



Implementation of a neurophysiological model of saccadic eye movements on an anthropomorphic robotic head

Luigi Manfredi, Eliseo Stefano Maini, Paolo Dario, Cecilia Laschi, Benoît Girard, Nicolas Tabareau, Alain Berthoz

► To cite this version:

Luigi Manfredi, Eliseo Stefano Maini, Paolo Dario, Cecilia Laschi, Benoît Girard, et al.. Implementation of a neurophysiological model of saccadic eye movements on an anthropomorphic robotic head. 6th IEEE-RAS International Conference on Humanoid Robots, Dec 2006, Gênes, Italy. pp.438-443, 10.1109/ICHR.2006.321309 . hal-01525198

HAL Id: hal-01525198

<https://hal.science/hal-01525198>

Submitted on 1 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Implementation of a neurophysiological model of saccadic eye movements on an anthropomorphic robotic head

Luigi Manfredi

Institute for Advanced Studies IMT
Lucca, Italy

luigi.manfredi@imtlucca.it

Eliseo Stefano Maini, Paolo Dario,
Cecilia Laschi

ARTS Lab - Polo Sant'Anna Valdera
Scuola Superiore Sant'Anna
Pontedera, Italy

smaini@arts.sssup.it,
dario@sssup.it,
cecilia@arts.sssup.it

Benoît Girard, Nicolas Tabareau,
Alain Berthoz

Laboratoire de Physiologie
de la Perception et de l'Action - UMR 7152
CNRS - Collège de France
Paris Cedex 05, France.

benoit.girard@college-de-france.fr,
nicolas.tabareau@college-de-france.fr,
alain.berthoz@college-de-france.fr

Abstract—In this paper we investigated the relevance of a robotic implementation in the development and validation of a neurophysiological model of the generation of saccadic eye movements. To this aim, a well-characterized model of the brainstem saccadic circuitry was implemented on a humanoid robot head with 7 degrees of freedom (DOFs), which was designed to mimic the human head in terms of the physical dimensions (i.e. geometry and masses), the kinematics (i.e. number of DOFs and ranges of motion), the dynamics (i.e. velocities and accelerations), and the functionality (i.e. the ocular movements of vergence, smooth pursuit and saccades). Our implementation makes the robot head execute saccadic eye movements upon a visual stimulus appearing in the periphery of the robot visual field, by reproducing the following steps: projection of the camera images onto collicular images, according to the modeled mapping between the retina and the superior colliculus (SC); transformation of the retinotopic coordinates of the stimulus obtained in the camera reference frame into their corresponding projections on the SC; spatio-temporal transformation of these coordinates according to what is known to happen in the brainstem saccade burst generator of primates; and execution of the eye movement by controlling one eye motor of the robot, in velocity. The capabilities of the robot head to execute saccadic movements have been tested with respect to the neurophysiological model implemented, in view of the use of this robotic implementation for validating and tuning the model itself, in further focused experimental trials.

Index Terms - Saccadic eye movements, robotic head, biomimetic implementation, neurophysiological model.

I. INTRODUCTION

A very relevant question that often arises in robotics research is why we should devote significant research efforts for developing 'human-like' robots. This question may result extremely appropriate at a stage where robotics technology is providing robust solutions to some technical problems that have constrained robot development over past years. The recent evolution of robotics technology has been demonstrated to solve some of the complex engineering issues that related

to the design of humanoid robots [1] [2]. One of the main motivation to humanoid robotics is the capability for the robot to operate in real-world environments and to establish some level of social and emotional interaction with human beings [3] [4]. In our vision, anthropomorphism in robotics is aimed not only to build an artificial human or to facilitate the interaction of the robotic artefact with human beings but it rather represents a compelling condition for taking advantage of the robotic artefact as a platform through which specific human behaviours can be investigated. In other words, an anthropomorphic robotic artefact may represent an excellent platform whereby computational models can be implemented and tested in order to better understand some human abilities [5][6]. According to this vision, in this paper we investigate the relevance of a robotic implementation in the development and validation of a neurophysiological model of the generation of saccadic eye movements. From neurophysiological studies it is known that saccadic eye movements are rapid eye movements occurring at approximately 800 deg/s that are used to position points of interest in the visual field into the high-resolution central part of the retina (fovea) for further examination. Eye saccades have been studied for more than 30 years by neuroscientists, using a wide spectrum of methods (neuroanatomy, electrophysiology, neuroimaging, behavior, etc.). Indeed, the saccadic system allows the study of motor and premotor brain mechanisms, in a mechanical system much simpler than, for example, limbs. Moreover, saccadic selection tasks are simple to teach to animals and to settle, and have thus been used in numerous studies of cognitive functions (working memory, sequence learning, etc.). In robotics, the investigation of saccadic movements have the purpose to implement fast eye movements, that have been implemented in industrial robots and commercial vision systems. Generally speaking, this purpose does not benefit from bioinspired approaches to the control of saccadic movements, according to a human-

439

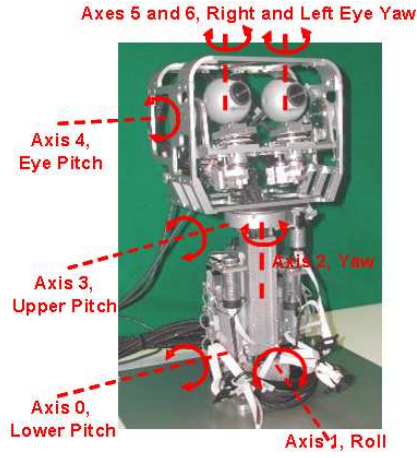


Fig. 2. Robotic Platform

order to limit computational load. As a result, no recurrent connections are used in the SC maps, as the input are forced to 2D Gaussians, and the visual bursts usually preceding the motor one in the motor layer are not modeled.

Collicular maps are modeled by 36×36 tables of leaky-integrator neurons including 5 neurons border. The system has the following hierarchical structure (for a formal description, see Appendix: Model) :

- 1) The retina *Ret* encodes the target's position by a 2D Gaussian with standard deviation $\sigma = 2.5$, centered around target's position.
- 2) The visual layer *Vis* receives the retinal signal with a 70ms delay. When its activity passes a given threshold, it is transmitted to the motor layer and the SBG OPNs are inhibited via Long-lead Burst neurons (LLB).
- 3) The motor layer *Mot* sends the command to the SBG, while its activity is integrated. When the integrator reaches a given threshold, the layer is inhibited and the saccade stops.
- 4) The SBG is first inhibited by the OPNs. The activity in *Vis* is transmitted to the LLBs which inhibit in turn the OPNs. When the activity in *Vis* is strong enough, OPNs are turned off and the EBNs/IBNs begin to receive the motor command from the *Mot* layer through a weighted sum. This command is then integrated by the couple of neurons TNs (tonic neurons). The activity of MNs is received by the eye plant (modeled by a second order differential equation) to generate the required eye's displacement.

III. IMPLEMENTATION OF SACCADIC MOVEMENTS ON THE HUMANOID ROBOT HEAD

A. The robotic platform

The robotic platform that we used to implement the proposed model is an anthropomorphic robotic head equipped with a stereoscopic vision system and endowed with 7 (DOFs) given by revolute joints.

As reported in Fig. 2, the DOFs are organized as follows:

TABLE I
PERFORMANCE OF THE ROBOTIC HEAD

| Operating | Max. range [deg] | Max. velocity [deg]/[s] | Max. acceleration [deg]/[s ²] |
|-----------|------------------|-------------------------|---|
| J0 | -28 32 | 20 | 200 |
| J1 | -32 26 | 25 | 200 |
| J2 | -108 108 | 120 | 750 |
| J3 | -20 30 | 100 | 750 |
| J4 | -25 53 | 400 | 4500 |
| J5 | -45 45 | 600 | 10000 |
| J6 | -45 45 | 600 | 10000 |

- 4 DOFs allow the neck movements (J0, J1, J2 and J3),
- 3 DOFs specifically allow the independent movements of eyes (J4, J5, J6).

The eye kinematic structure consists of one common pitch and two independent yaws. All these axes are actuated by a belt driven mechanism with a DC motor and a harmonic drive system. The 2 DOFs corresponding to yaw movements of the eyes (J5, J6) allow the vergence of the two cameras to foveation a visual target. The neck has 4 DOFs, for lateral flexion, rotation and ventral/dorsal flexion, this latter movement is obtained with 2DOFs, i.e. two pitches at different heights. The yaw axis is also driven by a belt driven mechanism with a DC motor and a harmonic drive system, while the two pitch, the roll are driven by a belt driven mechanism with a DC motor and a ball screw actuator. Each motor of the head is equipped with encoder to measure joint revolution and velocity. All the motors, gear-heads, belt reduction ratios and system actuators have been designed on the basis of the geometry and performance of the human head in terms of DOFs, ranges of motion, speeds and accelerations, allowing smooth pursuit movements (i.e slow tracking movements) as well as saccadic eye movements like in humans. Each joint has a switch-end to give the possibility to implement the system calibration. Table I reports features of the 7 DOFS.

B. The vision system

The vision system of the robotic head is a uncalibrated stereo-rig of two USB cameras with a maximal resolution of 640x480 pixel and a frame rate of 30 fps. The detection of the stimuli is performed with a simple procedure of colour recognition based on a target colour inserted by the user. The procedure implemented in this work makes use of the HSV colour representation, instead of the classical RGB, in order to improve the robustness with respect to light variations. Similarly to what happens in humans, during a saccadic movement the system provides the suppression of vision [13].

C. Implementation of the collicular mapping

The transformation from the retinotopic coordinates (α, β) of targets into coordinates on the SC surface (X, Y) is expressed by equation (1) (Fig. 3).

It is worth noticing that, a straight use of this transformation for obtaining the collicular mapping may introduce some relevant ambiguities in real-world applications. Indeed, the

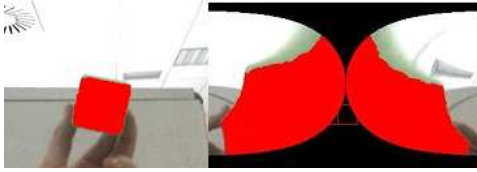


Fig. 3. An example of the mapping of one retino-topic image into the collicular coordinates, according to (1). Please note that the right and left side of the collicular mapping correspond to the right and left halves of the original image, respectively

CCD cameras have a discrete pixel resolution whereas the equation (1) is defined on a continuous domain. It is obvious that the logarithmic function maps the discrete points of the retinotopic plane in such a way that spreads the pixels in the central part of the collicular map. Conversely, in the peripheral part of the collicular map the density of pixels will increase with a consequent loss of information. In order to avoid this problem we operated as follows: we started from a void collicular map and for each pixel in the collicular map we back-projected the coordinates onto the retinotopic plane making use of the inverse equation of (1) as reported in [11]. The obtained coordinates (X,Y) will be - in the general case - not integers therefore the obtained point in the retinotopic coordinates will have no physical representation on the CCD matrix. In this case we associated to this virtual point an intensity value that was obtained by a weighted interpolation of the surrounding integer pixels. Finally we imposed the latter intensity value to the original point of the collicular map. By iterating this procedure for each point on the collicular plane we produced well smoothed collicular maps as the one reported in Fig. 7

D. Implementation of the velocity profile generation

For generating the velocity profiles of the saccadic movements we used the analytical model that is reported in the Appendix for the sake of clarity. The position, velocity and acceleration values of α and β were calculated every 1 ms making use of the MN firing rate and of the second order differential equation representing the eye plant which is reported and described in the Appendix. It is worth noticing that the computational time of the model was calculated and subtracted from the actual delays that are known to occur in a saccadic movement (approximately 200 ms) in order to find an appropriate delay to be feed to the motors. In this way the robotic artefact was endowed with the same response time of the biological system. In Fig. 4 we report a typical example of the obtained position and velocity profiles.

E. Experimental setup

The system architecture is composed of three fundamentals parts: the Peripheral Computer (PC) the Central Computer (CC) and the Robotic platform. Inside the CC, the low level control is implemented and it is directly connected to the Robotic Platform. It is intended to be used as a server, waiting for the command and executing it. Both object detection and

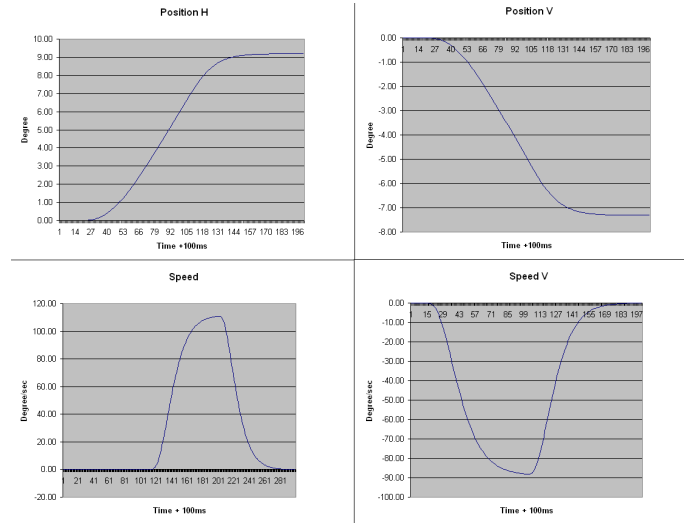


Fig. 4. Position and velocity profiles obtained with a sample stimulus in the periphery.

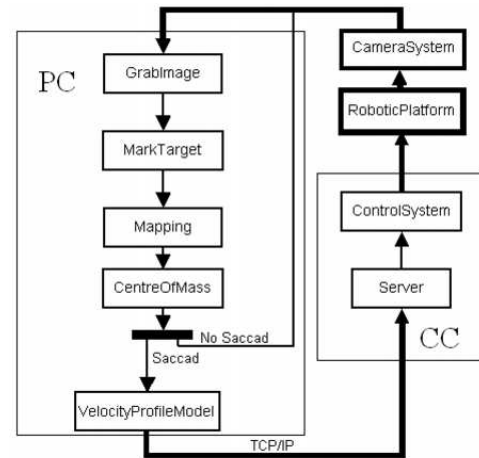


Fig. 5. Control System

collicular mapping are implemented on the PC. The steps to obtain a saccadic movement of the robotic eye, elicited by a visual stimulus in the periphery of the image, are executed on the PC. The first step is detecting the selected color. The second step is mapping the image from the retina to the colliculus. The third step is detecting pixels with a color similar to the selected one in the collicular map. Then, it is necessary to find a center of mass for the stimulus and to give its coordinates to the saccade generator model. The model generates a profile of velocity with a sampling time of 1 ms. The control system sends it, by TCP/IP protocol, from the CC to the PC.

The PC receives the velocity profile as input and executes all the required operations for the low level control of the DC motors.

IV. EXPERIMENTAL TRIALS

The experimental trials aimed at testing and assessing the execution of saccadic movements by the robotic head. The

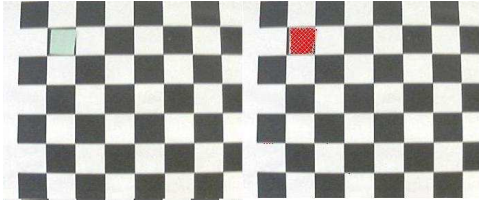


Fig. 6. Stimulus Detected

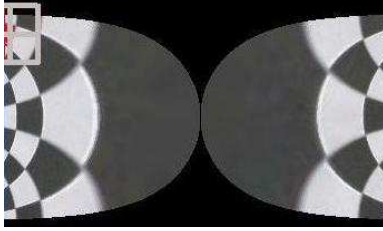


Fig. 7. Collicular mapping of the detected stimulus

saccadic movements were executed towards visual stimuli that were positioned in different points approximately spanning all the visual field. To this aim a grid was positioned in the visual field of the robot and a green coloured marker was moved onto the grid to simulate the visual stimuli. In Fig. 6 we report the visual field of the robot including the green marker and the same image after the detection of the stimuli (the detected area is coloured in red).

The system continuously maps the retinic image onto the colliculus plane and finds the stimulus coordinates in the colliculus reference system as previously described

For each position of the stimulus, its coordinates in the collicular map are then sent to the saccadic model in order to start the saccadic movement. The model generates the velocity profile that is fed to the CC for executing the saccadic movement of the robotic head. To assess the performance of the whole system in executing the saccadic movements we estimated the difference between the position of the stimulus and the center of the fovea of the robot eye after the saccadic movement. To this aim we calculated the center of mass of the object after the movement AND therefore the error was represented by the residual angle to be made by the camera for focusing exactly on the center of the object. 21 different positions of the robot's visual field were tested (corresponding to the black squares of the chessboard) and for each position we performed 10 consecutive saccadic movements. For each position on the chessboard the global error of the system was calculated as in (2):

$$\epsilon_i = \frac{\sum \sqrt{\Delta H^2 + \Delta V^2}}{N} \quad (2)$$

Where ϵ_i is the error estimated on the i -th position of the chess board, H and V are respectively the horizontal and vertical angle required to achieve the centre of mass of the stimulus and N is the number of saccades towards the i -th position. Furthermore we defined a global error calculated on

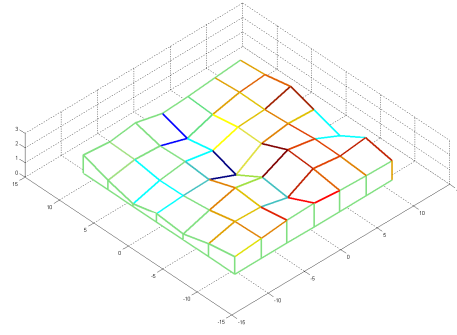


Fig. 8. Graphical representation of the error map in executing 10 saccades for each of the 21 targets

the whole visual field and that was obtained by averaging the errors estimated over the 21 different positions

$$E = \frac{\sum_{i=1}^M \epsilon_i}{M} = 1.57 \text{ degree} \quad (3)$$

V. RESULTS

In Table II we report the average error calculated over the 21 different positions of the visual field together with the standard deviations across the different trials.

TABLE II
DATA FROM EXPERIMENTAL TRIALS

| Number of Test | Point H [deg] | Point V [deg] | Average H [deg] | Average V [deg] | ϵ_i |
|----------------|---------------|---------------|-----------------|-----------------|--------------|
| 1 | -3,40 | 11,03 | -0,21 | -0,93 | 0,96 |
| 2 | 3,40 | 11,03 | 1,46 | -1,47 | 2,07 |
| 3 | -6,27 | 6,29 | -1,98 | 2,51 | 3,19 |
| 4 | 6,19 | 6,29 | 1,58 | 1,81 | 2,40 |
| 5 | -9,87 | 2,67 | -2,01 | 1,87 | 2,75 |
| 6 | -2,740 | 2,84 | -1,85 | 1,12 | 2,16 |
| 7 | 2,74 | 2,84 | 0,96 | 0,93 | 1,34 |
| 8 | 9,43 | 2,63 | 0,76 | 1,01 | 1,26 |
| 9 | -6,54 | 0 | -1,08 | 0,12 | 1,08 |
| 10 | 0 | 0 | 0,26 | 0,14 | 0,29 |
| 12 | 6,54 | 0 | 0,47 | 0,13 | 0,49 |
| 13 | -9,87 | -2,67 | -1,70 | -1,66 | 2,37 |
| 14 | -3,15 | -2,78 | -0,99 | -1,40 | 1,72 |
| 15 | 2,74 | -2,84 | 0,79 | -0,95 | 1,24 |
| 16 | 9,54 | -2,64 | 1,04 | -1,23 | 1,61 |
| 17 | -6,73 | -6,34 | -0,83 | -1,40 | 1,63 |
| 18 | -2,40 | -8,31 | 2,01 | 1,29 | 2,39 |
| 19 | 6,27 | -6,29 | 0,99 | -0,81 | 1,29 |
| 20 | -3,40 | -11,03 | -0,69 | -0,019 | 0,69 |
| 21 | 3,40 | -11,03 | 0,45 | 0,19 | 0,49 |

A graphic representation of the error map is provided in Fig. 8. In the error map the X,Y plane represents the discrete representation of the visual field whereas the Z axis is the estimated average error in each location. Finally, the global error that was previously introduced was estimated in 1.57 degree.

VI. DISCUSSION AND CONCLUSIONS

This work presented a robotic implementation of a neuro-physiological model of saccade generation in humans, with

the final goal of validation and tuning of the model itself. The experimental trials in which the robotic head performed 210 saccades to 21 different points showed that the error is low enough (1.57° in average) to consider this robotic implementation as adequate to reproduce the performance of the model. The results thus encourage the accomplishment of focused experimental trials in which the robot behaviour can help investigate the accuracy of the model, and to possibly tune it, with respect to human behaviour in saccade execution. Further work will include the implementation of the mechanisms for target selection, by using a model of the basal ganglia [14], which are, know to have this function.

ACKNOWLEDGMENT

This work is partly supported by the EU within the NEUROBOTICS Integrated Project (The fusion of NEUROscience and roBOTICS, FP6-IST-FET-2003-no. 001917).

APPENDIX: MODEL

The leaky-integrator rate neuron model used as building brick is as follows (τ : time constant in ms , I : input in mV) :

$$\tau \frac{da}{dt} = I - a \text{ and } y = [a]^+$$

where the transfer function $[]^+$ satisfies $[I]^+ = 0$ if $I < 0$ and $[I]^+ = I$ otherwise.

The input of the visual layer Vis is:

$$I_{Vis}(X, Y, t) = y_{Ret}(X, Y, t - t_0)$$

Long-Lead burst neurons (LLB) in charge of triggering saccades by inhibiting the OPN when the activity on Vis reaches the ϵ_{trig} threshold are modeled by:

$$I_{LLB} = w_{Vis}^{LLB} \sum_{X,Y} (y_{Vis}(X, Y)) - \epsilon_{trig}$$

$$I_{OPN} = -y_{LLB} + \epsilon_{OPN}$$

The activity in the motor layer Mot is gated by the OPNs and the Integrating-Saturating mechanism (note that saturation neurons have a longer time constant τ_{Sat}):

$$I_{Mot}(X, Y) = y_{Vis}(X, Y) - w_{OPN}^{Mot} y_{OPN} - w_{Sat}^{Mot} y_{Sat}$$

$$I_{Int}(X, Y) = w_{Mot}^{Int} \sum_{X,Y} y_{Mot}(X, Y)$$

$$I_{Sat}(X, Y) = y_{Int}(X, Y) - \epsilon_{stop}$$

The four SBG circuits (leftward, rightward, upward, downward) are identical, all of them are gated by OPN activity, and those operating in opposite directions are coordinated by the IBN crossed projections. The EBN and IBN activity is identical and defined by:

$$I_{BN}^D = \sum_{X,Y} (w_{Mot}^{BND} y_{Mot}(X, Y)) - w_{OPN}^{BN} y_{OPN}$$

| τ | 5ms | τ_{Sat} | 100ms | t_0 | 70ms |
|------------------|-------|-------------------|-------|-------------------|------|
| ϵ_{OPN} | 100 | ϵ_{trig} | 400 | ϵ_{stop} | 200 |
| w_{Vis}^{LLB} | 0.005 | w_{OPN}^{Mot} | 40 | w_{OPN}^{BN} | 40 |
| w_{Mot}^{Int} | 0.002 | w_{Sat}^{Mot} | 8 | w_{BN}^{TN} | 0.05 |
| w_{BN}^{MN} | 1.52 | w_{MN}^α | 4.07 | w_{MN}^β | 4.07 |

with $D \in \{U, D, L, R\}$, w_{Mot}^{BND} are obtained by inverting equation 1.

The tonic neurons are the only neurons modeled as perfect rather than leaky-integrators:

$$I_{TN}^D = w_{BN}^{TN} (y_{EBN}^D - y_{IBN}^{Dop})$$

$$I_{MN}^D = w_{BN}^{MN} (y_{EBN}^D - y_{IBN}^{Dop}) + y_{TN}$$

where D^{op} is the opposite direction of D .

The eye plant model used is modeled as a second order differential equation:

$$\begin{aligned} 0.003\ddot{\alpha} + 0.6\dot{\alpha} + 4\alpha &= w_{MN}^\alpha (y_{MN}^R - y_{MN}^L) \\ 0.003\ddot{\beta} + 0.6\dot{\beta} + 4\beta &= w_{MN}^\beta (y_{MN}^U - y_{MN}^D) \end{aligned}$$

The parameters are summed up in the following table

REFERENCES

- [1] S. K. Kaneko, F. Kanehiro, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi., "Humanoid robot hrp-2," in *Proceedings of the 2004 IEEE International Conference on Robotics & Automation*, 2004.
- [2] S. K. J.J. Kuffner, K. Nishiwaki, M. Inaba, and H. Inoue., "Motion planning for humanoid robots," in *Robotics Research: The Eleventh International Symposium*, vol. 15. Springer-Verlag, Aug. 2005, p. 365.
- [3] C. Breazeal, "Designing sociable robots," *The MIT Press*, 2002.
- [4] I. K. Atsuo Takanishi, Hideaki Takanobu and T. Umetsu, "Development of the anthropomorphic head-eye robot we-3rii with an autonomous facial expression mechanism," pp. 3255–3260, 1999.
- [5] M. Kawato, "Robotics as a tool for neuroscience: cerebellar internal models for robotics and cognition," in *Robotics Research, The Ninth International Symposium*, K. D. Hollerbach MH, Ed., Springer-Verlag, London, 2000, pp. 321–328.
- [6] E. G. P. Dario and C. Laschi, "Humanoids and personal robots: Design and experiments," in *Journal of Robotics Systems*, Dec. 2001, pp. 673–690.
- [7] M. U. James J. Clark Centre for Intelligent Machines, "Spatial attention and saccadic camera motion," pp. 3247–3252, 1998.
- [8] M. M. E. L. Rafael Barea, Luciano Boquete and L. Bergasa., "E.o.g. guidance of a wheelchair using neural networks," *IEEE*, pp. 668–671, 2000.
- [9] D. Robinson, "Eye movements evoked by collicular stimulation in the alert monkey," *Vision Research*, vol. 12, pp. 1795–1808, 1972.
- [10] F. Ottes, van Gisbergen J.A., and J. Eggermont, "Visuomotor fields of the superior colliculus: a quantitative model," *Vision Research*, vol. 26, no. 6, pp. 857–873, 1986.
- [11] B. Girard and A. Berthoz, "From brainstem to cortex: computational models of the saccade generation circuitry," *Progress in Neurobiology*, vol. 77, no. 4, pp. 215–255, 2005.
- [12] J. Groh, "Converting neural signals from place codes to rate codes," *Biol. Cybern.*, vol. 85, no. 3, pp. 159–165, 2001.
- [13] S. S. H. D. Stark L., Kong R. and B. B., "Saccadic suppression of image displacement," *Vision Research*, vol. 16, pp. 1185–1187, 1976.
- [14] B. Girard, N. Tabareau, J.-J. Slotine, and A. Berthoz, "Contracting model of the basal ganglia," in *Modelling Natural Action Selection: Proceedings of an international workshop*, J. Bryson, T. Prescott, and A. Seth, Eds. Brighton, UK: AISB Press, 2005, pp. 69–76.