How Can Robots Succeed in Unstructured Environments?

Blind Submission

***Abstract*— Roboticists are working towards the realization of autonomous mobile manipulators that can perform useful tasks in human environments. These environments pose a significant challenge because of their complexity and inherent uncertainty. They are characterized by having a high dimensional state space. Consequently, performing tasks in these unstructured environments remains a challenge. Recently, researchers have been successful in developing skills that can handle the complexity of unstructured environments. We hypothesize that those successes are due to a careful implementation that is able to reduce the complexity of the state space, and render the respective problems tractable. In this paper, we analyze this increasing body of literature, in an attempt to extract the common ideas that enable the reduction of the state space. Based on these commonalities, we propose a set of guidelines to facilitate progress for autonomous mobile manipulation in unstructured environments.**

I. INTRODUCTION

The realization of autonomous mobile manipulators will enable a variety of applications with significant societal, scientific, and economical impact. Motivated by the potential of these applications, researchers are beginning to address the challenges posed by unstructured environments (Fig. 1). In this paper, we will attempt to identify common characteristics of successful research efforts. We believe that the resulting insights will contribute to the understanding of the challenges of unstructured environments, and will accelerate the community's progress. Our goal is neither to survey the entire field of autonomous manipulation in unstructured environments nor to identify successful technical approaches and techniques. Instead, we try to uncover guidelines that will help focus our community's research efforts towards the successful deployment of robots in unstructured environments. We hope that this paper serves as a starting point for discussion.

The deployment of autonomous robots in unstructured and dynamic environments poses a number of challenges that cannot easily be addressed by approaches developed for highly controlled environments. In unstructured environments, for example, robots cannot rely on complete knowledge about their surroundings. In fact, perceiving the environment becomes one of the key challenges. Robots have to autonomously and continuously acquire the information necessary to support decision making. Moreover, robots cannot assume that their actions succeed reliably. Instead, they have to continuously monitor their effect on the environment and possibly react to undesired events. In contrast, many existing, well-established techniques in robotics rely on perfect knowledge of the world and perfect control of the environment.

The challenges associated with unstructured environments are a consequence of the high-dimensional state space and the inherent uncertainty in mapping sensory perceptions onto specific states. This fundamental premise will guide our examination of relevant work throughout the remainder of this paper. We believe that the high dimensionality of the state space represents the most fundamental challenge as robots leave the highly controlled environment of the factory floor and enter into unstructured environments.

Fig. 1. Examples of Mobile Manipulators. **Top:** Asimo (Honda), UMan (UMass Amherst), QRIO (Sony) **Bottom:** HRP-3 (Kawada Industries), ARMAR (University of Karlsruhe), WABIAN RIII (Waseda University Tokyo)

The main hypothesis in this paper is that to succeed in unstructured environments robots have to carefully select task-specific features and identify relevant real-world structure to reduce the state space without affecting the performance of the task. In the remainder of this paper, we will analyze existing work in autonomous mobile manipulation. We will show how these example exploit task-specific knowledge and inherent structure to reduce the complexity of problem solving in high dimensional state spaces. Ultimately, we hope, these and related ideas may render autonomous mobile manipulation computationally tractable, even within the high-dimensional state space associated with unstructured environments.

II. ROBOTS IN UNSTRUCTURED ENVIRONMENTS

We now analyze robotic research towards applications in unstructured environments. In our discussion, we will attempt to identify fundamental insights and ideas leveraged to address the problems associated with high-dimensional state spaces. Following our hypothesis that these problems can be addressed using task-specific structure inherent to the physical world, we will group relevant research according to the specific (sub)task they address. We begin with a discussion of robot motion generation, proceed with work in robot perception, and then examine relevant work in manipulation. Finally, we also discuss a task-independent method of providing structure to a robot, namely, through human/robot interaction.

A. Robot Motion

Robots perform tasks by moving through the environment. Given our emphasis on autonomous mobile manipulation, we focus on motions in service of manipulation, i.e., collision-free motion for end-effector placement. The problem of generating such motion is a specific instance of the motion planning problem. Motion planning for robots with many degrees of freedom is provably computationally difficult, even in highly structured environments, due to the high-dimensional configuration space [17].

Unstructured environments impose a number of additional difficulties for motion generation, when compared to the classical motion planning problem [15]. In unstructured environments, a robot can only possess partial knowledge of its surroundings, objects can change their state unbeknownst to the robot, and manipulation tasks may require the end-effector to move on a constrained trajectory rather than simply to reach a specific location. Each of these difficulties make the motion generation problem more difficult. The explicit coordination of planning and sensing necessary to handle dynamic environments further increases the dimensionality of the state space. Furthermore, the more complex task requirements impose stringent requirements for high-frequency feedback.

Existing motion planners make assumptions that are too restrictive for unstructured environments and are too computationally complex to satisfy the feedback requirements. These assumptions and the computational complexity are a consequence of a fundamental premise of motion planning: the assumption that the high-dimensional configuration space is the most suited solution space. Planners following this paradigm use workspace information solely for collision checking. Almost all real-world environments, however, contain significant amount of structure: buildings are divided into hallways, rooms, doors; outdoor environments contains paths, streets, intersections; objects, such as shelves, boxes, tables, chairs have favored approach directions. This information is ignored when a planner exclusively operates in configuration space. As a result, most motion planners have to assume that the environment is perfectly known and that it remains static during planning.

The structure of real-world environments can be used to identify regions of configuration space important to the solution of the planning problem. Compared to configuration space, workspace information is low-dimensional and its connectivity can be determined efficiently. Relevant workspace regions can then be mapped onto small subsets of configuration space. Effectively, the solution to a low-dimensional workspace problem is lifted into high-dimensional configuration space to provide a seed for the planner. The planner can now focus the search in configuration space on small areas and thereby alleviate the computational complexity of planning in a high-dimensional space. This general idea is known as decomposition [23, 3, 4, 24]. It uses an easily computed solution to a low-dimensional problem to simplify the solution to the high-dimensional problem.

The structure of real-world environments can also be used to collapse entire regions of configuration space onto a single state. This can be accomplished with the help of feedback controllers. For the purpose of this discussion, we view controllers as local planners that lead the robot from all state with the domain of attraction to the converged state or attractor. By adequately tiling a high-dimensional space with attractors and associated controllers, planning can be performed in a substantially reduced state space.

The elastic roadmap approach [25] combines the ideas of decomposition and tiling. Based on workspace information, the planner determines an appropriate tiling of configuration space with controllers (the tiling does not necessarily cover the entire configuration space). The tiling defines a discrete roadmap in which attractors are connected if the robot can transition between the respective states using the controller associated with the target state. The elastic roadmap planner can now determine global configuration space connectivity based on a simple graph computed using workspace information. The computation of the elastic roadmap is efficient because it only captures connectivity information and does not require the determination of specific paths that would be invalidated frequently in dynamic environments.

The gained efficiency comes at the cost of completeness guarantees for the planner. To maintain completeness guarantees for motion planning, it may be necessary to plan in configuration space. But even in this case it is possible to leverage the structure of real-world environments. The planning process can be viewed as search in configuration space. During this search, there is a classical trade-off between exploration and exploitation [22]. During search in configuration space, information about the local structure is acquired. This information can be used to deliberately balance exploration and exploitation. When relevant local structure has been identified, it can be used to perform exploitation. When such structure is not present, the planner performs exploration. Such deliberate balancing of exploration and exploitation has been shown to provide substantial performance improvements in motion planning [18].

The ideas of decomposition, tiling, and balancing of exploration and exploitation have proven effective at dealing with high-dimensional planning problems. Each of these ideas leverages information about structure in the environment to alleviate the computational burden associated with high-dimensional state spaces. We stipulate that taking advantage of structure present in the real world is key to achieving the performance and competence required for motion planning that is suited for applications of autonomous mobile manipulation in unstructured environments.

B. Robot Perception

To perform tasks in an environment that is not perfectly controlled and modeled, robots must have adequate perceptual capabilities. The process of perceiving the world and interpreting the acquired information enables robots to understand the state of the world, devise plans to alter the state, and observe the effects of their actions on the world.

The robot's environment can be controlled to varying degrees. In principle, environments that are less constrained are more challenging to perceive. In real-world unstructured and dynamic environments, perception has to address an intractable amount of information acquired by multiple sensor modalities. This sensor data is typically noisy and redundant. Moreover, even without the uncertainty introduced by the sensors, the world itself is often ambiguous: a lemon and a tennis ball may look the same from some perspective, a cup can be invisible if the cabinet's door is shut, and it may be difficult to distinguish between a remote control and a cell phone when they're both facing down. These factors all contribute to the difficulty of perceiving the state of the world.

Perception has been the target of several decades of research. Typical work in this field makes assumptions that are not valid in unstructured and dynamic environments. For example, work in face recognition often makes assumptions about the position and orientation of the person in the image, results in object segmentation are based on the ability to distinguish between object and background based on color differences, and object recognition is often reduced to computing similarities to a limited set of given objects. In unstructured environments, however, position and orientation cannot be controlled, assumptions about colors and shades are difficult to justify, and the range of possible objects the robot can encounter is intractable.

To address perception in unstructured environments, robots must be able to reduce the state space that needs to be analyzed. Sensors can be designed to facilitate some perceptual tasks by reducing uncertainty and therefore decreasing the dimensionality of the state space. For example, to compute the distance to objects in the environment, robots need to associate depth with visual information. This is typically done by using a stereo vision system and solving the correspondence problem between two static 2D images. Solving the correspondence problem, however, is difficult due to noise, multiple possible matches, and uncertainty in camera calibration. In [14] a system capable of capturing at least three viewpoints in a single image is introduced. This reduces the state space by collapsing a multi-sensor system down to one sensor.

In an unstructured environment, object recognition has been proven to be very difficult. With large amounts of sensor information and high variation within objects of the same category, object recognition is a high dimensional problem. Despite these difficulties, objects in the same category do share common characteristics. Using this insight, robots can to focus their attention to only a small subset of the state space that contains the most relevant features for classification. In face recognition, for example, specific relationships exist between the location of features such as eyes, nose, and mouth. In [10], this structure, which underlies the entire category of faces, is being exploited to increase the accuracy of pose estimation of faces. As a result, the dimensionality of the state space is dramatically reduced, and face recognition becomes tractable.

Obstacle avoidance is another hard perceptual problem. In order to avoid collisions robots must solve the high dimensional problem of distinguishing between objects and free-space, calculating how far away objects are, figuring out how they're positioned, etc. This large state space can be reduced by leveraging relevant knowledge about how the world behaves. For example, when a robot moves, optical flow is created by obstacles but not by free-space. This insight is used in [9] to create an insect-inspired vision system capable of measuring optical flow and turning away from obstacles. This reduces the state space by focusing only on features that are necessary for avoiding obstacles. Similarly, in [16] they avoid the common approach of calculating complete depth maps and instead build a learning algorithm to calculate steering angles directly from raw images.

Perceptual problems pose a significant challenge for robots in unstructured environment because of their high-dimensional state spaces. Designing sophisticated hardware, identifying common object characteristics and focusing on the goal are examples of approaches that deal with the complexity of perceptual problems. These techniques take advantage of existing structure in the world to reduce the state space, and therefore enable robots to solve perceptual tasks in unstructured environments.

C. Robot Manipulation and Grasping

Object manipulation requires both reliable motion capabilities and adequate perceptual capabilities. It is a prerequisite of many important applications for robotics such as planetary exploration, elder care, flexible manufacturing and construction in collaboration with human experts. The problem of manipulating the environment includes moving objects of varying dimensions by pushing or pulling, and prehensile and non-prehensile grasping of smaller objects. Manipulation is very challenging, even in structured environments, due to the complexity of the associated state space. This state space include the appearance, position, dimensions, and weight of objects in the scene, as well as many other relevant features indicating where to push or grasp, and how much force to apply. The addition of a rich set of actions further increases the complexity, as robots need to choose between many possible actions and determine the appropriate parameterizations for controllers.

Manipulation in unstructured environments faces several difficulties that are not present in structured environments. In unstructured environments, object properties required for manipulation cannot be known *a priori*. Information about objects has to be acquired through sensors, but those are

Fig. 2. Researchers often assume that *a priori* models, such as the CAD model of the kitchen on the left, are available. In practice, those models are usually difficult to obtain. Also, such environments are constantly changing and look more like the kitchen on the right.

often ambiguous, introduce uncertainty, and provide redundant information with respect to the manipulation task. Furthermore, manipulation in unstructured and dynamic environments typically requires responding in a timely fashion to a rapidly changing world.

Researchers typically make assumptions to reduce the complexity of manipulation in unstructured environments. For example, it is often assumed that complete models of objects in the environment are available *a priori* or can be acquired through sensors, and that the environment remains static during the interaction. In practice, it is impossible to provide manipulation with complete *a priori* models of the real world (Fig. 2). However, perfect models are not a prerequisite for successful manipulation in unstructured environments. Manipulation can be guided by the structure that exists in the world and which is oftentimes easy to perceive. By leveraging this structure, the complexity of manipulation in unstructured environments decreases significantly. For example, with the insight that most cups, coffee mugs, and teapots have handles, grasping such objects becomes simpler despite the absence of perfect models. Similarly, understanding the intrinsic degrees of freedom of objects such as scissors, staplers, doors, and books can also reduce the complexity of manipulation in unstructured environments.

In order to grasp arbitrary objects in unstructured environments robots have to search a very high dimensional state space. Grasping many real-world objects, however, requires considering only a small subset of that state space. When tasked with grasping a specific object, robots can focus their efforts on the relevant subset of the state space, thus simplifying the grasping problem. For example, grasping small rectangular objects can be accomplished by pinching, and does not require actuating many of the hand's degrees-of-freedom. Within the context of a specific grasping task, robots can use hardware to further decrease grasping's complexity. In [6], for instance, careful selection of joint compliance and coupling schemes enables grasping a large variety of real-world objects by actuating only a single degree-of-freedom.

Grasping can also be simplified by exploiting the structure that is inherent to human environments. Most objects in our world are designed to perform some function, and are intended to be used by humans. As a result, many real-world objects share common traits alluding to their intended use. By focusing on these task-related object properties, the complexity of grasping is reduced. For example, in [19] visual data is analyzed to identify a few points that correspond to good locations at which to grasp an object. Because grasping features are similar across multiple objects, robots can be trained to identify them. Consequently, the state space that needs to be explored in order to grasp objects is significantly reduced.

Perceiving structure in the world can assist manipulation. However, acquiring information about the state of the world can be very challenging in unstructured environments: objects may be partially obstructed, lighting conditions may be poor, and the purpose of an object may be difficult to perceive. This ambiguity in sensor information increases uncertainty about the world, and therefore increases the size of the state space. Closely integrating manipulation and perception can decrease the complexity of the state space. Manipulation can augment the robot's ability to perceive structure in the world, which in turn can benefit manipulation. Through interaction, robots can remove obstructions, reposition objects to improve lighting conditions and view point. Interaction with objects can also be used to facilitate the perception of kinematic structure [12, 11], which is then used to enable purposeful manipulation. In [5] interaction is used to determine kinematic and dynamic properties, which are then exploited to predict future interaction with objects. Interaction can also be used to generate motion which facilitates object segmentation [8, 13]. The integration of action and perception thus reduces complexity and renders manipulation in unstructured environments feasible.

Manipulation can ameliorate perception in unstructured environments. The converse is also true: manipulation depends on adequate perception. And yet, it can be very difficult for manipulation to use the right perceptual information. The complexity of unstructured environments results in an intractable amount of sensor data available for manipulation. This data is mostly redundant and irrelevant for the manipulation task at hand. By identifying the task's objective, the robot can focus its attention to the most task-relevant subset of its perceptual data. As a result, the state of the world is described with respect to the manipulation task, which decreases the size of the state space. For example, in [21] high dimensional streaming visual data is available for learning tool affordances. By focusing on motions that occur next to the end-effector, only a small portion of the visual information needs to be considered. Consequently, the robot learns tool affordances in a much lower dimensional and therefore tractable state space.

Robots have to solve high-dimensional problems in order to manipulate and grasp objects in unstructured environments. Techniques such as crossing boundaries between action and perception, exploiting *a priori* knowledge about objects in human environments, and focusing on task-specific perceptual features can be used to reduce the dimensionality of manipulation and grasping. As the state space is reduced, problems become tractable, thus enabling robots to perform grasping and manipulation in unstructured environments.

D. Human-Robot Interaction

Communication with humans is another resource robots can exploit to reduce the complexity of unstructured environments. Humans can point to interesting features, teach new skills by demonstration, or use language to transfer knowledge. Moreover, many real world tasks require cooperation between humans and robots.

Work towards understanding human communication and natural language is usually focused on analyzing text and speech. Also, researchers typically limit the domain to include only specific topics [20]. Human communication, however, includes more than just verbal or textual communication. It involves gestures and other actions with physical manifestation. Moreover, it is impractical to limit the domain of communication in unstructured and dynamic environments because those environments are, by definition, high-dimensional, rapidly changing, and unknown *a priori*.

In order to facilitate efficient Human-Robot Interaction, the dimensionality of the state-space has to be reduced without limiting the robot's performance. Robots can leverage the structure that exists in the world and consider the goal they are trying to achieve to focus their efforts on parts of the state space that are most relevant. For example, eye contact can be used to understand the intended audience of verbal instructions. Hand gestures can narrow the set of possible objects to which a person may refer. Also, the context of the task the robot or the person is performing can limit the objects and concepts included in the conversation, thus reducing the complexity of the state space.

Teamwork in the real-world often involves teaching and learning new skills. Communication can be used for teaching. For example, a skilled human worker can teach by demonstration or explain using verbal communication and hand gestures. For robots, learning new skills in unstructured environments requires reasoning in an intractable and inherently ambiguous state space: What object exactly is the person pointing to? Which one the "round" objects am I supposed to grasp? And what exactly does "come closer to me" mean? Humans often rely on expressive feedback for communication, and are experts in interpreting it. In [1, 2], expressive feedback is used during teaching. Robots express frustration, confusion and curiosity via facial expressions. The human teacher can easily interpret those cues and use them to accelerate and focus the teaching session.

Human-Robot cooperation in performing tasks requires communicating about objects, tools, and goals. Many tasks require the transition of objects between a person and a robotic collaborator. Using verbal communication to instruct the robot is challenging because of the complexity of the environment and the robot's mechanism: the robot has to decide where to position its hand, in what orientation, how to preshape its fingers, and how much force to apply. In [7] a human collaborates with a robot in the task of passing objects between them and placing them on a shelf. With the insight that humans usually hand objects in a configuration that is easy to grasp, robot grasping has to consider only a subset of the the state space related to grasping. Also, by considering the task, the robot only needs to track the human's hand to learn about the position of the object, thus decreasing the state space that needs to be explored. As a result of decreasing the dimensionality of the problem, both grasping and communication about grasping become tractable.

Creating successful interactions between humans and robots is difficult because of the high dimensionality imposed by communication. By using hand and eye cues and expressive communication, humans can direct a robot's focus toward relevant areas of the state space. This focuses attention on the task in order to reduce the overall size of the state space and makes Human-Robot Interaction possible in unstructured environments.

III. Conclusion

We began our discussion by hypothesizing that to succeed in unstructured environments robots have to carefully select task-specific features and identify relevant real-world structure to reduce the state space without affecting the performance of the task. We confirmed this hypothesis by analyzing successful examples of applications in motion planning, perception, manipulation, grasping and Human-Robot Interaction in unstructured and dynamic environments. Each one of the examples we discussed exploited structure present in the environment to reduce the size of the relevant state space. As a result, they were able to successfully solve complex tasks, despite the apparent complexity of the state space.

Encouraged by this positive evidence, we propose two guidelines that we believe will enable robots to uncover structure and exploit it. We believe that additional guidelines may be found. Our ultimate goal is to answer the question: "How can robots succeed in unstructured environments?".

To reduce the complexity of the state space, robots must exploit task-relevant structure. However, uncovering this structure may not be possible without crossing the boundaries of different technical areas that have governed robotics research for the last few decades. Some structure can only be revealed through the conjunction of methods from two or more technical areas. For example, active segmentation [8, 13] uses manipulation and vision to generate motion — the most pertinent signal for segmentation. Or, interactive perception [12] uses manipulation and vision to identify kinematic structure. In both cases, neither vision nor manipulation alone can reliably solve the problem. Our first guideline is therefore:

- To devise competent and robust skills for unstructured environments, skills must be task-centric and should consider all technical areas relevant for implementing the skill.

To make further progress, the field of autonomous mobile manipulation has to address complexity incrementally. Simple skills, such as the ones discussed in section II, provide a grounding for more abstract representations. Those representations, in turn, can reduce the state space for higher-level skills. As more and more skills become available, the dependencies among skills becomes more complex. The challenges thus becomes to resolve the complex dependencies among skills and discover which skills can facilitate the state space reduction of more complex skills. Our second guideline is therefore:

- Robust, autonomous, sophisticated behavior of embodied agents in unstructured environments will come about by a careful bottom-up development/design/learning of elementary to complex skills.

We believe that these two guidelines demonstrate how by leveraging the right structure for the right problem and building on top of other skills, high-level behavior becomes possible. With enough of these building blocks, we hope that robots will be able to perform tasks in high dimensional, unstructured, and dynamic environments.

Acknowledgments

We gratefully acknowledge support by the National Science Foundation (NSF) under grants CNS-0454074, IIS-0545934, CNS-0552319, CNS-0647132, and CNS-0812986.

References

[1] C. Breazeal. *Designing Sociable Robots (Intelligent Robotics and Autonomous Agents).* The MIT Press, September 2004.

[2] C. Breazeal, M. Berlin, A. Brooks, J. Gray, and A. L. Thomaz. Using perspective taking to learn from ambiguous demonstrations. *Robotics and Autonomous Systems*, 54(5):385–393, May 2006.

[3] O. Brock and L. E. Kavraki. Decomposition-based motion planning: Towards real-time planning for robots with many degrees of freedom. Technical Report 00-367, Rice University, 2000.

[4] O. Brock and L. E. Kavraki. Decomposition-based motion planning: A framework for real-time motion planning in high-dimensional configuration spaces. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1469–1474, Seoul, Korea, 2001.

[5] A. D. Christiansen, M. T. Mason, and T. M. Mitchell. Learning reliable manipulation strategies without initial physical models. *Robotics and Autonomous Systems*, 8(1-2):7–18, 1991.

[6] A. M. Dollar and R. D. Howe. Joint coupling and actuation design of underactuated hands for unstructured environments. *IEEE Transactions on Robotics*, 2008.

[7] A. Edsinger and C. Kemp. Human-robot interaction for cooperative manipulation: Handing objects to one another. In *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, 2007.

[8] P. Fitzpatrick and G. Metta. Grounding vision through experimental manipulation. *Philosophical Transactions of the Royal Society: Mathematical, Physical, and Engineering Sciences*, 361(1811):2165–2185, 2003.

[9] W. E. Green, P. Y. Oh, and G. Barrows. Flying insect inspired vision for autonomous aerial robot maneuvers in near-earth environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2004.

[10] G. B. Huang, M. Narayana, and E. Learned-Miller. Towards unconstrained face recognition. In *In The Sixth IEEE Computer Society Workshop on Perceptual Organization in Computer Vision IEEE CVPR*, 2008.

[11] D. Katz and O. Brock. Grounding relational reinforcement learning for manipulation in unstructured environments. In *Proceedings of Robotics: Science and Systems (RSS)*, Zurich, Switzerland, 2008.

[12] D. Katz and O. Brock. Manipulating articulated objects with interactive perception. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, USA, 2008.

[13] J. Kenney, T. Buckley, and O. Brock. Interactive segmentation for manipulation in unstructured environments. In *IROS*, 2008. under review.

[14] S. Kuthirummal and S. K. Nayar. Multiview radial catadioptric imaging for scene capture. In *Proceedings of ACM Sigraph*, 2006.

[15] S. M. LaValle. *Planning Algorithms.* Cambridge University Press, 2006.

[16] Y. LeCun, U. Muller, J. Ben, E. Cosatto, and B. Flepp. Off-road obstacle avoidance through end-to-end learning. In *In Proceedings of Conference on Neural Information Processing Systems (NIPS)*, 2005.

[17] J. H. Reif. Complexity of the movers problem and generalizations. In *Proc. of the Symposium on Foundations of Computer Science*, page 421427, 1979.

[18] M. Rickert, O. Brock, and A. Knoll. Balancing exploration and exploitation in motion planning. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, USA, 2008.

[19] A. Saxena, J. Driemeyer, J. Kearns, C. Osondu, and A. Y. Ng. Learning to grasp novel objects using vision. In *Proceedings of the International Symposium on Experimental Robotics*, 2006.

[20] L. Steels and F. Kaplan. Bootstrapping grounded word semantics. In *Linguistic evolution through language acquisition: formal and computational models*, 2002.

[21] A. Stoytchev. Behavior-grounded representation of tool affordances. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 3071–3076, Barcelona, Spain, 2005.

[22] S. Thrun. The role of exploration in learning control. In *Handbook for Intelligent Control: Neural, Fuzzy and Adaptive Approaches*, Florence, Kentucky, 1992. Van Nostrand Reinhold.

[23] Y. Yang and O. Brock. Viewing motion planning as disassmbly: A decomposition-based approach for non-stationary robots. Technical Report 04-108, Computer Science Department, University of Massachusetts Amherst, December 2004.

[24] Y. Yang and O. Brock. Efficient motion planning based on disassembly. In *Proceedings of Robotics: Science and Systems (RSS)*, pages 97–104, Cambridge, USA, June 2005.

[25] Y. Yang and O. Brock. Elastic Roadmaps: Globally task-consistent motion for autonomous mobile manipulation. In *Proceedings of Robotics: Science and Systems (RSS)*, pages 279–286, Philadelphia, USA, August 2006.