

Reaching and Grasping with NAO robot

N.Alberto Borghese^[0000-0002-0925-3448] and Eleonora Chitti^[0000-0003-4369-4615]

Laboratory of Applied Intelligent Systems, Department of Computer Science, University of Milan, Italy.

{alberto.borghese, eleonora.chitti}@unimi.it

Abstract. In this paper we propose a robust approach that combines classical computer vision with a smart setup that enables a NAO humanoid robot to play with children. Blocks are adopted as playing objects as they are widely used and are a highly versatile support. This approach addresses the challenges of object reaching and grasping, by allowing the robot to approach the table over which activity is carried out, orient appropriately, plan the reaching movement that is executed partly through direct and partly through inverse kinematics, grasp the block, moving it over the table and releasing it into its target position. By integrating visual information provided by the robot cameras into a closed-loop control system, no additional sensors are required. This approach enables not only children to play with a Humanoid robot in a richer way, but also to be used to treat cognitive / motor deficits inside what is termed Robot Assisted Therapy.

Keywords: NAO humanoid robot, reaching, grasping, Human-robot interaction, Robot Assisted Therapy (RAT)

1 Introduction

Recently, humanoid robots, and NAO robot in particular, have been successfully proposed as an effective companion especially for children with cognitive or social disabilities. These children generally find NAO enjoyable and non-threatening possibly due to its human-like shape and its predictable movements [1], [2], [3]. Such type of intervention, that goes under the generic name of Robot Assisted Therapy (RAT), is still at an initial stage and mainly relies on a Wizard of Oz approach in which NAO interacts through gestures and voice [2]. This is quite far from traditional interventions, in which the therapist plays with children with toys or table-top games.

We show here how combining classical computer vision methods with a smart setup enables NAO to play with children in an interactive way. To the aim we have selected blocks as these are the most used and versatile playing objects and a robust methodology to reach and grasp them on the playground, typically a table, and to move them to their target position over the table is described.

2 Related work

The ability to grasp objects of different sizes and shapes is one of the most important skills for a humanoid robot. A variety of reaching and grasping approaches has been described for NAO robots in different scenarios [4]. Proposed an adaptive grasping for the RoboCup¹ competition by replacing NAO's fingers with a flat gripper on both hands with a force sensor embedded. The grasping task is performed with both hands approaching the object from both sides through vision. Once the hands are positioned at the two sides of the object, grasping is obtained moving the two hands one towards the other until the required target force is measured by sensors. In [5] objects are recognized through stereo-vision obtained introducing two cameras inside NAO's eyes. Stereo-vision is used also to plan robot motion. These approaches are based on a modification of the hardware and firmware that makes them less attractive. [6] proposes a solution based on deep learning and NAOmarks [7]. The latter are azure markers on a white background (cf. **Fig. 1**). Each marker has its own shape and it is recognized robustly by NAO basic software. In [6] one NAOmark is placed frontally on the opposite side of a table with respect to NAO. The robot first moves towards the NAOmark thus walking towards the center of the table. Once arrived close to the table, object to be grasped is recognized through Yolo3 computer vision model and the 3D position of the object with respect to the robot is computed.

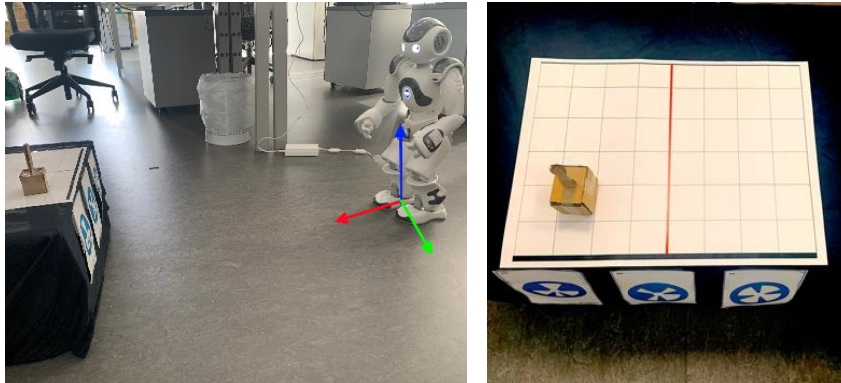


Fig. 1 A table covered with a checkerboard is used as playground (left) here. Blocks with a cylindrical handle are positioned and moved over the checkerboard squares (right). The borders of the checkerboard are marked with a thick black line and a thick red line is drawn centrally. These lines are used for fine centering and orienting NAO. On the table side three different NAOmarks are attached and are used by NAO to approach the table and for gross orientation. Pose data are expressed in NAO *RobotFrame*, centered between the two feet, with the Z axis vertical, the Y axis in the sagittal plane and the X axis in the coronal plane.

¹ <http://www.robocup.org>

3 Method

We have created a playground constituted of a table, with the child at one side and the NAO robot at the opposite side of the table. A checkerboard is painted over the table and the blocks have to be positioned inside its squares (**Fig. 1**). Given its limited workspace, NAO will be able to reach blocks in the first two lines. On one side of the table three different NAOmarks are attached.

Reaching and grasping by NAO can be subdivided into the following subtasks; A) move towards the table center. B) orient parallel to the table and raise right or left arm to get the hand over the table. C) Now NAO is ready for grasping: position the hand over the chosen block, orient the hand properly to have a stable grasp and close the fingers. D) translate the block to its target position and release it there, and E) get back to the starting position.

All these tasks are mostly based on vision information captured by the two NAO's cameras, thus realizing a closed-loop control of robot movement (cf. [7]).

3.1 Moving towards the table center and orienting parallel to it

As suggested in [6], [7], NAO computes its 3D position with respect to the central NAOmark from the images of its upper camera, and moves towards it. At each time step NAO measures its actual 3D position with respect to that of the marker and recomputes its path thus realizing a closed-loop gait control. The 3D NAO *Robot Frame* solid with the feet (cf. Fig. 1) is used here as this is the reference frame most frequently used for gait. Closed-loop control can be used until the marker can be viewed by the NAO camera, until a distance of about 300 mm from the table. Afterwards, the marker cannot be seen anymore and from then on the movement control becomes open-loop: the robot walks towards the last 3D position computed for the NAOmark. To avoid hitting its feet against the table, the robot stops at a distance of 150 mm from the marker, that is the length of its feet with respect to the NAO *Robot Frame* plus a margin.

To avoid hitting the table during subsequent arm reaching, NAO raises its arms above the table level, when it switches from closed to open-loop control gait. To this aim it moves arms from parallel to the body to the NAO *StandInit* position, with arms slightly folded (**Fig. 2a**, shoulder roll = 45 degrees and shoulder pitch = 15 degrees). From there, direct kinematics is applied to reach a final pose with shoulder roll = 45 degrees and shoulder pitch = 70 degrees to center the hands with respect to the working space (**Fig. 2b**)

The two lateral markers on the table side are used to orient the robot parallel to the table. To this aim, the yaw rotation angle (around z axis) is computed as $\theta = \arctan\left(\frac{x_R - x_L}{y_R - y_L}\right)$ where $\{x_R, y_R\}$ is the position of the right marker and $\{x_L, y_L\}$ the position of the left one. This rotation is applied to NAO before starting moving forward towards the table to start as parallel as possible with respect to the table.

To increase the accuracy of the final 3D position of the robot with respect to the table the stride length in the open-loop phase of gait is reduced by setting a maximum forward translation of 20 mm for each step.

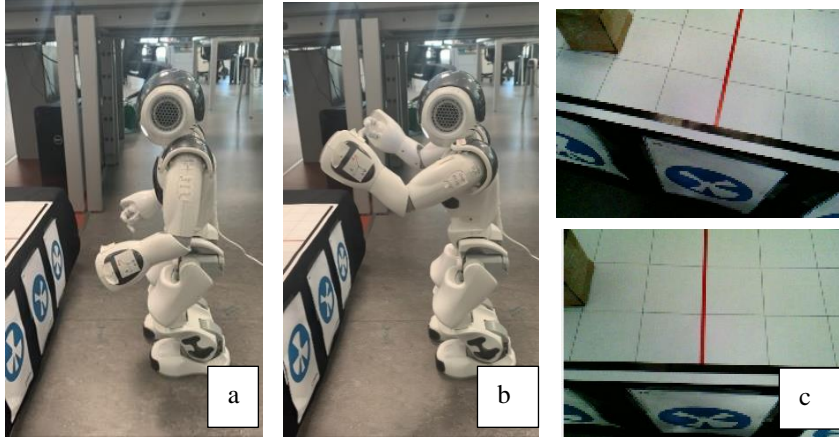


Fig. 2 NAO raises its hands at the end of its closed-loop walk from the NAO *StandInit* position (a) to the final position well over the table (b). When NAO is not parallel to the table it sees the checker-board at an angle (left). NAO rotates around the vertical Z axis to bring the thick horizontal line parallel to the horizontal axis on the image provided by its lower camera (column c). This allows orienting NAO parallel to the table.

3.2 Pose adjustment

Once the table has been reached the parallelism between NAO and the table may have got lost, so we perform a final adjustment of NAO orientation. To this aim, we first filter the image provided by the lower camera by a Canny edge detector [8]. We then apply the Hough transform to the edges detected to calculate the tilt angle of each line on the image. In this process we filter out all the lines close to vertical (a threshold of 50 degrees was used). We compute the NAO rotation as the median angle of the remaining lines and we rotate the robot around the vertical z-axis by this angle (Fig. 2c). After this rotation, the line becomes parallel to the horizontal axis and the robot becomes oriented correctly in front of the table and it can proceed with grasping.

3.3 Grasping

We start from the observation that given a 3D point on the checkerboard and the checkerboard 3D orientation, the 3D position of all the squares is defined. Therefore, the grasping and the release position of a block can be defined simply specifying the row and column of the desired square. To design an effective grasping, the structure of the NAO hand has been taken in consideration; this has three fingers: a shorter thumb and two longer fingers. It has only one motor that allows opening and closing synchronously the fingers, thus enabling a grasping that can be classified as Power grasp through palm opposition [9]. To allow a firm grasp, we have also worked on the design of blocks. These are generic blocks that mount a cylindrical handle on their upper face, slightly longer than the NAO hand width (Fig. 3).

To grasp a block, NAO first opens its fingers maximally, it then reaches the handle and positions its thumb and the two fingers on the opposite sites of the handle, and lastly it closes them around the handle. To grasp the handle, NAO hand is positioned with fingers and thumb parallel to the plane on the opposite sites of the handle. Such pose allows tolerance for height position error of the hand as the block can be grasped at any height of the handle. It is also tolerant to errors parallel to the plane on hand position as fingers are spaced apart 50mm before starting to close.

Grasping is planned and executed using the NAO *Frame_Torso* reference frame, located in the NAO torso, with axes parallel to the NAO *Robot_Frame*, as this is stable reference frame for arms and hands.

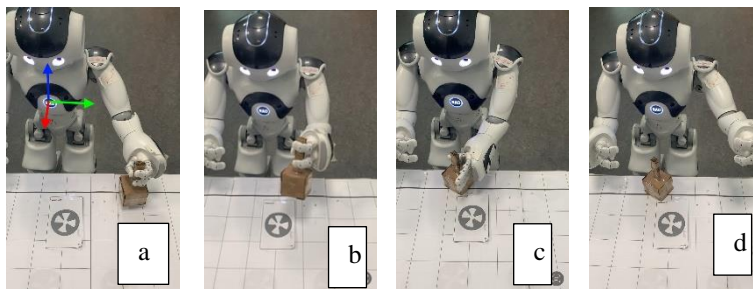


Fig. 3. Grasping a block (a). Raising the block (b). Translating it to the target position (c), and releasing it there (d). Afterwards NAO raises again its hand and moves it away from the playground. NAO *Torso Reference Frame*, shown in panel (a), is used to control this movement. The handle has a height of 50mm and a radius of 10mm

To get robust grasping we need to refine the position of the robot with respect to the block. In particular, the robot position can be offset laterally with respect to the table center at the end of gait. To correct this such offset has to be identified. This is the role of the thick vertical red line (**Fig. 1**), whose offset on the Y axis represents exactly such offset. Its detection is carried out with the same procedure used above to detect the thick horizontal border.

3.4 Translate the block in its target position and release it and go back

After the block has been grasped, NAO first raises it vertically, on the Z axis, by raising its hand 40 mm above the table plane and then moves the hand laterally on the X axis and Y axes to the target position using inverse kinematics. The vertical position remains constant in this latter phase.

When the target position has been reached, NAO lowers its hand by the same 40 mm to put the block over the table and opens it to release the block. After the block has been released, NAO raises its hand again, this time higher to get above the handle of the block, translating it vertically by 80 mm on the Z-axis. Afterwards, it translates the hand externally in the X and Y directions by 20 mm to move the hand away from the block. From this position the arm, through inverse kinematics, is brought back to its starting position, memorized before grasping was initiated.

4 Results

We tested this reaching and grasping approach with a NAO robot V6 and a host PC Dell with Intel core i7-1255U and 16 GB of RAM, and Windows 11 Pro.

The NAO robot was positioned on the floor approximately at 1 meter from the table at a small angle (< 20 degree) with respect to the table side. The table has a height of 25cm and a 8x8 checkerboard with squares of 5 cm was printed on an A3 sheet attached over its surface. The block prototype was realized with cardboard and can be positioned in any square of the first two rows of the board, as NAO is not able to reach further.

A total of 20 experiments were carried out. The robot succeeded in reaching and grasping a block 100% of the times. The mean error in hand grasping position was of $13,76 \text{ mm} \pm 23,1 \text{ mm}$ in the lateral, Y, direction that is the most critical direction, that explains the success rate and shows the robustness of the method.

Before final refinement of its orientation, NAO was rotated with respect to the table of an average angle of $-4,68 \text{ degrees} \pm 6,45 \text{ degrees}$. Open-loop gait started at an average distance of $347,3 \text{ mm} \pm 13,5 \text{ mm}$, that indeed requires a further refinement of its position and orientation.

5 Discussion

The methodology described here opens the door for enabling NAO to play with blocks in a robust way. This is particularly important as blocks are a very flexible playing tool that is also widely used in cognitive / physical treatment in clinical domains [2]. Their faces can show symbols or colors that call for ordering, matching or composition. They can also be put in an ordered sequence to represent a sequence of actions typical of everyday life. Blocks can also work as cover for objects or tiles and be therefore used in a memory like games [10].

Enabling children to carry out these tasks with a robot would allow them to decrease the stress level often induced in the interaction with a Human partner [11] on one side, and on the other side allows the therapist to concentrate on the direct interaction with the child. Lastly, as the children have shown large appreciation for NAO [2]; this would possibly raise the engagement level in the therapy session, that is a critical factor for therapy success [11]. Lastly, when carrying out these tasks with a NAO robot, we can also envisage a cooperative and/or competitive playing modality that allows addressing different interaction modalities [10].

The decomposition of the reaching and grasping task in sub-tasks over which we have full control allows to modulate each component of robot action to match actual task needs. We can for instance, increase / decrease the walking or reaching speed, we can insert pauses between one sub-task and another and so forth. This can be fundamental to match the state of the child [12], and in particular his/her stress level to provide an adequate challenge level [13].

The number of blocks that can be used is limited by NAO arms structure and size. Given a working space with a radius of 265 mm, we could design a 8x4 checkerboard, with squares of 50x50 mm, so that NAO reaches with each of its arms, 4x2 different

positions on its side of the table and the child will reach 8x4 position on his/her own side. Given this, blocks side is set to 40 mm. This is a compromise between blocks side length and their number: blocks of a smaller sizes would allow for more blocks, but their face area would be smaller and images attached on them more difficult to be understood.

Particular care has been put in obtaining a robust alignment of the robot with the table. In [6] this was obtained by putting one marker over the table but this would be intrusive as the marker would lay inside the checkerboard, the play area. We have thus preferred to put the markers on the side of the table and work open-loop from a certain distance onwards. Indeed, a drift in NAO position can take place as it approaches the table (we consistently measured a pitch of approximately 6 degrees when approaching the table for our NAO robot), but we can correct the error obtained through robot vision of the checkerboard when NAO reaches the table, as the path length of the open-loop walk is limited. We prefer to use the two thick lines for refining orientation and the lateral pairs of NAOmarks for orienting the robot parallel to the table for two reasons: NAOmarks can be detected only at a distance and the accuracy is limited. Nevertheless, this has allowed the robot to walk more easily towards the table center. Moreover, computation of the orientation through lines is most accurate when the lines cover a large portion of the NAO camera image and therefore are seen from a close distance. Accuracy was verified for a reasonable large set of NAO initial positions and orientations.

A critical issue is represented by charge level. A careful control of this level is required before the task is started to avoid that the robot shuts down before the end of the task or performs poorly: when NAO has low charge, it does not stand up completely. This would decrease the success rate and, in turns, the level of trust between the child and the robot [14].

6 Conclusion

This work proposes a robust approach for carrying out activities with blocks between a child and a humanoid robot, like NAO, without requiring to add special and costly hardware. However, this approach is not limited to this kind of applications but can be extended to all tasks that require interaction through objects that can mount a handle and use a board. It could be applied for instance to chess playing or to classical any board games, thus enlarging the domain of application and the users population.

It presently requires that the robot approaches the center of the table and from there it reaches and grasps different blocks. As the NAO arms are quite short, the working space is limited and thus few blocks can be used. This approach could be enhanced setting multiple starting position besides the table and allowing the robot to move laterally from one position to the other, thus increasing its working space.

Acknowledgment

This work was partially supported by H2020 project ESENCE (grant n. 101016112) and by PRIN2022 MIUR Italian project AIRCA (grant n. G53D23002860006).

References

1. W. A. Bainbridge, J. W. Hart, E. S. Kim, and B. Scassellati, “The benefits of interaction with physically present robots over video-displayed agents,” *Int J Soc Robot*, vol. 3, no. 1, pp. 41–52, 2011, doi: 10.1007/s12369-010-0082-7.
2. K. D. Bartl Pokorny et al., “Robot-Based Intervention for Children with Autism Spectrum Disorder: A Systematic Literature Review,” *IEEE Access*, vol. 9. Institute of Electrical and Electronics Engineers Inc., pp. 165433–165450, Dec. 03, 2021.
3. S. Baron-Cohen, “The hyper-systemizing, assortative mating theory of autism,” *Neuropsychopharmacol Biol Psychiatry*, vol. 30, no. 5, pp. 865–872, 2006,
4. H. Mellmann, M. Scheunemann, and O. Stadie, “Adaptive Grasping for a Small Humanoid Robot Utilizing Force-and Electric Current Sensors.”, vol. 1032, 2013.
5. J. Müller, U. Frese, and T. Röfer, Grab a Mug-Object Detection and Grasp Motion Planning with the NAO Robot. *Proc. IEEE-RAS Int. Conf. on Humanoid Robots*, 2012.
6. L. Zhang, H. Zhang, H. Yang, G. Bin Bian, and W. Wu, “Multi-target detection and grasping control for humanoid robot NAO,” *Int J Adapt Control Signal Process*, vol. 33, no. 7, pp. 1225–1237, Jul. 2019, doi: 10.1002/acs.3031.
7. H. Liu, C. Luo, and L. Zhang, “Target Recognition and Heavy Load Operation Posture Control of Humanoid Robot for Trolley Operation,” in *2018 IEEE-RAS 18th Int. Conf. on Humanoid Robots*, 2018, pp. 280–283.
8. R.C. Gonzales, E.R. Woods, *Digital Image Processing*, 4th edition, Person, 2019.
9. T. Feix, et al., “The GRASP Taxonomy of Human Grasp Types,” *IEEE Trans Hum Mach Syst*, vol. 46, no. 1, pp. 66–77, 2016.
10. J.C.J. Brok, E.I. Barakova. *Engaging Autistic Children in Imitation and Turn-Taking Games with Multiagent System of Interactive Lighting Blocks*. *Lecture Notes in Computer Science*, vol 6243. Springer, Berlin, 2010.
11. O. Rudovic et al. “Measuring engagement in robot assisted autism Therapy: A cross-cultural study,” *Frontiers Robotics AI*, vol. 4, July, 2017.
12. M. Pirovano, R. Mainetti, G. Baud-Bovy, P. L. Lanzi, and N. A. Borghese, “Intelligent Game Engine for Rehabilitation” *IEEE Trans of CIAIG* 8(1), 43–55, 2016.
13. S. Brambilla, G. Boccignone, N.A. Borghese, and L. Ripamonti, “Between the Buttons Stress Assessment in Video Games using Players’ Behavioural Data,” 59-69. 2022.
14. M. Luperto et al., “Seeking at-home long-term autonomy of assistive mobile robots through the integration with an IoT-based monitoring system,” *Robotics and Autonomous Systems*, Volume 161, 104346, ISSN 0921-8890, 2023.