

Biologically Inspired Multimodal Integration: Interferences in a Human-Robot Interaction Game

Eric L. Sauser and Aude G. Billard
Learning Algorithms and Systems Laboratory, LASA
Ecole Polytechnique Fédérale de Lausanne, EPFL
CH-1015 Lausanne, Switzerland
{eric.sauser, aude.billard}@epfl.ch

Abstract— This paper presents a biologically inspired approach to multimodal integration and decision-making in the context of human-robot interactions. More specifically, we address the principle of ideomotor compatibility by which observing the movements of others influences the quality of one’s own performance. This fundamental human ability is likely to be linked with human imitation abilities, social interactions, the transfer of manual skills, and probably to mind reading. We present a robotic control model capable of integrating multimodal information, decision making, and replicating a stimulus-response compatibility task, originally designed to measure the effect of ideomotor compatibility on human behavior. The model consists of a neural network based on the dynamic field approach, which is known for its natural ability for stimulus enhancement as well as cooperative and competitive interactions within and across sensorimotor representations. Finally, we discuss how the capacity for ideomotor facilitation can provide the robot with human-like behavior, but at the expense of several disadvantages, such as hesitation and even mistakes.

I. INTRODUCTION

Today’s humanoid robots are being endowed with an ever increasing range of sensory and motor capabilities. These capabilities are being developed in order to endow autonomous robots with a sense of touch, the ability to perceive and manipulate robots, and even with the ability to converse naturally with people. The aim is to design robots which are more and more similar to humans, and also capable of social interaction and aiding humans in their everyday life [1]–[6]. However, in order to produce coherent behavior, a huge amount of sensory and motor information has to first be processed and integrated. This creates the problem of taking into account the different levels of reliability, as well as processing speed, of each sensory modality. When dealing with environmental demands such as accuracy and reaction speed, a robotic control system should establish as quickly as possible a coherent view of it’s environment, and then execute the relevant motor responses in order to interact with the outside world.

In this paper, we present a biologically inspired approach to multimodal sensory integration and decision making, in the context of human-robot interactions. Our research activities try to link together robotics with neuroscience and human psychology, in order to understand cognitive abilities such as learning by imitation [7], [8]. Along these lines, we shall address here the principle of ideomotor compatibility.

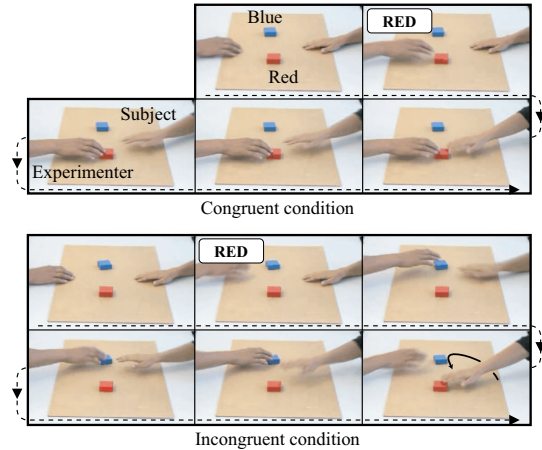


Fig. 1. A human subject is performing a stimulus-response task (see text for more details). At the top the named color and the target of the reach coincide, at the bottom they do not. A hesitating movement can be observed in the second case.

The principle of ideomotor compatibility states that *observing the movements of others influences the quality of one’s own performance* [9]. Experimental data shows that similarity between the observed and executed movements plays an important role. A close similarity between observed and executed movements produces a positive effect on the quality and initiation of the executed movement, whereas a difference between observed and executed actions generates interference. This implies that: the greater the similarity between observed and executed action, the easier the initiation and execution of the action will be. This fundamental human property is likely to be linked with human imitation abilities, social interactions, the transfer of manual skills, and probably even to mind reading [9], [10]. Such important capabilities may have emerged through human evolution, at the expense of several disadvantages such as hesitation, and reflexive errors.

Humans are also characterized by their mistakes. During everyday visuo-motor tasks, it is possible that we will hesitate or even fail in choosing the appropriate action to perform. Such cases express an ambiguity in our cognitive decision process, which may originate from conflicting environmental cues. Similarly, hesitation may also uncover late inhibition of reflex movements in favor of reflective movements, being the

result of longer mental processes. Such well known concepts, which appear frequently in experimental psychology literature [9], [11], [12], are often used to understand human cognitive pathways and behaviors, and hence to highlight design features which should be taken into account in human interfaces, in order to make them friendlier and more effective.

In order to illustrate the resultant effects of the ideomotor principle, we briefly describe a typical stimulus-response (SR) task. In such a task, the subject is instructed to respond, as quickly and accurately as possible, to an external stimulus. Usually a different response is associated with each of the stimuli. In addition, the stimuli are broken into two parts, a task-relevant cue, and a distractor. The idea behind such an experimental setup is to examine whether the distractor interferes with the task-relevant cue, and if so, to what extent. Here, two sensory modalities are considered: vision and speech. More specifically, goal directed gestures, such as reaching for an object, and color names which quantify the objects.

Two experimental trials are illustrated in Figure 1. The experimenter and the subject sit opposite to each other at a table. A blue and a red object are placed upon the table, at an equal distance from each of the two people. After a small initial waiting period, the experimenter names the color of one of the objects (red or blue), and also reaches for one of the two objects. However the color named and the object reached for may not necessarily be the same. The subject is asked to reach for the object which is reached for by the experimenter, as quickly as possible and with maximum accuracy. As can be seen, when the experimenter reaches for a different object than the one named, i.e. the gesture and the speech are incongruent, the subject either fails to reach for the correct object, or hesitates and later corrects the movement. Moreover, when a subject makes no error, the average reaction time is significantly faster in the case of a congruent (same gesture as speech) demonstration than in the incongruent case (see [9] for detailed results on a similar SR task).

In the following section, we present a robotic control model capable of integrating multimodal information, decision making, and replicating the ideomotor characteristics of human behavior. The model consists of a neural network based on a dynamic field approach [13], [14]. These dynamic systems possess several interesting properties, including stimulus enhancement, and cooperative and competitive interactions within and across neural representations.

II. EXPERIMENTAL SET-UP

The experiments are performed using a *Fujitsu* humanoid robot, HOAP2, with 25 degrees of freedom (DOFs). In our experiment, only the right arm (4 DOFs) is used. The arms DOFs consist of three rotations for the shoulder (flexion-extension, abduction-adduction, and humeral rotation), and one for the elbow. The torso and legs are set to a constant, stable position, in order to support the robot sitting with its torso in an upright position. Two external *Phillips* webcams track the location of colored objects, with a refresh rate of 15Hz and a resolution

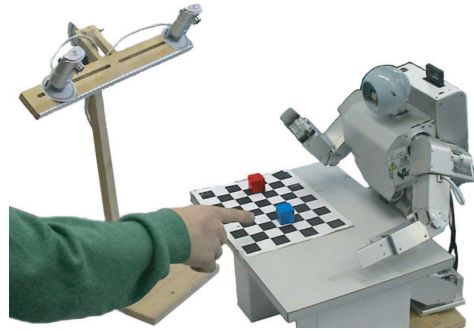


Fig. 2. The robotic set-up.

of 640x480 pixels. In order to provide the robot with speech recognition capabilities, a standard microphone is placed in close proximity to the robot. The Fujitsu HOAP2 controller library is used to control the robot, the *OpenCV* library is used to create stereoscopic vision from the two webcams [15], and *Sphinx-2* is used for speech recognition [16]. All of the pieces of software have access to their source-code, and run on the Linux platform. The robotic experimental paradigm is exactly the same as that used in the human experiments, except that the set-up shown in Figure 2 was adapted to the small size of the robot.

III. NEURAL NETWORK CONTROL

In this section we present the neural network which controls sensory integration and decisional processes. It is based on the neural fields approach [13], known to be a robust and smooth neural mechanism. The hypotheses behind this framework assumes a continuous encoding of the sensorimotor space [18], and also, as suggested by neuroscientific theories, the interactions of multiple information sources are performed in a locally common space [10], [17]–[19]. First we will describe the dynamics of the neural fields, followed by the model architecture, its sensory inputs and its motor outputs.

A. Dynamic Neural Fields

The neural controller of our system is based on the dynamic field approach [8], [13], [14], [19], which can account for temporal dynamics of stimuli interactions.

1) *Definition:* Formally, a neural field is composed of a continuous set of neurons, firing maximally, for a specific value θ in the parameter space. This unimodal type of activation is illustrated in Figure 3a. In order to avoid the problem of boundary effects, preferred values of θ are uniformly distributed along a circular space, such that $\theta \in [-\pi, \pi[$. Note that any kind of information can be encoded, indeed a unidimensional variable λ defined in an arbitrary domain can be transformed into the considered space using a mapping function $m(\lambda) \rightarrow \theta$. The network is fully connected by means of recurrent synaptic weights W_i^R , exhibiting symmetry, rotation invariance and center-surround excitation-

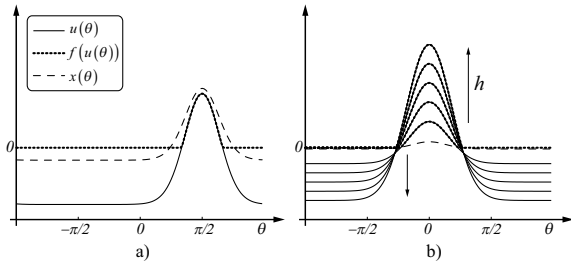


Fig. 3. a) Illustration of a neural field subjected to a spatial input localized at $\pi/2$. b) The population profile of activity can be modulated by an external homogeneous input h . This effect is shown for various values of h .

inhibition characteristics. The network's dynamic follows

$$\tau \dot{u}_i(\theta, t) = -u_i(\theta, t) + x_i(\theta, t) + h_i(t) + \oint W_i^R(\theta - \phi) f(u_i(\phi, t)) d\phi \quad (1)$$

where $u_i(\theta, t)$ is the membrane potential of the neuron with preferred value θ at time t , belonging to the population or network i . The non-linear activation function is defined as $f(y) = \max(0, y)$. $x_i(\theta, t)$ corresponds to the external input and $h_i(t)$ to a global homogeneous input. The weight linking two neurons, with preferred directions θ and ϕ , is given by a periodical Gaussian profile g defined as

$$g(\theta - \phi) = \frac{1}{\kappa} \exp\left(\frac{\cos(\theta - \phi) - 1}{2\sigma^2}\right) \\ W_i^R(\theta - \phi) = \alpha_i (g(\theta - \phi) - 1) \quad (2)$$

where α_i and σ are, respectively, the amplitude and variance of the weights. $\kappa = 1 - e^{-\frac{1}{\sigma^2}}$ ensures that the weights are inhibitory and bounded, i.e. $W_i^R(\theta - \phi) \in [-\alpha_i, 0]$. These recurrent connections define the metric and the interaction strength between spatially localized inputs, such that close by stimuli cooperate in the representation, whereas far away ones compete and interfere.

2) *External inputs*: Each network i can receive an external input $x_i(\theta, t)$ also consisting of a periodic Gaussian g localized at φ_i in the neural space, such that

$$x_i(\theta, t) = \beta_i(t) (g(\theta - \varphi_i) - \eta) \quad (3)$$

where $\beta_i(t)$ is the input amplitude, and η is a normalization factor, ensuring that $\int x_i(\theta, t) d\theta = 0$. As illustrated in Figure 3a, an external input produces a unimodal increase of the population potential.

3) *Network projections*: In addition to the external inputs, a neural field can be subjected to synaptic projections arising from other populations. They are made through synaptic weights $W_{i \rightarrow j}^P(\theta - \phi)$ between neurons with preferred directions θ and ϕ of the source and target population, i and j respectively. They are defined by

$$W_{i \rightarrow j}^P(\theta - \phi) = \gamma_{i \rightarrow j} (g(\theta - \phi) - \eta) \quad (4)$$

such that the external input $x_j(\phi, t)$ of the target population becomes

$$x_j(\phi, t) = \oint W_{i \rightarrow j}^P(\theta - \phi) f(u_i(\theta, t)) d\theta \quad (5)$$

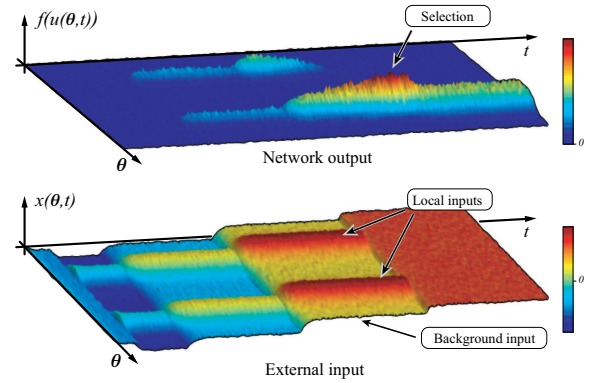


Fig. 4. Dynamics of input selection: Two equal and competing stimuli are presented to the network. The network converges towards a state where both stimuli are represented. Then the excitation level is increased and a small perturbation is induced in the inputs. This breaks the symmetry and leads to a single blob of activity which can be sustained.

where $\gamma_{i \rightarrow j} > 0$ is the strength of the projection weights.

4) *Stimulus enhancement*: An enhanced perception of a stimulus is characterized in the model by an increase in the stimulus amplitude. This can be induced by simply increasing the homogeneous excitatory input h_i in Equation (3). h_i acts as an attentional gain, which modulates the shape of the network's activity profile [8] (see Fig. 3).

5) *Selection, cooperation and interference*: The stimulus selection process is illustrated in Figure 4. First, two stimuli of equal strength are presented simultaneously at different locations in the resting network. The network will then converge towards a stage composed of two distinct regions of activity. Then, when the global background level of activity h_i reaches a certain threshold, the stable state of the neural field switches to an unstable state, causing the two active regions to compete. The interaction between the two regions will then cause even a small variation on intensity of one of the inputs to lead to a break in the symmetry, resulting in a single peak of activity.

B. System Architecture

The system's architecture, as depicted in Figure 5, can be separated into three major parts: a) sensory interfaces, b) the integrative and decisional network, and c) motor systems. As previously mentioned, it is assumed that, at the decisional level, each modality should be encoded into a common sensorimotor representation.

1) *Shared representation*: In order to choose a suitable common representation, we rely on a plausible cognitive method, involved in similar sensorimotor tasks [7]. The natural processing stages have been suggested to follow: a) visual perception, b) objects and gestures recognition, c) affordances and possible goals extraction, d) motor preparation, and e) execution. Since the proposed neural system involves intermediary steps c and d, the shared representation should, logically, reflect a mixture of motor and goal spaces. Furthermore, while considering experimental data [10], [11], [18], it seems likely that the brain uses a high level representation of goals, which combines spatial and motor reference frames. The data

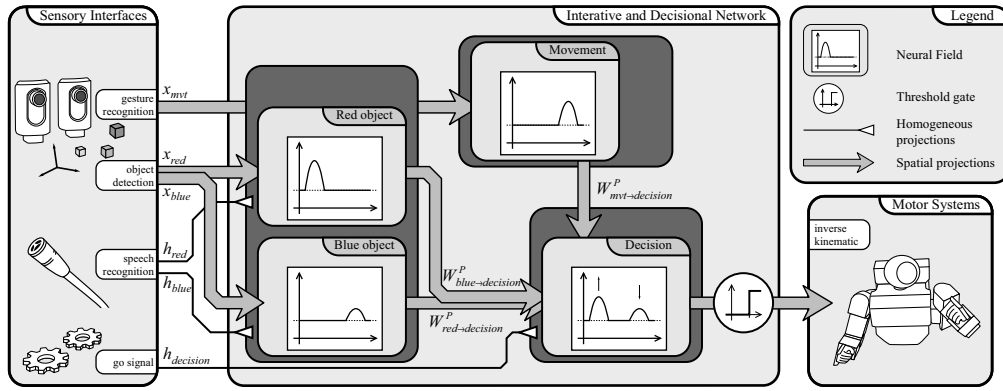


Fig. 5. Model architecture

supports movement direction and probability of selection as a crucial component of such a representation. Therefore, the θ variable (see Equ. 1) represents the angular direction of the targets relative to the robot, while the amount of network activity corresponds to the likeliness of a choice.

2) *Sensory systems*: The first modality is vision. Under our representational constraints, the locations of the two colored objects, i.e. the goals, are mapped onto the shared representation. As we see in Figure 5, the information relating to each possible target is fed into a distinct neural block. The 3D positions $\vec{r}_i, i \in \{red, blue\}$ of the targets are then mapped onto the network by setting the external inputs (Equ. 3) such that φ_i corresponds to the angular direction of \vec{r}_i relative to the robot body.

Another neural block is responsible for encoding observed movements. It works as follows: a significant displacement of the experimenter's hand towards an objects initiates the recognition of a reaching gesture, which mapped to a self-produced motor act, i.e. a self generated target at φ_i . The second modality, speech, is used to enhance the motor representation of the goal, to give it a greater importance. Thus, upon hearing one of the two color names, the corresponding modulatory input h_i is increased. Through the projections $W_{i,decision}^P$, the decisional system is influenced in favor of the corresponding object.

3) *Motor preparation and selection*: As the main task of the system is to perform a selection amongst different sensory sources, the information is then fed into a subnetwork which carries out the decisional process. Their common representation makes it easy to merge them, using the weights W^P . The go signal, triggering the initiation of the selection process, switches from 0 to $h_{decision}$ as soon as a gesture is observed or a word is heard. This value is also considered as the arousal level.

4) *Robot Control*: Finally, when the selection process has decided upon a goal, the robot movement is triggered. The triggering is carried out by the projection from motor preparation to motor execution, which is gated by a threshold mechanism. The robot initiates a movement only when the mean network

activity reaches a certain threshold δ , i.e. when:

$$\oint f(u_{decision}(\theta, t)) d\theta > \delta. \quad (6)$$

Then, since the network model uses a pseudo motor representation, coding for goals and movement direction, the action to be performed must be transformed into pure motor command, as given by the inverse kinematics function. The technique employed to resolve this problem uses the solution of the pseudo-inverse with optimization [4].

IV. RESULTS

The network was implemented on the robot¹ following the architecture depicted in Figure 5. Since our focus was not on explaining how to build such a network, but on demonstrating its use for dealing with multimodal inputs, the synaptic and inputs weights of the network were set manually, so as to reproduce human experimental data. Note that these weights can be learned by any kind of supervised learning, as shown in [8]. In order to illustrate the resulting robot behavior, two trials are depicted in Figure 6. These trials correspond to those presented earlier in an equivalent experiment between a human experimenter and a human subject (Fig. 1). It can be observed that the robot behaves similarly to a human, in both congruent and incongruent conditions. It even makes mistakes, in a similar way to the human subject.

We now consider the activity patterns produced by the network during such a task. In Figure 7, we see typical raster plots of the subnetworks' activity. We have displayed an incongruent trial as it represents the most interesting case. Initially, the two objects are on the table, which triggers an activation of the two object representations. This then leads to a small, but combined activation of the motor preparatory subnetwork. At *a*) the trial begins. The word *red* is spoken, and

¹Here is a short list of the parameters used in our experiments. The strengths of the recurrent weights are given by $\alpha_{\{blue,red\}} = 7.0$, $\alpha_{mvt} = 5.0$, $\alpha_{decision} = 9.0$, and the projections W^P by $\gamma_{\{red,blue\} \rightarrow decision} = 7.0$, $\gamma_{mvt \rightarrow decision} = 6.0$. The amplitude of the visual inputs $\beta_{\{blue,red,mvt\}} = 1.0$, the speech homogeneous inputs $h_{\{blue,red\}} = 2.0$, the excitable go signal $h_{decision} = 1.0$, and the execution threshold $\delta = 0.25$. Finally, the network variance $\sigma = 0.3$ and the ratio between the time constant of the object recognition subnetworks and the remaining is $\tau_{\{blue,red\}}/\tau = 5.0$.

the experimenter reaches towards the blue object. These events then cause, respectively, an enhancement of the corresponding object representation, and an activation of the movement recognition subnetwork. Since the auditory input takes longer to be integrated, the corresponding enhancement time is longer than that of the movement recognition. Therefore, *b)* the decisional network first balances its activity peak towards the visually induced movement choice. If the activity reaches the execution threshold, the visually induced movement is initiated, and the robots appears to be mistaken. Then, once the auditory integration activity surpasses that of the movement representation, *c)* the decision network modifies its choice, which now corresponds to the correct one.

Considering the arousal, or excitability level of the decisional network, simulations reveal that several operating modes can be attained through its variation. A typical example is shown in Figure 8. There are, respectively, two or three distinct operating modes in congruent and incongruent trials. In both cases, no response is initiated up to a certain threshold, the robot solely observes the scene. Then, as excitation increases, the robot begins to respond, but only in congruent trials, where the two sensory modalities positively cooperate. The reaction times shorten as excitation level increases. Furthermore, in the region marked as *d)* in Figure 8, only correct responses are given in incongruent trials. However, for high excitability levels, where the faster responses are found, the network can be mistaken and produces first a wrong answer and then corrects itself. In short, varying the arousal level modulates the speed-accuracy trade-off.

While in a computational and neuroscientific approach to human behavior, humanoid robots are a good method of grounding theory in practical experiments, however the reverse process should sometimes be considered with care. Consider the example in Figure 9, which is scenario adapted from that presented earlier, and also showing effects of the ideomotor facilitation. A human and a robot are walking towards each other, they both walk in a straight line, up until the point at which their trajectories must be modified in order to avoid a collision. In a similar way to what may happen between two people, the initiation of an avoidance movement by one person in one direction results in the facilitation by the other of the initiation of the same movement in the same direction, i.e. ideomotor facilitation. This may produce hesitation, and as can be seen here, sometimes collisions.

V. DISCUSSION AND CONCLUSION

In this paper, we proposed a biologically inspired approach to multimodal integration and decision making. This work is part of a multidisciplinary research project, in which we attempt to understand and model the cognitive mechanisms responsible for the human ability to imitate [7], [8]. Using a constructivist approach, we also attempt to ground our models in real physical robots. However, we should again stress that our goal here is not to find an efficient way to learn and perform such a task, but rather to look for a biologically plausible solution. Indeed, other technics such as

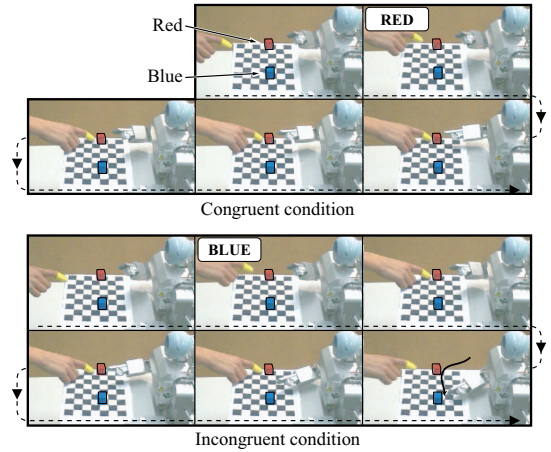


Fig. 6. The experiment depicted in Figure 1 was reproduced with the robot. Similar behavioral characteristics can be observed.

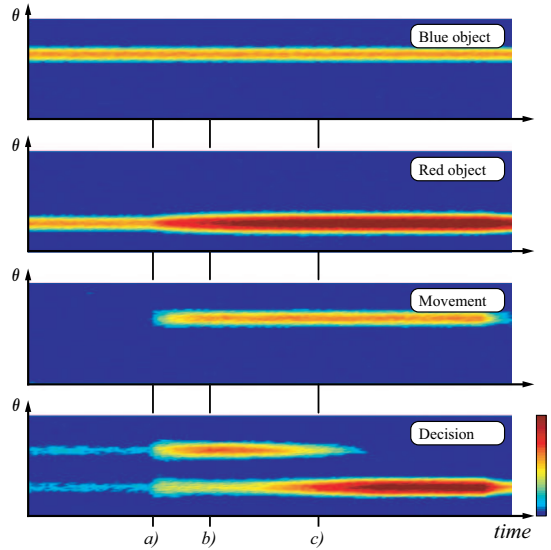


Fig. 7. Illustration of the neural field activity patterns during a trial consisting of an incongruent condition. *a)* The experimenter names the *red* word while simultaneously reaching for the blue object. At *b)* the movement execution threshold is attained. The robot is mistaken and initiates a reaching toward the wrong target. *c)* It finally corrects its decision and moves toward the right object. Please note that despite the activity patterns of the objects and movement networks are similar in *b)* and *c)*, the integration time within the decisional network is crucial and is at the origin of the delay in the movement switch.

bayesian networks or reinforcement learning would be more appropriate.

Our system consists mainly of an ensemble of neural networks known as neural fields, which possess several dynamical properties useful for sensorimotor integration [8], [13], [14], [19]. These networks allow other modalities to be added to our system, but at the cost of finding a relevant common representation. This issue is currently overcome by looking at the solution provided by human behavioral data and its neurophysiology. Indeed, similarly to the study of another model, also replicating experimental data [8], the architecture was successful in reproducing, in a robot, an experiment

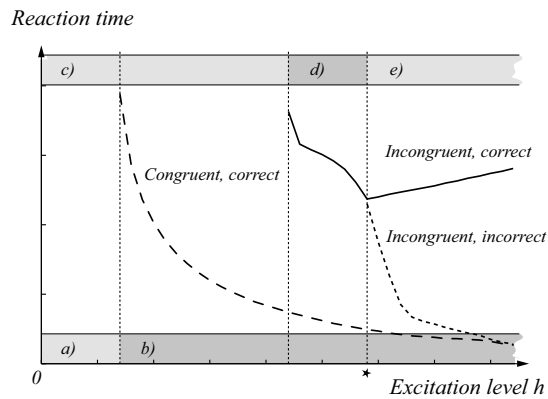


Fig. 8. Effect of the excitation level applied to the decisional network on the reaction times and on the correctness of the responses. In a congruent trial and a) at a very low excitation level, no responses are given. b) the response reaction time shortens as the excitation level increases. c) No responses range for an incongruent trial. d) A correct response is given, and e) the network produces first a wrong answer, then corrects itself. The * indicates the optimal arousal level for a no error with maximum speed policy.

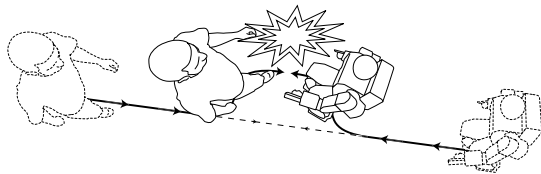


Fig. 9. Similarly to what happens between two people when crossing, a robot and a human may collide due to an ideomotor behavior.

originally designed to measure the ideomotor effect in humans. An interesting characteristic of this type of model is the ability to control the speed-accuracy trade-off by simply varying a network value: the arousal level. Robotic and automatic systems are generally developed under the constraint of being robust, i.e. no faults are tolerated, and a more continuous trade-off is rarely considered. However, such a capacity can be relevant for real world applications, where, for instance, computational processing speed and resources are limited, as well as in context-dependant situations.

To conclude, we previously mentioned an ergonomic problem in the interactions between humans and humanoid robots. It is supposed that artificial behaviors, which are coherent and similar to those of humans, can increase the likeness of a robot to a human, and facilitate social interactions [1], [3], [6]. However there is a drawback. Our model has been shown to imitate human behavior, which consists, in certain cases, of a facilitation of movement produced when observing similar movements. This, however, sometimes results in hesitation and even errors. In the crossing example (see Fig. 9), a collision between a robot, which has a much harder body with less elasticity, and a human can have worse consequences than if this occurred between two humans. Therefore, despite research into related areas such as social interactions, skill transfer, and theory of mind may provide us and robots with a deeper understanding of human behavior, the question remains as to whether we want to, or will accept robots making mistakes in

the same way that humans do.

ACKNOWLEDGMENTS

The authors would like to thank X. Perrin and C. Dubach, MSc students, for their contribution to the speech processing and the stereovision interfaces. Moreover, thanks to B. P. for her participation in the experiment shown in Figure 1. This work was supported by the Swiss National Science Foundation, through grant no 620-066127 of the SNF Professorships program.

REFERENCES

- [1] M. Asada, K. F. McDorman, H. Ishiguro and Y. Kuniyoshi, Cognitive developmental robotics as a new paradigm for the design of humanoid robots, *Robotics and Autonomous Systems*, Vol. 37, pp. 185-193, 2001.
- [2] A. Billard, S. Calinon and F. Guenter, Discriminative and Adaptive Imitation in Uni-Manual and Bi-Manual Tasks. *Robotics and Autonomous Systems*. In press, 2006.
- [3] C. Breazeal, Function meets style: Insights from emotion theory applied to HRI, *IEEE Transactions on Man, Cybernetics and Systems, Part C*, 2003.
- [4] S. Calinon, F. Guenter and A. Billard, Goal-Directed Imitation in a Humanoid Robot. In *Proceedings of the International Conference on Robotics and Automation, ICRA05*, 2005.
- [5] P. C. McGuire, J. Fritsch, J. J. Steil, F. Roethling, G. A. Fink, S. Wachsmuth, G. Sagerer and H. Ritter, Multi-Modal Human-Machine Communication for Instructing Robot Grasping Tasks, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS02*, pp. 1082-1089, 2002.
- [6] S. Woods, K. Dautenhahn, C. Kaouri, R. t. Boekhorst and K. L. Koay, Is this robot like me? Links between human and robot personality traits. *Proc. IEEE-RAS International Conference on Humanoid Robots*, pp.375-380, 2005.
- [7] M. Arbib, A. Billard, M. Iacoboni and E. Oztop, Mirror neurons, imitation and (synthetic) brain imaging, *Neural Networks*, Vol. 13, pp. 953-973, 2000.
- [8] E. L. Sauser and Aude G. Billard, Parallel and Distributed Neural Models of the Ideomotor Principle: An Investigation of Imitative Cortical Pathways, *Neural Networks*, Vol. 19(3), pp. 285-298, 2006.
- [9] M. Brass, H. Bekkering, A. Wohlschläger and W. Prinz, Compatibility between observed and executed finger movements: comparing symbolic, spatial and imitative cues, *Brain and Cognition*, Vol. 44, pp. 124-143, 2000.
- [10] G. Rizzolatti, L. Fogassi and V. Gallese, Neurophysiological mechanisms underlying the understanding of actions, *Nature Reviews Neuroscience*, Vol 2., pp. 661-670, 2001.
- [11] H. Bekkering, A. Wohlschläger, and M. Gattis, Imitation of gestures in children is goal-directed, *Quarterly Journal of Experimental Psychology*, vol. 53A, pp. 153-164, 2000.
- [12] J.R. Simon, P.E. Sly and S. Vilapakkam, Effect of compatibility of S-R mapping on reactions toward the stimulus source, *Acta Psychologica*, Vol. 1981, pp. 63-81, 1981.
- [13] S. I. Amari, Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics*, Vol. 27, pp. 77-87, 1977.
- [14] W. Erlhagen and G. Schöner, Dynamics field theory of movement preparation, *Psychological Review*, Vol. 109, pp. 545-572, 2002.
- [15] Intel Research, The opencv open source computer vision library, 2005, available at <http://www.intel.com/research/mrl/research/opencv/>.
- [16] X. Huang, F. Alleva, H.-W. Hon, M.-Y. Hwang, and R. Rosenfeld, The SPHINX-II speech recognition system: an overview, *Computer Speech and Language*, Vol. 7, pp. 137-148, 1993.
- [17] J. Decety and J. A. Sommerville, Shared representations between self and others: a social cognitive neuroscience view, *Trends in Cognitive Science*, Vol. 7, pp. 527-533, 2003.
- [18] P. Cisek and J. F. Kalaska, Neural correlates of reaching decision in dorsal premotor cortex: specification of multiple direction choices and final selection of action, *Neuron*, Vol. 45, pp. 801-814, 2005.
- [19] A. Pouget, P. and Dayan and R.S. Zemel, Inference and computation with population codes, *Annual Reviews Neuroscience*, Vol 26, pp. 381-410, 2003.