Science Robotics
AAAS

Manuscript Template

FRONT MATTER

Title

Can robots assemble an IKEA chair?

Authors

F. Suárez-Ruiz[1], X. Zhou[1], Q.-C. Pham[1]

Affiliations

[1]School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore.

Abstract

Exploring the limits of robotic manipulation by automatically assembling an IKEA chair

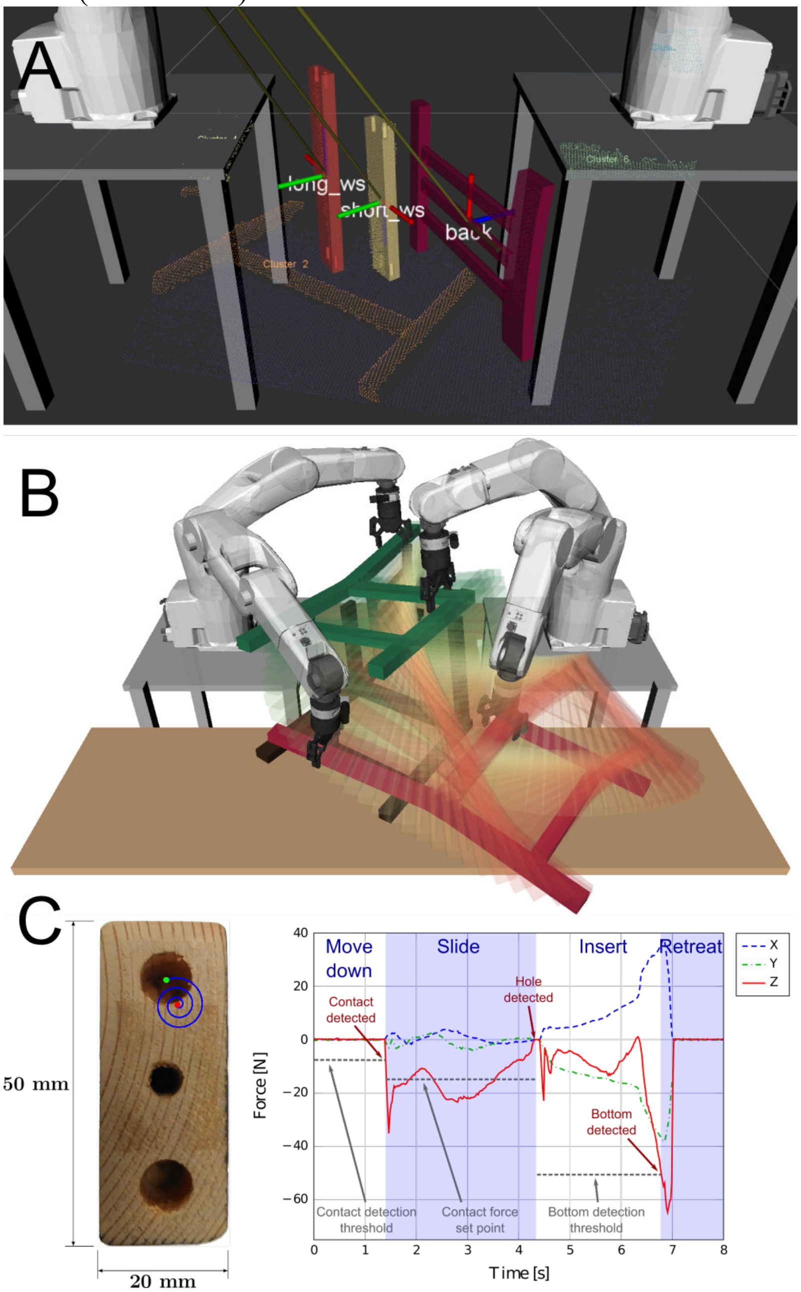
MAIN TEXT

In the past few years, rapid progresses in robotics and artificial intelligence (AI) have given rise to systems that display near or super-human skills in a variety of emblematic domains, from board games to autonomous driving. Dexterous manipulation is another marker of human intelligence. Yet, demonstrations of autonomous manipulation have been so far restricted to elementary tasks (*1*). A main reason is that complex manipulation tasks in *human environments* require mastering multiple skills – from visual and tactile localization, to motion planning, to force control, to bimanual coordination – and managing their complex interactions.

Here we wanted to assess whether, based on state-of-the-art robotic capabilities, it is possible to tackle a highly complex task that solicits all manipulation skills: the autonomous assembly of an IKEA chair, dubbed the "*moon-landing equivalent for robots*" (*2*). We procured a STEFAN chair kit from a local IKEA store (Fig. S1A). The chair parts are thus designed for human assembly, with no special provisions for robotic assembly. At the start, the parts were placed randomly within the environment. This is similar to human assembly settings and contrasts with most existing works, which rely on *structured environments* that constrain the parts to precisely-known locations. Unlike previous works (*3*), to emphasize the genericity of the setup, we used only commercial-off-the-shelf (COTS) hardware: industrial robot arms, parallel grippers, force sensors, 3D camera (Fig. S1B). This reflects the genericity of human "hardware": the same eyes and hands are used to assemble a large variety of objects.

We achieved the autonomous assembly of the full chair frame *in a single take* as shown in the accompanying video (Vid. S2) and Fig. S3. The generic manipulation pipeline was: localize the workpieces in the scene, plan the motions to reach and manipulate the workpieces in the simulation environment, execute the motions in the real environment, with force control, if necessary, to cancel discrepancies between planned and actual motions (Fig. 1). Overall, the assembly took 1219s (20min19s), of which 3s were for localization, 681s (11min21s) for motion planning, and 535s (8min55s) for execution.

Figure 1. Key components of manipulation framework. (A) Localization of workpieces by 3D vision. The estimated poses are indicated by the object models superimposed on the point cloud. (B) Planning a smooth, collision-free, bimanual motion in simulation, respecting the closed-chain constraint. (C) Force control to execute tight insertions. Left: Top surface of a rail with the superimposed spiral search pattern. Right: Graph showing the force readings during the execution of the insertion (see Text S4).



Generic manipulation framework

Localization was achieved by taking snapshots of the scene with a 3D camera and then matching the models of the objects to parts of the point cloud (Fig. 1A). The challenge is to quickly and reliably achieve a sufficiently precise localization for subsequent operations (eg. <2mm required by tight pin insertion), in a cluttered environment.

Motion planning was performed by using the Bidirectional Rapidly-exploring Random Tree (Bi-RRT) (*4, 5*) as the underlying algorithm (Fig. 1B). The challenge here is to quickly and consistently find fast, collision-free, motions in a highly cluttered environment, and to integrate with visuo-tactile perception, grasping, and execution.

Operations involving contacts require regulating the forces arising from the physical interactions between the robots and the environment. This is challenging because industrial robots, designed to be precise at positioning, are bad at regulating forces. We equipped them with force sensors mounted at the robot wrists, and used an indirect force control method (*6*): the force error obtained from the sensors is converted, through a proportional-derivative block, to a position error, which is then fed to the robot's position controller (Fig. S4). We were thus able to precisely and consistently detect holes by sliding the pin on the surfaces of the workpieces, and perform tight pin insertions (100% success across the 28 required insertions, Fig. 1C).

Planning bimanual motions (required for transporting large and heavy objects, such as the chair frames) is challenging because of the closed-chain kinematic constraint (*7*). We extended Bi-RRT as follows: at each step of the tree extension, plan a collision-free motion of the object, and then verify, by inverse kinematics, that this motion can be tracked by both arms (Fig. 1B). To *execute the planned bimanual motions*, we implemented a leader-follower approach (*7*): the leader robot was controlled in position, while the follower robot was made compliant – in order to cancel contact forces arising from the bimanual operation – by using the aforementioned force-control scheme.

Many difficulties arise from the *complex interactions between processes*: for example, small errors in the visual localization of a workpiece might be amplified during the grasping, which in turn might derail the spiral search for pin insertion. To tame that complexity, we abstracted and integrated all processes and hardware within a single framework based on ROS (Fig. S1C).

Conclusions

We have shown that robots can autonomously achieve a highly complex manipulation task: assembling an IKEA chair frame in an unstructured environment. This was possible thanks to (i) considerable progress in robotics research over the past two decades, spanning vision, planning, control, integration, and (ii) the recent trend emphasizing high-quality open-source robotics software (*5, 8, 9*).

Since the whole assembly process was performed using COTS industrial hardware, we argue that robotic manipulation in unstructured environments is ready for industrial deployment. This will truly unleash the full potential of robotics to address the needs for automation in such areas as electronics manufacturing, aircraft manufacturing, logistics, or generally high-mix, low-volume sectors, where sophisticated, rigid, assembly lines are not viable.

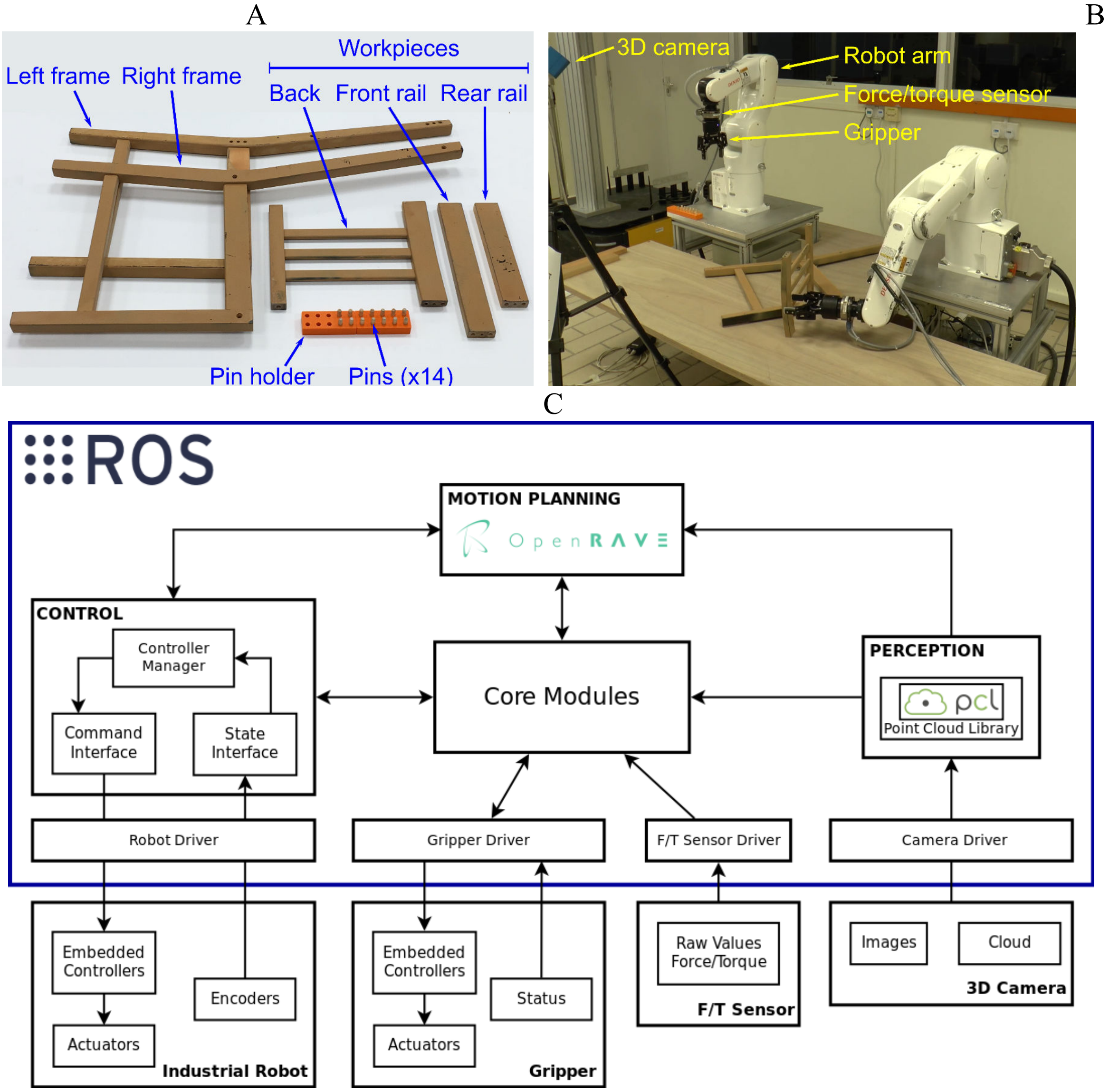
There is still an important limitation: while all the steps were automatically planned and controlled, their *sequence* was hard-coded through a considerable engineering effort. One can envision such a sequence being automatically determined from the assembly manual, through natural-language interaction with a human supervisor, or ultimately, from an image of the chair. Combining the capabilities and the framework developed here with the recent advances in AI could lead, in a near future (*10*), to such fully autonomous assembly.

References and Notes

1. F. Suárez-Ruiz, Q.-C. Pham, A Framework for Fine Robotic Assembly, in *2016 IEEE International Conference on Robotics and Automation* (Stockholm, Sweden, 2016), pp. 421–426.
2. MIT Technology Review, Moon-Landing Equivalent for Robots: Assembling an IKEA Chair (2015), (available at https://www.technologyreview.com/s/541781/moon-landing-equivalent-for-robots-assembling-an-ikea-chair/).
3. R. A. Knepper, T. Layton, J. Romanishin, D. Rus, IkeaBot: An autonomous multi-robot coordinated furniture assembly system, in *2013 IEEE International Conference on Robotics and Automation* (IEEE, 2013), pp. 855–862.
4. J. J. Kuffner, S. M. LaValle, RRT-connect: An efficient approach to single-query path planning, in *2000 IEEE International Conference on Robotics and Automation* (IEEE, 2000), vol. 2, pp. 995–1001.
5. R. Diankov, thesis, Carnegie Mellon University, Robotics Institute (2010).
6. H. Seraji, Adaptive admittance control: an approach to explicit force control in compliant motion, in *1994 IEEE International Conference on Robotics and Automation* (1994), pp. 2705–2712.
7. X. Zhou, P. Lertkultanon, Q.-C. Pham, Closed-Chain Manipulation of Large Objects by Multi-Arm Robotic Systems. *IEEE Robot. Autom. Lett.* **2**, 1832–1839 (2017).
8. R. B. Rusu, S. Cousins, 3D is here: Point Cloud Library (PCL), in *2011 IEEE International Conference on Robotics and Automation* (IEEE, 2011), pp. 1–4.
9. M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, A. Ng, ROS: an open-source Robot Operating System, in *ICRA Workshop on Open Source Software* (2009).
10. G.-Z. Yang, et al. The grand challenges of Science Robotics. *Science Robotics* 3.14 (2018): eaar7650.

SUPPLEMENTARY MATERIALS

Figure S1. Chair kit, hardware and software architectures. (A) STEFAN chair kit purchased from a local IKEA store. (B) Our hardware architecture, consisting uniquely of commercial-off-the-shelf hardware: two six-axis robot arms (Denso VS-060), two six-axis force sensors (ATI Gamma), two parallel grippers (Robotiq 2-Finger 85), one 3D camera (Ensenso N35-804-16-BL). (C) Our software architecture was built on top of open-source libraries and was designed to be modular and hardware-independent. We used OpenRAVE for collision-free motion planning, the Point Cloud Library (PCL) for 3D computer vision, the Robot Operating System (ROS) for integration. The hardware drivers, partially developed by our group, are open-source and can be accessed at our group's online repository https://github.com/crigroup/.



Video S2. Autonomous assembly of the full chair frame recorded in a single take.

Figure S3. Key steps in the assembly process. (A) Picking up workpiece. (B) Picking up dowel pin. (C) Simultaneously transporting workpiece and pin towards their pre-insertion positions. (D) Inserting pin into workpiece. (E) Inserting workpiece into left frame. (F, G, H) Transporting right frame using both robot arms. (I) Adjusting right frame. (J) Flowchart of the assembly process. Shaded boxes correspond to operations that require force control.

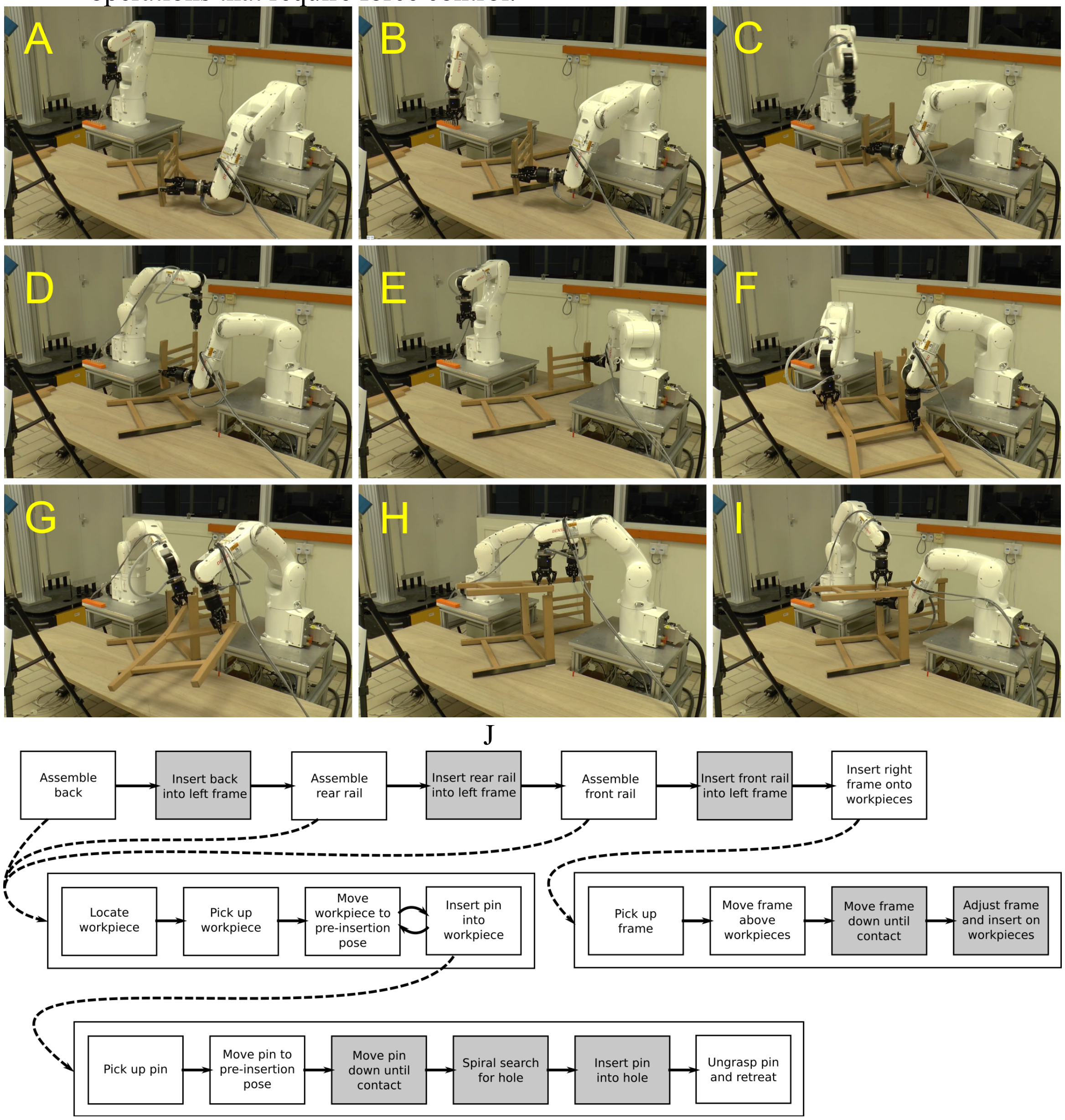
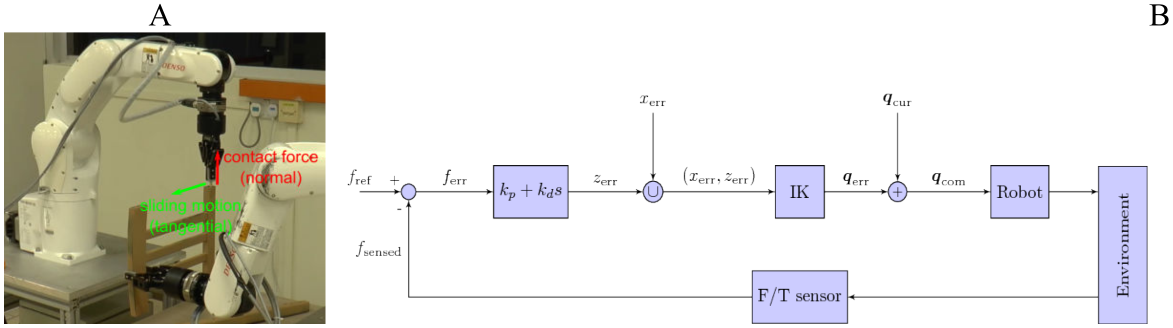
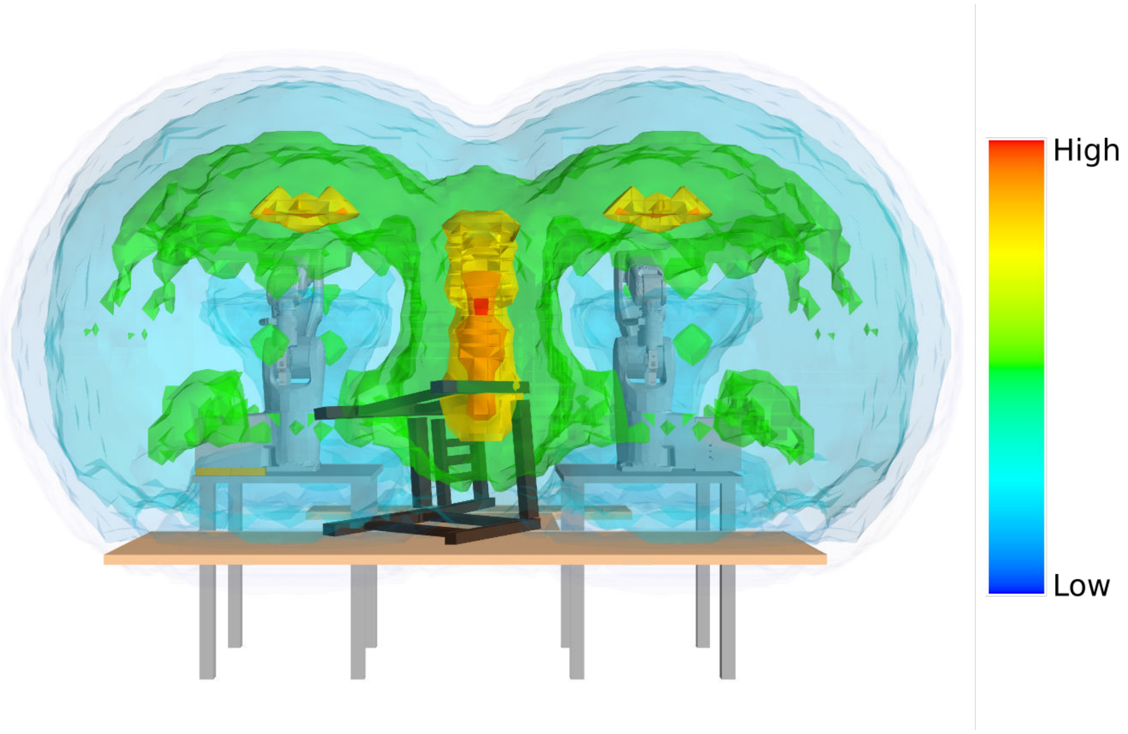


Figure S4. Hybrid position/force control. (A) Sliding a pin on a surface requires a hybrid position/force control scheme: force is regulated in the direction normal to the surface, while position is regulated in the tangential directions. (B) Block diagram of our hybrid position/force controller. Force set-points (f_{ref}) and position set-points (x_{ref}) can be specified in complementary directions, in a flexible manner. The left part of the controller implements the conversion of the force-error signal (f_{err}) to a position-error signal (z_{err}).



Text S5. Tight pin insertion using force control. Fig. 1(C), left, shows the top surface of a rail with the superimposed spiral search pattern. The red dot denotes the center of the pin at the beginning of the search. The green dot denotes the center of the pin when the hole is detected. The spiral search parameters were: pitch = 1.5 mm, linear velocity = 5 mm/s. Fig. 1(C), right, shows the force readings during pin insertion. Z corresponds to the direction normal to the workpiece surface, while X and Y span the workpiece surface. We took the following approach: (1) move the pin and the workpiece to their pre-insertion poses, computed based on the available estimates of the pin and workpiece positions with respect to the grippers; (2) from the pre-insertion pose, move the pin down until contact with the workpiece surface (detected by checking the normal contact force); (3) slide the pin on the workpiece surface using the hybrid position/force control scheme just presented: regulate the normal force to 15 N, regulate the tangential velocity to describe a spiral on the workpiece surface; (4) detect the hole by checking when the normal contact force vanishes; (5) push the pin down the hole using again hybrid position/force control: regulate the tangential force to zero, regulate the normal velocity to go downwards; (6) stop pushing when reaching the bottom of the hole (detected by checking the normal contact force), ungrasp the pin, and retreat.

Figure S6. Workspace optimization. The workspace was optimized to facilitate bimanual operations such as pin insertion or manipulation of the right chair frame. The figure shows the normalized combined reachability map of the bimanual setup. The goal was to install the manipulator bases as far as possible, to minimize possible mutual collisions, while maintaining a large enough collaborative volume (shaded in yellow).



Acknowledgments: We would like to thank H. Nguyen, I. Chew and P. Lertkultanon for their help in parts of the experiments. **Funding**: This work was supported in part by MOE Tier 1 grant RG109/14, NTUitive Gap Fund NGF-2016-01-028 and SMART Innovation Grant NG000074-ENG. **Author contributions**: F.S.R. developed the software framework including low and high-level interfaces for the devices. He also developed the modules for pick and place, hybrid position/force control and workpieces localization. X.Z. developed the module for bimanual closed-chain motions. Q.C.P contributed to the conception and execution of the project and led the writing of the paper. All authors contributed to the design of the experimental setup and the strategies to perform the assembly task. **Competing interests**: The authors declare that they do not have competing interests.