On The Early Research in Visuo-Motor Coordination based on Neural Networks

Tarcisio Passos Ribeiro de Campos

Universidade Federal de Minas Gerais
Departamento de Engenharia Nuclear

Av. Contorno, 842, 9.andar,
Belo Horizonte, MG, 30110-060, Brasil

March 30, 1994

Abstract

This paper presents various models that have approched the visuo-motor coordination in human beings. These selected models are based on neural networks coupled with intuitive ideas which try to explain how human beings solve the arm coordenation problem. Basically, the goals are the simulation of the circular reaction found in infants, the path planning or the solution of the inverse kinematics into a robot environment. A model proposed by the author is discussed and qualitativy analyses are made.

1 Introduction

Research in visuo-motor coordination is a large multidisciplinary field, embracing robotics, artificial intelligence, neuroscience, psychology, and physiology. Technological fields, such as robotics and artificial intelligence, expect to synthesize some aspects of human motor ability in order to assemble automaton machines [1, 2]. Juxtaposing the technological goals, psychology, physiology and neuroscience search for major understanding of brain functions, and consequently of human behavior [3].

Moreover, both the technological and scientifical research views have addressed distinct aspects of the visuo-motor coordination of an arm. Robotics, as a technological field, provides more exact mathematical formulations for the geometrical and time-dependent properties of motion, based on non-trivial algorithm techniques. Such formulations are essential to promote general and reliable motion in robot arms. In contrast, a scientific view, based on psychology, physiology and neuroscience, provides a description of the unconscious computation concerning the brain functions [4]. These descriptions take into consideration perception, adaptation and motivation, well-known characteristics of biological beings.

Indeed, in the study of visuo-motor coordination of an arm, no single mathematical formulation or description explains all mechanisms which generates arm movements. In this context, the research topic must be addressed by technological and scientific views. A technological engineering approach of representing robot motion, or a scientific description of the motor cortex and correlated neural areas, provides distinct representation and comprehension for arm visuo-motor coordination. Both views are scientifically acceptable, and must be carefully addressed in order to obtain a complete understanding of the processes addressed by the visuo-motor coordination of the arm.

At first, how learning can be artificially approached will be discussed. For the sake of understanding, a self-organizing neural mapping will be presented in which a competitive learning technique in an unsupervised fashion is employed.

After, methods that have recently been used for the visuo-motor control of robot arms will be examined. Such methods have integrated neural network models with intuitive ideas in order to simulate arm movements. These methods address the act of precisely positioning a non-redundant arm in its workspace, or the path-planning among obstacles, or the arm inverse kinematic.

The early research in visuo-motor coordination is presented based on goals, achievements, restrictions, basic ideas, applications of the methods. In the final part, the model proposed by the author [5, 6] is discussed, and its advantages are presented. The author stress that it not objective of the paper to present all models which have approach visuo-motor coordination, but only make qualitative analyses about some of them.

2 Learning: an artificial approach

The behavior of animals is not only determined genetically but also by a learning process which adapts the animal to its environment. Two styles of information processing which performs learning can be identified: (*i*) a logical processing through symbols, and (*ii*) an adaptive processing through a distributed parallel processing in a neural network fashion. In the second type of information processing, learning is carried out by modifying the synaptic weights in the neural circuits through a continuous change of the connectivity of the cells. We call such a process *competitive learning*.

Physiological interactions among excitatory and inhibitory neurons in a realistic neural circuit of the hippercampal field CA1 has been shown to operate a simple competitive mechanism, selecting the cell which receives the strongest activation to respond to the stimulus [7]. This mechanism is coined *winner-takes-all*. Artificial approaches which are biologically-inspired have been proposed to produce *competitive learning*. An important representative is the vector quantization algorithm [8, 9].

Artificial competitive learning implies that units in a matrix lattice compete in the process of activation in accordance with their relational distance measured over the lattice. The lattice is sheet-like, carrying a finite number of nodes. The units accommodate to a specific pattern of activation through an unsupervised learning process.

Competitive learning is characterized by only one group of units at a time being recalled in response to a current stimulus. This group has the tendency to become ordered in the location of the stimulus as if a meaningful, dynamic, local coordinate system is created in the place and time that the stimulus is applied. In an unsupervised fashion, disjointed clusters are formed and only one neuron is active in each cluster at a time. This characteristic is similar to the formation of a topological map governed by a local coordinate system [10].

To better understand this process, we consider the self-organizing feature map which makes use of the competitive learning in an unsupervised manner. This neural network will be described in greater detail in many other sections of this thesis. Other types of neural networks, such as backpropagation, can be used to learn the visual and motor maps which will be described here. However, the self-organizing neural network has been adopted due to the possibility of creating an abstract state-space associated with the vector units of the neural network. Such representation, useful as a mathematical tool for discribing the topology involved in the arm visuo-motor coordination, is not achieved by using feedforward or backpropagation neural networks.

The self-organizing feature map employs a finite set of $z \times z$ units of a two-dimensional lattice Γ, each unit being associated with a vector $\vec{w}_n$. The location of any unit in the lattice is specified by the index 'n' and a position vector $\vec{z}_n \equiv (z_x, z_y)$. At first the components of the vectors $\vec{w}_n$ are initialized randomly.

Let us consider random signals given by a k-dimensional vector $\vec{u}$ with the same dimen-

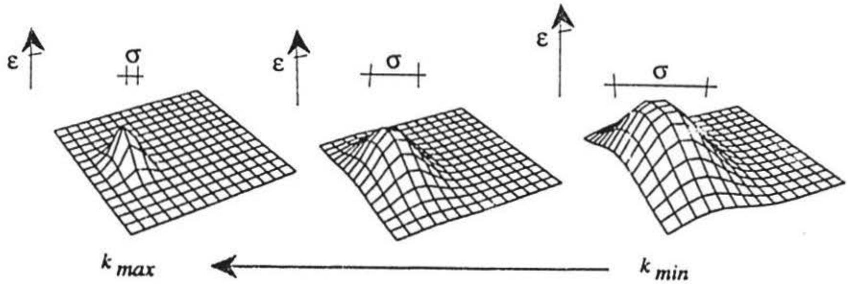

Figure 1: Learning-step-size parameters ϵ and σ. Both parameters alter in accordance with the adaptation step k. The parameter ϵ modulates the amplitude of the function $h_{n,c}$, and σ modulates the size of the neighborhood region defined by $h_{n,c}$.

angles to their respective visual end-effector position. In the first approach [11], given a target position in the workspace, an arm posture whose end-effector position is closed to the required target position is estimated by a "*winner-takes-all*" neural unit onto a three-dimensional neural network grid. The realization of this posture is denoted as gross movement. Indeed, the grid fills the arm workspace in which each node of the grid is associated with a matrix whose elements represent the partial derivatives of the displacement of joint angles in relation to visual displacement of the end-effector acquired by a pair of cameras (the local Jacobian matrix in that position). A fine movement following the gross movement is performed by using a specific linear correction involving the Jacobian matrix. The main difference between this approach and the other methods [21, 22, 16] is the use of the three-dimensional map over the workspace and the use of the local correction provided by the Jacobian. A variant of the previous approach has been successfully implemented in a real arm (PUMA 560) [20]. The positioning error achieved by one unit after the iterative learning is around $1\,mm$. Despite of this small error for positioning the arm end-effector, this approach is not able to reproduce all possible postures of a redundant robot arm since the grid response for a target position is a neural unit vector describing a set of joint angles which reproduce only a single posture for the desired position.

Neurobot is the designation of an approach which learns how to reproduce movements in a non-redundant arm based on a Kohonen neural network [8, 9] which also maps visual and motor information. This system simulates movements in a robot arm with two degree of freedom. The method develops a cognitive map which connects arm end-effector position and its respective joint angles set. Since the relationship between visual information (end-effector position) and arm posture (joint angle set) is one-to-one for the two degrees-of-freedom arm, a computational link between these two sets of information is established. Path planning is achieved by spreading activation passing signals among the neighboring units. For obstacle avoidance, Saxon and Mukerjee [15] suggest that when the obstacle is placed directly in the workspace, the neural units, whose receptive field lie in the region enclosed by the obstacle, are recruited onto the lattice responsible for tesselating the sionality of the weight vectors $\vec{w}_n \in \mathbb{R}^k$. The *winner* unit $\vec{w}_c$ is identified through

$$c : \| \vec{u} - \vec{w}_c \| \leq \| \vec{u} - \vec{w}_n \| \quad \forall n \in \Gamma \quad , \tag{1}$$

with reference to an arbitrary metric. Usually an Euclidian metric is adopted. Thus the winner unit $\vec{w}_c$ overtakes the representation of the input vector $\vec{u}$ in the lattice.

A symmetric neighborhood surrounding the winner unit c is defined by a function $h_{n,c}$. This function selects a group of units in which an adaptation of its component takes place.

Updating is thus applied to the units in the topological neighborhood defined by $h_{n,c}$ in accordance with

$$\vec{w}_n^{new} = \vec{w}_n^{old} + \epsilon h_{n,c}(\vec{w}_n^{old} - \vec{u}) \quad . \tag{2}$$

The best representation for the function $h_{n,c}$ is defined to be a Gaussian

$$h_{n,c} = \exp(-\| \vec{z}_n - \vec{z}_c \|^2 / 2\sigma^2) \quad , \tag{3}$$

where $\| \vec{z}_n - \vec{z}_c \|$ indicates a norm distance in the two-dimensional neural lattice between the elements 'n' and 'c'.

The parameter ϵ is the learning step size, and σ is the size of the neighborhood region. Both variables are functions of the adaptation step 'k' [11, 8], in conformity with

$$\sigma = \sigma_o (\frac{\sigma_f}{\sigma_o})^{\frac{e}{e_{last}}} \quad , \tag{4}$$

$$\epsilon = \epsilon_o (\frac{\epsilon_f}{\epsilon_o})^{\frac{e}{e_{last}}} \quad . \tag{5}$$

where $\sigma_o, \sigma_f, \epsilon_o$ and ϵ_f are parameters to be defined during a simulation, in order that the neural network may achieve convergence. Figure 1 illustrates how the ϵ and σ parameters are altered with the adaptation step k from values between k_{min} to k_{max}. Initially σ and ϵ are large for rapid but coarse adaptation but decrease during the learning time, achieving a fine tuning for k_{max}.

3 Adaptive methods for visuo-motor coordination

Recently, neural network models have been recognized as useful tools in creating theoretical approaches for reproducing arm movements. Ritter and Schulten [11, 12], Ritter *et. al* [13], Martinetz and Schulten [14], Saxon and Mukerjee [15], Graff and Lalond [16, 17], Kupperstein [18] and Groosberg [19], among others, have approached an artificial visuo-motor system integrating neural network and other computational techniques in order to conduct movements in an artificial arm.

Ritter *et al* [11, 12, 13] and Martinetz *et al* [14, 20] propose a system to simulate the visuo-motor coordination of an arm. This system learns arm postures relating their joint workspace. Those units immediately cause inhibition of the neural units onto the neural lattice which carries information about the joint angles. The *Neurobot* follows the same scheme proposed by Martinetz *et. al* [14] since it uses a Kohonen-type algorithm, and by Graf and LaLond [17] since it uses a similar inhibitory strategy linking the units of various lattices which represent the workspace and the configuration space of the arm. However, the literature does not clarify how the inhibition is really introduced into the neural network.

The geometric shape of an obstacle in the configuration space has geometrical deformations in comparison with its real shape in the workspace [1, 23]. In Saxon's method, the obstacle region in configuration space is not processed analytically; therefore, it must be pre-developed through links between the joint angles (arm configuration space) and the visual position (workspace) by trial and error during arm movements. As a consequence, the number of links to be processed among points in these two spaces is astonishingly large. Thus, the method is only useful in the case of non-redundant arms.

Neuroplan was developed by Graf and Lalonde [16, 17]. It is a well-designed approach for avoiding collisions using a non-redundant arm manipulator. It learns the topology of three maps: the obstacle, the arm configuration and the visual configuration. The inter-map connections, like Saxon's method, helps in generating the allowable postures onto the abstract space of the arm configurations. The implementation was performed for a two degrees-of-freedom and revolute joint arm robot. It is mentioned that the approach can be generalized for redundant robots. Graf [16] suggests that the system can be made arbitrarily precise by increasing the number of processors and connections. However, as in the case of *Neurobot*, by simple calculations it is verified that the number of inter-connection is prohibitive for simulating redundant arm movement in a space where obstacles are present. *Neuroplan* and *Neurobot* are historically important since they are the first approaches which have addressed distinct neural maps for the visual space, the configuration space, and the workspace.

Massoni and Bizzi [21] deal with the problem of generating time-trajectories of a limb from a starting posture moving toward a target position specified by sensory stimulus. A three-layer sequential network trained by a standard back-propagation procedure was used to learn a trajectory pattern given by a set of motor commands. Generalization over the trajectory space is successfully obtained. However, despite the generation of a set of profiles, the reproduction of arm movements for reaching any arbitrary goal position or for following a generic trajectory is not completely achieved. This is due to the programming of the network to respond to specific groups of trajectories defined by a path between an initial and final position as opposed to a generalized movement or path.

Another approach for learning visuo-motor coordination has been proposed by Grossberg and Kuperstein [19, 24, 25, 18]. Kuperstein's method groups information originated from a set of coupled topographical maps. Each map associated with each link of the arm learns a relation between joint angle and its visual position in the workspace. The information embedded on the combination of all these maps reproduces the complete set of joint angles which generates an arm posture.

The last method to be discussed has been developed by Campos [5]. The model adopts a self-organizing adaptive map algorithm to learn all possible postures for an artificial arm of arbitrary design placed in a three-dimensional workspace. Arm postures are represented through their projections onto a set of image planes. Based on the orientation and length of the links extracted from these images, a topological state space Q is generated. Arm kinematics is expressed as a transformation of topological hypersurfaces, the intersections of which represents the multiple postures of the arm in the workspace for a given end-effector position.

Implementation of the previous model for visuo-motor coordination of the SoftArm robot (Rubbertuator) had been developed [6]. The algorithm supported the learning of all possible postures of the redundant arm, reproducing multiple postures for any arbitrary end-effector position of the arm in the three-dimensional workspace. The computational algorithm accomplished three different artificial neural network components, responsible for visual and motor maps. These components: the motor components and the visual component were interconnected, forming a synergetic neural structure where the arm postures were memorized.

Gripper positioning presents a task of placing the arm gripper in the arm workspace based on the visual camera information stored into each neuron of the neural component which tessellates the visual space. After the learning phase, for each spatial position memorized in the visual neural component, many motor stimuli are possible to be generated by the motor components. Based on these stimuli, all memorized multiple arm postures may be synergetically reproduced by the motor components. Those arm postures reproduce various orientation for the gripper. As example, figure 2 shows three different orientation of the gripper, provided by the algorithm, for reaching a specific position in the arm workspace.

4 Conclusion

The major features of the methods presented by Graf and Lalond, Ritter, Saxon, Martinez, Massoni and Bizzi, Grossberg are: the vector units in the neural network must completely fill the workspace of the arm, the configuration space, and the visual end-effector position space (visuo-space) in order to

On the other hand, in the method proposed by the author, the mastery of learning multiple arm postures based on invariances in the configuration space is essential for performing the task of gripper positioning. During implementation of the method in an artificial arm, interconnection among neural structures were adopted, structuring a large adaptive data base. The neural structure and interconnection had obeyed a predefined design. In summary, components of the neural structure map the geometrical arm characteristics as a function of the end-effector position onto a set of two dimensional neural lattices, processing a loop between visual (light position in the retinas) and motor (pressure differential) stimuli.

The importance of the author's method has not been only to produce coordinated movements in a real robot, comparing accuracy in the end-effector positioning, since many other approaches discussed here have already been tried. Rather, the main goal has been to introduce a new philosophy in which a set of synergetic neural networks, each one mapping a particular subprocessor, is designed conjointly to promote general and reliable robot arm movements. Other important caracteristic is its learning based on the invariances of the configuration space, which is represented by the interception of hypersurfaces. Thus, the among of information memorized by the self-organizing neural network is reduced to a minimal.

In summary, despite of the success of the author's method in relation to the others presented here, the dynamics of the arm has been not addressed. As a conclusion, the visuomotor control of a redundant arm manipulator continues to be a challenging research topic since a complete method which can be used in order to guide redundant arm manipulators in a realistic space integrating kinematics, dynamics, control, vision, path planning has not been developed yet.

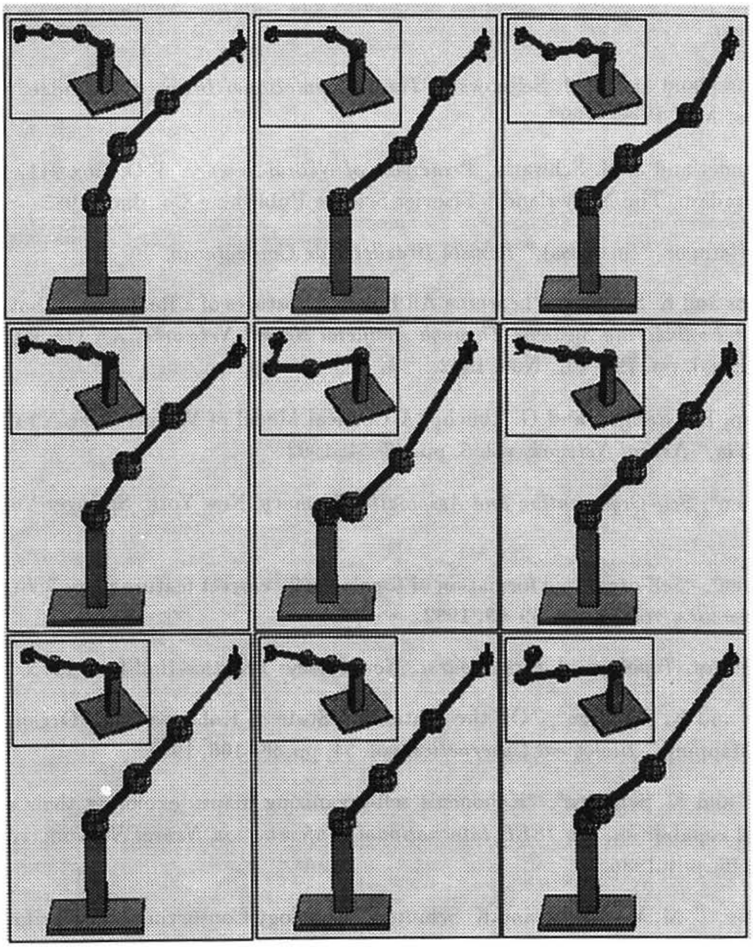

Figure 2: The skill of generating multiple arm postures provide the arm condition to reproduce different gripper's orientation for reaching a same position in the visual space (obtained by simulation).

References

[1] J. C. Latombe, *Robot Motion Planning.* Kluwer Academic Publishers, 1991.

[2] J. J. Craig[a], *Introduction to Robotics Mechanics and Control.* Addison-Wesley Publishing, 1986.

[3] P. S. Churchland and T. J. Seijnowshi, *The Computational Brain.* Cambridge, Massachusetts: MIT Press, 1992.

[4] E. R. Kandel and J. H. Schwartz, *Principles of Neural Science.* P.O. Box 211, 1000 AE Amsterdam, The Netherlands: Elsevier Science Publishing Co., Inc., 1985.

[5] T. P. R. Campos, "(in press)," *Revista Brasileira de Computacao.*

[6] T. Campos and K. Schulten, "Learning All Possible Postures of a Real Arm Robot," in *Intelligent Engineering Systems Through Artificial Neural Networks, ANNIE 92,* (St. Louis, Illinois), pp. 793–801, Nov. 1992.

[7] R. C. Trip, R. Granger, and G. Lynch, "A Cortical Model of Winner-Take-All Competition Via," *Neural Network,* vol. 5, pp. 47–54, 1992.

[8] T. Kohonen[a], *Self-Organization and Associative Memory.* New York: Springer-Verlag, 1984.

[9] T. Kohonen[b], "Self-organized formation of topologically correct feature maps," *Biological Cybernetics,* vol. 43, pp. 59–69, 1982.

[10] J. R. Munkres, *Topology: A First Course.* New Jersey: Prentice-Hall Inc., 1975.

[11] H. Ritter and K. Schulten[a], "On the Stationary State of Kohonen's Self-Organizing Sensory Mapping," *Biological Cybernetics,* vol. 54, pp. 99–106, 1986.

[12] H. Ritter and K. Schulten[b], "Kohonen's self-organizing maps: exploring their computational capabilities," *in IEEE International conference on Neural Network,* vol. 2, pp. 109–116, july 1986.

[13] J. H. Ritter, T. M. Martinetz, and K. Schulten, "Topology Conserving Maps for Learning Visuo-Motor Coordination," *Neural Networks,* vol. 2, pp. 159–168, 1988.

[14] T. Martinetz, H. J. Ritter, and K. Schulten, "Three-Dimensional Neural Net for Learning Visuomotor Coordination of a Robot Arm," *IEEE Transactions on Neural Networks,* vol. 1, pp. 131–136, march 1990.

[15] B. J. Saxon and A. Mukerjee, "Learning The Motion Map of a Robot Arm with Neural Networks," in *IJNN,* (San Diego), pp. 777–782, 1990.

[16] D. H. Graf, *A Neural Controller for Collision-Free Movement of Robot Manipulatiors.* Master's thesis, School of Computer Science, Carleton University, Ottawa, Ontario, August 1988.

[17] D. H. Graf and W. R. LaLonde, "A Neural Controller For Collision-Free Movement of General Robot Manipulators," in *Proc. IEEE First Intl. Conf. on Neural Networks*, (San Diego, Cal.), pp. 77–84, July 1988.

[18] M. Kuperstein and J. Rubinstein, "Implementation of an Adaptive Neural Controller for Sensory-Motor Coordination," *IEEE Control Systems Magazine*, pp. 25–30, april 1989.

[19] S. Grossberg and M. Kuperstein, *Neural Dynamics of Adaptive Sensory-Motor Control.* : North-Holland, 1986.

[20] J. A. Walter, T. Martinetz, H. J. Ritter, and K. Schulten, "Industrial Robot Learns Visuo-motor Coordination by Means of Neural-Gas Network," Technical Report TB91-07, Beckman Institute, UIUC, 1991.

[21] L. Massone and E. Bizzi, "A Neural Network Model for Limb Trajectory Formation," *Biological Cybernetics*, vol. 61, pp. 417–425, 1989.

[22] I. J. Michael, "Supervised learning and systems with excess degrees of freedom," *COINS Technical Report 88-27*, pp. 1–40, may 1988.

[23] V. K. Hawang, "Boundary Equations of Configuration Obstacles for Manipulators," *Proceed. of the IEEE Inter. Conf. on Robotics and Automation*, pp. 298–303, 1990.

[24] M. Kuperstein[a], "Adaptive Visual-Motor Coordination in Multijoint Robots Using Parallel Architecture," *in Proc. IEEE Int Robotics and Automation (Raleigh, NC)*, pp. 1595–1602, march 1987.

[25] M. Kuperstein[b], "Neural model of adaptive hand-eye coordination for single postures," *Science*, vol. 239, pp. 1308–1311, march 1988.

[26] J. Hyvarinen and A. Poranen, "Function of the parietal associative area 7 as revealed from cellular discharges in aler monkeys," *Brain*, vol. 97, pp. 673–692, 1974.

[27] W. B. M. *et al.*, "Posterior Parietal Association Cortex of the Monkey: command Functions for Operations within Extrapersonal Space," *Journal of Neurophysiology*, vol. 38, pp. 871–908, 1975.