Synergy-level Grasp Synthesis Learning

Miao Li, Sahar El-Khoury and Aude Billard

***Abstract*— Autonomous grasping is a complex task for robots. It is a high dimensional problem since it involves controlling for the hand position, orientation and joint angles to successfully grasp an object. In order to reduce the control complexity, we adopt a 3-step approach. In the first step, we compute several stable grasps that are adapted to the robotic hand using an optimization technique. In a second step, we extract postural synergies from this grasping data, project the grasps into these synergies subspace, and use this data representation to learn a distribution of the feasible grasps. The third step uses the learned model to generate quickly new grasps for the given object. Our approach was validated on the four degrees of freedom Barrett hand.**

I. INTRODUCTION

Autonomous grasping is a very complex problem mainly due to the high dimensionality of the hand configuration space, which includes the hand position and orientation, as well as the finger joint angles. Moreover, this configuration space is usually constrained by the object shape and the task requirements, resulting in a highly non-linear manifold of the feasible hand configurations.

However, it seems quite natural and trivial for humans to employ various kinds of grasps in every day tasks, from grasping a cup of water to drink to manipulating chopsticks to eat, despite the fact that the human hand has 21 degrees of freedom (DOFs) controlled by 29 muscles [1]. The key reason for this dexterity lies in the central nervous system that can integrate the information from different sources and coordinates the activity of the whole hand in order to reduce the control complexity. Studies on the configurations adopted by the human hand have already demonstrated the evidence of this coordination from different aspects, postural synergy considering the hand shape during grasping [2], force synergy regarding the pattern of coordination of the normal forces on each fingertip during pick and lift task [3], temporal postural synergies on joint angular velocities during rapid grasping movement [4], hand postural synergy generalization across different object manipulation tasks [5]. All these synergies on human hand are found by performing a Principal Component Analysis (PCA) of hand postures during grasping and manipulation experiments. As shown in [6], in more complex manipulation tasks, the number of synergies can vary from 4 to 10, but it is still much less than the number of recorded DOFs.

In robotic grasping, it would be very appealing to have a robotic hand with human-like grasp synergies. This would alleviate the complexity of controlling each joint while still keeping most of the dexterity of the hand. Recently, extensive research has been conducted regarding the usage of grasp synergies to control robotic hands: designing optimal gloves for hand pose sensing [7], assessing the role of hand synergies in grasping force optimization [8], implementation of compliant hand control in synergy level [9] and grasp synthesis in low-dimensional posture subspaces, i.e, grasp synergy space [10]. These studies map the synergies of human hand to that of robotic hands. However, there is no guarantee that the synergies obtained from human hands are optimal for the robotic hand, though they are, to some extent, similar to each other in functionality and kinematic structure.

In this paper, instead of mapping the synergies from human hands to robotic hands, we compute these synergies directly from different grasps generated for the corresponding robotic hand. These synergies are then used to control the hand position/orientation and joint angles to generate stable grasps for a given object.

The paper is organized as follows: section 2 describes the grasping data collection. In section 3 we discuss how to learn the joint distribution of hand position, orientation and the hand synergies. In section 4, new grasps are synthesised from the learned joint distribution. In section 5, we validate our approach on the four degrees of freedom Barrett hand.

II. TRAINING DATA COLLECTION

In order to demonstrate a large variety of different possible grasps for a given object, an optimization based grasp synthesis method is applied [11]. The constraints from the object surface, hand kinematics, grasp stability and collision avoidance are taken into account. A feasible grasp should satisfy all these constraints. As our goal here is to present the variety of the feasible grasps rather than just a few numbers of local optimal grasps as shown in [11], we do not take into account the grasp quality in the grasp generation procedure. Every stable grasp is thus eligible to be in the database.

Then grasp synthesis problem can thus be formulated as a feasibility problem as follows:

$$\begin{array}{ll} min: & \|\boldsymbol{\alpha}\|_2 \\ s.t. & g_i(\mathbf{h},\mathbf{o},\mathbf{q}) \leq \alpha_i, i=1...n_c; \\ & \alpha_i \geq 0, i=1...n_c; \end{array} \tag{1}$$

where $\mathbf{h} \in \mathbb{R}^3, \mathbf{o} \in \mathbb{R}^{3\times3}, \mathbf{q} \in \mathbb{R}^{n_j}$ denotes the hand position, orientation and the angles of finger joints respectively. $g_i, i = 1...n_c$ stands for the constraints that have been considered. $\boldsymbol{\alpha} = [\alpha_1...\alpha_{n_c}]$ is the slack variable, which should be $\mathbf{0}$ when the problem is feasible. n_c, n_j represent the number of constraints and the number of finger joints respectively.

M. Li, S. EL-Khoury and A. Billard are with LASA, Ecole Polytechnique Fédŕale de Lausanne, Switzerland miao.li@epfl.ch, sahar.elkhoury@epfl.ch, aude.billard@epfl.ch

III. Graspable Space Learning

With a large set of feasible grasps $\{\mathbf{h}, \mathbf{o}, \mathbf{q}\}$ obtained from the optimization (1), we first reduce dimension of finger joint space $\{\mathbf{q}\}$ to a lower dimension finger joint synergy space $\{\mathbf{s}\}$ using PCA, i.e., $\mathbf{s} = \mathbf{W}^T\mathbf{q}$, with $\mathbf{e}^j, j = 1...N$ the eigenvectors that compose the columns of $\mathbf{W}$. With PCA, we can find the most informative eigenvectors that can represent the data more efficiently.

Then a Gaussian Mixture Model (GMM) is used to model the joint distribution $p(\mathbf{h}, \mathbf{o}, \mathbf{s})$. A GMM is chosen for its ability to model the highly non-linear manifold of the feasible hand configurations [12]. A GMM is described as a sum of K Gaussian components:

$$p(\mathbf{h}, \mathbf{o}, \mathbf{s}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{h}, \mathbf{o}, \mathbf{s} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{2}$$

π_k is the prior of the kth Gaussian component and $\mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the Gaussian distribution with mean $\boldsymbol{\mu}_k$ and covariance $\boldsymbol{\Sigma}_k$. The number of Gaussian components K is determined by the *Bayesian Information Criterion* (BIC), which aims at maximizing the likelihood of the model while limiting the complexity of the model. The model parameters $\Omega = \{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ are selected using the *Expectation-Maximization algorithm* (EM) [13].

Given the joint distribution $p(\mathbf{h}, \mathbf{o}, \mathbf{s}|\Omega)$, a new grasp $(\mathbf{h}^*, \mathbf{o}^*, \mathbf{s}^*)$ is said to be feasible when its likelihood is above some threshold, i.e., $p(\mathbf{h}^*, \mathbf{o}^*, \mathbf{s}^*|\Omega) > L_{threshold}$. The value of the threshold is chosen such that the likelihood of 99% of the training data is above this value[1]. The graspable space is the constitution of all the feasible grasps.

IV. Grasp Synthesis using Synergies

In this section, the learned GMM model of the graspable space is used to find new feasible grasps. Starting from an initial configuration of the hand, the hand is moved in the gradient of likelihood ascent direction until the grasp belongs to the graspable space. For the sake of simplicity, we denote the current grasp configuration $(\mathbf{h}_t, \mathbf{o}_t, \mathbf{s}_t)$ as $\mathbf{x}_t$, then the next grasp configuration $\mathbf{x}_{t+1}$ is chosen as follows:

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \alpha \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\mathbf{x_t}} \tag{3}$$

Where α is the step size, which is set to 0.01. $\partial L(\mathbf{x})/\partial \mathbf{x}$ is the gradient of the likelihood of GMM at point $\mathbf{x}$, given by:

$$\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{\mu}_k - \mathbf{x}) \tag{4}$$

When the starting hand configuration is very far away from the graspable space, the gradient of likelihood in that point is almost zero. In this case, we simply move the hand towards the object center until the gradient is larger than some threshold and then move along the gradient ascent direction. As our algorithm will stop as soon as the likelihood of the grasp is higher than the threshold $L_{threshold}$ obtained from graspable space learning, the final grasps are not limited to the ones that are very close to the centers of the Gaussian components, which is the case in [12].

V. Experimental Results

Data collection: in order to collect a large number of feasible grasps, we solve the optimization problem for the flask spray object starting from 5 different initial hand positions and 637 different hand orientations sampled using Euler angle. In all, we have 3185 trials, from which the optimization framework found 2674 feasible grasps. In fact, there is infinite number of feasible grasps. Most of the time we can slide the fingertips a bit to get another feasible grasp. However, depending on the initial point, the optimization algorithm sometimes may not lead to a feasible solution. Part of these obtained grasps are shown in Fig. 1 [2].

Fig. 1: typical grasps for Barrett hand grasping a spray flask

Hand synergies: we apply the PCA dimension reduction technique on the finger joint angles of the previously computed grasps and find that the first three eigenvectors $\mathbf{e}^j, j = 1...3$ take up more than 90% of the whole joint variance. This means that we can represent the finger joint angles with only these three eigenvectors without loosing much accuracy. These eigenvectors are usually called grasp synergies or eigen-grasps, as shown in Fig. 2. These hand strategies are quite different from the previous work that uses the mapping from the human hand to the Barrett hand [10], as in our case all the finger joint angles are collected from feasible grasps rather than the joint angles of human hand collected during reaching and grasping movement.

Graspable space learning: given the hand synergies, the finger joint angles can now be described in the hand synergy space. Our training dataset is given by: $\{\mathbf{o_1}, \mathbf{o_2}, \mathbf{h}, \mathbf{s}\}^{j=1...2674} \in \mathbb{R}^{12}$. $\mathbf{o_1}, \mathbf{o_2}$ are the first two columns of hand orientation matrix as the third column can

[1] In practice, we compute the likelihood for each training data point and sort the computed likelihood in an ascend order. Then the one located near 1% of the data points is chosen as the threshold.

[2] In some grasps, the fingertips are not exactly contacting with the object surface, which is due to the fact that we relax several constraints in the optimization to obtain more feasible grasps.

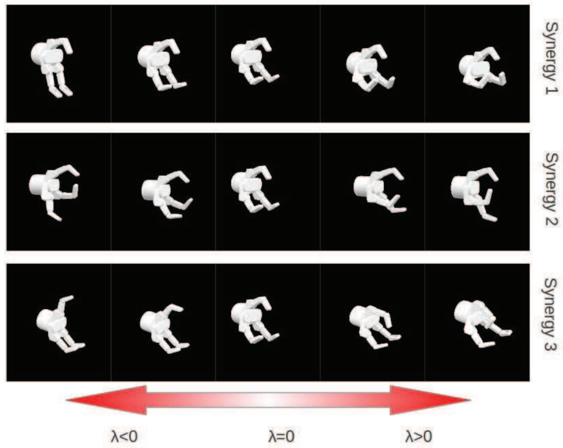


Fig. 2: The first three hand synergies. The finger joint angles are plotted using $\mathbf{q}_{mean} + \lambda \mathbf{e}^j, j = 1...3$, $\mathbf{q}_{mean}$ is the mean of the finger joint angles and from left to right, λ is varying from -1 to $+1$.

be determined from the first two. We use BIC to choose the number of Gaussian components. BIC can make a trade-off between the improvement of likelihood and the complexity of model. As shown in Fig. 3, the number of Gaussian components is selected as $K = 19$ because the BIC will not improve a lot when K is bigger than 19. The GMM is trained using EM algorithm, which is initialized with K-means.

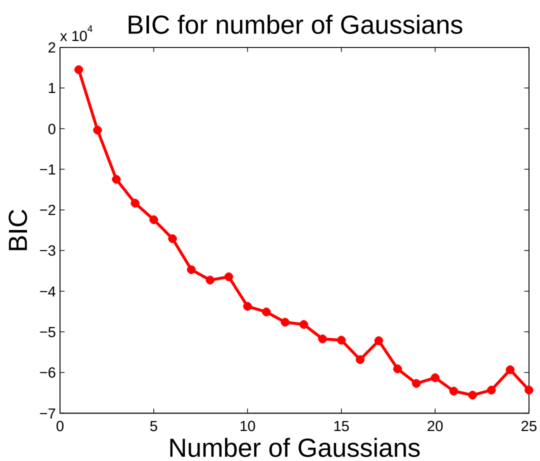
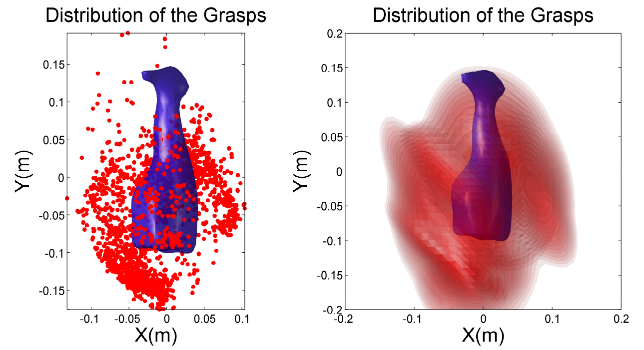


Fig. 3: The number of Gaussians in GMM selected by BIC. The formula for the BIC is: $BIC = -2\ln(L) + K\ln(n)$, where L is the likelihood of all the training data points, K is number of Gaussian components, and n is the number of training data points. The smaller the criteria is, the better the model fits the data.

Given the learned model of the graspable space, we can show how the graspable space looks like in some lower dimensions. Here, we will use the hand positions $\mathbf{h}$. Figure 4(a) shows the hand position of these different grasps relative to the object. There are more points at the bottom and middle part of the spray flask, which means that there is a larger chance to find a feasible grasp in the middle and bottom part. In Fig. 4 (b), the graspable space is successfully learned by GMM with a higher likelihood in the part that more feasible grasp can be found.



(a) (b)

Fig. 4: (a) The distribution of hand position for all the grasps obtained from optimization, i.e., $\{\mathbf{h}\}^{j=1...2674}$. The red points represent the hand positions. (b) The graspable space projection in the hand position subspace. The red area stands for the subspace of hand position that belongs to the graspable space.

Fig. 5: some of the new grasps generated from 12 different initial points. The grasps in the upper row are the initial grasp configurations and the lower row are the found final grasps. As the learning is conducted in hand synergies space, the final grasps are rarely precise grasps, i.e., the fingertips are not exactly contacting with the object surface. However, these obtained grasps can be considered as potential preshapes that will lead to stable grasps.

Grasp synthesis: in order to test the generalization performance of our method, we choose 100 different hand positions which are randomly distributed inside a sphere with radius $0.4m$ englobing the object. As for the hand orientation, the initial normal direction of the palm is always oriented towards the object center. In all these trials, our approach can always lead to final grasps of high likelihood in between 15 and 33 iterations defined in eq. (3). The average time for each trial is $47.17 \pm 0.55ms$. Only 12 of the final grasps are shown in Fig. 5. As our graspable space are trained in the hand synergy space, the final grasps are not necessary exactly contacting with object surface, instead, it is a grasp configuration that is very close to a feasible one in the training data set. In addition, we check the kinematical feasibility of the 100 final grasps, only 2 of them are not kinematically feasible, i.e., the finger joint angle is not in a valid range.

Compared with the results in [12], where the GMM model is learned in the finger joint space directly, the advantage of using synergies are hence two-folds: a) It can significantly increase the number of final grasps that can be found, see Fig. 5. In [12], the final feasible grasp is found in two steps: projecting the current hand position and orientation to a new valid query point and then employing the Gaussian Mixture Regression (GMR) to find the corresponding finger joint angles. Due to the projection step, the final grasps are mostly very close to the centers of Gaussian components, which greatly reduce the number of new feasible grasps that can be found. This step is not necessary for our approach as it can learn a dense graspable space model using synergies and there is no need to project to a valid query point. b) The final grasp in our approach may be more plausible in the sense that it is close to the initial grasp. This is due to the fact that the grasp is found in an iterative way that each step moves along the gradient of likelihood ascent direction until reaching a threshold, see Fig. 6. In all the 100 trials, the average distance between the initial hand position and the final hand position is $0.2697 \pm 0.0260m$. In [12], the projection of the initial grasp to the nearest Gaussian component may result in a final grasp that is far away from the initial grasp. To compare, in our experiment the average distance between initial hand position and the nearest center of Gaussian component is $0.4061 \pm 0.0838m$. However, due to the iterative behaviour, the computation time of our approach is a bit longer than that of [12], which is around $10ms$.

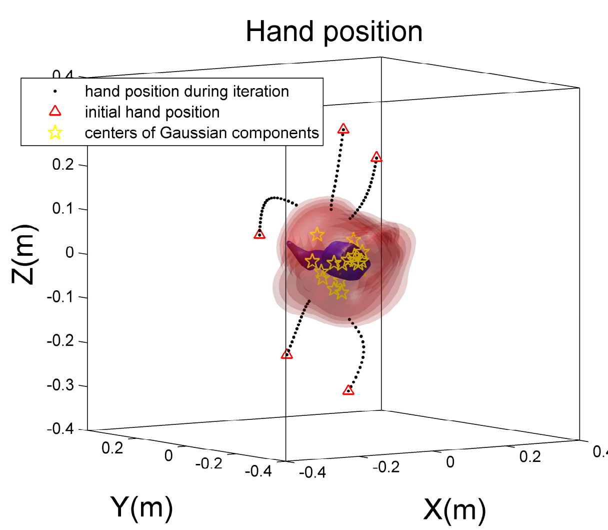


Fig. 6: The red area stands for the subspace of hand position that belongs to the graspable space. The hand will move along the gradient ascending direction until the grasp has a likelihood higher than given threshold. The final grasps will not always close to the centers of Gaussian components (yellow stars), which can increase the diversity of final grasps.

VI. Conclusion and Future Work

In this paper, we proposed a method to learn the graspable space using hand synergies. The training data of feasible grasps were obtained from an optimization framework. We demonstrated that the learned graspable space can be used to generate a variety of new grasps quickly given a staring hand configuration. Future work will focus on learning the graspable space using synergies for different object shapes and sizes.

References

[1] L. A. Jones and S. J. Lederman, *Human Hand Function*. Oxford University Press, USA, 1 ed., Apr. 2006.

[2] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of Neuroscience*, 1998.

[3] M. Santello and J. F. Soechting, "Force synergies for multifingered grasping," *Experimental Brain Research*, vol. 133, no. 4, pp. 457–467, 2000.

[4] R. Vinjamuri, M. Sun, C.-C. Chang, H.-N. Lee, R. J. Sclabassi, and Z.-H. Mao, "Temporal postural synergies of the hand in rapid grasping tasks," *Trans. Info. Tech. Biomed.*, vol. 14, pp. 986–994, July 2010.

[5] P. H. Thakur, A. J. Bastian, and S. S. Hsiao, "Multidigit movement synergies of the human hand in an unconstrained haptic exploration task.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 28, pp. 1271–1281, Feb. 2008.

[6] E. Todorov and Z. Ghahramani, "Analysis of the synergies underlying complex hand manipulation," in *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 4, pp. 4637–4640, IEEE, 2004.

[7] M. Bianchi, P. Salaris, and A. Bicchi, "Synergy-based optimal design of hand pose sensing," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 3929 –3935, oct. 2012.

[8] M. Gabiccini, A. Bicchi, D. Prattichizzo, and M. Malvezzi, "On the role of hand synergies in the optimal choice of grasping forces," *Auton. Robots*, vol. 31, pp. 235–252, Oct. 2011.

[9] T. Wimbock, B. Jahn, and G. Hirzinger, "Synergy level impedance control for multifingered hands," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pp. 973 –979, sept. 2011.

[10] M. Ciocarlie and P. Allen, "Hand posture subspaces for dexterous robotic grasping," *The International Journal of Robotics Research*, vol. 28, pp. 851–867, 07/2009 2009.

[11] S. El Khoury, M. Li, and A. Billard, "Bridging the gap: One shot grasp synthesis approach," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 2027 –2034, oct. 2012.

[12] B. Huang, S. El-Khoury, M. Li, J. J. Bryson, and A. Billard, "Learning a real time grasping strategy," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 6–10 2013. Accepted for publication.

[13] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, "Active learning with statistical models," *Journal of Artificial Intelligence Research*, vol. 4, pp. 129–145, 1996.