Discovery, Segmentation and Reactive Grasping of Unknown Objects

David Schiebener, Julian Schill and Tamim Asfour
Karlsruhe Institute of Technology (KIT)
Institute for Anthropomatics (IFA)

{schiebener,schill,asfour}@kit.edu

***Abstract*—Learning the visual appearance and physical properties of unknown objects is an important capability for humanoid robots that are supposed to be working in an open environment. We present an approach that enables a robot to discover new, unknown objects, segment them from the background and grasp them. This gives the robot full control over the object and allows its further multimodal exploration.**

In order to discover an unknown object in a cluttered scene and segment it from the (likewise unknown) background, we generate hypotheses based on visual input and try to verify one of them by pushing it. The induced motion solves visual ambiguities and allows a clear object-background segmentation.

The acquired estimation of the object position and extent allows the robot to try grasping it. As we do not have exact shape information, we apply a reactive grasping approach. Based on tactile sensor feedback of the hand, we execute correction movements until the object can be grasped in a stable manner.

I. Introduction

Humanoid robots that are supposed to work in a human-centered, uncontrolled environment need to be flexible and adaptive in many different respects. One key prerequisite to their usefulness in everyday tasks is the capability to manipulate and interact with common household objects. In particular, they need to be able to deal with objects that they have not encountered before. This requires the ability to autonomously acquire knowledge about at least some important properties like the visual appearance, and sufficient information about their shape and extent to manipulate them.

To this end, several problems have to be solved: First, the unknown object must be discovered. Second, it must be segmented from the rest of the environment. This will already allow the robot to learn its visual appearance, at least from one side. The third step is to grasp the unknown object. Once this has been done, the robot has full control over it and can explore it completely, learning all its visual and physical properties. Thus, it will have become a known object allowing all kinds of interactions that the robot is able to perform.

In this paper, we describe our integrated approach that enables the humanoid robot ARMAR-III [1] to solve these three problems consecutively, i.e. discover, segment and grasp completely unknown objects in full autonomy, using its visual, manipulative and haptic capabilities. First, the robot generates hypotheses of possible objects based on the camera images of its active stereo vision system. To test these hypotheses and reliably segment them from the background, it then tries to move them. If thereby an object is discovered, its motion relative to the rest of the scene is exploited to segment it exactly and completely. Based on the now known object position and an estimation of its extent, we initiate an attempt to grasp it. The grasping itself is realized in a reactive manner using haptic feedback, allowing it to correct the grasp and adapt it to the actual shape of the object. Figure 2 gives a schematic overview of our approach.

The following section will give a short synopsis of previous research work concerning the subproblems of the our task. Section III describes our interactive approach to discovering and segmenting unknown objects, and section IV the reactive grasping exploiting the knowledge that has been obtained up to that point. In section V we test the validity of our approach with experiments on ARMAR-III, and finish with a conclusion and a perspective for future work in section VI.

II. Related Work

Some research efforts have been focused on learning the visual appearance of beforehand unknown objects. Besides creating an appropriate visual description, the central problem here is the segmentation of the unknown object from the rest of the observed scene. In [2], the authors

Fig. 1: The humanoid robot ARMAR-III in a kitchen, pushing an unknown object to segment it from the rest of the scene.

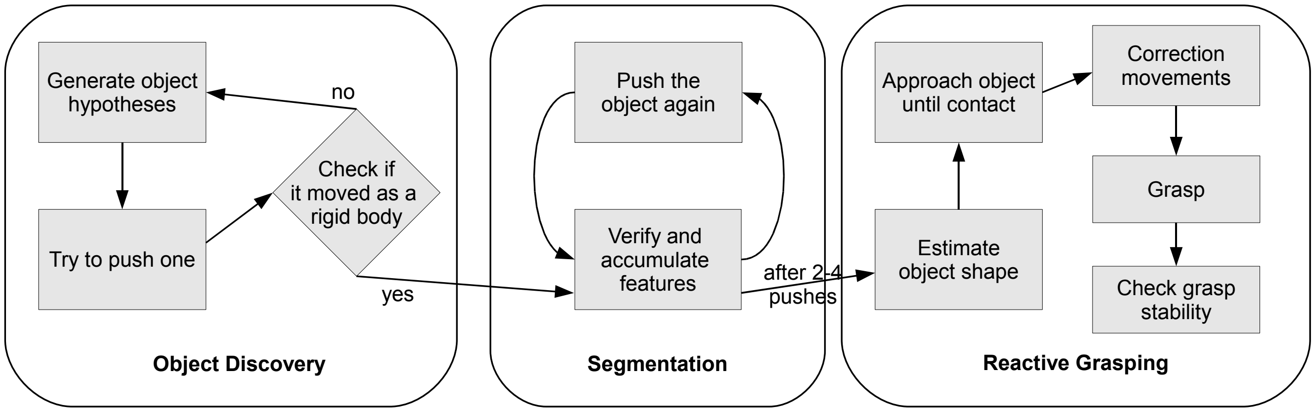

Fig. 2: System overview showing the different stages of our approach for discovering, segmenting and grasping completely unknown objects.

demonstrate that pushing an object does in principle allow such segmentation. This idea has been followed in [3] where object hypotheses are generated by detecting planar surfaces amongst local features. These hypotheses are pushed and the features tracked. In [4] and [5], this has been extended to allow for almost arbitrary shapes of the initial hypotheses and different types of local features, and the completely autonomous learning of object representations has been demonstrated on different humanoid platforms. The basic idea from [2] has also been pursued in [6], assuming the initial position of the object is known, and in [7], which is restricted to symmetric objects.

Another approach to the problem of segmentation and learning of unknown objects is to assume that the robot has already grasped them, so they can be moved in front of the camera. This idea has been pursued e.g. in [8], [9] and [10]. The method presented in [11] aims at the same objective as we do; the main difference is that they use a 3-D range sensor to detect the novel object assuming that it is standing solitary on a flat table, and grasp it based on the resulting 3-D point cloud. The focus of that work is rather on learning a large variety of object properties (visual appearance, haptic properties, sound) once it has been grasped by the robot.

Several approaches have been tried in the field of reactive grasping using different sensors to adjust the grasp trajectory during grasp execution. Hsiao et al. [12] presented an approach using IR proximity sensors to adjust the position of the hand to the object, while other approaches use haptic information, such as tactile sensors [13] or tactile and force information [14] to adapt the grasp. These methods increase the success rate of the grasp by dealing with the uncertainties. In this work we use a similar approach.

To ensure that a grasp was successful, a grasp stability check can be executed after execution of a grasping attempt. There have been several approaches for analyzing the grasp stability in an analytical way, which are reviewed in detail in [15]. Other approaches are using machine learning techniques to estimate the stability of a grasp [16]. In this work we use the method presented in [17], which extends this approach to check the grasp stability continuously instead of just checking it once after the hand is closed.

III. Object Discovery and Segmentation

A. *Hypothesis generation*

The question of what exactly is an object is subject to ongoing discussions since many centuries [18], and no satisfying definition has been generally agreed on up until today. Throughout this work, we pragmatically define an object as a rigid coherent physical body that is not fixed to anything else and has an appropriate size and weight so that a humanoid robot can move it.

The first step in discovering unknown objects is to hypothesize which parts of an observed scene might constitute one. We do that based on the images provided by the stereo vision system in the head of ARMAR-III. We determine Harris Interest Points [19] and Maximally Stable Extremal Color Regions [20] [21] as (relatively) local features in the whole camera images and, based on a stereo calibration, use epipolar geometry [22] to reliably match them in the stereo images and determine their 3D positions. Thus, we obtain a set of 3D points covering most parts of the observed scene.

Based on these points, we generate hypotheses of possible objects. We hereby follow the intuition that points belonging to the same object will lie on a (at least partly) smooth surface and be relatively close to each other. Therefore, we search for regular surfaces amongst the 3D points. As we have to expect that only a small subset of all points belongs to one regularly shaped surface patch, we use Random Sample Consensus (RANSAC) [23] to find planes, spheres and cylinders. We search for all three shapes simultaneously, select the one that contains the maximal number of points, and remove it from the set. We then search again amongst the residual points until no more surface containing at least a minimal number of points can be found (12 points in our experiments). Details on the implementation of RANSAC for planes, spheres and cylinders can be found in [4] and [5]. To

Fig. 3: Initial object hypotheses that have been generated based on the stereo images provided by the cameras in the head of ARMAR-III. One of these hypotheses is then interactively validated by pushing it and analyzing its motion. The motion allows to determine which visual features belong to the object and which do not. In the course of 3 pushes, a complete visual description of the visible object surface is obtained.

avoid subsuming several objects into the same hypothesis, we apply x-means [24], a variant of k-means that also estimates the parameter k, to the resulting point sets.

This way, we obtain all regular surface patches in the set of 3D points, and consider each of them a possible object. In addition, we cluster the remaining points using x-means. If we find a cluster that contains many points within a compact area, we consider it a hypothesis too, so we can also detect objects with very irregular surfaces but a sufficient number of local features. The first image of figure 3 shows hypotheses that were generated by our approach; each consists of a set of local visual features and their 3D positions.

B. *Object segmentation by pushing*

Having obtained hypotheses of possible objects, we need to verify if one of them actually constitutes an object or is at least part of one. Although a lot of research effort has been spent on the segmentation of unknown objects in camera images, it seems that this problem can not reliably be solved by passive observation only. Segmentation becomes fairly easy when the object moves relative to the background, therefore we let the robot try to move the hypothetical object that is to be examined. The easiest and least fault-sensitive way to move an object whose shape and physical properties are unknown is to just push it. By inducing motion to the object, we are able to segment it from the background.

When pushing the object, we want to make sure that on the one hand it will move significantly enough to allow clear segmentation from the rest of the scene. On the other hand, its visual appearance should not change too much to allow relocalizing its visual features, and it should stay in an area where it is visible for the cameras and well reachable for the hands to allow for subsequent pushes later (see next subsection). We therefore define a central point in front of the robot towards which the object will be pushed. By setting the intended length of the push to a fixed value (10cm in our experiments), we make sure that the object will not converge towards this central point but rather “oscillate” around it, which guarantees sufficient motion for object segmentation.

Initially, the robot hands are outside the field of view of the cameras to avoid unnecessary occlusions. We choose the hand for pushing which is better suited for the intended trajectory by inspecting the versatility along the planned path using the reachability analysis of the cartesian space presented in [25]. The hand approaches the object on a trajectory significantly above it to avoid collisions with other objects and is then lowered to the object’s height besides it. The push towards the central point is executed, then the hand is lifted again and moves back to its initial pose. Unexpected collisions of the hand with obstacles are detected using the force-torque sensor in the wrist of ARMAR-III. If there are other objects close to the one the robot wants to push, the hand may collide with one of them during the lowering phase besides the object. In this case, it is lifted again and lowered a bit closer to the estimated object position. If there is another collision, the robot tries a closer position again etc. If the hand comes very close to the object position, it is lowered over the estimated object center until contact and a sliding movement is executed instead of a push. With this flexible reactive approach we are able to move the object even in very cluttered scenes.

After trying to push the object, we check all original hypotheses for having moved as a rigid body, i.e. having undergone a translation and rotation in 3D space. To this end, we determine correspondences for all features between the camera images before and after the push. We use SIFT descriptors for the Harris Interest Points and a compact color and shape descriptor for the color MSERs. A transformation in 3D is (over)determined by 3 point correspondences, therefore we can apply RANSAC in a straightforward manner to estimate the transformation that each hypothesis has undergone. If a hypothesis moved significantly, i.e. the weighted sum of the length of the translation and the angle of rotation is above a certain threshold (an eqivalent of 3cm or 15 in our experiments), we consider it a validated object. If none of the hypotheses has moved, we generate new hypotheses and try to push one of them again.

C. *Feature verification and accumulation*

We can now choose one of the validated object hypotheses for further investigation. We have to assume that our initial hypotheses did not correspond exactly to an actual object, but that it may have included features that do not belong to the underlying object and, very probably, that significant parts of the object are not included in the hypothesis yet. The induced motion allows us to correct that: By determining correspondences for all point and region features between the camera images before and after the push, we can now check for each of them if they moved coherently with the object. Those features in the hypothesis that did not are rejected. Other features that did not belong to the original hypothesis but moved concurrently with it are added as possibly belonging to it. Now the robot pushes the object again to verify or discard these new candidate features.

This step of pushing the object and adding and verifying features belonging to it can be repeated several times. Thus, we obtain a complete set of all visible point and region features as can be seen in figure 3. Based on these confirmed features, we can learn a visual description of the object as shown in [4] [5]. This allows object recognition and localization, at least from the sides of the object that become visible during the repeated pushing. But to get a full multiview description, and also to obtain other information about the object like its weight, rigidity etc, we need to grasp it.

IV. Reactive Grasping

A. Determination of hand orientation

Classical grasping approaches require a complete and exact 3D model of the object that is to be grasped. Since we do not have that, we resort to a reactive grasping approach instead. The fact that we already know that there actually is an object, where it is, and which parts of the image belong to it, gives us a significant head start compared to completely blind grasping. Based on the positions of the features we know to belong to the object, we calculate its center and principal axes. We can not determine the depth of the object in the viewing direction, but we can estimate the extent in the two dimensions parallel to the image plane. As this is approximately the direction from which we will approach the object when trying to grasp it, this information is sufficient. In fact, we essentially need the position of the object and the direction of its maximal extent with relation to the approach direction, so that we can grasp it around the shorter side.

To further support the grasping attempt, in the end of the pushing stage we modify the pushing direction in such a way that, instead of pushing the object towards a central point, we move it to a position on the right side in front of the robot that facilitates grasping. Once we are there, we estimate the object position and the orientation of its largest extent and use that information to initiate the reactive grasping.

B. Correction movements

After estimating the object position and pose, we approach the position of the object from the top with the robot hand. We use the methods presented in the following subsection (IV-C) to detect a contact with the object. According to the estimated contact point, a correction movement towards the

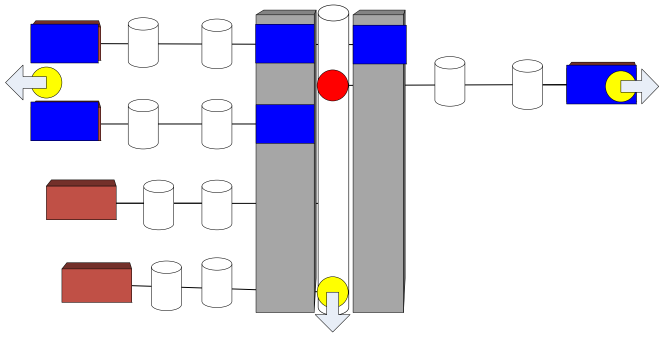

Fig. 4: Kinematic overview of the hand. On contact on the yellow contact points correction movements are executed, the hand is closed when a contact is detected on the red point. The positions of the tactile sensors are shown as blue rectangles.

contact point is executed. Therefore the robot lifts the hand, corrects the position towards the assumed contact point and starts to lower the hand again. This is repeated until a contact in the palm is detected after which the hand is closed and thus an attempt to grasp the object is executed. After closing the hand the stability of the grasp is checked with an SVM classifier using the proprioceptive and tactile information, which is presented in [17].

C. Contact detection

To detect a contact of the hand with an object, we use different sensors: Tactile sensors in the finger tips and the palm, a 6DoF force/torque sensor mounted in the wrist and the force measurement of the hand controller, which uses the joint encoder and the air pressure to estimate the torque of the finger joint actuators [26]. With these sensors we can detect four virtual contact points: One between the finger tips of the index and the middle finger, one at the finger tip of the thumb, one at the inner side of the palm and one at the outer side of the palm. Figure 4 shows the positions of these points. For detecting a contact with the tactile sensors, we simply sum up the values of the taxels and threshold the result. A contact can also be detected by the force/torque sensors, when the force measured in direction of the arm is above a specified value. The contact point is then determined by evaluation of the measured torques. The torque direction points to the most likely point of contact. If the torques are too small or non-existent, we assume that the contact took place in the palm of the hand. To detect a contact with the force measurement of the hand controller, the hand has to be slightly closed. As the pneumatic actuators are compliant, the finger will be bent back when a contact happens, resulting in a higher air pressure and a decreased joint angle, which can be interpreted as an increased torque in the actuator. By summing up the measured torques of each finger, an approximation of the force pushing against the finger can be made, which allows us to detect a contact. Because of the high resolution of the joint encoders, this measurement method is very sensitive, especially when the contact takes place at one of the finger tips.

By combining all these measurement methods, we can detect contacts reliably almost everywhere on the hand, even if the object is not touched with one of the tactile sensor pads, although the location of the contact can not be determined very accurately then.

V. Experimental Evaluation

We evaluated the system on the humanoid robot ARMAR-IIIb [1], which is equipped with a stereo camera system, two arms with 7 DoF each and two pneumatic actuated anthropomorphic robot hands [27]. The right hand is equipped with joint encoders and air pressure sensors which allows us to control the joints with a force position controller [26]. Additionally we mounted six tactile pressure sensors from Weiss Robotics [28] on the finger tips of the thumb, index and middle finger and in the palm.

For the experiments, we positioned the robot in front of a table on which we placed several objects. As outlined in figure 2, the robot first looks at the table and creates object hypotheses based on the stereo images taken with the cameras in the head. One of the potential objects is chosen for verification and pushed. When the hypothesis has been validated, it is pushed again to accumulate all visible features belonging to it. During the later pushes we also move it towards a position that is well suited for the grasping attempt. Using the 3D positions of the features, we estimate the object position and the direction of its maximal extent.

The robot moves its right hand to a position above the verified points and turns the hand according to the orientation of the main object axis. Then it starts to move the hand downwards towards the object, while continously checking for a contact with the sensors. If a contact is detected, the movement is stopped and a correction movement is executed, after which the robot starts moving the hand downwards again. This is repeated until a contact in the palm is detected. After that the hand is closed and the stability of the grasp is determined. If the grasp is stable the robot lifts the object and moves it to a position above a box, where it drops the object by opening the hand again. If the grasp is detected as not stable, the robot opens the hand again, and starts from the beginning. Figure 5 shows ARMAR-IIIb executing the different stages of our approach. The first three images show the object discovery and segmentation through pushing, the other images the reactive grasping with correction movements.

We evaluated our approach using 30 different objects. For each trial, 3-10 of them were arranged on a table in front of the robot. The main criterion for assessing our approach is arguably the success rate of the overall process of detecting, segmenting and grasping unknown objects. As it turned out, there are essentially two points in the process where it may fail. The first one is the verification of an object hypothesis through pushing. Here, it sometimes happens that after the push the object can not be relocated because it has moved (in particular: turned) so much that a large part of its visual features can not be rediscovered. This happens

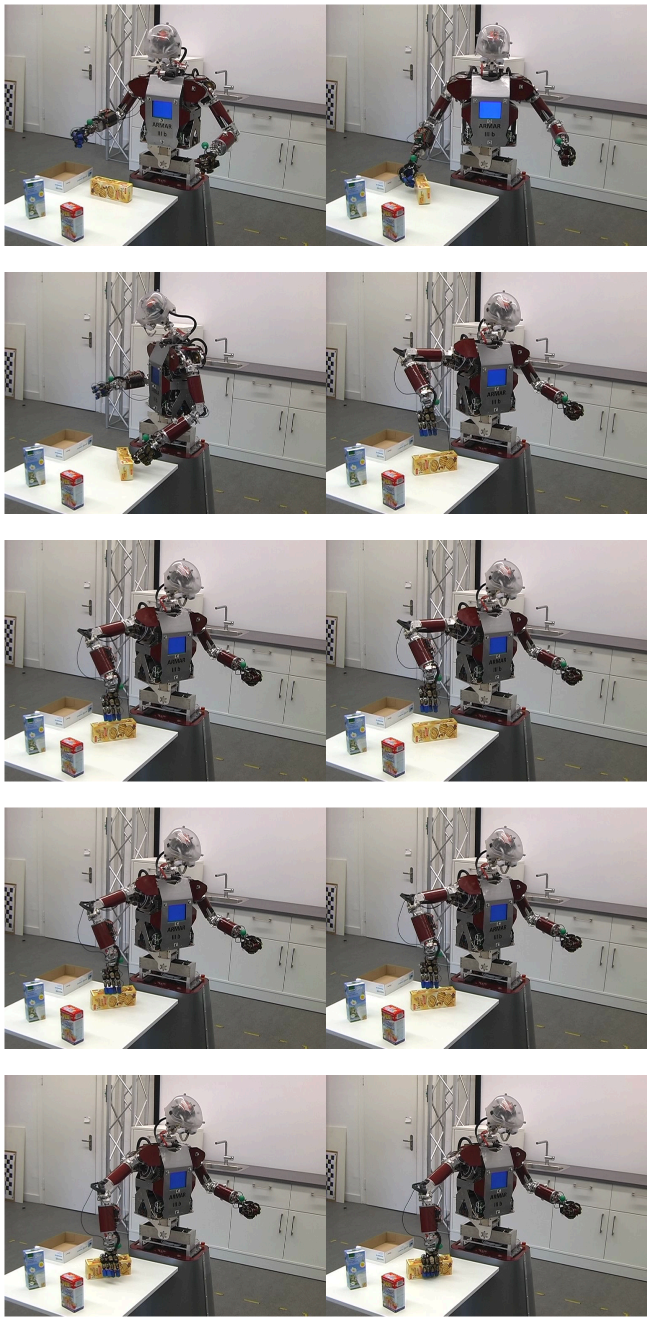

Fig. 5: The robot ARMAR-IIIb during the different stages of our approach: First, object hypotheses are generated based on stereo images (picture 1). One of these hypotheses is then confirmed, corrected and completed through several pushes that cause it to move and allow clear object-background segmentation (pictures 2-3). Based on the estimated object position and extent, a grasp attempt is initiated (picture 4). The robot reactively grasps the object using haptic feedback from different sensors in its hand and wrist (pictures 5-10).

in about 10-20% of the cases, depending on the number and manifestness of the local visual features on the object surface. This problem seems to be unavoidable as initially we have only a very rough and probably incomplete hypothesis about the object's position and extent, therefore the pushing is inevitably rather uncontrolled. But as this failure is in any case detected by the robot, we can just repeat the first step and start over without having spent much time and effort yet, therefore we consider this problem not a serious one.

Once a hypothesis has been validated and more features belonging to it have been accumulated, its relocalization becomes easier and due to the better estimation of its center the pushes cause smaller rotations, therefore at this stage we almost never loose track of it. Bringing it approximately to the desired grasping position also works reliably. The second point where our approach may fail is the reactive grasping itself. Due to the usually good estimation of the object position and the direction of its maximal extent, in about 60% of the cases the grasp is successfull after $0 - 3$ correction movements. In 30% of the cases, it takes more than 3 correction movements, and in 10% the grasping fails. This can happen either because we loose contact with the object (e.g. because we accidentally pushed it away) and do not find it again, or because a part of it (e.g. a corner) touches the palm and thus triggers the grasp, but the hand does not close completely around it and therefore it slips out of the hand.

All in all, the whole approach turned out to be very solid. We expect imprecisions in each step of the process, and all algorithms are designed in a way that they can deal with it. Thus, the system mostly recovers from unfortunate situations. A failure in the first step is not costly and therefore does not hurt, while a failed grasp attempt requires us to repeat the whole process. It would probably make sense that in such a case we try to relocalize the object and immediately initiate another grasp attempt instead of starting all over again. Anyways, as the implementation is now, the robot usually completes the task successfully and otherwise can recover from almost all failures on its own.

VI. Conclusion and Future Work

We have presented an approach that enables humanoid robots to autonomously discover, explore and grasp unknown objects in an uncontrolled and complex environment. To this end, we exploit the visual, manipulative and haptic capabilities of the robot. A realisation of the proposed concept has been demonstrated on the humanoid robot ARMAR-III, showing that autonomous object exploration is possible and that we can already achieve a high success rate at the very difficult task of grasping these beforehand completely unknown objects.

One serious limitation of our approach is that so far we need the objects to be at least somewhat textured to allow the detection of regular surfaces at the object discovery stage and the hypothesis matching after having pushed an object. For this reason, we can not handle unicolored objects yet, which is something we intend to try in the near future. Promising steps in that direction have been made in [29].

The method to grasp reactively is limited by the sensitivity and accuracy of the contact detection. To avoid movement of the object during grasping, we approach the object from the top. To approach the object from the side, which would allow grasping of more complex objects, we will need sensors which are more sensitive and accurate, so that a contact can be detected before the object is pushed away. This would also allow us to improve the correction movements, by taking into account the more precise location of the contact point.

We believe that it is only a question of time, improved sensors and some more research effort to overcome these limitations, and that autonomous interactive object exploration is a skill that humanoid robots will soon be able to master, which will significantly increase their autonomy and adaptability.

VII. Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement 270273 (Xperience).

References

[1] T. Asfour, K. Regenstein, P. Azad, J. Schröder, A. Bierbaum, N. Vahrenkamp, and R. Dillmann, "ARMAR-III: An integrated humanoid platform for sensory-motor control," in *6th IEEE-RAS International Conference on Humanoid Robots*, 2006.

[2] G. Metta and P. Fitzpatrick, "Early integration of vision and manipulation," *Adaptive Behavior*, vol. 11, no. 2, 2003.

[3] E. S. Kuzmič and A. Ude, "Object segmentation and learning through feature grouping and manipulation," in *10th IEEE-RAS Int. Conf. Humanoid Robots*, 2010.

[4] D. Schiebener, A. Ude, J. Morimoto, T. Asfour, and R. Dillmann, "Segmentation and learning of unknown objects through physical interaction," in *11th IEEE-RAS Int. Conf. Humanoid Robots*, Bled, Slovenia, 2011.

[5] A. Ude, D. Schiebener, N. Sugimoto, and J. Morimoto, "Integrating surface-based hypotheses and manipulation for autonomous segmentation and learning of object representations," in *IEEE Int. Conf. Robotics and Automation*, St. Pauls, Minnesota, 2012.

[6] J. Kenney, T. Buckley, and O. Brock, "Interactive segmentation for manipulation in unstructured environments," in *IEEE Int. Conf. Robotics and Automation*, 2009.

[7] W. H. Li and L. Kleeman, "Segmentation and modeling of visually symmetric objects by robot actions," *Int. Journal of Robotics Research*, vol. 30, no. 9, 2011.

[8] A. Ude, D. Omrčen, and G. Cheng, "Making object learning and recognition an active process," *Int. Journal of Humanoid Robotics*, vol. 5, no. 2, 2008.

[9] K. Welke, J. Issac, D. Schiebener, T. Asfour, and R. Dillmann, "Autonomous acquisition of visual multi-view object representations for object recognition on a humanoid robot," in *IEEE Int. Conf. Robotics and Automation*, Anchorage, Alaska, 2010.

[10] M. Krainin, P. Henry, X. Ren, and D. Fox, "Manipulator and object tracking for in-hand 3d object modeling," *Int. Journal of Robotics Research*, vol. 30, no. 11, 2011.

[11] T. Araki, T. Nakamura, K. Funakoshi, M. Nakano, and N. Iwahashi, "Autonomous acquisition of multimodal information for online object concept formation by robots," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2011.

[12] K. Hsiao, P. Nangeroni, M. Huber, A. Saxena, and A. Y. Ng, "Reactive grasping using optical proximity sensors," in *IEEE International Conference on Robotics and Automation*, may 2009, pp. 2098 –2105.

[13] K. Hsiao, S. Chitta, M. Ciocarlie, and E. Jones, "Contact-reactive grasping of objects with partial shape information," in *IEEE/RSJ International. Conference on Intelligent Robots and Systems*, oct. 2010, pp. 1228 –1235.

[14] J. Felip and A. Morales, "Robust sensor-based grasp primitive for a three-finger robot hand," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, oct. 2009, pp. 1811 –1816.

[15] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *IEEE International Conference on Robotics and Automation*, 2000, pp. 707–714.

[16] Y. Bekiroglu, J. Laaksonen, J. A. Jørgensen, V. Kyrki, and D. Kragic, "Assessing grasp stability based on learning and haptic data," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 616 –629, 2011.

[17] J. Schill, J. Laaksonen, M. Przybylski, V. Kyrki, T. Asfour, and R. Dillmann, "Learning continuous grasp stability for a humanoid robot hand based on tactile sensing," in *4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, Rome, Italy, 2012, pp. 1901 – 1906.

[18] J. Feldman, "What is a visual object?" *Trends in Cognitive Sciences*, vol. 7, no. 6, 2003.

[19] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, 1988, pp. 147–151.

[20] J. Matas, O. Chum, M. Urba, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *Proc. British Machine Vision Conference*, 2002.

[21] P. Forssen, "Maximally stable colour regions for recognition and matching," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2007.

[22] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.

[23] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," in *Communications of the ACM*, vol. 24, 1981.

[24] D. Pelleg and A. Moore, "X-means: Extending k-means with efficient estimation of the number of clusters," in *Proc. 17th Int. Conf. Machine Learning*, San Francisco, CA, 2000.

[25] N. Vahrenkamp, D. Berenson, T. Asfour, J. Kuffner, and R. Dillmann, "Humanoid motion planning for dual-arm manipulation and re-grasping tasks," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2009.

[26] A. Bierbaum, J. Schill, T. Asfour, and R. Dillmann, "Force position control for a pneumatic anthropomorphic hand," in *IEEE-RAS International Conference on Humanoid Robots*, Paris, France, 2009, pp. 21 – 27.

[27] I. Gaiser, S. Schulz, A. Kargov, H. Klosek, A. Bierbaum, C. Pylatiuk, R. Oberle, T. Werner, T. Asfour, G. Bretthauer, and R. Dillmann, "A new anthropomorphic robotic hand," in *IEEE-RAS International Conference on Humanoid Robots*, Daejeon, Korea, 2008.

[28] Weiss Robotics, "Tactile sensor module, type: DSA 9335," accessed 7/7/2012. [Online]. Available: http://www.weiss-robotics.de/

[29] C. Bersch, D. Pangercic, S. Osentoski, K. Hausman, Z. Marton, R. Ueda, K. Okada, and M. Beetz, "Segmentation of textured and textureless objects through interactive perception," in *RSS Workshop on Robots in Clutter: Manipulation, Perception and Navigation in Human Environments*, 2012.