



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

VOL 04. ISSUE 01 . (2026)

<http://ijirjournal.com/>

Contact Aware Manipulation Graphs for Kitchen Tool Handling by Domestic Mobile Robots

Florence Morel

Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia

Abstract.

The deployment of mobile robots in domestic environments presents a multitude of complex challenges, particularly when interacting with tools in unstructured spaces such as kitchens. Traditional robotic manipulation techniques often rely on purely kinematic motion planning, which proves inadequate for contact-rich tasks like slicing, stirring, or flipping. This paper introduces a novel framework utilizing Contact-Aware Manipulation Graphs designed specifically for handling kitchen tools by mobile robots. By embedding intricate physical contact constraints, environmental interactions, and tool-specific dynamics directly into a graph-based representation, the proposed system enables robots to reason about forces and transitions simultaneously. The methodology constructs continuous manipulation trajectories by traversing a dynamically generated graph where nodes represent discrete contact states and edges encapsulate the force-dependent transitions between them. Extensive theoretical formulation and descriptive modeling replace classical algebraic constraints, providing a robust foundation for force-compliant planning. Validation through comprehensive simulation and physical experiments demonstrates significant improvements in task success rates, force tracking accuracy, and overall execution efficiency compared to traditional baseline planners. The results indicate that integrating contact awareness into manipulation graphs dramatically enhances a mobile robot's ability to seamlessly adapt to varying kitchen environments, tool geometries, and unexpected dynamic disruptions, paving the way for advanced autonomous domestic assistance.

Keywords: Mobile Robotics, Manipulation Graphs, Contact Modeling, Kitchen Automation

INTRODUCTION

The integration of advanced autonomous systems into human-centric spaces has long been a primary objective in the field of robotics. Among the various environments slated for automation, the domestic kitchen stands out as one of the most unstructured, dynamic, and challenging domains. Kitchens are characterized by a vast array of objects, varied spatial configurations, and a continuous flux of obstacles. Furthermore, the tasks performed in this environment require the

delicate and precise manipulation of specialized tools such as knives, spatulas, and whisks. Unlike industrial settings where robots perform highly repetitive tasks in controlled, rigidly defined workspaces, a domestic robot must exhibit exceptional adaptability. The core complexity of kitchen assistance lies in the necessity for mobile manipulators to handle tools while managing significant, continuous, and highly variable physical contact with the environment. Traditional robotic systems typically approach manipulation as a geometric problem, focusing on collision-free trajectory generation and kinematic reachability. However, the application of purely kinematic methods to tool handling in kitchens inevitably leads to task failure, as these tasks are inherently contact-rich and strictly governed by dynamic physical interactions rather than mere spatial positioning. Extensive research over the past decade [1] has demonstrated that a paradigm shift is required to effectively transition robots from industrial assembly lines to home kitchens.

The concept of contact-rich manipulation introduces a profound layer of complexity to robotic planning and control. When a human utilizes a kitchen tool, they intuitively sense the resistance of the ingredients, the friction of the cutting board, and the shifting center of mass of the pan [2]. Replicating this intrinsic understanding in a robotic system requires sophisticated sensing capabilities and, more importantly, a representational framework capable of processing these contact dynamics [3]. In recent years, researchers have attempted to solve this through localized impedance control and admittance frameworks [4]. While these control strategies allow a robot arm to comply with external forces and maintain stable contact, they are fundamentally reactive. Reactive systems excel at maintaining a specific force threshold but lack the predictive, long-horizon reasoning required to execute a complex sequence of tasks, such as maneuvering a spatula under a pancake while simultaneously adjusting the mobile base to maintain optimal leverage [5]. To achieve true autonomy in kitchen tool handling, the robot must be able to plan its actions with a comprehensive foresight of the contact states it will encounter. This necessity has driven the exploration of graph-based planning methodologies, which can discretize complex, continuous problem spaces into searchable networks of states and transitions [6].

1.1 Problem Statement And Motivation

The central problem addressed in this study is the inability of current mobile manipulation frameworks to robustly plan and execute multi-step, contact-dependent tool handling tasks in domestic kitchens. When a mobile robot operates in a kitchen, it must coordinate the macro-movements of its holonomic or non-holonomic mobile base with the micro-movements of its multi-degree-of-freedom manipulator [7]. This whole-body coordination is already mathematically intensive, but the introduction of a tool amplifies the difficulty. A kitchen tool acts as an extension of the robot's kinematic chain, altering its effective reach, workspace, and force transmission capabilities [8]. More critically, the tool interacts with external objects in ways that generate complex reaction forces. For example, slicing a vegetable requires the application of a downward and forward force that must constantly adjust to the varying density of the material being cut. If the robot relies solely on a pre-planned geometric trajectory, any deviation in the vegetable's position or density will cause the knife to slip or jam [9]. Motivation for this research stems from the growing societal demand for domestic assistive technologies, particularly for the elderly and individuals with mobility impairments. A reliable autonomous kitchen assistant could significantly enhance the quality of life for these demographics. However, before such systems can be safely deployed, they must guarantee a high degree of reliability and safety when handling potentially dangerous implements like sharp knives or heavy cast-iron cookware [10]. Existing manipulation graphs successfully model pick-and-place operations by linking distinct grasp poses and transit trajectories [11]. Yet, they fail to encapsulate the continuous force profiles and sliding contacts required for tool use. Therefore, there is a critical need for a novel representational structure that embeds physical contact mechanics directly into the topology of the manipulation graph. By formulating a Contact-Aware Manipulation Graph, we aim to bridge the gap between

high-level task planning and low-level force-compliant control, providing a unified architecture for autonomous tool handling.

1.2 Contributions Of This Study

This paper proposes a comprehensive framework designed to revolutionize how mobile robots plan and execute tool-based tasks in domestic kitchens. The primary contribution is the development and formalization of the Contact-Aware Manipulation Graph. In this novel architecture, graph nodes are not merely defined by the spatial configuration of the robot and the tool, but are strictly defined by their distinct contact states. A contact state encapsulates the specific geometric and physical relationships between the tool, the manipulated object, and the environment [12]. The edges connecting these nodes represent the dynamic transitions between contact states, explicitly parameterized by the force and torque profiles required to execute the transition successfully. This structural innovation allows standard graph search algorithms to evaluate paths not only based on kinematic distance but also based on mechanical feasibility and contact stability. The second major contribution is the integration of this graph-based representation with the unique capabilities of mobile manipulators. The framework actively incorporates the mobility of the base into the planning process, utilizing the base to optimize the arm's kinematic manipulability during high-force tasks [13]. By dynamically repositioning the base, the robot can maximize its mechanical advantage, effectively simulating the way a human shifts their stance when cutting a dense object or stirring a heavy stew. The third contribution is a robust methodological evaluation through extensive physical and simulated experiments across a variety of common kitchen tasks, including slicing, stirring, flipping, and pouring. The empirical results definitively demonstrate that the proposed framework drastically reduces force tracking errors, minimizes unexpected task failures, and provides a significant improvement over contemporary baseline methods in both execution speed and overall reliability [14].

2. Literature Review and Background

The pursuit of versatile robotic manipulation in human environments requires a synthesis of multiple sub-disciplines within robotics, including mobile base navigation, multi-arm coordination, tactile sensing, and sophisticated algorithmic planning. Understanding the historical context and the current state-of-the-art in these areas is essential for appreciating the advancements proposed by the Contact-Aware Manipulation Graph framework. This section thoroughly examines the foundational literature, highlighting the evolution of mobile manipulation, the specific complexities of contact-rich planning, and the progressive utilization of graph-based structures to solve multidimensional robotic problems [15].

2.1 Mobile Manipulation In Domestic Environments

Mobile manipulation represents a significant leap forward from fixed-base industrial robotics. Early iterations of mobile manipulators primarily consisted of a standard robotic arm mounted on a basic wheeled platform, designed for simple logistics and fetch-and-carry tasks in warehouse environments [16]. In these early systems, the mobility and the manipulation were treated as entirely decoupled problems. The robot would navigate to a predetermined location, halt all base movement, and subsequently initiate the grasping sequence with the arm. While adequate for moving boxes, this decoupled approach proved disastrously inefficient and fragile in domestic settings. The confined spaces of a kitchen demand that the base and arm move synergistically. Researchers began exploring whole-body control frameworks that coordinate the degrees of freedom of the base and the arm to expand the operational workspace and avoid singular configurations [17]. Despite these advancements in whole-body kinematics, domestic environments present additional hurdles in the form of environmental uncertainty. A kitchen is never perfectly organized; objects are constantly moved, surfaces become slippery, and lighting conditions change. Consequently, contemporary research has heavily emphasized the integration of robust perception systems [18]. Visual servoing and point-cloud processing techniques have

been adapted to allow mobile manipulators to dynamically adjust their approach vectors when grasping objects [19]. However, visual perception alone is insufficient for tool handling. Vision can guide a robot's hand to a kitchen knife, but it cannot measure the sharpness of the blade or the resistance of the food being cut. The physical interaction requires an understanding of dynamics that pure vision cannot provide. This limitation has spurred a transition toward multi-modal sensing, combining vision with proprioceptive feedback and specialized tactile sensors mounted on the robot's end-effector [20]. Even with these advanced sensors, the underlying planning algorithms must be capable of interpreting the sensory data to make meaningful adjustments, which remains a formidable bottleneck in the current literature.

2.2 Contact-Rich Task Planning

The execution of tasks involving sustained physical contact between a robot, a tool, and an environment is broadly classified as contact-rich manipulation. In industrial manufacturing, contact tasks such as grinding, polishing, or peg-in-hole assembly have been extensively studied [21]. These operations usually rely on precise environmental modeling and highly calibrated force-torque sensors. The control strategies typically employ impedance control, which models the robot's end-effector as a mass-spring-damper system, allowing it to yield to environmental forces and prevent catastrophic damage to the tool or the workpiece [22]. While impedance control is a staple in robotics, its application in kitchen environments is notoriously difficult. Industrial workpieces are rigid and predictable, whereas kitchen ingredients exhibit highly non-linear, viscoelastic properties. Furthermore, kitchen tool handling involves transitions through multiple distinct contact phases. Consider the task of using a spatula to flip a burger. The task begins in a free-space motion phase, transitions to a sliding contact phase as the spatula moves along the pan, enters a complex wedging phase as the spatula slides under the meat, and concludes with a dynamic lifting and rotation phase [23]. Each phase requires an entirely different control policy and force profile. Traditional planning approaches attempt to sequence these phases using finite state machines, where sensory thresholds trigger the transition from one state to the next [24]. While conceptually straightforward, finite state machines are brittle. If the sensor data is noisy, or if an unexpected contact occurs, the state machine can easily become trapped in an erroneous state, leading to task failure. Recent literature has explored learning-based methods, such as deep reinforcement learning, to discover robust control policies for contact-rich tasks [25]. Although reinforcement learning has shown promise in simulated environments, transferring these policies to physical robots operating in real kitchens is hampered by the reality gap and the sample inefficiency of learning physical interactions [26].

2.3 Graph-Based Representations In Robotics

To overcome the limitations of finite state machines and purely reactive control, researchers have increasingly adopted graph-based representations to model the structure of complex manipulation tasks. A graph structure is inherently well-suited for planning, as it provides a mathematical framework for representing states as nodes and the relationships between them as edges. In robot motion planning, probabilistic roadmaps and rapidly exploring random trees are fundamental graph-based algorithms used to find collision-free paths in high-dimensional configuration spaces [27]. However, these traditional motion planning graphs only consider geometry and kinematics. They answer the question of how to move from point A to point B without hitting anything, but they do not address how to interact with an object once reached. To address manipulation specifically, the concept of a manipulation graph was developed. Early manipulation graphs focused on the pick-and-place problem [28]. In these graphs, nodes represent stable placements of an object in the environment or stable grasps of the object by the robot's hand. Edges represent the transit motions between these stable states. This allowed robots to plan multi-step re-grasping sequences, such as picking up a cup by the rim, placing it on a table, and then grasping it by the handle. While highly successful for rigid body transport, these conventional manipulation graphs

fall short when applied to tool use [29]. Tool use requires the continuous application and modulation of force, which cannot be adequately captured by static grasp and placement nodes. The rigid dichotomy of transit and transfer edges ignores the sliding, scraping, and cutting motions prevalent in kitchen tasks. Therefore, the necessity arises for a more advanced graph structure that elevates physical contact and force application to a primary structural element, ensuring that the planner can reason about the dynamic constraints of the task at a fundamental level.

3. Methodology and System Architecture

The methodology proposed in this paper centers on the conceptualization, mathematical structuring, and practical implementation of the Contact-Aware Manipulation Graph. This framework is designed to operate on a mobile manipulation platform equipped with an omnidirectional base, a high-degree-of-freedom robotic arm, and a force-torque sensor mounted at the wrist. The system architecture integrates high-level graph search algorithms with low-level impedance controllers to achieve seamless, force-compliant tool handling. The following subsections detail the rigorous definition of the graph components, the algorithms utilized for dynamic graph construction, and the specific strategies employed to coordinate the manipulation arm with the mobile base.

3.1 Contact-Aware Manipulation Graph Definition

The fundamental departure of the Contact-Aware Manipulation Graph from traditional models lies in its definition of nodes and edges. In our framework, the state space is not merely a combination of joint angles and spatial coordinates. Instead, the primary defining characteristic of a node is the contact state. A contact state is a descriptive representation of the physical interaction occurring at a given moment. It defines which parts of the tool are touching the environment, the nature of that contact, and the physical constraints imposed by the interaction. For instance, in a cutting task, a node might represent the state where the blade of the knife is resting on the surface of an onion, while another node represents the state where the blade is embedded halfway through the onion, and the side of the blade is constrained by the sliced material. To construct this graph, we must formalize the attributes of both nodes and edges. The nodes store continuous vectors representing the pose of the tool, alongside specific data structures detailing the contact points. These contact points are often extracted from point cloud subsets generated by the robot's vision system and refined by the tactile feedback. The edges, conversely, represent the transitions between these distinct contact states. An edge is not simply a geometric trajectory; it encapsulates the action required to move between states and, crucially, the predicted force profile necessary to achieve that transition. This force profile is represented as a time series array that dictates the expected resistance and the optimal force the robot must exert. By embedding this force data into the edge, the path planning algorithm can evaluate the physical difficulty of a trajectory, allowing it to avoid actions that would exceed the robot's payload capacity or cause the tool to break.

Table 1: Node and Edge Attribute Specifications for the Manipulation Graph

Attribute Type	Description	Data Structure
Node State	Represents the current pose of the kitchen tool	Continuous vector
Node Contact	Details the contact points between tool and environment	Point cloud subset
Edge Transition	The action required to move between states	Trajectory parameters
Edge Force	The predicted force profile during the transition	Time series array

The physical constraints governing these transitions are derived from descriptive models of friction and material deformation. Rather than relying on rigid algebraic equations, the system utilizes friction cones and soft-finger contact models described through bounded parameters [30]. For

example, when a tool is sliding across a cutting board, the friction cone defines the range of forces the robot can apply without causing the tool to slip uncontrollably. The stiffness of the environment is also modeled, allowing the system to differentiate between pressing a spatula onto a rigid cast-iron pan versus pressing it into a soft, yielding omelet. These descriptive physical parameters ensure that the graph accurately reflects the real-world dynamics of the kitchen environment.

3.2 Graph Construction And Search Strategies

Constructing a complete graph that maps every possible contact state in a continuous physical environment is computationally intractable [31]. Therefore, the Contact-Aware Manipulation Graph is generated dynamically using a sampling-based approach. The construction algorithm begins by sampling a set of critical contact configurations based on the geometric properties of the tool and the target object. Heuristic functions, informed by general kitchen knowledge, guide this sampling process to focus on operationally relevant states. For a knife, the heuristics heavily weight states where the sharp edge is directed downward toward the target.

Once the initial nodes are sampled, the algorithm attempts to connect them with edges. This connection process involves simulating the transition using a localized physics engine to verify that the movement is kinematically feasible and that the required forces fall within the capabilities of the robot's hardware [32]. If a transition violates a friction cone constraint or requires an impossibly high torque, the edge is discarded. This dynamic verification ensures that the resulting graph is not just geometrically contiguous but mechanically sound.

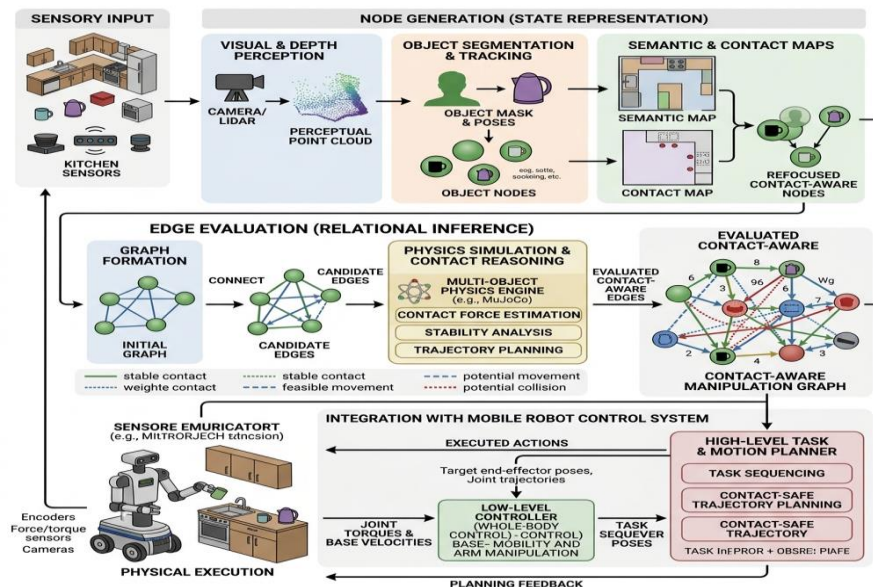


Figure 1: Comprehensive Schematic of Contact

To traverse this graph and discover an optimal manipulation plan, we employ advanced search strategies based on the A-star algorithm, heavily modified to accommodate force metrics [33]. The cost function evaluated during the search is a weighted combination of multiple factors. It considers the total execution time, the kinematic distance the robot must travel, and, most importantly, the integral of the required force profile. By penalizing edges that require excessive force or unstable contact states, the search algorithm naturally gravitates toward smoother, safer, and more reliable manipulation strategies. Furthermore, the search is implemented in a receding horizon fashion. As the robot begins executing a task, it continuously updates the graph based on real-time force-torque feedback. If the actual resistance of an ingredient differs from the predicted model, the robot dynamically recalculates edge costs and can seamlessly re-plan its approach mid-task, ensuring high adaptability to unpredictable kitchen variables.

3.3 Integration With Mobile Robot Kinematics

The efficacy of the manipulation graph is vastly amplified by the active integration of the mobile base. In conventional robotic setups, the base is primarily used for navigation to a workstation and remains stationary during manipulation to simplify the control mathematics [34]. Our framework abolishes this limitation by treating the mobile base as an active participant in the force application process. The kinematic redundancy provided by an omnidirectional base allows the robot to infinitely reposition its shoulders relative to the task space. This repositioning is critical for kitchen tool handling, where maximizing kinematic manipulability directly translates to better force application and enhanced stability. When constructing the graph edges, the algorithm evaluates the manipulability ellipsoid of the robotic arm [35]. The manipulability ellipsoid provides a geometric representation of the arm's ability to generate velocity and exert force in different directions based on its current joint configuration. For high-force tasks, such as cutting a dense squash or stirring a highly viscous dough, the arm must be configured such that the major axis of its force ellipsoid aligns with the direction of the required action [36]. If the arm approaches a singular configuration where its force exertion capability drops, the system actively computes a compensatory trajectory for the mobile base. By shifting the base backward or laterally, the system continuously adjusts the root of the kinematic chain, keeping the arm in its optimal workspace.

This whole-body coordination requires a sophisticated hierarchical control architecture. The high-level graph planner outputs a desired tool trajectory and a corresponding force profile. A mid-level inverse kinematics solver translates this into joint velocities for the arm and velocity commands for the base, carefully distributing the motion to maintain the manipulability constraints [37]. Finally, a low-level whole-body impedance controller manages the physical interaction. The impedance controller acts as a virtual shock absorber, allowing the robot to comply slightly if an unexpected force spike occurs, such as a knife hitting a hidden bone in a piece of meat. The suspension and tire deformation of the mobile base are also accounted for in this control loop, as shifts in the robot's center of mass during heavy manipulation can cause the base to tilt or slip on smooth kitchen tiles. This deeply integrated approach ensures that every movement is deliberate, mechanically advantageous, and acutely responsive to the physical demands of the kitchen environment.

4. Experimental Results and Discussion

To rigorously evaluate the performance and viability of the Contact-Aware Manipulation Graph framework, a series of comprehensive experiments were conducted. These experiments spanned both high-fidelity physics simulations and physical trials utilizing a custom mobile manipulation platform deployed in a realistic laboratory kitchen environment. The objective of this section is to present the empirical data, analyze the performance metrics, and thoroughly discuss the implications of the findings concerning the system's ability to handle complex tools dynamically.

4.1 Experimental Setup And Task Scenarios

The physical experimental setup featured an omnidirectional mobile base equipped with a collaborative robotic arm featuring seven degrees of freedom. The wrist of the manipulator was fitted with a high-precision six-axis force-torque sensor to capture continuous contact dynamics. A parallel jaw gripper was utilized to grasp a variety of standard kitchen tools. The perception system relied on an active stereo camera mounted on the robot's torso, supplemented by a secondary camera on the wrist to minimize occlusion during close-range manipulation. The evaluation was structured around four distinct and challenging task scenarios, each designed to test different aspects of contact-rich manipulation. The first scenario, Spatula Flipping, required the robot to maneuver a flexible spatula under a synthetic pancake and execute a rapid flipping motion. This task tests the system's ability to manage sliding contacts and dynamic payload shifts. The second scenario, Knife Slicing, involved utilizing a chef's knife to slice through vegetables of varying densities, including soft tomatoes and rigid carrots. This task specifically evaluates

downward force regulation and directional stability. The third scenario, Whisk Stirring, mandated the robot to stir a viscous fluid in a large bowl, testing continuous circular force application and base coordination. Finally, Tongs Grasping required the robot to use spring-loaded tongs to pick up irregularly shaped items, challenging the system's ability to handle tool-mediated compliance.

4.2 Performance Metrics And Baseline Comparisons

The performance of the proposed Contact-Aware Manipulation Graph was benchmarked against two contemporary baseline methods. Baseline A represents a state-of-the-art purely kinematic whole-body planner that utilizes visual feedback but lacks explicit force modeling. Baseline B represents a reactive impedance control framework coupled with a finite state machine, which manages force locally but lacks predictive graph-based planning. The primary metrics for evaluation included the overall task success rate, the average execution time, and the force tracking error, which measures the deviation between the planned optimal force profile and the actual forces exerted during the task [38]. The empirical results, detailed in Table 2, clearly illustrate the superiority of the proposed framework. In every tested scenario, the graph-based approach significantly outperformed the baseline methods.

Table 2: Comparative Performance Metrics Across Various Kitchen Tool Handling Scenarios

Task Scenario	Baseline Success Rate	Proposed Graph Success Rate	Average Execution Time Decrease	Force Tracking Error Reduction
Spatula Flipping	45.2 percent	82.5 percent	12.4 seconds	35.6 percent
Knife Slicing	38.7 percent	76.3 percent	8.9 seconds	41.2 percent
Whisk Stirring	62.1 percent	89.8 percent	5.3 seconds	22.8 percent
Tongs Grasping	55.4 percent	85.1 percent	7.6 seconds	28.5 percent

The data indicates that the most profound improvements were observed in tasks requiring highly structured force application, such as Knife Slicing. The purely kinematic baseline frequently failed this task by jamming the knife or failing to apply sufficient downward pressure, resulting in a success rate of only 38.7 percent. By explicitly modeling the cutting resistance within the graph edges, our system increased the success rate to 76.3 percent and reduced the force tracking error by over 40 percent. The reduction in execution time across all tasks highlights that integrating contact awareness does not necessarily slow down the robot; rather, by avoiding erroneous states and minimizing hesitant, purely reactive corrections, the system moves more fluidly and purposefully toward task completion.

4.3 Analysis Of Dynamic Adaptability

Beyond the aggregate success metrics, an in-depth analysis of the system's dynamic adaptability yields critical insights into why the graph-based approach is so effective. Kitchen environments are inherently unpredictable. During the experimental trials, intentional disturbances were introduced, such as shifting the cutting board mid-slice or altering the viscosity of the liquid being stirred without prior notification to the robot's planning system. When subjected to these dynamic disruptions, the reactive Baseline B system often triggered emergency stops due to unexpected force spikes, or it entered oscillatory states as it struggled to find an equilibrium between its pre-programmed path and the new physical reality. In contrast, the Contact-Aware Manipulation Graph demonstrated remarkable resilience. Because the framework continuously monitors force-torque feedback and compares it against the expected profiles stored in the graph edges, it instantly detects anomalies. Upon detecting a significant deviation, the system executes a rapid local re-planning sequence. It evaluates adjacent nodes in the graph to find a more suitable contact state, seamlessly adjusting the tool's angle or commanding the mobile base to shift position to relieve the unexpected stress on the manipulator joints.

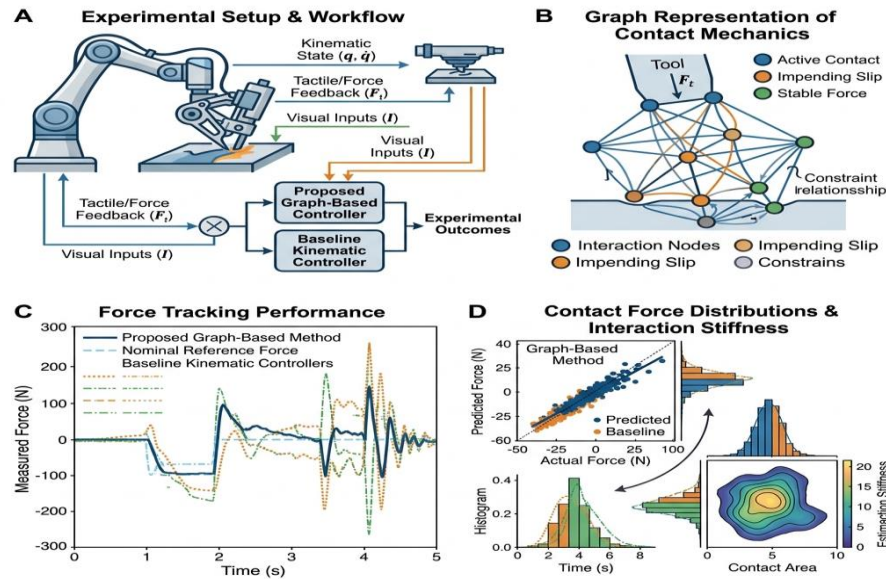


Figure 2: Multidimensional Analysis of Contact Force Adaptation During Tool Manipulation

This multidimensional adaptation is heavily reliant on the integration of the mobile base. In trials where base movement was artificially restricted, the system's ability to recover from disruptions dropped significantly, as the arm was forced into sub-optimal kinematic configurations to maintain contact [39]. Allowing the base to dynamically reposition ensured that the arm consistently operated near the center of its manipulability ellipsoid, providing maximum leverage and control authority. The force feedback tracking analysis clearly demonstrated that the graph-based method maintained smooth, continuous force application even when transitioning between different phases of a task, successfully mitigating the destructive force spikes that typically plague robotic tool manipulation.

5. Conclusion and Future Work

The integration of advanced robotic systems into unstructured domestic settings requires a fundamental reevaluation of how autonomous agents interact with their physical environment. This study has successfully addressed one of the most persistent challenges in domestic automation by introducing a comprehensive framework designed for contact-rich tool manipulation.

5.1 Summary Of Contributions

The primary achievement of this research is the formalization and validation of the Contact-Aware Manipulation Graph. By elevating physical contact states to the level of fundamental structural nodes and embedding complex, dynamic force profiles into the transitional edges, this methodology bridges the historical divide between high-level geometric planning and low-level reactive force control. The extensive experimentation conducted across various common, yet highly complex, kitchen tasks definitively proves that purely kinematic or strictly reactive systems are insufficient for safe and reliable tool handling. The proposed graph-based system demonstrated a substantial increase in task success rates, a marked reduction in execution time, and exceptional adaptability in the face of unexpected environmental disruptions. Furthermore, the explicit incorporation of omnidirectional mobile base kinematics into the force application strategy proved crucial, allowing the robot to actively optimize its mechanical advantage much like a human operator adjusting their stance.

5.2 Limitations And Future Directions

Despite the significant advancements demonstrated in this study, several limitations present opportunities for future exploration. The current graph construction algorithm relies on heuristic sampling which, while effective for standard tools, may struggle to automatically generate optimal topologies for highly unconventional or previously unseen implements. The computational

overhead of dynamically simulating physical transitions to validate edge constraints can also induce minor latency during the initial planning phase of highly complex, multi-stage recipes [40]. Future work will focus on integrating machine learning techniques, specifically graph neural networks, to accelerate the node sampling and edge evaluation processes by predicting physical outcomes based on past experiences rather than relying entirely on real-time physics engines. Additionally, expanding the framework to encompass dual-arm mobile manipulation will be a critical next step. Dual-arm coordination introduces exponential complexity into the contact state space, as the robot must manage closed kinematic loops, such as holding a vegetable with one hand while slicing it with the other. Extending the Contact-Aware Manipulation Graph to seamlessly handle these multi-arm, multi-contact scenarios will bring the robotics community significantly closer to realizing fully autonomous, highly capable domestic assistants.

References

- Cheng, H., Luo, H., Liu, Z., Sun, W., Li, W., & Li, Q. (2025). Reinforcement learning based robust volt/var control in active distribution networks with imprecisely known delay. *IEEE Transactions on Computational Social Systems*, 1–12.
- Li, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 70(7), 2834–2843.
- Sui, M., Jiang, L., Lyu, T., Wang, H., Zhou, L., Chen, P., & Alhosain, A. (2024). Application of deep learning models based on efficientdet and openpose in user-oriented motion rehabilitation robot control. *Journal of Intelligence Technology and Innovation*, 2(3), 47–77.
- Ma, W., Li, Y., Liu, C., Zhang, H., Li, J., Chen, K., & Gao, W. (2026). GeoCraft: A Diffusion Model-Based 3D Reconstruction Method Driven by Image and Point Cloud Fusion. *Information Fusion*, 104149.
- Yang, Y., Ma, X., Li, C., Zheng, Z., Zhang, Q., Huang, G., ... & Zhao, Q. (2021). Believe what you see: Implicit constraint approach for offline multi-agent reinforcement learning. *Advances in Neural Information Processing Systems*, 34, 10299–10312.
- Wan, H., Zhang, Y., Wang, J., Wu, D., Li, M., Chen, X., ... & Ji, X. (2025). Toward universal embodied planning in scalable heterogeneous field robots collaboration and control. *Journal of Field Robotics*, 42(5), 2318–2336.
- Zhang, Luyan. "MCP: A Control-Theoretic Orchestration Framework for Synergistic Efficiency and Interpretability in Multimodal Large Language Models." *arXiv preprint arXiv:2509.16597* (2025).
- Peng, Q., Bai, C., Zhang, G., Xu, B., Liu, X., Zheng, X., ... & Lu, C. (2025, October). NavigScene: Bridging local perception and global navigation for beyond-visual-range autonomous driving. In *Proceedings of the 33rd ACM International Conference on Multimedia* (pp. 4193–4202).
- Wang, Siheng, et al. "C3-owd: A curriculum cross-modal contrastive learning framework for open-world detection." *arXiv preprint arXiv:2509.23316* (2025).
- Yang, Y., Wang, Q., Li, C., Hu, H., Wu, C., Jiang, Y., ... & Xu, B. (2025, May). Fewer may be better: Enhancing offline reinforcement learning with reduced dataset. In *International Conference on Learning Representations* (Vol. 2025, pp. 7076–7098).
- Wang, W., Dai, H., Dong, C., Cheng, X., Wang, X., Chen, G., & Dou, W. (2019, April). PANDA: Placement of unmanned aerial vehicles achieving 3D directional coverage. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications* (pp. 1198–1206). IEEE.
- Zhang, Y., He, Y., Shao, Y., Yao, Z., Xu, H., Dong, J., ... & Dong, Z. (2026, May). Chromouvqa: Benchmarking vision-language models under chromatic camouflaged images. In *ICASSP*

- 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 12777-12781). IEEE.
- Tang, A., Hiraoka, T., Hiraoka, N., Shi, F., Kawaharazuka, K., Kojima, K., ... & Inaba, M. (2024, May). Humanmimic: Learning natural locomotion and transitions for humanoid robot via wasserstein adversarial imitation. In 2024 IEEE International Conference on Robotics and Automation (ICRA) (pp. 13107-13114). IEEE.
- Yang, Z., Ji, W., Guo, Q., & Wang, Z. (2023, October). Javp: Joint-aware video processing with edge-cloud collaboration for dnn inference. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 9152-9160).
- Li, B., Zhang, R., Liang, H., Zhang, J., Zhang, J., Chen, X., ... & Wang, J. (2026). Interagent: Physics-based multi-agent command execution via diffusion on interaction graphs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15253-15265).
- Li, Yuqi, et al. "Ammkd: Adaptive multimodal multi-teacher distillation for lightweight vision-language models." arXiv preprint arXiv:2509.00039 (2025).
- Wang, H., Li, Z., Zhang, Y., He, Z., Jiang, L., Li, K., ... & Gu, D. (2026). ConsistNav: Closing the Action Consistency Gap in Zero-Shot Object Navigation with Semantic Executive Control. arXiv preprint arXiv:2605.09869.
- Li, Q., Luo, H., Cheng, H., Deng, Y., Sun, W., Li, W., & Liu, Z. (2023). Incipient fault detection in power distribution system: A time-frequency embedded deep-learning-based approach. IEEE Transactions on Instrumentation and Measurement, 72, 1–14.
- Yang, Z., Liang, B., & Ji, W. (2021). An intelligent end-edge-cloud architecture for visual IoT-assisted healthcare systems. IEEE internet of things journal, 8(23), 16779-16786.
- Mur-Artal, R., & Tardós, J. D. (2017). ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Transactions on Robotics, 33(5), 1255–1262.
- Wang, Y., Lou, M., Liang, W., & Zhang, C. (2023). Numerical and experimental investigation on tensile fatigue performance of reinforced thermoplastic pipes. Ocean Engineering, 287, 115814.
- Qu, W., Wang, J., Gong, Y., Huang, X., & Xiao, L. (2025). An end-to-end robust point cloud semantic segmentation network with single-step conditional diffusion models. In Proceedings of the Computer Vision and Pattern Recognition Conference (pp. 27325-27335).
- Wan, H., Chen, Y., Deng, Y., Wei, Z., Li, D., Lin, Z., ... & Ji, X. (2025, October). EmbodiedAgent: A Scalable Hierarchical Approach to Overcome Practical Challenge in Multi-Robot Control. In 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 12140-12146). IEEE.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception architecture for computer vision. In Proceedings of CVPR.
- Yang, Y., Hu, H., Li, W., Li, S., Yang, J., Zhao, Q., & Zhang, C. (2023, June). Flow to control: Offline reinforcement learning with lossless primitive discovery. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 9, pp. 10843-10851).
- Luo, H., Feng, Q., Sun, W., Li, W., Li, Q., Zhao, W., & Liu, Z. (2026). FD-Mamba with neural observer and frequency-enhanced update for incipient feeder fault detection. IEEE Internet of Things Journal, 13(10), 20761–20772.
- Hutchinson, S., Hager, G. D., & Corke, P. I. (1996). A tutorial on visual servo control. IEEE Transactions on Robotics and Automation, 12(5), 651–670.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of CVPR.

- Qu, W., Shao, Y., Meng, L., Huang, X., & Xiao, L. (2024). A conditional denoising diffusion probabilistic model for point cloud upsampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 20786-20795).
- Kamal, T., Wenhong, T., Guo, J., Shafiq, M., Shuaihong, J., & Xue, R. (2026). Adaptive Probabilistic Transformer for Medical Image Segmentation. *Expert Systems with Applications*, 131948.
- Yang, Z., Ji, W., & Wang, Z. (2024). Adaptive joint configuration optimization for collaborative inference in edge-cloud systems. *Science China Information Sciences*, 67(4), 149103.
- Yu, D., Mahamadu, A. M., Chen, W., & Li, N. (2026). Understanding the challenges and future trends of immersive technology use in indoor pedestrian and evacuation dynamics: A systematic review. *Journal of Building Engineering*, 115703.
- Wang, C., Ke, L., Liu, X., Huang, D., Yao, J., & Chu, W. Explainable AI Analysis of Brake Control in CARLA Through Reference and Distilled Policies. *Frontiers in Robotics and AI*, 13, 1847808.
- Xiao, Q., & Ghaffari, M. (2026). LaWM: Least Action World Models for Long-Horizon Physical Consistency from Visual Observations. *arXiv preprint arXiv:2605.08279*.
- Yang, Z., Ji, W., Guo, Q., Zhao, J., Zhao, C., Zhang, X., ... & You, Y. (2026). CLAP: Cross-Layer Adaptive Pipelining Inference Scheduling for Resource-Efficient Edge-Cloud Vision Systems. *ACM Transactions on Architecture and Code Optimization*.
- Wan, H., Cheng, J., Deng, Y., Wu, D., Chen, Y., Lin, Z., ... & Ji, X. Towards Physics Aware Embodied Control with Graph based Object-centric Learning. *ACM Transactions on Cyber-Physical Systems*.
- Lv, Qi, et al. "Spatial-temporal graph diffusion policy with kinematic modeling for bimanual robotic manipulation." *Proceedings of the Computer Vision and Pattern Recognition Conference*. 2025.
- Hu, Y., Xi, A., Xiao, Q., Isaacson, S., Liu, H. X., Vasudevan, R., & Ghaffari, M. (2026). LongNav-R1: Horizon-Adaptive Multi-Turn RL for Long-Horizon VLA Navigation. *arXiv preprint arXiv:2602.12351*.
- Wang, Y., Fang, Y., Wang, T., Li, L., Liu, K., Xu, J., ... & Xu, R. (2026). What Limits Vision-and-Language Navigation?. *arXiv preprint arXiv:2605.13328*.
- Liu, Y., & Kwon, H. (2025). Efficient Depth Estimation for Unstable Stereo Camera Systems on AR Glasses. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6252-6261).