Augmenting the Reachable Space in the NAO Humanoid Robot

Marco Antonelli, Beata J. Grzyb, Vicente Castelló and Angel P. del Pobil
Robotic Intelligence Lab
Universitat Jaume I, Castellón, Spain
{antonell,grzyb,castellv,pobil}@uji.es

Abstract

Reaching for a target requires estimating the spatial position of the target and to convert such a position in a suitable arm-motor command. In the proposed framework, the location of the target is represented implicitly by the gaze direction of the robot and by the distance of the target. The NAO robot is provided with two cameras, one to look ahead and one to look down, which constitute two independent head-centered coordinate systems. These head-centered frames of reference are converted into reaching commands by two neural networks. The weights of networks are learned by moving the arm while gazing the hand, using an on-line learning algorithm that maintains the covariance matrix of weights. This work adapts a previously proposed model that worked on a full humanoid robot torso, to work with the NAO and is a step toward a more generic framework for the implicit representation of the peripersonal space in humanoid robots.

Introduction

Humans live surrounded by objects. Reaching for an object is one of the most common tasks of a human's everyday life. As robots are expected to be active participants in humans' daily life, they also need to have good reaching skills. Moreover, the robots need to be able to constantly learn and improve their reaching abilities autonomously so as to act on unknown objects in new environments.

Reaching for a target, however, is not an easy task. It requires to estimate the spatial position of the target and to convert it into an arm motor command. Estimation of the object position is problematic on its own as a three dimensional object is projected into two dimensional surface of camera sensor which in turn causes the distance to a target to be lost. Typically, the common solution is to employ stereopsis to reconstruct the depth of the scene. However, stereopsis is not the only cue that can be used to estimate the distance and, indeed, humans can perceive distance even with a single eye. A more flexible solution would be to use multiple cues to estimate the target position and then combine them in order to maximize the accuracy of the distance estimation.

Another challenge in reaching action, is the conversion of the object's spatial location into the arm position that allows reaching the target. The conventional approach here is to compute these transformations analytically by using the known geometric properties of the robotic system. Usually, such information is provided by the manufacturer or is estimated by applying calibration procedures. This approach to reaching allows achieving good performance, but only under the assumption that the parameters of the system are time invariant. In practice, it is not always the case, and the system needs to be re-calibrated periodically in order to keep working correctly. Therefore, it is convenient to develop a framework that continuously adapts the sensorimotor mapping to the constantly changing robot parameters.

In previous works we developed a framework that allowed the implicit sensorimotor mapping of the peripersonal space on a full humanoid robot torso (Chinellato et al. 2011; Antonelli, Chinellato, and del Pobil 2011). Instead of using the classical cartesian space, the spatial position of the target was encoded by the gaze direction and by the angular position of the arm joint. Indeed, these variables were implicit because they were directly provided by proprioception cues (encoders). The plastic maps were encoded by radial basis function networks because of their biological plausibility (Pouget and Sejnowski 1997) and their ability to approximate any kind of non-linear function (Park and Sandberg 1991).

This paper presents our reaching framework extended and adapted to work on a monocular robotic setup. Although, the NAO robot has two cameras, their location does not allow the use of vergence to estimate the distance of the target. Thus, the first challenge is to modify our radial basis function framework to work with the distance estimation provided by one monocular camera, as our previous framework was based only on the depth information provided by stereo cameras. Herein, the target position is represented by the gaze direction by which the target is viewed in the central region of the image, and by the estimated target distance. Alternatively, the same position is expressed in terms of the arm posture that allows for reaching the target. The transformation between these frames of reference is encoded by a radial basis function network (RBFN).

The visual fields of NAO's cameras intersect just in a limited range, thus the use of both cameras extends greatly the field of view of the robot. On the other hand, the existence of the common field of view raises an interesting issue of how to combine different cues to perform a more accurate reaching. Thus, the second challenge is identifying which camera provides a better estimation of the reaching movement, or how combine the result obtained from the two cameras. In our framework, we make use of both cameras to have a more coherent and augmented representation of the space.

The paper is structured as follows. The next section provides a background of the related work and briefly presents the neuroscientific findings that inspired our work. The subsequent section describes how the target can be implicitly encoded by the robot sensorimotor maps, which is then followed by the description of the computational model and learning algorithm. The next sections show our experimental setup and the results obtained from both computer simulation and real robot experiments. We close the paper with the discussion of the results and future work.

Background

Our approach to the sensorimotor transformation problem is inspired by neuroscientific findings, mainly concerned with human and primates' brain. Two types of visual processing exist in the brain, that is visual processing to obtain information about the features of objects such as color, size, shape ("vision for perception") in the ventral stream of visual cortex, and visual processing needed to guide movements such as reaching and grasping ("vision for action") in the dorsal stream of visual cortex (Goodale and Westwood 2004). The main cortical areas related to reaching action are V6A and MIP (Galletti et al. 2003; Fattori et al. 2001; Dechent and Frahm 2003; Caminiti, Ferraina, and Mayer 1998), both located in the parietal lobe. Findings in V6A neurons showed neurons that encoded the gaze directions and the distance of the target (Fattori et al. 2005; Marzocchi et al. 2008). Moreover, some neurons seemed to be involved in the execution of reaching movements (Galletti et al. 2003). These findings indicate that V6A is in charge of performing the sensorimotor transformations required for reaching for a given target in the 3D space.

The radial basis function networks are suitable for modeling the parietal cortex neurons as they are able to naturally reproduce the gain-field effects often observed in parietal neurons (Salinas and Thier 2000). Moreover, it was suggested that locations of objects in the peripersonal space are coded through the activity of parietal neurons that act as basis functions, and any coordinate frame can be read out from such population coding according to the task requirements (Pouget and Sejnowski 1997).

In robotics, even though extensive literature describes the problem of learning eye-hand coordination (Martinetz, Ritter, and Schulten 1990; Jones and Vernon 1994; Han, Okada, and Kondo 2006; Fuke, Ogino, and Asada 2009; Hoffmann, Schenck, and Möller 2005; Nori et al. 2007; Marjanovic, Scassellati, and Williamson 1996; Sun and Scassellati 2005), to the best of our knowledge only few papers describes the use of RBF networks (Marjanovic, Scassellati, and Williamson 1996; Sun and Scassellati 2005). The main differences of our model with respect their approach

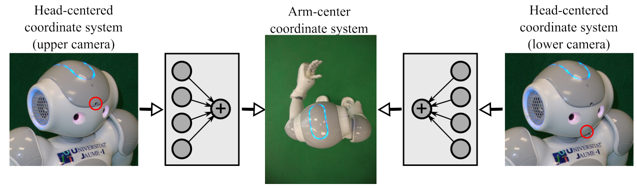


Figure 1: Computational framework of the sensorimotor integration model. A transformation for each camera allow converting the head position (pan-tilt-distance into arm-motor position (shoulder pitch, shoulder roll and elbow).

can be pointed out. Marjanovic et al. learned the transformation only on a surface of the space, that is the target distance was not explicitly taken into account (Marjanovic, Scassellati, and Williamson 1996). Sun et al. used a stereo system to compute the cartesian position of the target, while our system employes implicit variables (Sun and Scassellati 2005).

Model

In our proposed framework, the spatial position of the target object is maintained by three global frames of reference (f.o.r.), two centered in the cameras and one centered in the left arm. The two head-centered f.o.r.s (one for each camera of the robot) consist of a spherical-like coordinate system in which the azimuth and the inclination angles are replaced by the gaze direction, while the radius is the estimated distance of the target.

One important remark should be made about the use of the distance in the RBFN framework. Indeed, the distance is not directly observable by the robot, that is, it is not an implicit variable. However, primates have access to several cues that can be used to estimate the distance, such as stereopsis, familiar size, motion parallax and so on (Landy et al. 1995). These variables are implicit and are inversely proportional to the distance and could be used in our framework in place of the distance. For example, our previous work, used vergence alone (Chinellato et al. 2011). However, when multiple cues are available, it seems more reliable to integrate the cues altogether before calculating the arm position. Such a computation can be performed by a three layer neural network with reward-mediated learning similarly to what is done in (Karaoguz et al. 2011). Thus, in our framework, it is possible to replace the distance with the output of another computation as long as it provides neural activation which is related with the distance of the target. In this way, the framework becomes more general and can be used independently of the cues available to estimate the distance.

The arm position also provides the spatial position of the target when the robot is touching the object. Herein, we use just one arm, the left one. In this case, the target position is encoded by the angles of the joints which are provided by the proprioceptive signals. Usually the arm-centered f.o.r. is redundant in the representation of the position, because many joint configurations can bring the hand in the same spatial position. The implication is that the mapping between the head-centered f.o.r. and the arm-centered f.o.r. is not a single-valued function.

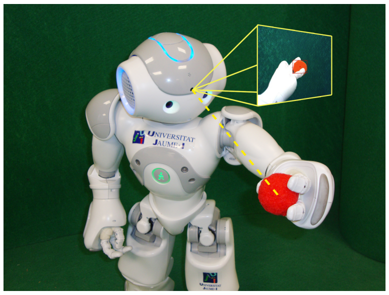
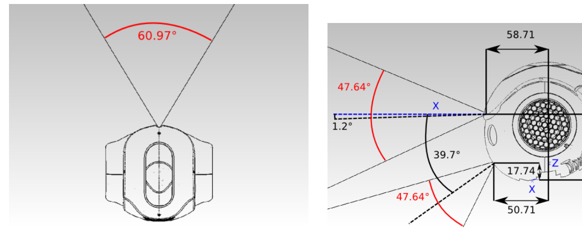


Figure 2: Association between the oculomotor and arm-motor signals. When the robot moves its hand and gazes towards the same point, it can update its sensorimotor representation to locally reduce the error of the transformations.



(a) Top view (b) Lateral view

Figure 3: Sketch of the NAO robot's cameras taken from the manufacturer website.

As the main focus of this work is put on the learning of the sensorimotor transformations, the redundancy problem is bypassed by simplifying the experimental setup. Therefore, only three joints of the arm, two for the shoulder and one for the elbow are used. In this way, the two transformations became functions, enabling to employ a RBFN for encoding each sensorimotor transformation (see Fig. 1).

Both maps are updated when the hand position and the gaze direction are pointing to the same spatial location. The robot autonomously verifies such a condition by checking whether the visual position of the hand is in the center of the visual field (see Fig. 2). If the hand is visible but it is not in the center of the image, the robot gazes the hand to reinforce the head-arm association. The gazing is performed using a hard-coded proportional controller based on the visual input, but a saccadic map can be implemented as proposed in the previous work (Antonelli, Chinellato, and del Pobil 2011).

In principle, the mapping between the distinct sensorimotor modalities can be learned during the interaction with the environment, through reaching and gazing movements. After each performed movement, visual feedback is used to check the coordination of gaze and arm. At the beginning, the system does not have any previous knowledge of the sensorimotor transformation, thus, random movements can be introduced to begin the exploration of the environment. Successively, these random movements can be suppressed and the system can keep adapting during the goal-directed exploration.

Radial Basis Function Networks

Learning of the sensorimotor transformations, as stated in the previous section, can be seen as a function approximation problem. Radial basis function networks can potentially approximate any function with the desired precision, so they are especially suitable for encoding sensorimotor transformations (Pouget and Sejnowski 1997).

Basis function networks are three-layer feed-forward neural networks whose hidden units perform a non-linear transformation of the input data, whereas the output is computed as a linear combination of the hidden units. Let $\mathbf{y}$ be the output of the network, it can be expressed as:

$$\mathbf{y} = \mathbf{h}^{\mathbf{T}}(\mathbf{x}) \cdot \mathbf{W} \tag{1}$$

where, $\mathbf{h}$ is the activation of the hidden layer for the input $\mathbf{x}$ and $\mathbf{W}$ is the matrix of the weights. Radial basis functions are a particular case of basis functions which values depend on the euclidean distance from a point, called *center of activation*.

Learning in the context of the radial basis function networks can be divided into two phases. An unsupervised phase allows setting the parameters of the network, such as the number of hidden units and the position of their centers, whereas a supervised phase allows for finding the values of the weights.

In the proposed framework, we employed fixed centers, whose receptive fields could not move according to the input data. Gaussian functions were used as non-linear transformation of the input. Hidden units were characterized by their center of activation $\mathbf{c}_i$, whereas the width of the "bell" curve Σ was the same for every units:

$$h_i(\mathbf{x}) = h(||\mathbf{x} - \mathbf{c}_i||) = e^{-(\mathbf{x}-\mathbf{c}_i)^T \Sigma^{-1} (\mathbf{x}-\mathbf{c}_i)} \tag{2}$$

Using this setup, the learning process is reduced to find the weights that better approximate the sensorimotor transformation. Given a new input-output sample of the target function, the weights $\mathbf{W}$ can be updated to reduce the error ($\mathbf{e}$) of the network by minimizing sum of the square error. Among the available techniques, we employed the recursive least square algorithm (RLS) because of its fast convergence rate. Moreover, the RLS keeps trace of the covariance matrix $\mathbf{P}$ which can be used to evaluate the accuracy of the transformation. The algorithm is the following:

$$S = \mathbf{h}^T \mathbf{P} \cdot \mathbf{h} + R \tag{3}$$

$$\mathbf{K} = \mathbf{P} \cdot \mathbf{h} \cdot S^{-1} \tag{4}$$

$$\mathbf{W} = \mathbf{W} + \mathbf{K} \cdot \mathbf{e} \cdot \mathbf{P} \tag{5}$$

$$\mathbf{P} = \mathbf{P} - \mathbf{K} \cdot \mathbf{h}^T \cdot \mathbf{P} \tag{6}$$

where R is the variance of the observation error and $\mathbf{K}$ is the Kalman gain. The matrix $\mathbf{P}$ is initialized to large values ($10^4 \cdot \mathbf{I}$), whereas $\mathbf{W}$ is initialized to zero.

Experimental Framework

Robotic Setup

Aldebaran's commercially available humanoid robot NAO was used as platform for testing the proposed framework. The robot is provided with 25 degrees of freedom (d.o.f.s)

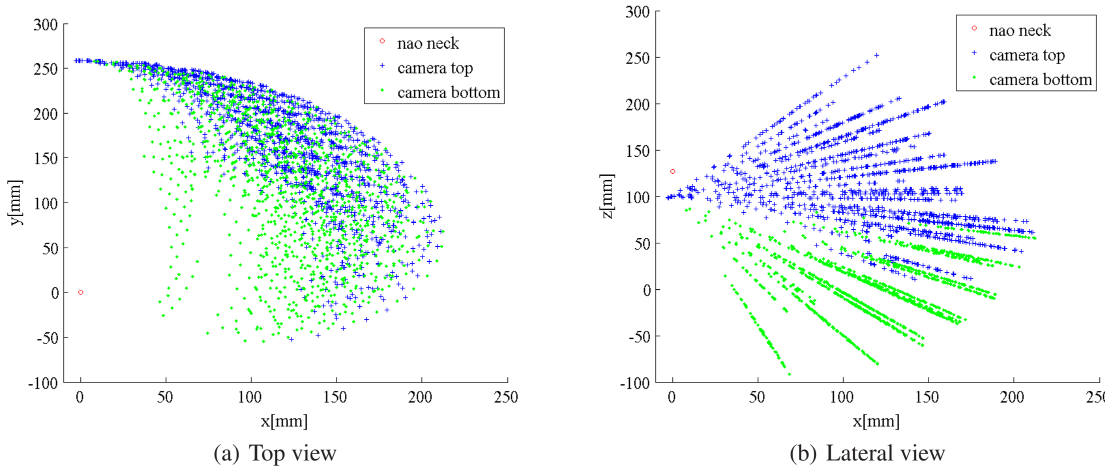


Figure 4: Distribution of the training points in the cartesian space. The x-z axis define the sagittal plane while the y-z axis define the coronal plane. The ground is the plane $z = 0$.

among which two are placed in the head (pan and tilt) and five in each arm. Herein, we have used just three d.o.f.s for the arm. The vision system of the robot is composed by two identical cameras, but not in the classical stereo setup (see Fig. 3).

Training data collection

Learning of the sensorimotor transformation is essentially approximating the function through training samples of the form $(d, \theta_{head}, \theta_{arm})i = 1, 2, ..., n$ where d is distance from the camera to the robot's hand, $\theta_{head}, \theta_{arm}$ are the joint angles of the head and arm, respectively, and n is the size of the training set. Such a training set was generated by moving the arm, while the robot was gazing the hand. Herein, a visual marker (a ball) was used to facilitate the recognition of the hand in a visual field of view. The distance of the hand was computed using the familiar size of the ball. That is, knowing the physical size of an objects ($S_{physical}$), its absolute depth (d) was calculated by using equation (7):

$$d = f \times \frac{S_{physical}}{S_{observed}} \tag{7}$$

where and $S_{observed}$ is size of the object observed in the image, while f is the focal length. Both $S_{observed}$ and f are expressed in terms of pixels. Equation (7) assumes that the observed feature is presented orthogonally to the camera sensor. Such a restriction is avoided by the choice of the ball as visual marker.

The training points shown in Fig. 4 were collected during the exploration of the arm-joint space. As it can easily be observerd, the fields of view of the cameras intersect just within a small area. These training points were inherently noisy, due to small imprecisions in identifying the center and the distance of the marker in the image and to the proprioceptive errors. However, the ability of averaging–typical of the feed-forward neural networks–overcome the problem.

The structure and parameters of the RBFNs were the same for both cameras and were chosen using a heuristic search on a simulated model of the robot. We employed Gaussian receptive fields, uniformly distributed on a lattice composed of 7x7x7 neurons. The spread of the Gaussian activation Σ was a diagonal matrix $\sigma \mathbf{I}$ with σ was set to 0.28. The input space of the two RBFNs was the pan-tilt-distance space normalized between 0 and 1, while the output space was the shoulder (pitch and roll)-elbow space. The weights of the networks were learned using the recursive least square algorithm on the training samples.

Results

Learning the Sensorimotor Transformations

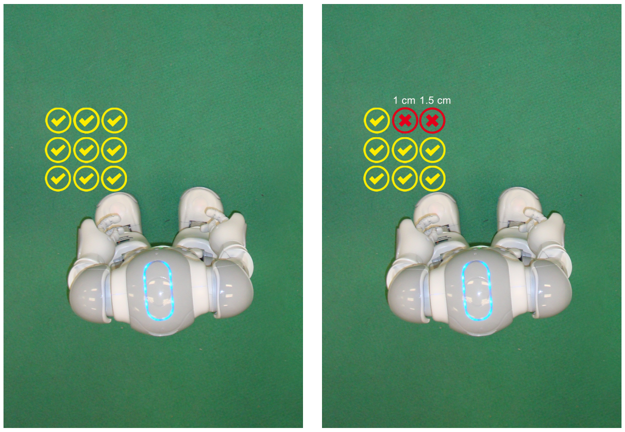
At the end of the data collection process, the networks were tested off-line on the acquired sample points using the K-Fold cross validation with K set to 5. Once calculated the error in the joint-space, it was converted into the cartesian space using the kinematics of the arm. Thus, in order to simplify the evaluation of the results, the error was shown as euclidean distance in the cartesian space between sampled and computed values (see Table 1). The magnitude of the error was small enough to allow the robot to reach a target.

Grasping Task

The performance of the system trained with the whole dataset was tested on a grasping task. The robot had to localize and to grasp a ball easily identificable by its color. The ball was placed on two lattices of 3 by 3 points that covered a region of 5 cm by 8 cm (x,y) on left side of the robot (see Fig. 5). The grasping was executed using the two cameras independently to evaluate the single performance. In both cases, the arm began every movement from a safe position that allowed reaching the ball without any collision. During the training of the transformations, the ball was put in the center of the hand, so we expected that a correct arm move-

Table 1: Parameters of the RBFNs and their performances on the training set using the K-Fold cross validation (K=5). Mean error and standard deviation ($\mu \pm \sigma$) are expressed in mm in the cartesian space.

Camera	RBFN par. centers	RBFN par. radius(σ)	N. points	K=1 $\mu \pm \sigma$	K=2 $\mu \pm \sigma$	K=3 $\mu \pm \sigma$	K=4 $\mu \pm \sigma$	K=5 $\mu \pm \sigma$
Top	7x7x7	0.28	1458	7.32 ± 7.55	8.44 ± 14.28	7.31 ± 9.10	7.65 ± 10.49	6.41 ± 6.71
Bottom	7x7x7	0.28	1254	7.52 ± 8.22	7.75 ± 6.48	7.27 ± 8.47	7.09 ± 5.26	6.91 ± 5.90



(a) Upper camera (b) Lower camera

Figure 5: Experimental setup. A ball was put on a lattice on the left side of the robot to test the grasping task. Yellow markers indicate a correctly grasping task while the red ones indicate a failure.

ment would bring the center of the hand near the ball. The robot grasped correctly the ball 9 times out of 9 when using the upper camera and 7 out of 9 when using the lower one. When the failure occurred, distance of the target was measured and it was about 1 cm in one case and 1.5 in the other case (see Fig. 5).

Integrating different fields of view

As was mentioned before, the sensorimotor transformations allowed for grasping a target ball in the 100% of the trials using the upper camera and in the 77.8% using the lower one. During this experiment, however, the information provided by the covariance matrix was not taken into account.

The covariance matrix can be exploited to calculate the variance σ^2 of the transformation for the new observed input:

$$\sigma^2 = \mathbf{h}^T \cdot \mathbf{P} \cdot \mathbf{h} \tag{8}$$

where $\mathbf{h}$ is the activation of the hidden layer for the new input and $\mathbf{P}$ is the covariance matrix. The variance obtained by the transformations of the two cameras can be used, within their small intersecting field of view, to compare or merge the output of the RBFNs.

Preliminary results in this direction was obtained by positioning the ball in 50 different positions visible by both cameras. For each position, we calculated the errors made by: 1) each transformation alone; 2) the average of the two transformations; 3) the average weighted by the variance value; 4) the transformation with the minimun variance.

Table 2: Mean error [mm] during the reaching task using to upper camera (Top), the lower camera ($Bottom$), the average between the two camera (AV), the average weighted by the variance (WAV) and by taking the transformation with the minimun variance (MV).

Top	$Bottom$	AV	WAV	MV
8.78	8.47	6.75	5.44	6.35

The results presented in Table 2 are quite promising and show that the use of the covariance matrix can improve the accuracy of the reaching movement.

Discussion and Future Work

This work is focused on the encoding of the visuomotor transformations that allows for eye-hand coordination. The RBFNs were trained with real data collected while the robot was gazing its hand. The training was performed with an on-line algorithm that keeps trace of the covariance matrix P of the transformation, which can be used to evaluate the accuracy of the transformation.

In the currently implemented framework, distance was calculated using the familiar size of the object. Such a distance, however, can be estimated by others cues, e.g. motion parallax, kinetic depth effect and so on, which can be integrated together in the spirit of the Bayesian theorem in order to obtain a reliable distance estimation. Our future work will focus on the integration of the proposed sensorimotor framework with another model that implicitly encodes the distance perceived by several cues.

This work is a part of a larger framework that is inspired by infant development. The final goal is to provide the robot with a coherent near and far space representation. The visuomotor knowledge of the peripersonal and extrapersonal space should be built in a dynamical way, through the active interaction with the environment in a similar way as infants do. Following this approach, the robot has to be able to keep learning during its normal behavior, by interacting with the world and continually update the representation of the world itself. Moreover, the learning process should be self-supervised in order to avoid the need of an external teacher. That is, the robot should be able to improve its capabilities by observing the outcome of its actions.

Conclusions

This paper presented a framework for sensorimotor transformations that is inspired by neuroscientific findings. The plausibility of our framework was tested with the NAO humanoid robot. Although the robot's cameras were not located in a typical stereo-vision setup, we demonstrated how they can complement each other augmenting the robot's effective reachable space. We showed that these space representations are very plastic as the robot were able to update and to improve its performance during interaction with the environment. Moreover, the adaptation of our framework on the NAO robot, further supports the extendability and generality of our approach.

Acknowledgements

This research was partly supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Grant No. R31-2008-000-10062-0), by Ministerio de Ciencia e Innovación (FPU grant AP2007-02565, FPI grant BES-2009-027151, DPI2011-27846), by Generalitat Valenciana (PROMETEO/2009/052) and by Fundació Caixa Castello-Bancaixa (P1-1B2011-54).

References

Antonelli, M.; Chinellato, E.; and del Pobil, A. 2011. Implicit mapping of the peripersonal space of a humanoid robot. *IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain* 1–8.

Caminiti, R.; Ferraina, S.; and Mayer, A. B. 1998. Visuomotor transformations: early cortical mechanisms of reaching. *Current opinion in neurobiology* 8(6):753—761.

Chinellato, E.; Antonelli, M.; Grzyb, B.; and del Pobil, A. 2011. Implicit sensorimotor mapping of the peripersonal space by gazing and reaching. *Autonomous Mental Development, IEEE Transactions on* 3:45–53.

Dechent, P., and Frahm, J. 2003. Characterization of the human visual V6 complex by functional magnetic resonance imaging. *European Journal of Neuroscience* 17(10):2201–2211.

Fattori, P.; Gamberini, M.; Kutz, D.; and Galletti, C. 2001. Arm-reaching neurons in the parietal area V6A of the macaque monkey. *European Journal of Neuroscience* 13:2309–2313.

Fattori, P.; Kutz, D.; Breveglieri, R.; Marzocchi, N.; and Galletti, C. 2005. Spatial tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey. *European Journal of Neuroscience* 22(4):956–972.

Fuke, S.; Ogino, M.; and Asada, M. 2009. Acquisition of the head-centered peri-personal spatial representation found in vip neuron. *Autonomous Mental Development, IEEE Transactions on* 1(2):131–140.

Galletti, C.; Kutz, D.; Gamberini, M.; Breveglieri, R.; and Fattori, P. 2003. Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Experimental Brain Research* 153(2):158–170.

Goodale, M., and Westwood, D. 2004. An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Current opinion in neurobiology* 14(2):203–211.

Han, M.; Okada, N.; and Kondo, E. 2006. Coordination of an uncalibrated 3-d visuo-motor system based on multiple self-organizing maps. *JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing* 49(1):230–239.

Hoffmann, H.; Schenck, W.; and Möller, R. 2005. Learning visuomotor transformations for gaze-control and grasping. *Biological Cybernetics* 93(2):119–130.

Jones, M., and Vernon, D. 1994. Using neural networks to learn hand-eye co-ordination. *Neural Computing and Applications* 2(1):2–12.

Karaoguz, C.; Weisswange, T. H.; Rodemann, T.; Wrede, B.; and Rothkopf, C. A. 2011. Reward-based learning of optimal cue integration in audio and visual depth estimation. In *The 15th International Conference on Advanced Robotics, Tallinn, Estonia*.

Landy, M. S.; Maloney, L. T.; Johnston, E. B.; and Young, M. 1995. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research* 35(3):389 – 412.

Marjanovic, M.; Scassellati, B.; and Williamson, M. 1996. Self-taught visually guided pointing for a humanoid robot. *From Animals to Animats 4: Proc. Fourth Int l Conf. Simulation of Adaptive Behavior* 35—44.

Martinetz, T. M.; Ritter, H. J.; and Schulten, K. J. 1990. Three-dimensional neural net for learning visuomotor coordination of a robot arm. *IEEE T Neural Networ* 1(1):131–136.

Marzocchi, N.; Breveglieri, R.; Galletti, C.; and Fattori, P. 2008. Reaching activity in parietal area V6A of macaque: eye influence on arm activity or retinocentric coding of reaching movements? *European Journal of Neuroscience* 27(3):775–789.

Nori, F.; Natale, L.; Sandini, G.; and Metta, G. 2007. Autonomous learning of 3d reaching in a humanoid robot. *IEEE/RSJ IROS, International Conference on Intelligent Robots and Systems* 1142–1147.

Park, J., and Sandberg, I. W. 1991. Universal approximation using radial-basis-function networks. *Neural Comput.* 3(2):246–257.

Pouget, A., and Sejnowski, T. J. 1997. A new view of hemineglect based on the response properties of parietal neurones. *Philosophical Transactions of the Royal Society B: Biological Sciences* 352(1360):1449–1459.

Salinas, E., and Thier, P. 2000. Gain modulation: a major computational principle of the central nervous system. *Neuron* 27(1):15–21.

Sun, G., and Scassellati, B. 2005. A fast and efficient model for learning to reach. *International Journal of Humanoid Robotics* 2(4):391–414.