From Animals to Animats 11 – International Conference on the Simulation of Adaptive Behavior
Lecture Notes in Computer Science, 2010, Volume 6226/2010, 156-166, DOI: 10.1007/978-3-642-15193-4_15
Author post-print. The original publication is available at www.springerlink.com.

Adapting Preshaped Grasping Movements using Vision Descriptors

Oliver Krömer, Renaud Detry, Justus Piater, and Jan Peters

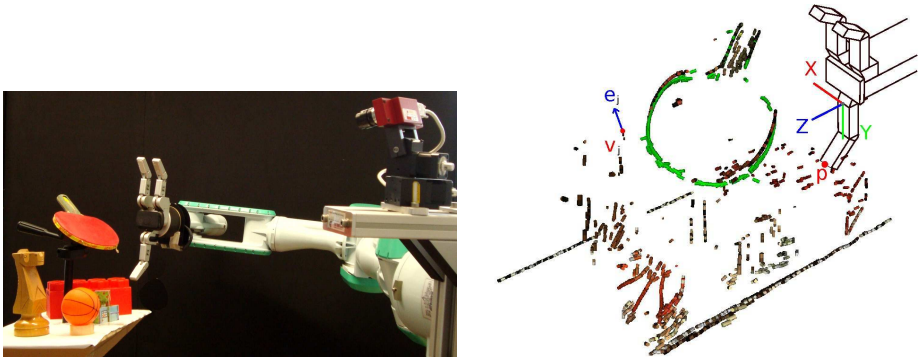
Max Planck Inistitute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany
{oliverkro,Jan.Peters}@tuebingen.mpg.de
{Renaud.Detry,Justus.Piater}@ulg.ac.be

Abstract. Grasping is one of the most important abilities needed for future service robots. In the task of picking up an object from between clutter, traditional robotics approaches would determine a suitable grasping point and then use a movement planner to reach the goal. The planner would require precise and accurate information about the environment and long computation times, both of which are often not available. Therefore, methods are needed that execute grasps robustly even with imprecise information gathered only from standard stereo vision. We propose techniques that reactively modify the robot's learned motor primitives based on non-parametric potential fields centered on the Early Cognitive Vision descriptors. These allow both obstacle avoidance, and the adapting of finger motions to the object's local geometry. The methods were tested on a real robot, where they led to improved adaptability and quality of grasping actions.

1 Introduction

Consider grasping an object at a specific point in a cluttered space, a common task for future service robots. Avoiding collisions is easy for humans, as is preshaping the hand to match the shape of the object to be grasped. Most adults perform these actions quickly and without excessive planning. All of these actions occur before the hand comes into contact with the object, and can therefore be accomplished using stereo vision [1,2]. In contrast, robots often struggle with executing this task, and rely on specially designed sensors (e.g., laser scanner, ERFID) to get accurate and complete representations of the object and environment [3,4], followed by lengthy planning phases in simulation [5].

To avoid excessive planning, a robot can employ a sensor-based controller, which adjusts its motions online when in the proximity of obstacles or other external stimuli [6]. Sensors such as time-of-flight cameras, ultrasonic sonar arrays, and laser range finders are favored for these purposes due to their relatively dense sampling abilities [7,8]. Stereo vision systems, while usually giving sparser readings, have also been used for obstacle detection, especially in the field of mobile robots. However, these methods often rely on task-specific prior knowledge (e.g, assume the ground is flat) and are designed to avoid obstacles completely [8,9], while the robot must get close to the object for grasping tasks. In terms of robot



A. Scene B. ECV Representation

Fig. 1. A) The robot used in our experiments and an example of a grasping task in a cluttered environment. **B)** The green ECVDs represent the object to be grasped, while the surrounding ECVDs in the scene are clutter. The coordinate frame of one of the robot's fingers and variables used in section 2 are shown. The x-y-z coordinate system is located at the base of the finger, with z orthogonal to the palm, and y in the direction of the extended finger. The marked ECVD on the left signifies the j^{th} descriptor, with its position at $\mathbf{v}_j = (v_{jx}, v_{jy}, v_{jz})^T$, and edge direction $\mathbf{e}_j = (e_{jx}, e_{jy}, e_{jz})^T$ of unit length. The position of the finger tip is given by $\mathbf{p} = (p_x, p_y, p_z)^T$.

manipulators, the research has focused on coarse object representations of novel objects [10–13] and using additional sensor arrays when in close proximity to the object [14, 15].

In this paper, we propose a sensor-based robot controller that can perform human inspired grasping motions, including preshaping of the hand, smooth and adaptive motion trajectories, and obstacle avoidance, using only stereo vision to detect the environment. The controller uses potential field methods [6], which treat the robot's state as a particle in a force-field; i.e. the robot is attracted to a goal state, and repelled from obstacles.

The system uses the dynamical system motor primitive (DMP) framework [16, 17] for the attractor field, which are capable of encoding complex trajectories and adapting to different grasp locations. These DMPs are implemented as a passive dynamical system superimposed with an external force; i.e.,

$$\ddot{y} = \alpha_z(\beta_z \tau^{-2}(g - y) - \tau^{-1}\dot{y}) + a\tau^{-2} f(x), \tag{1}$$

where α_z and β_z are constants, τ controls the duration of the primitive, a is an amplitude, $f(x)$ is a nonlinear function, and g is the goal for the state variable y. The variable $x \in [0, 1]$ is the state of a canonical system $\dot{x} = -\tau x$, which ensures that the different hand and arm motions are synchronized. The function $f(x)$ is used to encode the trajectory for reaching the goal state, and takes the form $f(x) = (\sum_{i=1}^{M} \psi_i)^{-1} \sum_{j=1}^{M} \psi_j(x) w_j x$, where $\psi(x)$ are M Gaussian basis functions, and w are weights. The weights w can be programmed through imitation learning [18]. The DMPs treat the goal state g as an adjustable variable and ensure that this final state is always reached.

The scene's visual representation is used to augment the DMP motions and form the basis of the repelling field. The scene description needs to be in 3D, work at a fine scale to maintain geometric details, and represent the scene sparsely to reduce the number of calculations required per time step. The Early Cognitive Vision system of Pugeault et al. [19,20] (see Fig. 1) fulfills these requirements by extracting edge features from the observed scene. The system subsequently localizes and orientates these edges in 3D space [21], with the resulting features known as early cognitive vision descriptors (ECVD) [19]. By using a large number of small ECVDs, any arbitrary object/scene can be represented.

The methods for generating the DMP and ECVD based potential fields are detailed in Section 2. In Section 3, the system is tested on a real robot and shown to be capable of avoiding obstacles and adapting the fingers to the local geometry of the object for improved grasps using only stereo vision.

2 Methods for Reactive Grasping

The methods proposed in this section were inspired by human movements. Human grasping movements can be modeled as two linked components, transportation and preshaping, synchronized by a shared timer or canonical system [22,23]. Transportation refers to the actions of the arm in moving the hand, while the preshaping controls the opening and subsequent closing of the fingers [24].

Humans perform the reaching/transportation component in a task-specific combination of retina and hand coordinates [25], which allows for easier specification of object trajectories in a manipulation task than joint coordinates would and also results in a reduction in dimensionality.

Similar to the transportation component, the main purpose of the finger posture component is to preshape the hand by extending the fingers sufficiently for them to pass around the object upon approach, and then close on the object simultaneously for a good grasp [22,24]. Over-extending the fingers is undesirable as it makes collisions with the environment more likely and is usually restricted to situations where the shape of the object is uncertain [22,26].

The DMP and ECVD based potential field implementations are described in Sections 2.1 and 2.2. Section 2.3 proposes methods that improves the interpolation of grasping movements to new grasp locations.

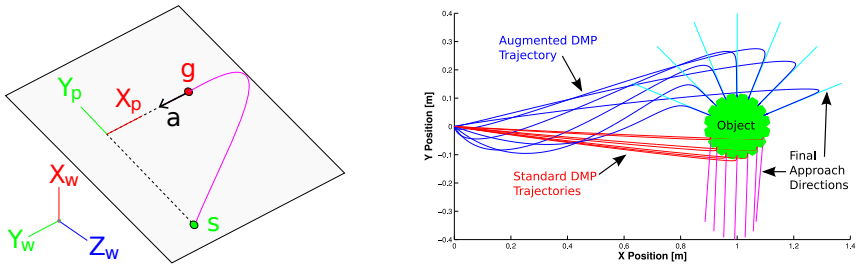
2.1 Regular Dynamical Motor Primitives for Grasping

The first step towards specifying the grasping movements is to define an attractor field as a DMP that encodes the desired movements given no obstacles. The principal features that need to be defined for these DMPs are the goal positions, and the generic shape of the trajectories to reach the goal.

Determining the goal posture of the hand using the ECVDs has been investigated in a previous paper [27]. Possible grasp locations were hypothesized from the geometry and color features of the ECVDs, and subsequently used to create a kernel density estimate of suitable grasps. It was then refined by evaluating grasps on the real system. However, this grasp synthesizer only gives the desired location and orientation of the hand and not the exact finger locations.

Using the ECVDs, the goal position of each finger is determined by first estimating a local contact plane for the object in the finger coordinate system shown in Fig. 1. If the region to be grasped is not planar, it can still be linearly approximated as such for each finger to give good results. To ensure the approximation is accurate in the proximity of the finger, the influence of the i^{th} ECVD is weighted by $w_i = \exp(-\sigma_x^{-2} v_{ix}^2 - \sigma_y^{-2} v_{iy}^2 - \sigma_z^{-2} v_{iz}^2)$, where σ_x, σ_y, and σ_z are length scale constants that reflect the finger's length and width, and $\mathbf{v}_i$ is the position of the ECVD in the finger reference frame. The hand orientation was chosen such that the Z direction of the finger should be approximately parallel to the contact plane, which reduces the problem to describing the plane as a line in the 2D X-Y space. The X-Y gradient of the plane is approximated by $\phi = (\sum_{i=1}^{N} w_i)^{-1} \sum_{i=1}^{N} w_i \arctan(e_{iy}/e_{ix})$, where N is the number of vision descriptors, and $\mathbf{e}_i$ is the direction of the i^{th} edge. The desired Y position of the fingertip is then given by $\tilde{p}_y = (\sum_{i=1}^{N} w_i)^{-1} \sum_{i=1}^{N} (w_i v_{iy} - \tan(\phi) w_i v_{ix})$, which can be converted to joint angles using the inverse kinematics of the hand.

Many of the beneficial traits of human movements, including smooth motions and small overshoots for obstacle avoidance [23, 24, 28], can be transferred to DMPs through imitation learning. To demonstrate grasping motions, we used



A. Proposed DMP coordinate system B. Example Trajectories

Fig. 2. A) The above diagram shows the the coordinate systems for the transportation DMPs. The axes X_w-Y_w-Z_w are the world coordinate system, while X_p-Y_p-Z_p is the coordinate system in which the DMP is specified. The trajectory of the DMP is shown by the **pink** line, starting at the **green point**, and ending at the **red point**. Axis X_p is parallel to the approach direction of the hand (the **black arrow a**). Axis Y_p is perpendicular to X_p, and pointing from the start **s** towards the goal **g**.
B) The plot shows reaching trajectories, wherein the x and y values are governed by two DMPs sharing a canonical system. The standard DMPs and the augmented DMPs described in Section 2.3 are presented along with their respective final approach directions.

a VICON motion tracking system to record the movements of a human test subject during a grasping task. It is not necessary for the object used for the demonstration to match that grasped by the robot later. VICON markers were only required on the back of the hand and finger tips. As the reaching trajectories are encoded in task space rather than joint space, the correspondence problem of the arm was not an issue for the imitation learning step. Details for imitation learning of DMPs using locally weighted regression can be found in [18].

As DMPs are provably stable [17], they are safe to execute on a robot and also ensure that the final arm and finger postures will also always be achieved when physically possible. The repelling field must maintain this stability.

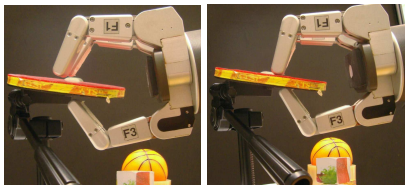
2.2 Adapting the Motor Primitives with Vision Descriptors

Having specified the basic grasping movements, a repelling field refines the motions in order to include obstacle avoidance for the transportation and ensure that the finger tips do not collide with the object during the hand's approach.

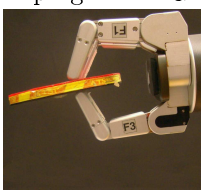
The repelling field is based on ECVDs, which can be understood as small line segments of an object's edges localized in 3D (see Fig. 1). The repelling potential fields for ECVDs are characterized by two main features; i.e., the repelling forces of multiple ECVDs describing a single line do not superimpose, and the field should not stop DMPs from reaching their ultimate goals. The system therefore uses a Nadaraya-Watson model [29] of the form

$$u_a = -s(x)\frac{\sum_{i=1}^{N} r_i c_{ai}}{\sum_{j=1}^{N} r_j},$$

to generate a suitable repelling field, where r_i is a weight assigned to the i^{th} ECVD, s is the strength of the overall field, x is the state of the DMPs' canonical system, and c_{ai} is the repelling force for a single descriptor. Subscript a specifies if the detractor field is for the finger motions "f" or the reaching movements "h".



A. Preshaping B. Grasping



C. Lifting

Fig. 3. The three main phases of a basic grasp are demonstrated. The preshaping of the hand (**A**) tries to pose the fingers to match the object's geometry. The grasping (**B**) then closes the three fingers at the same rate until they secure the object. Finally (**C**) the object is lifted. The objects on the bottom **A** and **B** are clutter that had to be avoided.

The weight of an ECVD for collision avoidance is given by $r_i = \exp(-(\mathbf{v}_i - \mathbf{p})^{\mathrm{T}}\mathbf{h}(\mathbf{v}_i - \mathbf{p}))$, where $\mathbf{v}_i$ is the position of the i^{th} ECVD in the local coordinate system, $\mathbf{h}$ is a vector of width parameters, and $\mathbf{p}$ is the finger tip position, as shown in Fig. 1. A suitable set of width parameters are $\mathbf{h} = 2[w,\ l,\ l]^{\mathrm{T}}$, where w and l are the width and length of the finger respectively.

The reaching and finger movements react differently to edges and employ different types of basis functions c_{fi} and c_{hi} for their potential fields. For the fingers, the individual potential fields are logistic sigmoid functions about the edge of each ECVD of the form $\rho(1+\exp(d_i\sigma_c^{-2}))^{-1}$, where $d_i = \left\|(\mathbf{p} - \mathbf{v}_i) - \mathbf{e}_i(\mathbf{p} - \mathbf{v}_i)^{\mathrm{T}}\mathbf{e}_i\right\|$ is the distance from the finger to the edge, $\rho \geq 0$ is a scaling parameter, and $\sigma_c \geq 0$ is a length parameter. Differentiating the potential field results in a

force term of $c_{fi} = \rho \exp\left(d_i \sigma_c^{-2}\right) \left(1 + \exp\left(d_i \sigma_c^{-2}\right)\right)^{-2}$. As the logistic sigmoid is monotonically increasing, the repelling always forces the fingers open further to move their tips around the ECVDs and thus ensure that they always approach the object from the outside. Similarly, a symmetrical potential function can be employed to force the hand closed when near ECVDs pertaining to obstacles.

The reaching motion uses basis functions of the form $\varrho \exp(-0.5 \mathbf{d}_i^{\mathrm{T}} \mathbf{d}_i \sigma_d^{-2})$, where $\mathbf{d}_i = (\mathbf{q} - \mathbf{v}_i) - \mathbf{e}_i (\mathbf{q} - \mathbf{v}_i)^{\mathrm{T}} \mathbf{e}_i$ is the distance from the end effector position, $\mathbf{q}$, to the edge, and $\varrho \geq 0$ and $\sigma_d \geq 0$ are scale and length parameters respectively. Differentiating the potential with respect to $\mathbf{d}_i$ gives a force term in the Y direction of $c_{hi} = \varrho(\mathbf{d}_i.\mathbf{Y}) \sigma_d^{-2} \exp(-0.5 \mathbf{d}_i^{\mathrm{T}} \mathbf{d}_i \sigma_d^{-2})$, which can be interpreted as a radial force from the edge with an exponentially decaying magnitude.

To synchronize the repelling field with the DMPs and ensure the repelling strength is zero at the end of a motion, the strength s is coupled to the canonical system of the DMPs. Hence, $s(x) = (\sum_{j=1}^{M} \psi_j(x))^{-1} \sum_{i=1}^{M} \psi_i(x) w_i x$, where x is the value of the canonical system, ψ are the DMP basis functions, and w specify the varying strength of the field during the trajectory. To reflect the human tendency towards more precise movements during the last 30% of a motion [28], the strength function was set to give the highest strengths during the first 70% of the motion for the reaching trajectories, and the last 30% for the finger movements.

The repelling fields of both the grasping and reaching components have now been defined, and can be superimposed into the DMP framework as

$$\ddot{y} = \left(\alpha_z(\beta_z \tau^{-2}(g - y) - \tau^{-1}\dot{y}) + a\tau^{-2} f(x)\right) - \tau^{-2} u_a,$$

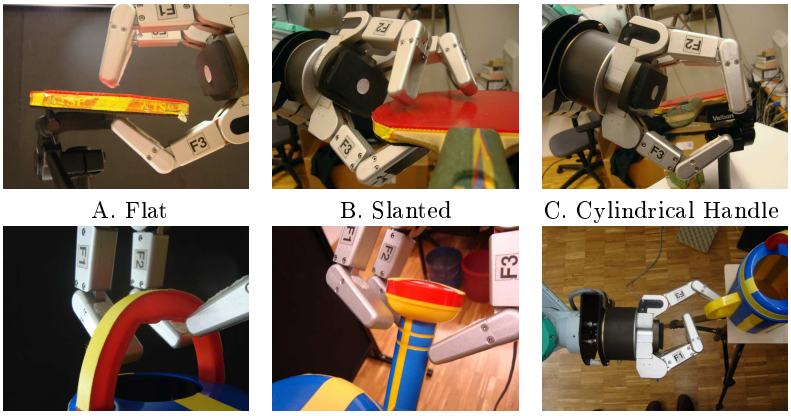
which then represents the complete ECVD and DMP based potential field.

2.3 Generalizing Dynamical Motor Primitives for Grasping

Having defined the potential field for a single grasping motion, we must generalize the movements to new target grasps. By interpolating the trajectories in a task-specific manner, the number of example trajectories required from the demonstrator for imitation learning can be greatly decreased. While the goal states of DMPs can be set arbitrarily, the approach direction to the grasp cannot be easily defined and the amplitude of the trajectory can be unnecessarily sensitive to changes in the start position y_0 and the goal position g.

Fig. 4. Examples of different approach directions are presented, all based off of a single human demonstration.

The correct approach direction can be maintained by using a task-specific coordinate system. We propose the X_p-Y_p-Z_p coordinate system



A. Flat B. Slanted C. Cylindrical Handle

D. Arched Handle E. Knob F. Extreme Point

Fig. 5. Pictures **A** and **B** show the system adjusting to different plane angles. Images **C** and **D** demonstrate the preshaping for different types of handles. Picture **E** shows the preshaping for a circular disc structure, such as a door knob, and manages to get its fingers closely behind the object. Picture **F** shows a preshape where the object was too far away to be reached by two of the fingers, but still hooks the object with 1 finger.

shown in Fig. 2, which dedicates one axis $\mathbf{x}_p$ specifically to the approach direction. The majority of the unobstructed reaching motion will lie in a plane defined by the starting point, the goal location, and the final approach direction, which we use to define our second axis $\mathbf{y}_p$. The final axis $\mathbf{z}_p$ is given by $\mathbf{z}_p = \mathbf{x}_p \times \mathbf{y}_p$.

The second problem relates to the sensitivity of scaling motions with ranges greater than $\|y_0 - g\|$, which grasping motions require to move around the outside of objects. The system can be desensitized to variations in $y_0 - g$ by employing the amplitude term $a = \|\eta(g - y_0) + (1 - \eta)(g_T - y_{0T})\|$ instead of the standard $a = (g - y_0)$ [16], where g_T and y_{0T} are the goal and start positions of the training data respectively, and $\eta \in [0, 1]$ is a weighting hyperparameter that controls how conservative the generalization is. By taking the absolute value of the amplitude, the approach direction is specified solely by the choice of X_p-Y_p-Z_p coordinate system and not the amplitude term. This amplitude term is a generalization of the amplitude proposed by Park et al. [12], which corresponds to the special case of $\eta = 0$. Example interpolations of a transportation trajectory can be seen in Fig. 2.

3 Grasping Experiments

The methods described in Section 2 were implemented and evaluated on a real robot platform consisting of a Videre stereo camera, a Barrett hand, and a 7-degrees-of-freedom Mitsubishi PA10 arm, as shown in Fig. 1.

3.1 Grasping Experiment Procedure

To test the system's obstacle avoidance ability, the robot was given the task of grasping an object without hitting surrounding clutter (see Fig. 1). Each trial begins with an estimate of the pose of the object relative to the robot [30] and

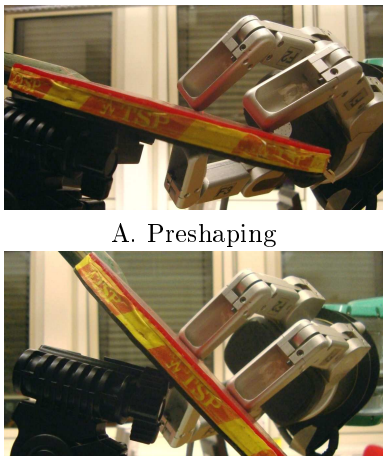
setting the desired grasp location. The model's ECVD are then projected into the scene, and the robot attempts to perform the grasp and lift the object off the table.

If the hand collides with an obstacle or knocks the object down during its approach, the trial is marked as a failure. Grasp locations on the object were predefined, and all successful trials had to lift the object from its stand (see Fig. 3). After each grasp attempt, the hand reverses along the same approach direction, but with a static preshaping of the hand in order to determine if collisions would have occurred if the proposed controller had not been used. The experiment consisted of 50 trials and were varied to include different approach directions and locations around the object.

Additional trials were performed on another object to further explore how the system's preshaping ability adapts to different object geometries. The clutter was removed in these trials to demonstrate the range of grasps that a single human demonstration can easily be generalized to.

3.2 Experimental Results

The repelling field and preshaping of the hand allowed the system to handle the cluttered environment that the object had been placed in, which was not a trivial task. The hand came into contact with the clutter for an estimated 8% of the grasp attempts, but never more than a glancing contact. When the proposed controller was deactivated and a static preshape was used, the hand collided with one or more pieces of clutter in 86% of the trials. Thus, the proposed sensor-based controller led to a factor of ten decrease in the number of contacts with the clutter. The few instances when the hand did collide with the obstacles were the result of obstacles being partially occluded, and thus not fully represented by the ECVDs. This problem represents the main restriction of the current method, which can be overcome by simply using multiple views to accumulate the ECVD representation of the scene, as described in [19, 20]. The repelling fields of the fingers ensured that the hand always opened sufficiently to accept the object without colliding with it.



A. Preshaping

B. Grasping

Fig. 6. The preshaping allows for more controlled grasping. (A) The preshape has matched the local geometry of the object. When grasping, the two fingers on the left immediately pinch the paddle, while the finger on the right turns the paddle about the pinched point. (B) The grasping ends when the paddle has become aligned with all three finger tips.

Using only a single human demonstration, the robot could perform a wide range of reaching movements with varied approach directions, as demonstrated in Fig. 4. Requiring fewer demonstrations hastens the imitation learning process, while still allowing the robot to perform smooth and natural reaching motions.

The incorporation of ECVDs allowed the fingers to adapt to a wide variety of different object geometries, as shown in Fig. 5, and place the finger tips very close to the object before applying the grasp. This close proximity to the object restricts how much the object can move during the final grasping phase, as the fingers make contact with the object at approximately the same time, and leads to grasps being applied in a more controlled manner. An example of a controlled grasp is shown in Fig. 6, which would not be possible without the proposed preshaping, as the finger on the right would have made first contact with the paddle and simply knocked it down.

The results ultimately show that our hypothesis was correct and the proposed methods represent a suitable basis for avoiding obstacles without relying on a complicated path planner, and using only stereo vision information.

4 Conclusions

The proposed methods augment dynamical system motor primitives to incorporate Early Cognitive Vision descriptors by using potential field methods, and represent important tools that a robot needs to execute preshaped grasps of an object in a cluttered environment using stereo vision. The techniques allow for preshaping the fingers to match the geometry of the object and shaping the trajectory of the hand around objects. The controller was tested on a real robot, and was not only successful at performing the task, but also requires very few demonstrations for imitation learning, improves obstacle avoidance, and allows for more controlled grasps to be performed.

References

1. C. Bard, J. Troccaz, and G. Vercelli, "Shape analysis and hand preshaping for grasping," in *IROS proceedings*, 1991.
2. T. Iberall, "Grasp planning for human prehension," in *ICAI proceedings*, 1987.
3. A. Morales, T. Asfour, P. Azad, S. Knoop, and R. Dillmann, "Integrated grasp planning and visual object localization for a humanoid robot with five-fingered hands," in *IROS*, pp. 5663–5668, 2006.
4. Z. Xue, A. Kasper, J. M. Zoellner, and R. Dillmann, "An automatic grasp planning system for service robots," in *procceedings of International Conference on Advanced Robotics (ICAR)*, 2009.
5. D. Bertram, J. Kuffner, R. Dillmann, and T. Asfour, "An integrated approach to inverse kinematics and path planning for redundant manipulators," in *ICRA*, pp. 1874–1879, 2006.
6. M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*. WSE, 2005.
7. M. Khatib, "Sensor-based motion control for mobile robots," 1996.
8. K. Sabe, M. Fukuchi, J.-S. Gutmann, T. Ohashi, K. Kawamoto, and T. Yoshigahara, "Obstacle avoidance and path planning for humanoid robots using stereo vision," in *ICRA*, pp. 592–597, 2004.

9. S. L. And, "Visual sonar: Fast obstacle avoidance using monocular vision," 2003.
10. J. Tegin, S. Ekvall, D. Kragic, J. Wikander, and B. Iliev, "Demonstration based learning and control for automatic grasping," in *Demonstration based Learning and Control for Automatic Grasping*, 2008.
11. A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, "Automatic grasp planning using shape primitives," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2003.
12. D.-H. Park, H. Hoffmann, P. Pastor, and S. Schaal, "Movement reproduction and obstacle avoidance with dynamic movement primitives and potential fields," in *IEEE International Conference on Humanoid Robots(HUMANOIDS)*, 2008.
13. F. Bley, V. Schmirgel, and K.-F. Kraiss, "Mobile manipulation based on generic object knowledge," in *proceedings of Robot and Human Interactive Communication (ROMAN)*, 2006.
14. K. Hsiao, P. Nangeroni, M. Huber, A. Saxena, and A. Ng, "Reactive grasping using optical proximity sensors," in *ICRA Proceedings*, 2009.
15. J. Steffen, R. Haschke, and H. Ritter, "Experience-based and tactile-driven dynamic grasp control," in *IRS proceedings*, 2007.
16. A. J. Ijspeert, J. Nakanishi, , and S. Schaal, "Learning attractor landscapes for learning motor primitives," in *NIPS*, 2003.
17. S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," in *ISRR proceedings*, 2003.
18. A. J. Ijspeert, J. Nakanishi, , and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *ICRA*, 2002.
19. N. Pugeault, *Early Cognitive Vision: Feedback Mechanisms for the Disambiguation of Early Visual Representation*. Vdm Verlag Dr. Mueller, 2008.
20. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
21. N. Krueger, M. Lappe, and F. Woergoetter, "Biologically motivated multimodal processing of visual primitives," *The Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour*, 2004.
22. S. Chieffi and M. Gentilucci, "Coordination between the transport and the grasp components during prehension movements," 1993.
23. E. Oztop and M. Kawato, *Sensorimotor Control of Grasping: Physiology and Pathophysiology*, ch. Models for the control of grasping. Cambridge University Press, 2009.
24. M. Jeannerod, *Perspectives of Motor Behaviour and Its Neural Basis*, ch. Grasping Objects: The Hand as a Pattern Recognition Device. 1997.
25. M. S. Graziano, "Progress in understanding spatial coordinate systems in the primate brain," *Neuron*, 2006.
26. E. Oztop, N. S. Bradley, and M. A. Arbib, "Infant grasp learning: a computational model," 2004.
27. R. Detry, O. Kroemer, M. Popovic, Y. Touati, E. Baseski, N. Krueger, J. Peters, and J. Piater, "Object-specific grasp affordance densities," in *ICDL*, 2009.
28. M. Jeannerod, *Sensorimotor Control of Grasping: Physiology and Pathophysiology*, ch. The study of hand movements during grasping. A historical perspective. Cambridge University Press, 2009.
29. C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
30. R. Detry, N. Pugeault, and J. Piater, "Probabilistic pose recovery using learned hierarchical object models," in *International Cognitive Vision Workshop*, 2008.