor (S_f) for f' was 0.15.

Figure 4 shows the interaction and synchronization between Nao and the automaton. Figure 4(a) shows the motion signals of both automaton and Nao's arm. Similarly (like for human robot interaction), Automa-

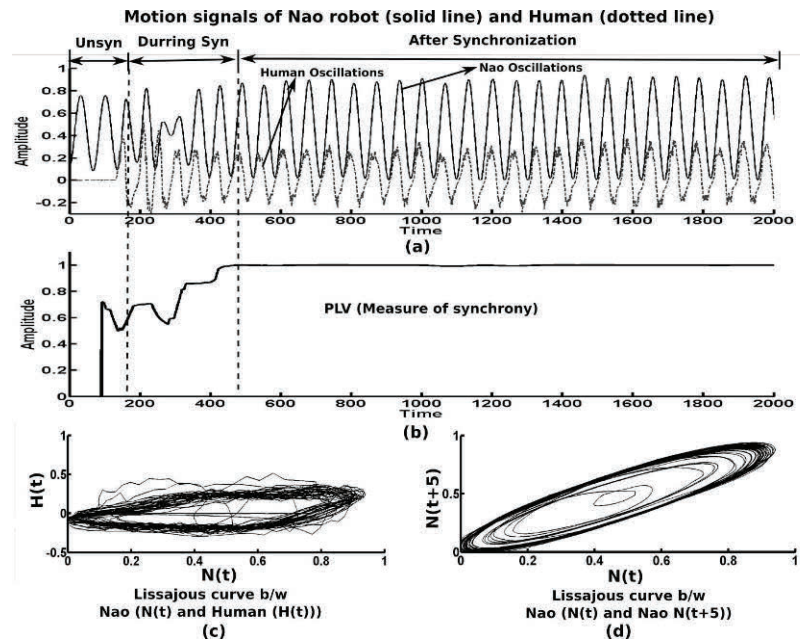


Figure 3. Robot (Nao)-Human interaction: (a) Motion signals of the human (dotted line) and the Nao's modifiable oscillator (solid line). (b) PLV measurement. (c) Lissajous curve $N(t)$ (Nao's oscillation) and $H(t)$ (Human's movements). The elliptic shape of the curve indicates that both signals are almost identical. (d) Lissajous curve between two different time values of Nao ($N(t)$ and $N(t+5)$).

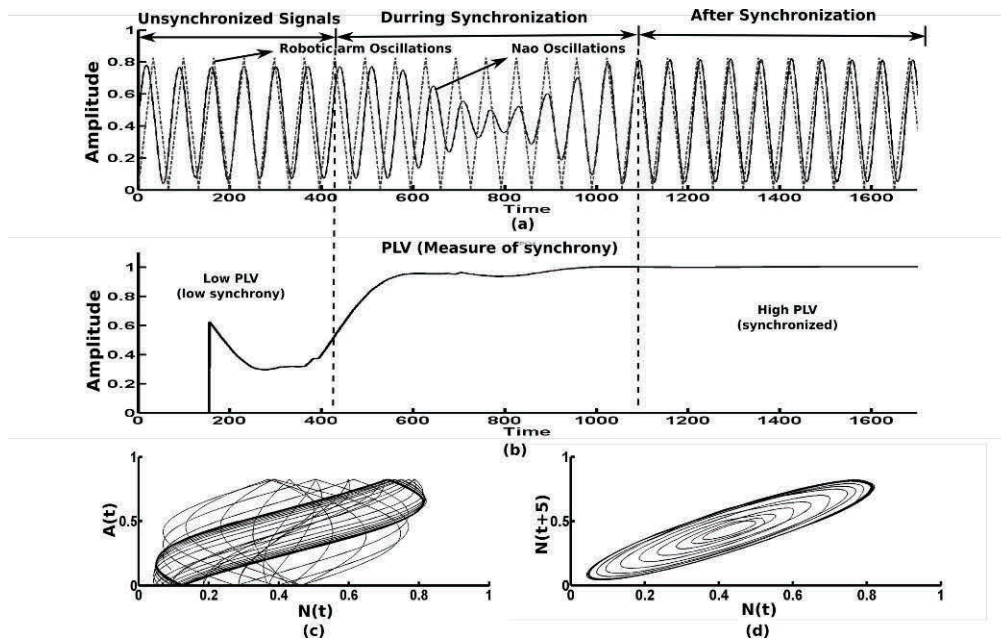


Figure 4. Nao-automaton interaction: (a) Motion signals the automaton and the Nao's modifiable oscillator (b) PLV measurement (1 for complete synchronization and 0 for completely asynchrony) (c) Lissajous curve between $N(t)$ (Nao's oscillation) and $A(t)$ (automaton's movements). (d) Lissajous curve between two different time values of Nao ($N(t)$ and $N(t+5)$).

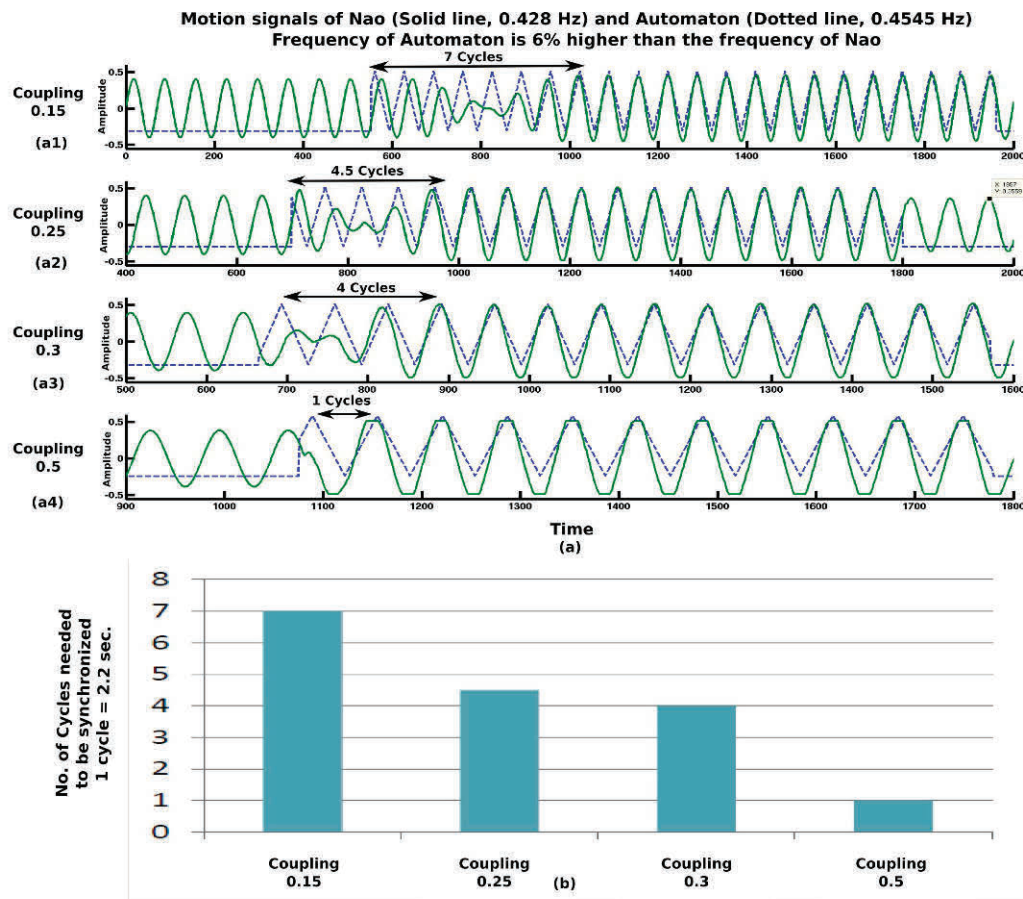


Figure 5. (a) and (b) detail the relationships between the quantity of coupling energy and the synchronization time of two agents. Figure 5(a) shows the motion signals of Nao (solid line, 0.428 Hz) and Automaton (dotted line, 0.4545 Hz), it can be observed that with low scaling for coupling energy (0.15) two agents take a long time (4.5 cycles) to be synchronized (fig 5(a1)). As the scaling factor for coupling energy increases to 0.25 and to 0.3, synchronization time decreases to 4.5 cycles to 4 cycles respectively (fig 5(a2,a3)). Further increases in coupling energy (0.5) reduce the synchronization time to only 1 cycle. However, this higher energy induces saturation (clipping) in the Nao's signal (fig 5(a4)).

ton and Nao are unsynchronized initially. Nao gradually gets the rhythm and finally synchronize after some time.

The human robot interaction seems to be more dynamic because human changes his frequency and amplitude continuously. Consequently Nao has also to adopt new frequency continuously. In fact, in the case of human-robot interaction, it is clearly observed that both agents got rhythmic motion in less time compared to the Nao-automaton interaction. It is due to the fact that the automaton has a fixed frequency, however, in a human-robot interaction both agents modify and correct themselves in order to be synchronized.

Interesting facts are observed during the experiments, some of these observations were also made by Pantaleone in his study of metronome synchronization [6]:

1) Our experiments show that the synchronization time between two agents is directly proportional to the coupling energy (f'). Optimized synchronization time can be achieved by scaling (by S_f) the coupling energy (f'). Figure 5(a) and Figure 5(b) details the relationships between the strength of coupling energy and the synchronization duration

of the two agents. Figure 5(a) shows motion signals Nao (solid line) and Automaton (dotted line). Initially, Nao is oscillating with its default frequency. As the automaton starts to move with a different frequency (6% higher than the Nao), Nao starts synchronizing with it. It can be observed that with low scaling (S_f) of the strength of coupling energy (0.15), the two agents take a long time to be synchronized (7 cycles of 2.2 seconds as in fig 5(a1)). As the scaling factor (S_f) for the strength of coupling energy increases to 0.25 and to 0.3, synchronization time decreases to 4.5 cycles and 4 cycles respectively (fig 5(a2,a3)). Further increases in coupling energy ($S_f = 0.5$) reduce the synchronization time to only 1 cycle. However, this higher energy induces saturation (clipping) in the Nao's signal (fig 5(a4)).

2) It is also interesting to know that if the frequency of the two agents (in Pantaleone's case, two pendulums) differs by more than a certain limit, synchronization will not occur. It is worth noticing that by increasing the coupling energy (by scaling S_f) feeding the Nao's oscillator, the range of interacting frequencies (that can be synchronized with Nao) can be expanded. With a low scaling of the strength of coupling energy, both agents can be synchronized if their natural frequency differs by no more

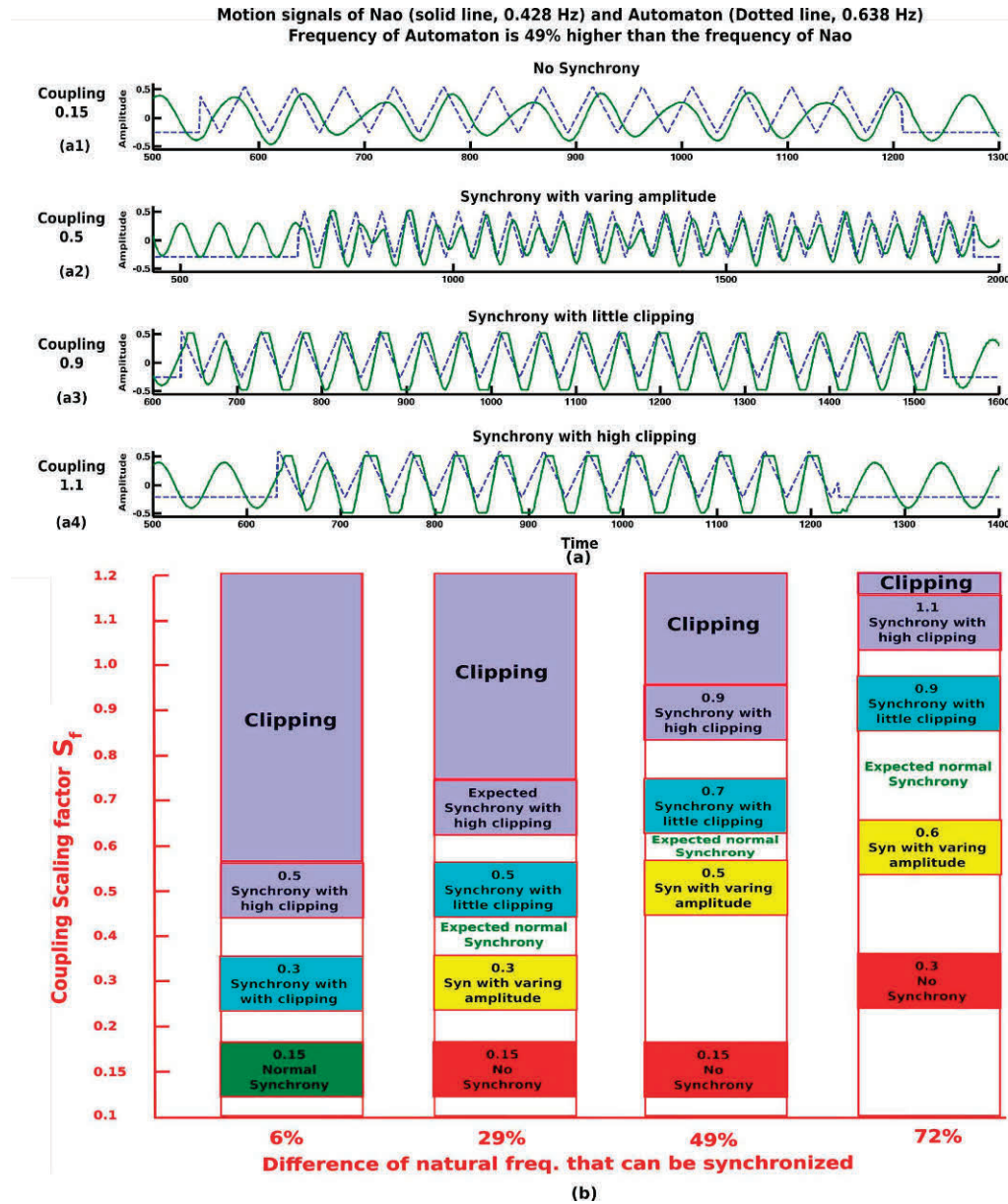


Figure 6. (a) and (b) demonstrate the fact that how the range of interacting frequencies Δf (difference of interacting frequencies) can be expanded by scaling the coupling energy. (a) motion signals of Nao (solid line, 0.428 Hz) and Automaton (dotted line, 0.638), frequency of the automaton is 49% higher than the frequency of Nao. It details that how the coupling scaling factor influences the interaction. Similarly, (b) compares the effect of increasing scaling factor (coupling) on 4 different Δf (6%, 29%, 49% and 72%).

than a few percent. Similarly, a high value of the scaling factor is needed if the difference of natural frequencies is larger.

Figure 6(a) and Figure 6(b) illustrate how the range of interacting frequencies (that could be synchronized) can be expanded by scaling the coupling energy. Figure 6(b) demonstrates the influence of increasing the coupling energy to four different Δf , the interactant frequency is then 6%, 29%, 49% and 72% higher than Nao's frequency. It can be observed that a scaling of $S_f = 0.15$ synchronizes only the agents having

a frequency which is very close to Nao's frequency ($\Delta f = 6\%$). As we increase the scaling factor to 0.3, for $\Delta f = 6\%$, the agents are synchronized with little saturations (clipping) because this energy is higher than the one required for this small difference of frequencies; for $\Delta f = 29\%$ the agents are synchronized with a little varying amplitude. The robot Nao needs a little more energy to be synchronized perfectly. However, others having Δf greater than 29% remain unsynchronized. Now, the coupling scaling is augmented to 0.5. This induces high saturation for

$\Delta f = 6\%$ and low saturation for $\Delta f = 29\%$, a synchrony is established with a varying amplitude for $\Delta f = 49\%$.

For the experiment shown in Figure 6(a), the automaton frequency is 49% higher than the frequency of Nao. The plots of the Nao and Automaton signals illustrate the fact that a small scaling (0.15) of the coupling strength can not synchronize the agents. As we increase the coupling strength by scaling of 0.5, the agents synchronize with a varying amplitude (little imperfection). A little more coupling strength (between the scaling (S_f) of 0.5 and 0.7) could establish perfect synchrony as shown in fig 5(a). However, If the strength of coupling energy is augmented (by scaling of $S_f = 0.9$ or $S_f = 1.1$) Nao synchronizes with a clipping (saturation). Higher values for the strength of coupling energy lead to high saturation.

3) For the same parametric conditions if the natural frequencies of both agents are the same, no phase lag was observed but as the Δf increases to a certain limit, the phase lag increases too. Beyond the limits corresponding to a given scaling factor, the interaction ends up with an asynchronous state. We experienced 0° to 90° of phase shift in our experiments. This fact of phase lagging can be easily noticed in Figure 5(a) where the difference of frequencies is less ($\Delta f = 6\%$), both agents are synchronized almost with a same phase. However, for $\Delta f = 49\%$, a small phase shift could be observed in Figure 6(a) while the system is in a synchrony state.

5. Selection of Partner

After developing a basic architecture initiating automatically a human-robot interaction by synchronizing agent's movements (in an imitating framework), we developed an architecture capable of choosing an interacting partner among various interacting agents.

We propose a neural network architecture (Figure 7(b)) selecting an interacting partner on the basis of synchrony detection among various interacting agents. It can be segregated into two parts. The first one is the dynamical interactions model (presented in the previous section) and the other one is the signal-prediction part. In the dynamical interaction module (section 4.2, Figure 2(b)), the visual stimuli f' (optical flow) was directly connected to the oscillator that controls the Nao's arm motion. Now the oscillator is fed indirectly by the signal-prediction block (f'') (Figure 7(b)). This indirect coupling of f'' is made to ensure that our algorithm will choose an interacting agent that moves with a frequency approximately similar to the robots inner dynamic (learned by the signal-prediction block). The equation 4 can be rephrase as

$$N_1(n+1) = N_1(n) - \beta N_2(n) + \alpha 1 + f'' \quad (5)$$

Where, f'' is the coupling strength feed by the signal-prediction block. The other variables remain unchanged.

The signal-prediction block (represented by y') is linked to the robot oscillator (represented by y) with a non modifiable link while the image of the visual activities (represented by X) is linked with a modifiable link. The signal-prediction (y') module learns the robot's oscillation as a weighted sum of image pixels (X). The neuron activity in the signal-prediction (y') corresponding to the predicted future value can be computed using the $X \rightarrow y'$ synapses:

$$y'_i(t) = \sum_{k \in X} W_{X_k \rightarrow y'_i} X_k \quad (6)$$

The learning of $X \rightarrow y'$ synaptic weights can be computed by the equation 7 and is based on the NLMS (Normalized Least Mean Square) algorithm [51].

$$W_{X_j \rightarrow y'_i}(t+dt) = W_{X_j \rightarrow y'_i}(t) + \alpha \eta \cdot \frac{y_i(t) - y'_i(t)}{\sum_{k \in X} X_k(t)^2 + \sigma 1} \cdot X_j(t) \quad (7)$$

Where y' stands for signal-prediction, X for the image of visual activities and y for the Nao's arm oscillator. α is the learning rate and $W_{X_j \rightarrow y'_i}$ represents the synaptic weights from the image neuron j to **signal – prediction** neuron i . y_i is the activity transmitted to the neuron i by the oscillator, it is a teaching signal for the Least Mean Square (LMS) algorithm [52]. To improve the LMS convergence during online learning, learning modulation η has been introduced. It modulates randomly the learning speed by introducing a randomization effect that suppresses the negative effects of the temporal regularities of the input data. The normalization term $\sum_{k \in X} X_k(t)^2 + \sigma 1$ is specific to the NLMS, $\sigma 1$ is a small value used to avoid the divergence of the synaptic weights if the visual activities (X image values) are too small. The use of the NLMS is motivated by the fact that the normalization term suppress the effect of rapid variations in the input data during the online learning. A faster convergence is then obtained.

For the selection of partner, the architecture works in two successive phases: a learning phase and a testing phase. During the learning phase, Nao oscillates according to its standard frequency. To learn its own dynamics, Nao looks at its own hand. The signal-prediction module which was zero due to the non availability of visual stimuli starts now learning the robot's modifiable oscillator signal as a weighted sum of the visual stimuli induced by its own actions. More precisely, the robot learns the association between its movements (sensori-motor information) and the induced visual velocity vectors (optical flow). As a consequence, as described in section 3, it also modifies the Nao's oscillator (Nao's arm movement). This process of modifying, learning and adapting continues and converge after some time. This adjustment can be assumed as a basic process by which infants gain self reflective abilities as underlined by Rochat [53]. In the same line, Gold and Scassellati proposed a probabilistic methods for learning the robot to recognize its own motor controlled body parts or its reflections [54]. After this learning phase, Nao can predict oscillatory movements similar to its own movement. When an agent interacts with a frequency close to the learnt one, weights already learnt are associated with the visual activities induced by the human movements. Nao's modifiable oscillator adopts the interacting frequency and phase. If the interacting frequency is different from the learnt one, the weights (modifiable links) could not be associated with the visual stimuli and Nao continues to move at its default frequency. The same is true for the multiple agents case. Among two interactants, only the agent having a similar frequency as Nao is selected.

In this experiment, three agents are involved, in addition to Nao and human, a basic automaton is introduced (Figure 1(d)). The coupling factor was 0.07, Nao's default frequency was 0.407 Hz, the automaton synchronized frequency was 0.4318 (6% higher) and the human synchronized frequency was 0.36 (11% low) to 0.38 (6% less than default frequency). When a subject interacts with a frequency close to the learnt one, this selection of partner algorithm selects this agent as a good interacting partner and Nao's modifiable oscillator synchronizes with it. Initially, both agents move with close frequencies (within an allowable range) but after some time of interaction Nao adopts the human movements and both oscillate with exactly the same frequency corresponding to the human motion. Good results are obtained with this architecture, these are collectively shown in the next sections.

In the presented work, our aim is to use the notion of unintentional synchrony to automatically initiate the human-robot interaction leading to learn complex tasks in a development way (during imitation games).

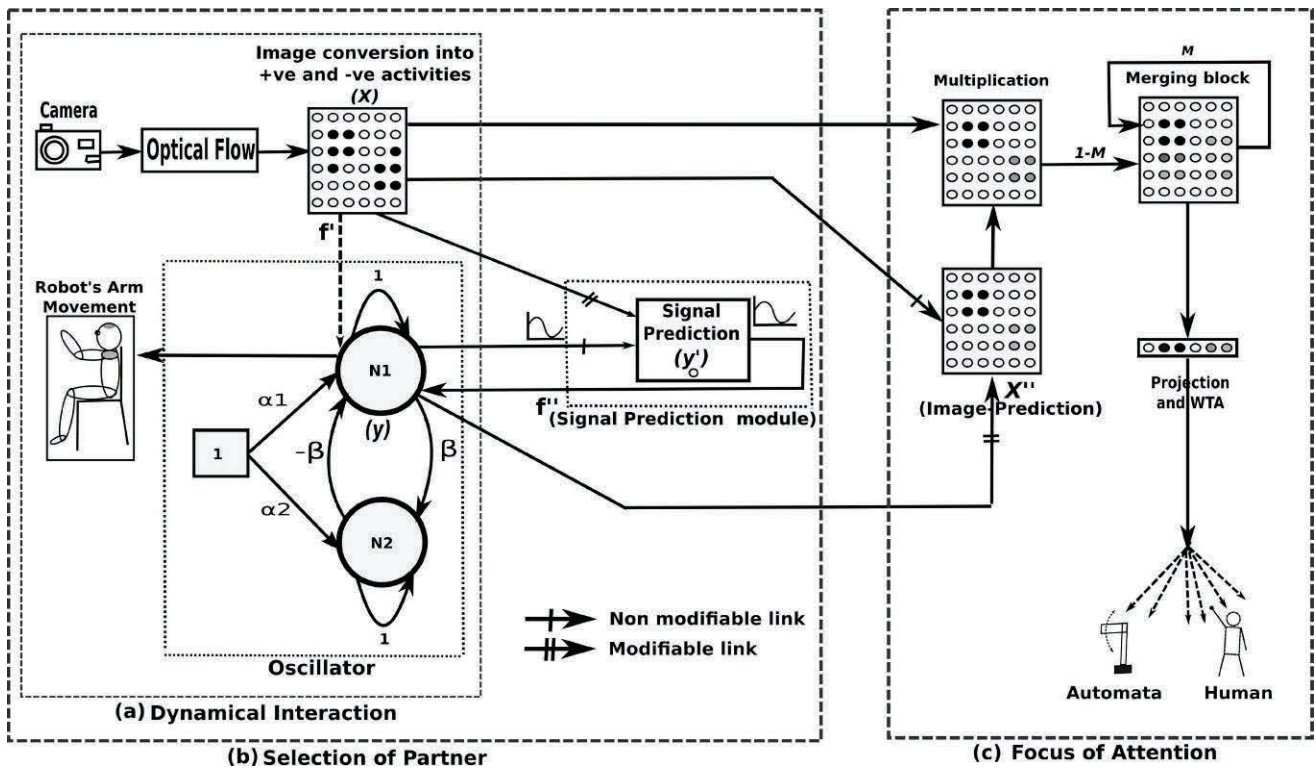


Figure 7. The global architecture (a) Dynamical Interaction model (b) Selection of Partner: select an interacting partner on the basis of synchrony detection among various interacting agents. (c) Focus of Attentional: this architecture capable of locating and pointing the synchronized interacting agent defined by the selection of partner algorithm.

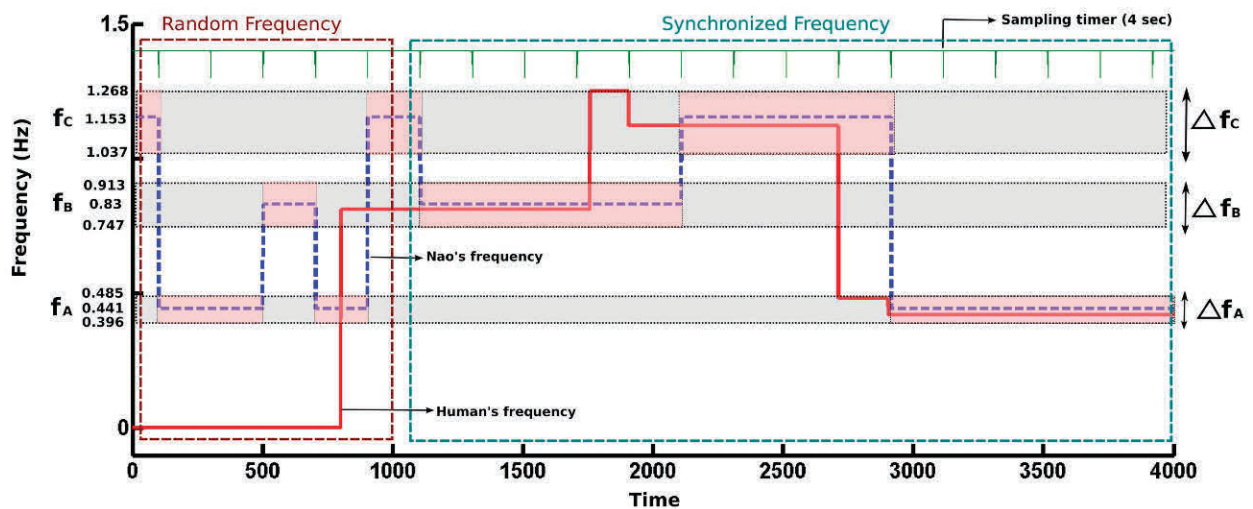


Figure 8. Shows how the different oscillators (A, B and C) are selected according to the visual stimuli. Due to the lack of interactant, random selection of the oscillators is observed. After 800 (26.6 sec) time units, oscillators are chosen according to the frequency of the agents. Nao's oscillator selection (dotted line) according to human frequency (solid line)

Hence, to be capable to deal with more complex gestures, Nao should be able to select partners in a larger range of frequencies.

In order to test a possible generalization of our model to more complex tasks we introduced (using the same model described previously) three different oscillators A, B and C with the following frequencies $f_A = 0.441\text{Hz}$, $f_B = 0.83\text{Hz}$ and $f_C = 1.153\text{Hz}$ respectively. These oscillation frequencies are learnt by three different signal-prediction modules. Nao's oscillating frequency is selected among these three oscillators depending on the visual stimuli. In the absence of visual stimuli, the oscillators controlling Nao's arm are selected randomly (every 4 seconds). As the human interacts with a certain frequency, one of the signal-prediction module close to the interactant's frequency synchronizes with it. Our architecture selects among three oscillators, the one having the minimal error with the visual stimulus. Experimental results are shown in Figure 8. In the onset, there is no visual stimulus (till 26.6 seconds or 800 time units). Consequently, one of the three oscillators are selected randomly. After some time (800 time units), an agent starts interacting with the oscillations of 0.81 Hz near to the oscillator B which is, consequently, selected as the robot's current dynamic. After 58 seconds (1750 time units), the agent changes his oscillations close to the frequency of the oscillator C, so the oscillator C is selected. The oscillator A is selected in a similar way.

It's worth noticing that during experiments with naive interactants, we asked the participants to move their arm at their own preferred frequency and style, some of the interactants moved with complex gestures implying the presence of multiple frequencies. In this case, if the fundamental frequency of the interacting agent was close to the robot's internal dynamics Nao was able to synchronize with him (the other harmonics are neglected) otherwise, Nao was moving at its own default frequency. These results are available for partner distant from the robot by 0.5 to 3.5 meters.

6. Attentional Mechanism

One of the major concerns of interactive robotics is how to focus on salient features among various visual stimuli. In fact, focusing attention and discriminating useful data from the others reduce significantly the big amount of incoming information from sensors and keep computational resources available for other important tasks.

Current approaches of attentional mechanisms are usually based on the sole visual information. We propose here to control the attentional mechanism from a low level motor controller.

Using the selection of partner architecture, in the presence of two visual stimuli in the Nao's visual field, the robot will synchronize with the "interacting" partner having a motion frequency close to nao's own dynamic. However, Nao will not be able to locate the good interacting partner in its visual field, because our algorithm works on the perceived energy irrespective of the spatial information (agent location).

Figure 7(c) shows the architecture to control the Focus of Attention (FOA). The FOA architecture functions in two steps: learning and testing. When there is no agent, the FOA moves around randomly depending on the noise. As the human partner start moving his arm with a frequency close to Nao's dynamics, the image-prediction module (X'') learns the spatial location (in its visual field) of the interacting partner as a weighted sum of Nao's synchronized frequency. The architecture is able to predict the location of the synchronized partner. If another agent comes and interacts with a different frequency (lower or higher than Nao, as shown in Figure 1(c)), X'' which already learnt the spatial position of the synchronized rhythmic movements strongly predicts the location of the synchronized agent even in the presence (in Nao's visual

field) of an unsynchronized one (because the prediction is made by the weighted sum of the learnt frequency).

To determine the correct interacting partner and to discriminate between multiple stimuli, our algorithm modulates the current visual stimuli with the image-prediction of the moving areas X'' . A merging block is used to calculate a weighted average (modulation) of these current visual motion and the predicted ones. The higher values of this merging block are then correlated to the location of synchronous movements. All the pixels of the merging block are projected on the x-axis (i.e all pixels in each column are added). A Winner Takes All (WTA) selects the highest activated column. This selected column indicates the location of the synchronized movement and the robot can point to the synchronized region to show the current Focus Of Attention (FOA). The robot always focus with the synchronized stimulus even if the partner changes his location. For this experiment, the resolution of the predicted image is 32×24 (32 columns or location), these 32 possible locations are in a range of 60° angles (-30° to 30° , precision of 2°). In our experiments, Nao's head direction is used to point the current focus of attention.

The learning rule for the image-prediction (X'') module is near to Hebbian learning with a normalization of weights. The weight normalization avoids the divergence and allows to forget less used associations. The neurons activities in (X'') can be computed using the $y \rightarrow X''$ synapses (8). The learning of $y \rightarrow X''$ synaptic weights can be computed by the equation 9.

$$X''_i(t) = \sum_{k \in y} W_{y_k \rightarrow X''_i} y_k \quad (8)$$

$$W_{y_j \rightarrow X''_i} = \frac{W_{y_j \rightarrow X''_i}^*}{\sum_{k \in y} W_{y_j \rightarrow X''_i}^*} \quad (9)$$

with:

$$W_{y_j \rightarrow X''_i}^*(t + dt) = W_{y_j \rightarrow X''_i}(t) + \epsilon y_j U_i \quad (10)$$

and:

$$U_i = \sum_{m \in X} W_{X_m \rightarrow X''_i} X_m \quad (11)$$

Where X''_i is the neuron activity in X'' group, y_k is a neuron of the conditional group (Oscillator) and U_i is the unconditional stimulus (Image (X)). X_m is the activities in the unconditional group (X). $W_{y_j \rightarrow X''_i}$ represents the synaptic weights from the oscillator (y) neuron j to the image-prediction (X'') neuron i, unnormalized weights are shown with " * ".

7. Experimental results

7.1. Synchrony based focus of attention on a static platform (fixed camera)

The focus of attention mechanism has been tested in a simple way and results are shown in Figure 9. In (a) Nao's oscillations along with the interacting agent motion signal are shown. (b) indicates the angle of FOA according to the rhythmic motion. And (c) displays the interactant's locations in front of Nao. At the beginning, no visual stimuli is presented to Nao, the FOA moves randomly between -30° and 30° . After 500 time units (16.6 sec.), a synchronized motion is introduced from the left side (-28.3°) of Nao. This interaction results in turning the

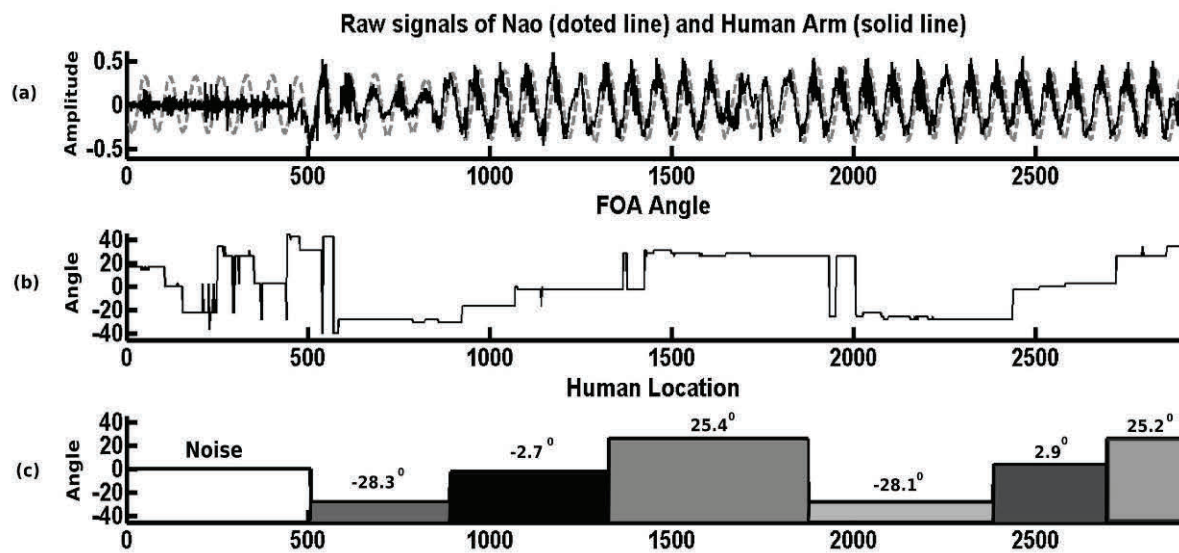


Figure 9. (a) Nao's oscillations along with interacting agent motion signal (b) Fluctuating angle of FOA according to the rhythmic motion (c) It describes the interactant's location in front of Nao.

FOA to the synchronized location as shown in fig. 9. As the interacting agent moves a little to the right side (-2.7°), our architecture force the FOA to relocate itself in the direction of the synchronized motion. Next, the agent moves to the left side 25.4° of Nao and FOA again follows the agent. The same sequence is repeated again to verify that FOA always follow the interactant. The Figure. 9 corresponds to the experiment shown on the video available on our website.

After this simple experiment, we examine our selection of partner algorithm along with FOA architecture by extending the study to the case of two agents: an automaton (one DoF) and a human (only one of them is synchronized at a time). Results show that when two agents interact simultaneously with Nao, the one moving similarly (in terms of frequency) as the robot, will be selected as an interacting partner, the FOA rotates the robot's head towards the synchronized partner. Likewise, if the interactants roles are switched (switch of the moving frequency) consequently, FOA and selection of partner inverse the selected partner.

The experimental results are shown in Figure 10. Figure 10(a1) sketches the curves of the automaton and Nao's motion signal. Initially, the automaton is moving from the left side of Nao (about -20° , see Figure 10(a)). Both agents synchronizes after a short period by using our partner selection architecture. Figure 10(a3) describes the quality of synchrony between Nao and the automaton in terms of PLV. At the beginning, PLV was at its lowest value but as the interaction continues, it increases slowly to the highest one during the synchronization phase between the agents. Initially, there is no other agent in front of robot except the Automaton therefore, FOA turns towards the Automaton (Figure 10(a4)). After 700 time units (23.33 seconds), the human starts interacting from the right side of the robot with a different frequency but he fails to disturb Nao-Automaton interaction (the PLV continues to its higher values for the automaton) and FOA continues to point out to automaton.

Now, the roles are switched. Human is advised to make similar movements as Nao while the automaton is adjusted to a lower frequency

(Figure 10(b1) and (b2)). Consequently, Nao also flips its role by synchronizing with the human and selects him as a partner. PLV increases for the human and decreases for the automaton (Figure 10(c3)). As the synchrony emerges between human and Nao, FOA also turns from automaton to human (Figure 10(c4)). To validate our experiment we again switch the roles of the two interacting agents after 2650 time units (88.3 sec). Consequently, this induces a switch of the focus of attention and the synchronized agent (Figure 10(b)).

7.2. Synchrony based focus of attention on a mobile platform (moving camera)

To study the possible generalization of the previously detailed model to the case of a mobile platform, we embedded the architecture on a mobile Robosoft Robulab 10 equipped with four wheels, two for directions and two for stabilization, a proximity sensors for obstacle avoidance, an embedded computer, and for the visual perception, a pan-tilt camera controlled with a SSC-32 card through a serial communication (see Figure 11). The experiments were performed in an unconstrained indoor environment.

Our aim here is to use our synchrony based model to focus the attention of the mobile robot on a preferred partner to initiate the interaction for starting to leaning and to gain new knowledge by using synchrony detection as excitatory and inhibitory signals.

More precisely, in this application, after focusing its attention on a selected partner on the basis of synchrony detection, the mobile robot must enable or disable the learning of the partners shape relative to the maintaining or not of a synchronized interaction in order to ensure that the robot's visual attention is focused on the preferred partner.

As we are interested by locating a partner in the visual field (using synchrony), a major problematic in the case of a mobile robot is to deal with the ego-motion induced by the moving camera. In fact, the robot must be able to predict and compensate the ego-motion induced by

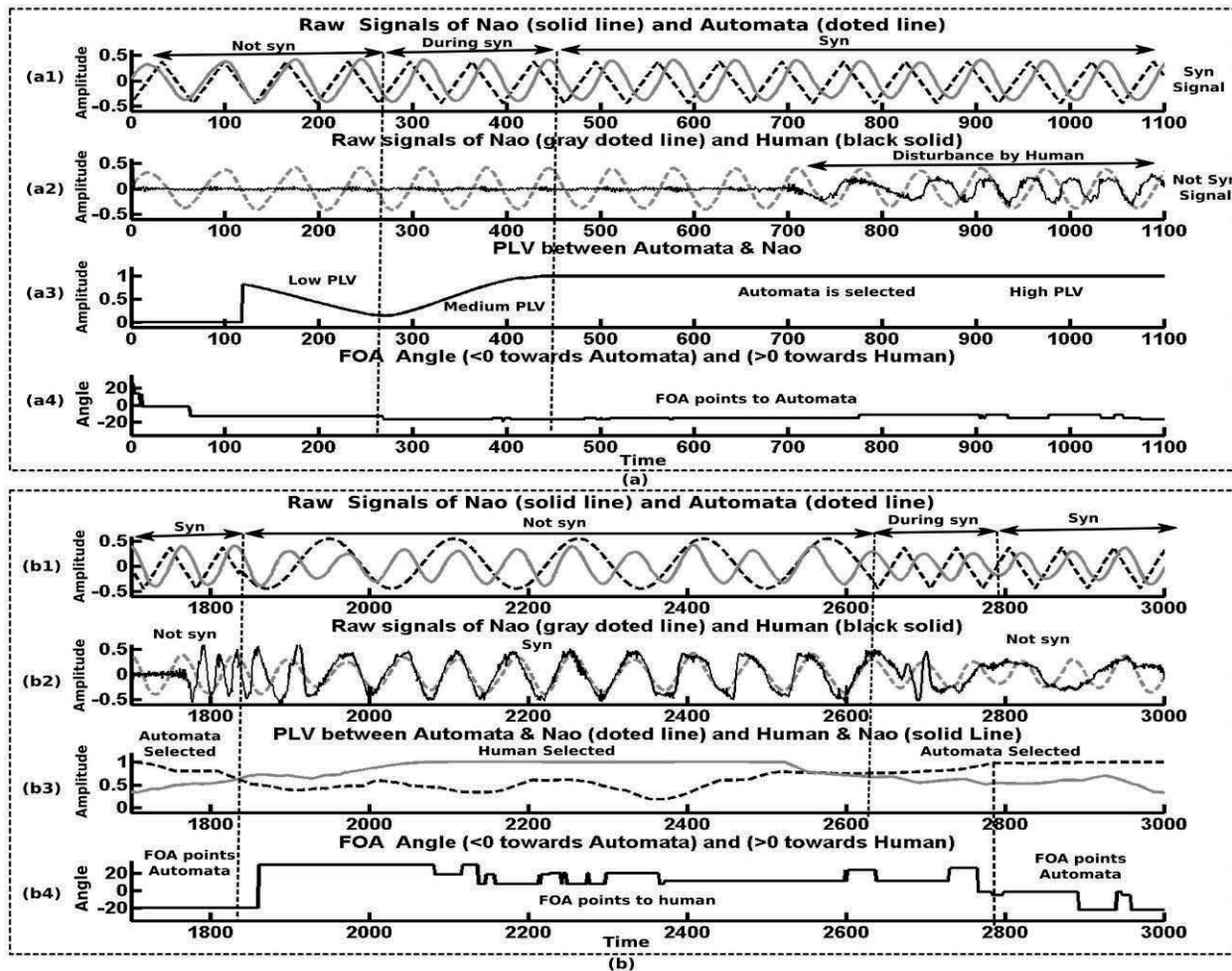


Figure 10. Results: Every set has 4 graphs with same sequence where first time series of every set shows the raw signals of Nao oscillations along with robotic arm while second contains raw signals of Nao along with human and third time series shows the PLV (quality of synchrony) for the pairs of interacting agents and finally, fourth shows the FOA angle of Nao which follows the synchronized region. (a) shows start of experiment with single agent and then disturbed by the other agent. (b) Multiple agent having different frequencies interact (one of them with same as Nao) and Nao always selects similar frequency partner.

its own movements to be capable of segmenting the moving humans and objects in the visual field and differentiating them from the static background.

To do so, we propose a simple bio-inspired model permitting the robot to learn the cross-modal link between the motor controller (velocities of the Robulab) and the induced visual stimuli (optical flow) while the robot is moving. The objective is to be able, after this learning phase, to predict and compensate the correct optical flow corresponding to a given robot motion velocity.

As illustrated in Figure 13, the same optical flow algorithm [49] as for the previous model is used to extract velocity vectors from the images acquired by the embedded camera of the Robulab. The computed velocity vectors are modulated by the image gradient to take into account only highly contrasted regions of the image. The reason behind this modulation is the unavailability of accurate and salient visual stimuli (optical flow) in uniform area of the images. This is due to the limitations

of the used optical flow algorithm even more with the small number of iterations used in this experiment (set to a fixed value of 4) for saving computational time.

The modulated optical flow feed the *component direction-selective neurons*. These neurons simulate those of V1 and MT brain areas which are designed by neurobiological records as sensitive to preferred motion directions (called *component direction-selective neurons* by Movshon [55]).

These component direction-selective neurons filter the modulated optical flow, the firing of each of these neuron (A_i) is proportional to the angular distance between the visual stimulus (optical flow) and its preferred direction weighted by the motion intensity as below:

$$A_i = \exp\left(-\frac{(\beta - \beta_i)^2}{2\sigma_i^2}\right) \cdot \left(1 - \exp\left(-\frac{V^2}{2\sigma_i^2}\right)\right) \quad (12)$$

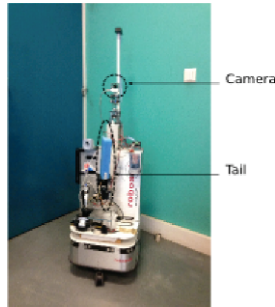


Figure 11. Experimental Setup for the Mobile Platform

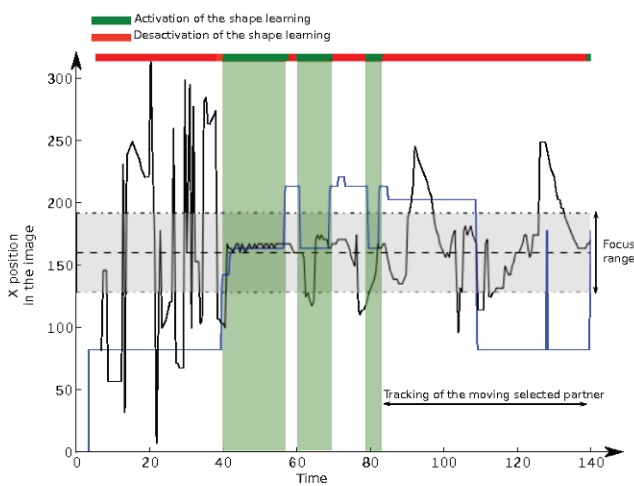


Figure 12. Partner detection and tracking: Experimental results in real conditions by combining synchrony based and shape based strategies

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 (set to a value of 20) are the coefficients regulating the dynamics of the neuron activation for the motion direction and the motion intensity, respectively.

In this experiment, we use only 2 different types of direction-selective neurons sensitive to the directions 0° and 180° . This simplification was made due to the fact that the robot, in this specific application, turns right or left to focus its attention on the partner inducing different intensities but only in the two horizontal directions of the optical flow. The outputs of these neurons are then integrated in time by neurons which are sensitive to different visual motion intensities for a given direction of motion. They are similar to a simplified version of the *pattern direction-selective neurons* which integrate the component direction-selective neurons in the MT area of the human brain [55]. Each of these neurons can be seen as an histograms representing the occurrence in time of visual velocity vectors having a given motion direction (here 0° or 180°) with a given motion intensity (in practice, a given small range of intensities). For each motion direction we use 20 pattern direction-selective neurons in our model. A weighted mean of the outputs of the 20 neurons is then computed representing, for each

direction, the mean visual motion intensity extracted from the visual field over time.

During the learning phase, we rotate the robot with 5 different speeds on the left side and 5 different ones on right side. The speeds are controlled by 10 motor neurons (5 for each direction) as shown Figure 13. These neurons represent the unconditional inputs of two LMS network (one for each direction) which learn to associate the motor controller orders (motor velocities) to the mean visual motion intensity extracted from the pattern direction-selective neurons during the experiment.

Consequently, after this learning step, when the robot start moving with a given velocity, the LMS will be able to trigger (for each direction) the correct visual motion intensity mean value which can be subtracted (see Figure 13) from the current optical flow in order to compensate the induced ego-motion. Therefore, the filtered visual stimuli can be used as an Input for the previous model to select and locate a partner in the case of a moving robot.

Lets now consider the complete experimental scenario. First a human partner moves his arm in front of the robot to teach a preferred frequency of interaction (the robot is in a static position). After this learning phase, the robot starts moving randomly. The ego-motion induced by the robot's movements are compensated as explained above. The learnt frequency can be observed on the robot's pendulum (tail in the Figure 11). Using the previously described models, when a human interacts with the robot with a dynamic close to the learnt one, the robot selects and locates this preferred partner on the basis of synchrony detection.

Consequently, the synchrony based focus of attention enables the learning of the selected partner shape using the algorithm described in [56].

The results of this experiment are illustrated in Figure 12. First the mobile robot starts moving by focusing its attention (black line) on random regions of the visual field due to the lack of salient regions of interest. At time $t = 40$ seconds, a human starts interacting with a frequency close to the one learned by the robot. Consequently, the synchrony based focus of attention selects and predicts the location of the human partner (blue line in Figure 12). A neural Field controlling the robot's movements turns the Robulab toward the partner and keeps him in the center of the visual field (black line in Figure 12). When the robot attention is focused on the partner on the basis of synchrony detection, the shape learning is activated (green areas in Figure 12). From time $t = 40$ to $t = 80$, we can notice that the shape learning is stopped (red areas in Figure 12) and re-engaged relatively to the maintaining or not of the synchrony based focus of attention. A refined learning of the partner shape is consequently obtained. As we can see Figure 12, starting from $t = 80$, the shape learning is strong enough, even if the robot loses the synchrony. If the human move to the left side or the right side the robot tracks its partner and moves toward his direction in the visual field using shape recognition.

Videos of our all experiments can be found on <http://www.etis.ensea.fr/neurocyber/Videos/synchro/>

8. Conclusion

In this paper, we presented a new model allowing the robots to select an interacting partner among multiple agent based on synchrony detection. We demonstrated that prediction of synchrony (for spatial position) could be used as a tool to locate the Focus Of Attention. Our experimental results showed that when several agents interact with Nao and one of them moves in synchrony with the robot, Nao will select it as a partner.

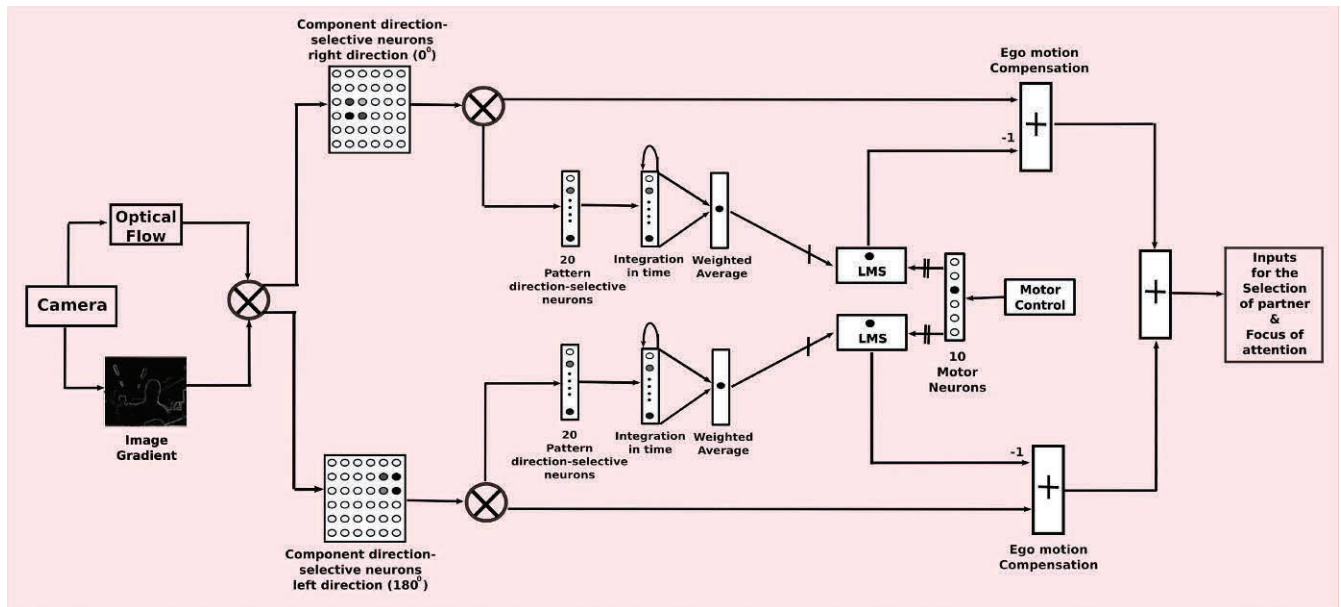


Figure 13. Architecture for a mobile robot.

From a psychological point of view, we were inspired by the unintentional communications between humans. The synchronous exchanges during social interactions are directly associated to the sensorimotor information of the two agents. According to the experiments in neurosciences; when two human interact by doing spontaneous exchanges (between two participants) of intransitive bi-manual movements, researchers have revealed the emergence of inter-brain synchronization across multiple frequency bands. Low frequency physical synchronized movements reflect low frequency inter-brain synchronization. However, inter-brain networks in high frequencies do not correspond to the physical movements of the interactants. We suppose this synchronization in (inter-brain) high frequencies is related to some higher-level behavior. In our case, we assume that low frequency behaviour is related to our arm synchronization and high frequency behaviour may be related to the Focus of attention [34].

We are actually investigating three applications of synchrony detection in human-robot interactions. The first and the most important one is to extend the architecture to a developmental learning of more complex tasks (complex interactions). In our study, synchrony detection and partner selection architectures allow to sustain the interaction by synchronizing low fundamental frequencies of interaction. Consequently, complex gestures (presenting higher temporal frequencies) could be taught to the robot during imitative interactions. Likewise, we are also intended to use the proposed architecture for navigation tasks. A mobile robot can choose a synchronized interacting human partner (for instance a human having his legs periodic movements close to the robot's dynamics), therefore, while moving in synchrony with the robot, the selected human partner can play a role of a tutor to teach navigational tasks. Finally, as we assume that synchrony detection is not only a starting point for social interactions but also as a tool to maintain and re-engage the interaction, we plan to use the selection of partner and FOA neural models for driving turn-taking games in HRI.

Furthermore, in order to validate our assumptions, on going psycho-experimental studies analyze the influence of unidirectional and bi-

directional interactions on unintentional synchronization during interactions between naive human subjects and: (i) Nao moving with a fixed frequency of interaction (Unidirectional interaction) (ii) Nao having minimal abilities (by using our synchrony based neural model) to adopt the human partner frequency of interaction (Bidirectional interaction).

References

- [1] N. Otero, S. Knoop, C. L. Nehaniv, D. Syrdal, K. Dautenhahn and R. Dillmann. Distribution and recognition of gestures in human-robot interaction. IEEE symposium on robots and human interactive communication (RO-MAN06), Hatfield, UK September 6-8, 2006.
- [2] B. Gates. A robot in every home. Scientific American 296(1):5865, 2007.
- [3] K. Dautenhahn, The art of designing socially intelligent agents: Science, fiction, and the human in the loop, Applied Artificial Intelligence, vol. 12, pp. 573617, 1998.
- [4] M. M. Haque, D. Das, T. Onuki, Y. Kobayashi and Y. Kuno. Model for controlling a target human's attention in multi-party settings. IEEE RO-MAN-2012, IEEE Symposium on robots an human interactive communication, Paris 2012.
- [5] A. Pikovsky, M. Rosenblum and J. Kurth. Synchronization: a universal concept in nonlinear sciences, Cambridge, 2001.
- [6] J. Pantaleone. Synchronization of metronomes. American Journal of Physics, Vol. 70, Nr. 10AAPT, p. 992–1000, 2002
- [7] M. Rosenblum, A. Pikovsky and J. Kurth. Synchronization approach to analysis of biological system. Fluctuation and noise letters, vol. 4, No. 1, L53-L62, World scientific publishing company, 2004.
- [8] S. Strogatz and I. Stewart. Coupled Oscillators and Biological Syn-

- chronization. *Scientific American*, 269:102-109, Dec. 1993.
- [9] I. I. Blekhnman. *Synchronization in Science and Technology*. ASME, New York, 1988.
 - [10] R. E. Mirollo and S. H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics* 50, 1645-1662, 1990.
 - [11] T. J. Walker. Acoustic synchrony: two mechanisms in the snowy tree cricket. *Science* 166, pages 891-894, 1969.
 - [12] M. Rosenblum, A. Pikovsky. Synchronization: from pendulum clocks to chaotic lasers and chemical oscillators. *Contemporary Physics*, volume 44, number 5, pages 401-416, September-October 2003.
 - [13] A. J. W. Ward, S. Axford, and J. Krause. Mixed-species shoaling in fish: the sensory mechanisms and costs of shoal choice. *Behavioral Ecology & Sociobiology*, 52, 182187, 2002.
 - [14] L. J. Cator, B. J. Arthur, L. C. Harrington and R. R. Hoy. Harmonic convergence in the love songs of the dengue vector mosquito. *Science*, 323, 10771079, 2009.
 - [15] D. C. Michaels, E. P. Matyas and J. Jalife. Mechanisms of sinoatrial pacemaker synchronization: a new hypothesis. *Circulation Res.*, 61, pp. 704-714, 1987.
 - [16] E. Delaherche, M. Chetouani, A. Mahdhaoui C. Saint-Georges, S. Viaux and D. Cohen. Interpersonal Synchrony: A Survey Of Evaluation Methods Across Disciplines. *IEEE Transactions on Affective Computing*. Vol 3 No 3 Pages 349 - 365, 2012.
 - [17] L. Glass. Synchronization and rhythmic processes in physiology. *Nature* 410: 277284, 2001.
 - [18] R. Hari and M. V. Kujala. Brain basis of human social interaction: From concepts to brain imaging. *Physiological Reviews* 89: 453479, 2009.
 - [19] V. Muller and U. Lindenberger. Cardiac and Respiratory Patterns Synchronize between Persons during Choir Singing. *PloS ONE*, Volume 6, Issue 9 e24893, September 2011.
 - [20] F. Cummins. Periodic and aperiodic synchronization in skilled action. *Frontiers in Human Neuroscience*: 5:170; doi: 10.3389/fnhum.2011.00170, 2011.
 - [21] R. Feldman. Parent/Infant Synchrony: Biological Foundations and Developmental Outcomes. *Current Directions in Psychological Science*, Volume 16, issue 6, p. 340-345. ISSN: 0963-7214 DOI: 10.1111/j.1467-8721.2007.00532.x Blackwell Publishing Inc, December 2007.
 - [22] A. Kendon. *Conducting Interaction: Patterns of behaviors in focus encounters*. Cambridge University press, Cambridge, UK. 1990.
 - [23] R. C. Schmidt and M. J. Richardson. *Coordination Neural behavioral and social dynamics*, Volume: socialdyna, Issue: 4, Publisher: Springer, Pages: 1-53, 2008.
 - [24] J. Nadel, I. Carchon, C. Kervella, D. Marcelli, D. Reserbat. Expectancies for social contingency in 2 months old. *Developmental science* 2, 164-173, 1999.
 - [25] W. S. Condon and L. W. Sander. Neonate movement is synchronized with adult speech: Interactional participation and language acquisition. *Science*, 183, 99101, 1974.
 - [26] C. G. Prince, G. J. Hollich, N. A. Helder, E. J. Mislivec, A. Reddy, S. Salunke, and N. Memon. Taking synchrony seriously: a perceptual-level model of infant synchrony detection. *Proceedings of the Fourth International Workshop on Epigenetic Robotics*. Lund University Cognitive Studies, (2004)..
 - [27] L. J. Gogate and L. E. Bahrick. Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, 69, 133-149, 1998.
 - [28] J. S. Watson. Smiling, cooing, and The Game. *Merrill-Palmer Quarterly*, 18, 323-339, 1972.
 - [29] G. Gergely and J. S. Watson. Early socio-emotional development: Contingency perception and the social-biofeedback model. In: P. Rochat (Ed.), *Early Social Cognition: Understanding Others in the First Months of Life* (pp. 101-136). Mahwah, NJ: Lawrence Erlbaum, 1999.
 - [30] P. Rochat and T. Striano. Perceived self in infancy. *Infant Behavior & Development*, 23, 513-530, 2000.
 - [31] J. A. S. Kelso, G. Dumas and E. Tognoli. Outline of a general theory of behavior and brain coordination. <http://dx.doi.org/10.1016/j.neunet.2012.09.003>, 2012.
 - [32] U. Hasson, Y. Nir, I. Levy, G. Fuhrman, R. Malach. Intersubject synchronization of cortical activity during natural vision. *Science*; 303(5664):1634-40, 2004.
 - [33] G. J. Stephens, L.J. Silbert, U Hasson. Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci USA*; 107(32):14425:30, 2010
 - [34] G. Dumas, J. Nadel, R. Soussignan, J. Martinerie, L. Garnero. Inter-Brain Synchronization during Social Interaction. *PLoS ONE* 2010;5(8):e12166, 2010.
 - [35] J. Issartel, L. Marin, M. Cadopi. Unintended Interpersonal Coordination: Can we march to the beat of our own drum?. *Neuroscience Letters*, 411:174-179, 2007.
 - [36] Z. Nda, E. Ravasz, T. Vicsek, Y. Brechet, and A.-L. Barabási, The sound of many hands clapping, *Nature*, 403, 850, 2000.
 - [37] M. Varlet, L. Marin, S. Raffard, R.C. Schmidt, D. Capdevielle et al. Impairments of Social Motor Coordination in Schizophrenia. *PLoS ONE* 7(1): e29772. doi:10.1371/journal.pone.0029772, 2012.
 - [38] P. Andry, A. Blandchard, P. Gaussier. Using the rhythm of non-verbal human-robot interaction as a signal for learning. *IEEE Transactions on Autonomous Mental Development* 3(1): 30-42, 2011.
 - [39] K. Prepin and P. Gaussier. How an agent can detect and use Synchrony parameter of its own interaction with a human. *COST 2102 Training School*, 50-65, 2009.
 - [40] A. Blanchard and L. Caamero. Using visual velocity detection to achieve synchronization. *Procs 3rd Int Symposium on Imitation in Animals and Artifacts (SSAIB)*, 2005.
 - [41] Q. Shen, H. Kose-Bagci, J. Saunders, K. Dautenhahn. The Impact of Participants' Beliefs on Motor Interference and Motor Coordination in Human/Humanoid Interactions. *Autonomous Mental Development*, *IEEE Transactions*, 04/2011;
 - [42] L. Marin, J. Issartel and T. Chaminade. Interpersonal motor coordination, from human/human to human/robot interactions. In: Dautenhahn, Kerstin (ed.), *Robots in the Wild: Exploring human-robot interaction in naturalistic environments: Special Issue of Interaction Studies* 10:3, 2009.
 - [43] M. P. Michalowski, S. Sabanovic and H. Kozima. A dancing robot for rhythmic social interaction. *Proceedings of the ACM/IEEE international conference on Human-robot interaction* ACM New York, NY, USA 2007.
 - [44] T. Ikegami and H. Iizuka. Joint attention and Dynamics repertoire in Coupled Dynamical Recognizers, the proceedings of the AISB 03: the Second International Symposium on Imitation in Animals and Artifacts, pp.125-130, UK, April 2003.
 - [45] C. Crick, M. Munz, B. Scassellati. Robotic drumming: Synchronization in social tasks. *Proceedings of IEEE International Symposium on Robot and Human Interactive Communication*, 2006.
 - [46] M. Rolf, M. Hanheide and K. J. Rohlfing. Attention via synchrony. Making use of multimodal cues in social learning. *IEEE Transactions on Autonomous Mental Development* 1: 55-67. 2009
 - [47] V. V. Hafner and F. Kaplan. Interpersonal Maps: How to map affordances for interaction behaviors. *Proceedings of international conference on Towards affordance-based robot control*, 2006.

- [48] J. P. Lachaux, E. Rodriguez, J. Martinerie and F.J. Varela. Measuring phase synchrony in brain signals. *Human Brain Mapping*, 1999.
- [49] K. P. B. Horn and B. G. Schunck. Determining optical flow. *ARTIFICIAL INTELLIGENCE*, vol:17, 185–203, 1981.
- [50] A. Revel, P. Andry. Emergence of Sturctured Interactions: From theoretical model to pragmatic robotics. *Neural Networks*, vol. 22, no. 2, pp. 116–125, 2009.
- [51] J. Nagumo. A learning method for system identification. *IEEE Trans. Autom. Control* 12(3), 282-287 NLMS reference, 1967.
- [52] B. Widrow, M. E. Hoff. Adaptive switching circuits. In *IRE WESCON*, pages 96-104, New York, Convention Record, 1960.
- [53] P. Rochat. Ego functions of early imitation. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind* (pp. 85-97). Cambridge University Press, 2002.
- [54] K. Gold and B. Scassellati. Using Probabilistic Reasoning over Time to Self-Recognize. *Robotics and Autonomous Systems*. Vol. 57(5), p. 384-392. 2009.
- [55] J. A. Movshon, E.H. Adelson, M.S. Gizzi, and W.T. Newsome. The analysis of moving visual patterns. *Pattern recognition mechanisms*, 54:117151, 1985.
- [56] S. Lepretre, P. Gaussier and JP. Cocquerez. *From navigation to active object recognition*. Citeseer. 2000