ESANN'2005 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 27-29 April 2005, d-side publi., ISBN 2-930307-05-6.

Using generic neural networks in the control and prediction of grasp postures

F. Carenzi [1/2], P. Gorce [1], Y. Burnod [2], M.A. Maier [2]

1 - LESP EA 3162, University of Toulon-Var,
Avenue de l'université, La Garde, France.
2 - INSERM U.483, University Pierre et Marie Curie,
9 quai S[t] Bernard, Paris, France.

Abstract. We have developed a neural network model that learns the kinematics of object-dependent reach and grasp tasks with a simulated anthropomorphic arm/hand system. The network learns to combine multi-modal arm-related information as well as object-related information such as object size, location and orientation. We will first describe the learning of the finger inverse kinematics and then the learning of the grasp configuration. Finally, to illustrate the performance of the network, we will simulate object-dependent grasp configurations by the use of a 5-digit hand (4 DoF per digit) linked to a 7 DoF arm.

1 Introduction

Humans show an extraordinary capability for manipulating and using objects. For successful grasp of an object, the hand has to be shaped appropriately to match the form, size and orientation of the object. It has been proposed that specific parieto-frontal circuits provide a sensorimotor transformation of the object's visual properties into a set of motor commands that adapts the shape of the hand to the object to be grasped [1] [2]. Reach and grasp movements are learned behaviors and they change their kinematic characteristics during development. It was shown that the number of action units in reach decreases with age, that their sequencing becomes more systematic and that reach trajectories get straighter [3]. Similarly, precision grip skill increases during normal development up to about 10 years of age when control of grasp forces becomes increasingly anticipatory [4] and when visual information on object characteristics is fully incorporated [5]. These data suggest that increasingly complex hand-object interaction progresses through successive stages of development. The hand is a very complex system as indicated by its high number of degrees of freedom and biomechanical constraints [6]. In recent years, neural network based techniques have been used to learn the mapping from objects properties to grasp configuration [7] [8] [9]. This technique was successfully used in the prediction of anthropomorphic hand postures for the grasp of various objects [10]. Some of these neural networks were inspired from neuroscience, either in terms of architecture or in terms of learning rules [11].

We first present, in part two, the kinematics of the arm-hand model, then in part three, the neural network and the learning of the finger inverse kinematics as well as the learning strategies. Finally, we illustrate the learning model and present the simulation results of hand postures.

2 Kinematics of the arm-hand model

2.1 Arm model

The 7 degrees of freedom (DoF) arm model is composed of two segments (arm and forearm) linked by three joints as shown in Figure 1. The shoulder joint has 3 DoF, the elbow joint has 1 DoF and the wrist joint has 3 DoF.

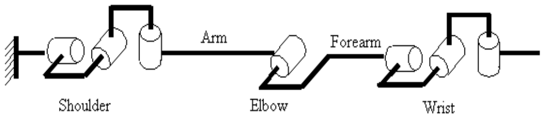

Figure 1. Seven degrees of freedom arm model

2.2 Hand model

The hand model is composed of 5 articulated rigid chains representing the fingers, which are connected to a common body representing the palm. Each finger has 4 DoF. For the index, middle, ring and little finger, the joint coordinates represent the following DoF: φ: abduction-adduction at the proximal joint, θ_1, θ_2 et θ_3: flexion-extension at the proximal, middle and distal joints. For the thumb, the DoF are φ: thumb rotation at the proximal joint, θ_1: abduction-adduction at the proximal joint, θ_2 et θ_3: flexion-extension at the middle and distal joint. The anthropomorphic geometrical parameters of the hand were taken from [12].

3 Neural network architecture and learning strategies

3.1 Neural network architecture

The neural network learning algorithm is based on the Locally Weighted Projection Regression (LWPR), used for incremental learning of nonlinear functions [13] [14]. It uses locally linear models, spanned by a small number of univariate regressions in selected directions in input space, to achieve a piecewise linear function approximation.

Learning object-dependent grasp configurations requires multiple matching units (i.e. generic neural networks). Matching units correlate (match) multi-modal information through Hebbian mechanisms [15]. The proposed architecture uses matching units for the learning of the finger inverse kinematics (one matching per finger) associated with additional matching units dedicated to the learning of grasp configurations.

3.2 Learning of the inverse kinematics

We first determine randomly the 4 DoF of one finger $\Theta = (\varphi, \theta 2, \theta 3, \theta 4)$ by taking into account the joint limits. Then by forward kinematics, we obtain the 3D end effector position V = (X, Y, Z). For the learning of the inverse kinematics, the input vector is called V^T and the output vector Θ^T contains the angular configuration of the digit (Fig. 2). This represents one learning procedure for each digit: Thumb, Index, Middle, Ring and Little finger.

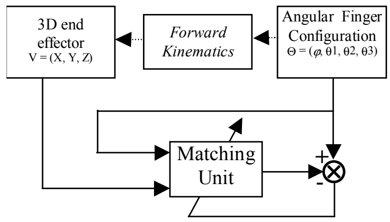

Figure 2. Algorithm for learning the inverse kinematics of a single finger by a single matching unit.

3.3 Learning of the grasp configuration

We have simulated three types of objects: blocks, cylinders and spheres of different size. The learning of grasp configurations is based on the previous learning step that determines the angular configuration of the digits for a given target or contact point.

The learning of the grasp configurations (angular digit configuration) depends on object-specific and heuristically determined contact points on the surface of the three objects. Three novel matching units 'MU Shape', one for each object type, have been introduced to establish the relation between object-related properties (type and size) and angular grasp configuration (Fig. 3). In order to learn the relation between the object-related properties and the grasp configuration of the five fingers, we define an input vector 'Shape' [Dim $= (L, l, h)^T$]. For the block, *L*, *l* and *h* correspond respectively to the length, the width and the height. For the sphere, *h* corresponds to its radius, L and l are not given. For the cylinder, *L* determines its length, *h* its radius, and l is not given. Minimal and maximal size varied between 40 and 80 mm. The output of the model is a vector of size 20, which contains the grasp configuration of the five fingers.

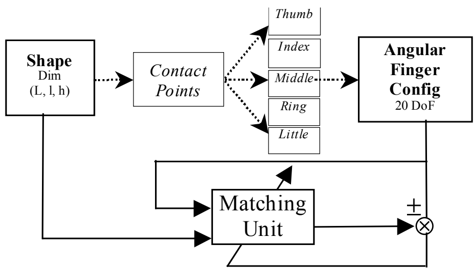

Figure 3. Algorithm for learning grasp configurations.

4 Simulation results

4.1 Finger inverse kinematics results

Here we present the simulation results which concern the learning of the inverse kinematics for each finger. Fig. 4 shows the five learning curves (Mean Squared Error, MSE) corresponding to each of the five matching units. Learning saturates at about 5000 iterations.

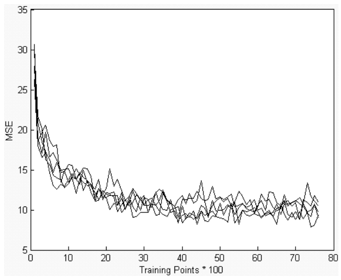

Figure 4. Learning curves.

Finger	Positional error (mm)	SD
Thumb	3.9	1.6
Index finger	2.2	1.0
Middle finger	2.4	1.1
Ring finger	2.3	1.0
Little finger	2.3	1.0

Table 1. Positional error (mm) for each finger

To evaluate the efficiency of the neural network, we used the following procedure. We input to the neural network the desired fingertip position and calculate from the obtained joint angles the actual fingertip position. The positional error corresponds to the norm of the vector from the desired position to the actual position. Table 1 shows the mean positional error in mm and the standard deviation (SD) obtained with a test set of 1000 different configurations. The results are acceptable to have a good approximation of joint angles (small positional error).

4.2 Results on grasp configurations

Figure 5 shows the learning curves for the three matching units 'MU Shape' that relate object properties to the grasp configuration for the block, the cylinder and the sphere. Learning converges quickly to sufficient error levels (at ~2000 iterations).

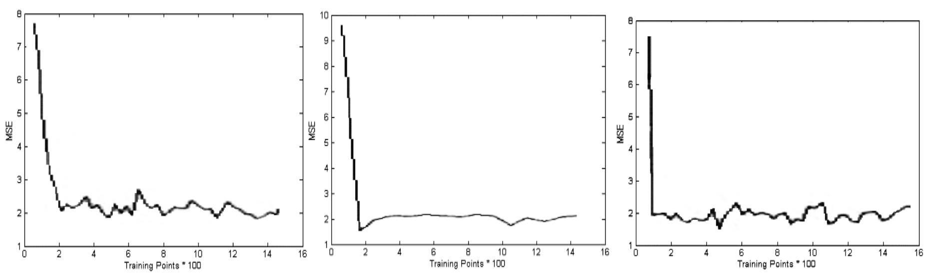

Figure 5. Learning curve for the grasp configuration of the Cube, Cylinder and Sphere.

To illustrate grasp configurations, Fig. 6 gives several examples of an appropriate hand shape for grasping a sphere, a block and a cylinder. The wrist position and orientation relative to the object have been pre-defined.

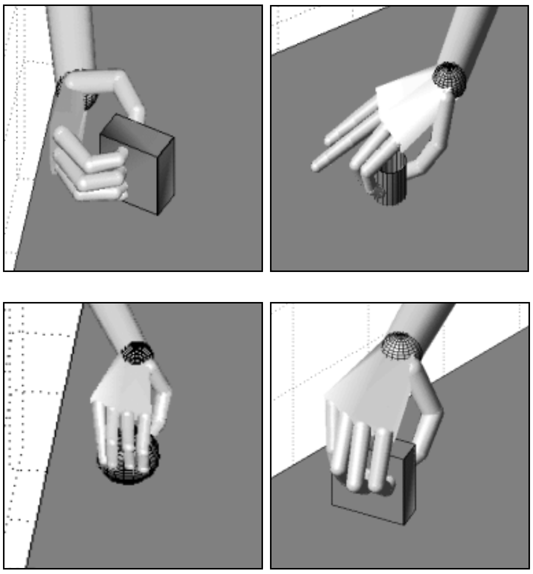

Figure 6. Examples of learned object-dependent grasp configurations.

5 Conclusion

We have shown that a modular neural network architecture composed of generic 'matching units' based on Hebbian learning can learn object-dependent grasp configurations for a simulated 20-DoF kinematics hand system. This is done in two successive learning steps. The network is based on 5 matching units, each of them learns independently the inverse kinematics of one finger. Then in a second step, the acquired inverse kinematics is used in combination with three novel matching units to learn in relatively few training iterations the appropriate grasp configurations for three different objects. This mimics in a highly simplified way the step-wise ontogenetic development of human grasp.

Acknowledgements

This work was in part supported by a European 5th framework program grant: IST-2001-33073.

References

[1] Jeannerod M, Arbib MA, Rizzolatti G and Sakata H (1995) "Grasping objects: the cortical mechanisms of visuomotor transformation". in *Trends Neurosci* 18: 314-20.

[2] Burnod Y, Baraduc P, Battaglia-Mayer A, Guigon E, Koechlin E, Ferraina S, Lacquaniti F, Caminiti R (1999) "Parieto-frontal coding of reaching: an integrated framework." in *Exp Brain Res* 129: 325-46

[3] von Hofsten C (1991) "Structuring of early reaching movements: a longitudinal study." in *J Mot Behav* 23: 280-92.

[4] Forssberg H, Kinoshita H, Eliasson AC, Johansson RS, Westling G, Gordon AM (1992) "Development of human precision grip. II. Anticipatory control of isometric forces targeted for object's weight." in *Exp Brain Res* 90: 393-398.

[5] Gordon AM, Forssberg H, Johansson RS, Eliasson AC, Westling G (1992) "Development of human precision grip. III. Integration of visual size cues during the programming of isometric forces." in *Exp Brain Res* 90: 399-403.

[6] Kapandji I. A. (1977) "The upper limb, logistic support of the hand" in *Ann Chir.*, vol. 31, no. 12, pp. 1021-1030.

[7] Z. Taha, R. Brown, D. Wright (1997) "Modelling and simulation of the hand grasping using neural network" in *Medical Engineering and Physiology,* vol. 19, pp. 536-538.

[8] Y. Uno, N. Fukumura, R. Suzuki, M. Kawato (1995) "A computational model for recognizing objects and planning hand shapes in grasping movements" in *Neural Networks,* vol. 8, pp. 839-851.

[9] T. Iberall (1987) "Grasp planning for human prehension" in *Proceedings of the International Joint Conference on Artificial Intelligence,* Milan, pp. 1153-1156.

[10] N. Rezzoug, P. Gorce (2001) "A neural network architecture to learn hand postures definition" in *IEEE International Conference on System Man and Cybernetics,* pp. 3019-3025.

[11] T. Iberall, A. Fagg, (1994) "Neural network models for selecting hand shapes" in *Hand and Brain: The Neurophysiology and Psychology of Hand Movements,* NY: Academic press.

[12] B. Buchholz, T. J. Armstrong, S. A. Goldstein (1992) "Anthropometric data for describing the kinematics of the human hand" in *Ergonomics,* vol. 35, pp. 261-273.

[13] S. Vijayakumar, S. Schaal (2000) "LWPR : An *O(n)* Algorithm for Incremental Real Time Learning in High Dimensional Space" in Proc. *of Seventeenth International Conference on Machine Learning, ICML2000,* Stanford, pp. 1079-1086.

[14] A. D'Souza, S. Vijayakumar, S. Schaal (2001) "Learning Inverse Kinematics" in *Proc. International Conference on Intelligence in Robotics and Autonomous Systems, IROS 2001*, pp. 298-303.

[15] Carenzi F, Bendahan P, Roschin VY, Frolov AA, Gorce P, Maier MA (2004) "A generic neural network for multi-modal sensorimotor learning", in *Neurocomputing* 58-60: 525-533.