Abstract—The ability to grasp and manipulate objects can be regarded as a distinctive feature of humans in the animal world. The human hand is studied for different purposes, such as the control of anthropomorphic robotic hands. One of the fundamental features of the human hand is the thumb opposition, not always considered as active degree of freedom in the design of robotic hands. The purpose of this study is to investigate the role of thumb opposition during cyclic manipulation tasks through the interaction with different objects and a bio-inspired control architecture based on reinforcement learning. The control architecture has been implemented in simulated environment on two robotic hands with different thumb features, i.e. the iCub hand and the DLR/HIT Hand II, interacting with objects of different sizes and shapes. Furthermore, a systematic analysis of the working areas of the two hands during the manipulation of a sphere located at different positions has been carried out. The achieved results show that thumb opposition, characterizing human hands, plays a key function in fine manipulation and allows increasing the hand working area.

I. INTRODUCTION

The ability of finely manipulating and using objects as tools can be regarded as a distinctive feature of humans in the animal world. The complex structure of the human hand is at the basis of these manipulation capabilities [1].

The human hand is studied in robotics for different purposes, such as design of the new mechanical components and control. In particular, the analysis of the role of the kinematic structure of the human hand, and the role it plays in manipulation capabilities, is fundamental for the realization of multifingered anthropomorphic robotic hands [2]-[6]. The literature shows how human manipulation capabilities are heavily related to a peculiar feature of the human hand: thumb opposition [7] [8]. Thumb opposition particularly attracts attention because of its fundamental role in hand functions [8] [9]. In this respect, clinical and biomechanical studies show that the thumb is responsible for 50% of hand functions [10]. The thumb opposition enables the human hand to carry out prehensile movements and hold objects firmly. It is also important for everyday life movements involving object grasping and manipulation, such as writing, turning a key, etc.

The human thumb has been described by different kinematic models [11]-[13] with different characterization of the thumb joints. One of these models represents the thumb with 5 degrees of freedom: flexion/extension of the IP joint, considered as hinge joint, flexion/extension and adduction/abduction of the Metacarpophalangeal (MCP) joint, and flexion/extension and adduction/abduction of the Carpometacarpal (CMC) joint. Despite the literature on robotic hands presents a wide variety of configurations and kinematic models of the thumb, only a few of them have thumb opposition as an active DOF [4] [14]. For instance, in the Utah/MIT hand the thumb is mounted directly opposed to the other fingers in a fixed position [15], while the UB hand is characterized by an opposable thumb with 4 actuated degrees of mobility [16].

The purpose of this work is to investigate the role of thumb opposition in tasks of cyclic manipulation acquired through exploration and learning, based on the bio-inspired neural architecture proposed in [17]. We used this model to learn controlling two different anthropomorphic robotic hands engaged with manipulation of different objects: the iCub hand [3] and the DLR/HIT Hand II [5]. The main distinction between them concerns the thumb: the iCub hand has thumb opposition as an active DOF whereas the DLR/HIT Hand II has a fixed thumb opposition.

The study was conducted using two 3D simulated robotic hands interacting with 3D simulated objects. The hands were controlled with the bio-inspired neural architecture proposed in [17]. It is based on two key elements: a hierarchical reinforcement learning neural network and Central Pattern Generators (CPGs). The neural network implemented the trial-and-error mechanisms guiding the system learning and the hierarchical model permitted to search parameters of different CPGs. The CPGs allowed producing rhythmic patterns to move the fingers during the manipulation tasks. This hierarchical bio-inspired control was applied to the two simulated robotic hands interacting with 9 different objects of different shapes and dimensions. In particular, the task required the hands to rotate each object around a fixed axis.
as much as possible during a fixed amount of time. The speed execution of the task was not fixed to a specific value; it was established by the controller so as to allow the hand to maximize the rotation of the object.

The task allowed the comparative analysis of the two hands with the same objects and of each hand with different objects. The control architecture was also used to systematically measure the viable working areas of the two hands in the manipulation of a sphere in different locations in space.

The approach implemented in this paper permits to evaluate the role of the thumb in an innovative way: the automatic search of good solutions, given the physical features of the actuators and the task demands, of different hands engaged in cyclic manipulation tasks. This method allows the quantification of two important features: the quality of the learning process and the final performance it produces, and a precise definition of the hand working area. The most important expected results are a better performance and a larger working area in the presence of an active thumb opposition. This information is used to identify which of the two simulated hand best fitted with the execution of the desired tasks, and to pave the way to the definition of “quantitative” criteria for the design of new thumbs of robotic hands.

The rest of the paper is organized as follows. The second section presents the system hierarchical architecture, the simulated hand and objects, the performance tests, and the procedure used to measure the hand working area. The third section presents the results of the tests, in particular the quality of learning for the 9 different objects and a comparative analysis of the effects on the working area produced by the availability of the active thumb DOF. The final section discusses the results and draws the conclusions.

II. METHODS

A. System Architecture

The architecture of the computational bio-inspired model used to investigate the hand working area [17] is shown in Fig.1.

The system is composed of:
1) Three neural “experts”;
2) One neural “selector”;
3) Three CPGs;
4) Proportional Derivative controller (PD);
5) Simulated hand and object.

The reinforcement learning neural network has the aim of calculating the CPGs parameters with a trial-and-error process and selecting the action that maximizes the rotation of the object. The actor-critic version of the reinforcement learning is used because it has a structure and functioning that captures some of the features of basal ganglia, a brain structure implementing trial-and-error in organisms [18]-[21].

The neural “experts” have as input the object size and generate the parameters of the CPGs; the neural “selector” sets the weights to gate CPG output trajectories.

The system components are: actor-critic reinforcement learning experts and selector, CPGs, PD controller, and simulated hand and object.

Three CPGs were implemented with different complexity, in terms of numbers of oscillators and number of searching parameters (Fig. 2):
- CPG-C (complex CPG) having 4 oscillators and requesting 12 parameters;
- CPG-M (medium complexity CPG) having 2 oscillators and requesting 5 parameters;
- CPG-S (simple CPG) having 1 oscillator and requesting 2 parameters.

The desired joint angles were obtained weighting the CPGs output trajectories with the “selector” output. The actor, the critic and the selector receive the same input: the object dimension, as shown in Fig. 2.

The output of the “experts” $y_j$ and the input unit $x_i$ generate activation potential $PA_j$ by means of connection weights $w_{ji}$:

$$PA_j = \sum_i \left(w_{ji} * x_i \right)$$

and an activation function using a sigmoidal function:

$$y_j = 1/(1 + e^{-PA_j})$$

A random noise $N$ is added to the activation to foster exploration:

$$y_j^N = y_j + N$$

1093
The selector generates three outputs, with Eq. (1), used to gate the “experts” output as follows:

$$\mathbf{out} = \mathbf{s}_j (O_{\text{CPG-C}} + s_2 O_{\text{CPG-M}} + s_3 O_{\text{CPG-S}})$$

$$s_j = \frac{y_{sj}}{\sum_i y_{si}}$$

where $\mathbf{out}$ is the vector of the desired joint angles, $O_{\text{CPG-C}}$, $O_{\text{CPG-M}}$, and $O_{\text{CPG-S}}$ are the desired joint angles of the CPGs, $y_{sj}$ is the activation of the selector output unit $j$ (calculated as in Eq. 3), and $s_j$ is the normalized selector output unit $j$.

The critic estimate the evaluation $E$ of the currently perceived state:

$$E = \sum_i (w_i \cdot x_i)$$

The critic update the connection weights $w_j$ on the basis of the outcome expectation, the TD-error ($TD$) [18]:

$$\Delta w_j = \eta E TDx_j$$

where $\eta$ is a learning rate.

The Central Pattern Generator is used to produce rhythmic trajectories for each finger joints and it is modeled as coupled oscillators. The neural rhythm generator can be mathematically represented by the model in [22], properly modified to regulate the center of oscillation:

$$\dot{\theta}_i = 2\pi v_i + \sum_j w_{ij} \sin(\theta_j - \theta_i - \phi_j)$$

$$\dot{r}_i = a_i \left( \frac{a_i}{4} (R_i - r_i) - \dot{r}_i \right)$$

$$\ddot{c}_i = b_i \left( \frac{b_i}{4} (C_i - c_i) - \ddot{c}_i \right)$$

$$z_i = r_i (1 + \cos(\theta_i))$$

where $\theta_i$ is the phase of the CPG oscillator $i$, $r_i$ is the amplitude, $z_i$ is the controlled variable (e.g., a joint angle), $v_i$ is the intrinsic frequency, $R_i$ is the desired amplitude, $a_i$ determines how quickly $r_i$ converges to $R_i$, $\phi_j$ is the desired phase difference between oscillators, $w_{ij}$ is the coupling strength of $i$ with $j$, $C_i$ is the desired center of oscillation of the oscillator $i$, and $c_i$ is the actual center. The Proportional Derivative controller generates the joint torque on the basis of the desired and current angles.

### B. The task and the robotic hands

The robotic hand is required to perform a cyclic manipulation task consisting of rotating an object through the cooperation of thumb and index fingers. The reward for the reinforcement learning architecture is represented by the normalized rotation angle. Thus, the fingers start exploring the task and learning how to coordinate each other on the base of the achieved reward.

The system was trained with 9 different objects for both simulated hands:

- Sphere with a radius of 0.028 m (small sphere);
- Sphere with a radius of 0.032 m (medium sphere);
- Sphere with a radius of 0.036 m (large sphere);
- Cylinder with a radius of 0.028 m (small cylinder);
- Cylinder with a radius of 0.032 m (medium cylinder);
- Cylinder with a radius of 0.036 m (large cylinder);
- Cube with a base diagonal size of 0.028 m (small cube);
- Cube with a base diagonal size of 0.032 m (medium cube);
- Cube with a base diagonal size of 0.036 m (large cube);

Neural architecture in Fig. 2 was implemented on two different simulated anthropomorphic robotic hands: iCub hand (Fig. 3 (a), (b)) and DLR HIT-hand II (Fig. 3 (c), (d)) and a comparative analysis was carried out during manipulation tasks involving all objects. In particular, system capability to learn rotating the object and hand working area, in the case of one handled object (i.e. the sphere with radius of 0.032 m), were assessed.

Both simulated hands rely on kinematic and dynamic models of the real hands. The 3D hands and the environment were simulated using the open-source physical-engine software interface “OPAL” (Open Physics Abstraction Layer).

The iCub robotic hand (in Fig. 3(a)) has 5 fingers with in total 20 joints and 9 actuated degrees of freedom (DOFs) [5]. In particular, the thumb has four joints: two uncoupled joints (opposition and abduction/adduction) and two coupled joints (PIP and DIP flexion/extension). The index and middle fingers have one uncoupled joint (MCP flexion/extension) and two coupled joints (PIP and DIP flexion/extension); the abduction/adduction joint of index, ring and little finger is actuated by a single motor. The joints of the other fingers (ring and little) are coupled by a single motor (MCP, PIP and DIP flexion/extension). The iCub hand has the same size of a 2 years old child hand. The fingers are 0.068 m length and have a diameter of 0.012 m. The 20 joints are actuated using 9 DC brushed motors (2 in the hand and 7 in the forearm).

The DLR HIT-Hand II (in Fig. 3(c)) is composed of 5 identical fingers with 20 DOFs in total and 15 motors. Each finger (including the thumb) has 4 joints with 3 actuated DOFs: adduction/abduction, MCP flexion/extension and PIP flexion/extension. PIP and DIP joints are mechanically coupled with 1:1 ratio [7]. The thumb is constrained to have a fixed opposition of 35° in the xy-plane and an inclination, respect to z-axis, of 44° (Fig. 3 (c)). This hand is slightly bigger than an adult human hand with finger length of 169 mm and diameter of 20 mm. The DOFs are actuated by flat brushless DC motors embedded in the fingers and the palm. The two hands have two important differences: the size (as explained above) and the thumb degree of freedoms. In particular, in the iCub hand the thumb opposition is an active DOF [0°, 120°] actuated by a single motor; on the other hand, in the DLR/HIT Hand II the thumb has a fixed opposition.

This paper focuses the attention on these two main features and tries to find new insights into the design of new...
robotic hands by investigating the fundamental role of the thumb opposition in hand manipulation capabilities.

In the simulation trials only 4 DOFs were controlled: 2 DOFs of the index finger and 2 DOFs of the thumb; all other DOFs were kept to fixed values.

The 4 DOFs controlled on the iCub hand are:
- for the index finger: PIP flexion/extension (F/E) joint with a range of motion (ROM) of [0°, 90°] (PIP and DIP are mechanically coupled) and MCP adduction/abduction (A/A) with a ROM of [-15°, 15°] (MCP and IP are mechanically coupled);
- for the thumb: MCP flexion/extension joint with ROM of [0°, 90°] (MCP and IP are mechanically coupled) and opposition (O) of the thumb [0°, 130°].

The 4 DOFs controlled on DLR/HIT hand II are:
- for the index finger: PIP flexion/extension (F/E) joint with ROM of [0°, 90°] (PIP and DIP are mechanically coupled 1:1) and MCP adduction/abduction (AA) with ROM of [-15°, 15°];
- for the thumb: PIP flexion/extension (F/E) joint with range of motion of [0°, 90°] and MCP adduction/abduction (AA) with ROM [-15°, 15°]. In this phase of preliminary explorations coupling between PIP and DIP joints of the DLR thumb was not considered. This issue will be extensively investigated in the future for understanding if and how it constrains the learning mechanism of fine manipulation tasks.

C. The tests and the working area

The hierarchical bio-inspired model was tested on the two simulated robotic hands interacting with the 9 different objects described above.

The training phase consisted of 500 simulation trials (each trial was made of 3000 cycles with integration steps of 0.001s). Each hand was tested with each object and the learning capability was studied through the observation of the object rotation angle, imposed by the hand.

The role of the thumb in performing the manipulation task was further investigated through the characterization of the working areas described by the two hands when handling the same object. The hand working area was determined by testing the neural architecture with the same sphere located in different positions of the space. In particular, the object location was moved along x and y axes (in Fig. 4) with a step of 0.01m while z-coordinate was kept constant.

The initial object position was fixed in front of the index finger, in a central location between thumb and index, in order to facilitate reaching to both fingers. As a consequence, the starting position was different for the two hands, in relations to the hand dimensions.

Three types of tests were carried out:
1) the sphere position was moved along x (y is fixed)
2) the sphere position was moved along y (x is fixed)
3) the sphere position was moved along x and y.

Performance achieved by the DLR/HIT Hand II and iCub hand during interaction with the 9 different objects is summarized in Table I. This reports the reward values at the end of the 500 learning trials. The results show that for each shape, the hand performance decreases when the size of the object increases.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Reward (iCub hand)</th>
<th>Reward (DLR/HIT hand II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>3.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Cylinder</td>
<td>2.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Cube</td>
<td>1.3</td>
<td>1.2</td>
</tr>
</tbody>
</table>

The iCub and DLR/HIT Hand II index joint trajectories for adduction/abduction and PIP flexion/extension are presented in Fig. 5 and Fig. 6, respectively.

The analysis of the hand working area was carried out with the medium sphere. Figure 7 shows the resulting working areas for each hand. Note that, to compare the working areas, the starting position of the object for each hand was shifted in a central position (x,y)=(0;0). All the other locations of the object were shifted accordingly.

The black dot on the plot represents sphere position during the manipulation task. For each position of the sphere the corresponding reward value (rotation of the object) is indicated. NO indicates there was no rotation of the sphere. The shaded area indicates the hand working area. The iCub hand working area is extended of 0.07 m along x and y axes. The DLR/HIT Hand II working area is extended of 0.05 m along x and 0.04 m along y.

Fig. 8(a) and Fig. 8(b) show a comparison of the rewards obtained by different positions of the object for each hand. Every curve represents the performance of the hand.
executing the manipulation of the object in the position indicated in the legend.

![Diagram](image)

Fig. 5 (a) The index adduction/abduction joint trajectory of iCub hand; (b) The index PIP flexion/extension joint trajectories iCub hand.

![Diagram](image)

Fig. 6 (a) The index adduction/abduction joint trajectory of DLR/HIT hand II; (b) The index PIP flexion/extension joint trajectories DLR/HIT hand II.

D. Discussion

Observing the reward values in Table I it is apparent that both robotic hands can learn to manipulate all the tested objects. However, looking at the achieved rewards it is possible to note that task difficulty increases with the size of the object for both hands. In any case, the reward obtained by the simulated iCub hand is always bigger than the reward achieved by the simulated DLR/HIT Hand II, despite the hand dimensions. The slimmer size of the iCub seems to play a beneficial role in the maximization of the rotation impressed to the object.

On the other hand, the systematic analysis of the hand working areas for a given object in Fig. 7 has shown that the working area of the iCub hand is nearly twice that of the DLR/HIT Hand II, despite the smaller hand size. In particular, the main difference is obtained in the left half-plane where the thumb is expected to significantly contribute to the task. The DLR/HIT hand II cannot learn to rotate the object located in the positions with high values of x and y; on the contrary, the iCub hand can advantageously exploit the thumb opposition to explore a bigger area and, globally, outperform the DLR/HIT Hand II.

In particular, performance achieved by the iCub hand is higher than the DLR/HIT Hand II in two main cases:

1) when thumb opposition has a central role in the execution of the task, e.g. for low values of x and y coordinates of the object.

2) when the index finger has a central role in the manipulation task, e.g. in the upper area of the second quadrant. In this case, the finger touches the object from below and the hand benefits from the slimmer size with respect to the DLR hand.

In general, however, it is possible to observe that the reward changes with the object position and the best performance for both hands is achieved when the two fingers cooperate during the task execution. For the DLR/HIT Hand II this corresponds to locations in the first quadrant (Fig. 7 (b)), instead, for the iCub hand this happens for positions in the first and third quadrants (Fig. 7 (a)). In these object locations the system learns to use both fingers and alternate them to impress major rotation. In the other locations, one of the two fingers is too distant from the sphere and the rotation is committed at the nearest finger.

III. Conclusion

This paper has proposed a study on the role of the thumb in cyclic manipulation tasks, with special attention to thumb opposition, regarded as a fundamental feature of the human hand but often neglected in robotic hands.

A bio-inspired neural architecture is tested on two robotic hands with different kinematic structures and sizes: iCub hand and DLR/HIT Hand. The robotic hands were required to learn rotating 9 different objects. The achieved results showed that the task difficulty increased with object size and that the iCub always outperformed DLR/HIT hand II.

In addition to learning capabilities, also a systematic analysis of the working areas of the two robotic hands was conducted. The results showed again that the iCub structure advantageously influence the manipulation capability in the performed cyclic tasks. It is clearly demonstrated that the best performance of the iCub hand is due to the presence of thumb opposition that permits the exploration of a bigger area and allows touching the object several times thus...
maximizing the rotation.

Fig. 7 (a) The working area of iCub hand. (b) The working area of DLR/HIT Hand II. Every point on the plot represents a different position of the sphere center. The origin of the reference system is the shifted starting position of the sphere. The sphere is moved with a step of 0.01 m along x and y directions. The red label represents the reward number obtained at the end of the learning for the corresponding position of the object.

Fig. 8 The reward obtained in 500 trials of (a) iCub hand and (b) DLR/HIT Hand II for different positions of the sphere. The legend indicates the position of the object for every curve.

Future works will be addressed to include in the model of the robotic hands the coupling between PIP and DIP joints, existing in the real systems, and to test the learning capabilities of the two hands in a real experimental scenario.

REFERENCES