Learning Grasping Force from Demonstration

Yun Lin Shaogang Ren

Matthew Clevenger Yu Sun

Abstract—This paper presents a novel force learning framework to learn fingertip force for a grasping and manipulation process from a human teacher with a force imaging approach. A demonstration station is designed to measure fingertip force without attaching force sensor on fingertips or objects so that this approach can be used with daily living objects. A Gaussian Mixture Model (GMM) based machine learning approach is applied on the fingertip force and position to obtain the motion and force model. Then a force and motion trajectory is generated with Gaussian Mixture Regression (GMR) from the learning result. The force and motion trajectory is applied to a robotic arm and hand to carry out a grasping and manipulation task. An experiment was designed and carried out to verify the learning framework by teaching a Fanuc robotic arm and a BarrettHand a pick-and-place task with demonstration. Experimental results show that the robot applied proper motions and forces in the pick-and-place task from the learned model.

I. INTRODUCTION

Learning from demonstration (LfD) has been a powerful mechanism to reduce the complexity and burden of searching or generating a successful action for tasks. For most applications using LfD, a number of human movements are recorded during a task, then analyzed and modeled with machine learning algorithms. Motion elements are decomposed and learned. Relations between motion elements are modeled from the distributions of the demonstrated movements. With the learning results, a robot can mimic the human motions by reproducing similar movements as in the demonstration [1]. For example, grasping and releasing movements during a pick-and-place task demonstrated by a human worker can be tracked with vision-based motion tracking system and decomposed to transport and grasp phases [7]. For some cases, a force sensing glove was used to better segment the movement [19], [10].

The existing LfD frameworks have not yet included force elements that provides important additional dimensions for human dexterous grasping and manipulation demonstration and learning. The force is usually treated as a feedback control element for optimal control problem [13], instead of as flexible distributions for learning. In neuroscience, many research studies have been devoted to understanding human grasping force and gripping strategies. Soechting and his team [14] have measured the contact forces at the digits and decomposed the grip force into two components: a manipulating force responsible for accelerating the object and a grasping force responsible for holding the object steady. Johansson and his team [6] have discovered that humans attempt to avoid horizontal tangential forces even at a small cost in total force and slight object tilt to keep tangential torques small and to compensate for variations in digit contact positions in multidigit manipulative tasks. Many research results have shown that there are complicated patterns in the multidigit grasping and manipulation forces, which are related to task, hand motion and gesture.

The fingertip force is not always correlated to the finger motion that is regularly tracked in LfD frameworks since exerting fingertip forces do not necessarily require any obvious finger motion. Since applying force is an important aspect of the interacting with the environment, and essential for many tasks, it is not surprising that humans exhibit complicated force strategies for different tasks. Measuring the force along with the hand motion, and understanding human force strategies is important for learning grasping and manipulation. However, unlike motion tracking system that are usually not intrusive, most grasping force measurement approaches require force sensors being placed on the grasping points, either on the object or on the demonstrator's fingertip. In many neuroscience studies, specially designed objects have to be fabricated to incorporate force sensors and the points of grasping have to be defined beforehand, which poses big limitation for grasping demonstration.

Previously, we have presented a force imaging approach to estimate fingertip force from the images of the back of the fingertip. That approach does not encumber a subject and there is no need for sensor fabrication or embedded sensors on the objects so that everyday objects can be used for studies and demonstration. The existence of low-cost cameras and image processing methods readily performed on PCs makes the instrumentation of such an approach relatively inexpensive. We have shown that by imaging the coloration changes in the fingernail and surrounding skin with an external camera, normal and shear forces can be estimated with an accuracy of 5-10% for a force range of up to 8 N. However, individual calibration was required, and a generalized least squares (GLS) estimator was used [18], [15]. We also have estimated the dynamic features of the approach and found that the time constants were different for different force levels and directions (loading and unloading) and the typical time constant is around 0.2 second [16]. The slow dynamic features of the approach may not pose a big limitation for many grasp studies and applications, as the finger force frequency is fairly low. For example, during surgical training, the majority of the frequency content is below 5 Hz; during tissue grasps, the average grasp force frequency is approximately 1 Hz [2].

Yun Lin, Shaogang Ren, and Yu Sun are with the Department of Computer Science and Engineering, University of South Florida, Matthew Clevenger is with the Department of Mechanical Engineering, University of South Florida. ({yunlin,shaogangren,matthewc}@mail.usf.edu, yusun@cse.usf.edu)

In this paper, we propose a novel grasping and manipulation learning framework that trains a learning model with both motion and force data so that with our force imaging approach, a robot can learn the demonstrator's fingertip motion and force for a grasping and manipulation process. The recorded motion and force from demonstrations were processed with Gaussian Mixture Model (GMM) based machine learning approach which was presented by Calinon et al in [3]. To apply the learning results to a robotic system, a force and motion trajectory was generalized with Gaussian Mixture Regression (GMR) from the trained GMM, and was mapped to the robotic system. We have implemented the grasping and manipulation learning framework in our lab, analyzed the measurement and learning results, and successfully applied a learning result to a robotic hand-arm for a pick-and-place task.

II. GRASPING AND MANIPULATION LEARNING FRAMEWORK

A. Framework Overview

We propose a full grasping and manipulation learning framework that is capable of measuring and learning from human grasping and manipulation processes. The design of the framework is illustrated in Figure 1. It is composed of three systems: a fingertip force estimation system, a grasping and manipulation learning from demonstration (LfD) system, and a robotic system that applies the learning results. Among them, the motion and force LfD system is the center piece of the framework. It measures not only motion, but also force in demonstrated grasping and manipulation processes. The captured motion and force data is then used to train a Gaussian Mixture model that represents the joint distribution of the data. The learned motion and force motion representing skills is then mapped to the robotic system to generate proper motions and force to carry out learned tasks. The fingertip force estimation system measures the fingertip force in grasping and manipulation process from the color in the corresponding fingernail and surrounding skin based on the trained models of the relation between the force and the color, so that the motion and force LfD system can measure the fingertip force without a force sensor embedded on the object or fingertip.



Fig. 1. Presented Framework of grasping and manipulation LfD system.

B. Fingertip Force Estimation System

The fingertip force is measured remotely from the color of the corresponding fingernail and its surrounding skin. The relation between the color and the force can be modeled with a generalized least-squares model [18]:

$$\hat{\mathbf{f}} = (\mathbf{B}^T \mathbf{\Sigma}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{\Sigma}^{-1} (\mathbf{h} - \mathbf{a})$$
(1)

where **h** is vector with the color reading of all the pixels in the fingernail and surrounding skin. Parameter vectors **a** and **B** are learned linear parameters relating the color response to the force. The covariance matrix Σ is estimated from the data which represent the weights of all pixels contributing to the estimation of force [18].

The least-squares model needs to be trained individually for every finger that the system wants to measure force on. A training process is designed to automatically apply a serial of force on fingertip and record force and the fingernail images at the same time. The sequence of force is designed to cover the potential 3D force cone the finger can exert with reasonable density. The fingernail images are segmented, aligned and normalized with orientation compensation [17]. The pixels on the fingernail and surrounding skin in the processed fingernail images are then organized into a color vector for the estimation model Equation 1. With enough training data, the generalized least-squares model can estimate fingertip force fairly accurately – usually has error at around 8% [18].

The trained fingers can be used at the motion and force demonstration station. In the proposed grasping and manipulation learning framework, multiple cameras are setup to continually monitor the fingernail images during a dynamic grasping and manipulation process. Since the illuminations are not identical between the training station and the demonstration station, the HSV color space is used.

C. Motion and Force LfD system

With the remote force measurement approach and a fingertip tracking system, demonstrated hand motion and force can be captured and recorded. The recorded demonstrated motions and forces are then modeled with a Gaussian Mixture Model (GMM) that summarizes a probabilistic representation of multiple demonstrations [3]. Different from previous approaches, the proposed motion and force LfD system has extra three dimensions that represent a 3-dimension force vector. At any time point, a 6-dimension motion vector and a 3-dimension force vector are combined to represent an action state of a task.

The demonstrated motion and force can be encoded together by GMM. Given a set of data points of dimensionality D, one dimension is time steps, while the other dimensions are motion and force trajectories. The dataset is defined by $\xi_j = \{\xi_{t,j}, \xi_{m,j}, \xi_{f,j}\}$, where j=1, 2, ... N is the number of trials, $\xi_{t,j}$ represents the time step, $\xi_{m,j} \in \Re^6$ is the position and orientation vector, $\xi_{f,j} \in \Re^3$ is the force vector. The dataset is modeled by a mixture of Gaussian distributions of K components. K is defined as 4 in the pick-and-place task, for we segmented the whole grasping process into 4 stages - grasp the object, lift the object, place the object back, and release the object. The GMM model is defined by the function:

$$p(\xi_j) = \sum_{k=1}^{K} p(k) p(\xi_j | k)$$
 (2)

where p(k) is the prior, and $p(\xi_j|k)$ is the conditional probability of the jth data given the kth Gaussian distribution. They are further defined as:

$$p(k) = \pi_k$$
(3)

$$p(\xi_j|k) = N(\xi_j; \mu_k, \Sigma_k)$$

$$= \frac{1}{\sqrt{(2\pi)^D |\Sigma_k|}} e^{-\frac{1}{2}((\xi_j - \mu_k)^T \Sigma_k^{-1}(\xi_j - \mu_k))}$$
(4)

for k = 1 to K. The parameters of prior π_k , mean μ_k , covariance matrix Σ_k of the Gaussian Mixture Model are estimated by an expectation-maximization (EM) algorithm [5], which maximize the likelihood of $P(\xi|\pi)$.

D. Mapping Learning Results to Robotic System

Once the demonstration model is built by GMM, smooth force and motion trajectories are retrieved from the model for the robot using Gaussian Mixture Regression (GMR) [4]. Given a joint probability distribution $p(\xi_t, \xi_m, \xi_f)$ of the dataset modeled by a GMM, GMR estimates the conditional expectation $E[p(\xi_m, \xi_f | \xi_t)]$. The regression thus produces the expectations of the motion and force at each time step ξ_t , which provide smooth motion and force trajectories along the time space.

The generated motion and force trajectory is then applied to a robotic system with the trajectory as its control input. Since the motion is represented with the location and orientation of the hand, an inverse kinematics approach is used to compute the joint angles of the robotic hand to generate the desired fingertip motion. A hybrid motion and force controller is applied to ensure that the robot hand keeps contacting an object with a certain force, i.e., the force learned from the user. The hybrid motion and force control allows the separation of the position and force in two independent subspaces. Hence the position and force are controlled simultaneously.

In this paper we aim to illustrate our LfD framework, so we make two simplifications to the problem. First, objects are held vertically, which makes sure that the contact force are uniformly distributed on each finger. Another simplification is that we only control normal force, due to the limited DOF of the Barrett hand used in our experiment.

The motion controller transfers the error between desired and actual position of the fingertip to a joint angle using inverse Jacobian matrix. The force controller transfers the error between desired force and actual force into a joint torques using transposed Jacobian matrix. Both outputs are combined together at the end and converted to joint angles.

III. FRAMEWORK IMPLEMENTATION

A preliminary grasping and manipulation learning framework is constructed in our lab. Figure 2 shows a robotic work cell on the left and the motion and force LfD on the right. We setup two cameras on two circular rails to track the fingertip position and measure the fingertip force from the color in the corresponding fingernails and surrounding skin. The horizontal and vertical position is extracted from the 2D image taken by the camera. Figure 3 shows our training platform for fingertip force and color modeling. For a human demonstrator, the force and color calibration is executed once and does not need to be re-calibrated for each demonstrated task.



Fig. 2. The robot grasping and manipulation station (left) and the motion and force demonstration station (right) in the grasping force learning from demonstration framework.

A. Fingertip Force Modeling Setup



Fig. 3. (A) The 5-DOF calibration stage based on two Novint Falcon Haptics Controllers; (B)The hand rest with finger restraint; (C) the kinematic model of the calibration station.

An inexpensive automated calibration system is designed to apply calibration force trajectory on a human fingertip with high precision and take calibration images and force data simultaneously (Figure 3A). The calibration system is composed of a 5-DOF actuation device, a 6-axis force sensor, a video camera, a hand-rest stage, and a finger restraint.

The actuation device is an integration of two Novint Falcon devices linked by two universal joints and a rigid bar to provide 5-DOF motion and force, with feedback from an ATI Nano 17 force/torque sensor. The kinematic model of the actuation device is shown in Figure 3(C). A force controller is designed with an inner position control to meet

the calibration goal and requirement. The system was capable of controlling force with a settling time of less than 0.25 seconds, and tracking force trajectories with an interval of 0.3 seconds and step sizes of 0.1 N and 1 N·mm. Root-mean-squared errors are 0.02 - 0.04 N for forces and 0.39 N·mm for torques. The design and implementation of the actuation device was described in details in Ref. [9].

A Point Grey Flea video camera is used to take training images with the force reading from the force sensor. The force sensor reading is sampled at 1 kHz and the video camera works at 30 frame-per-second (fps).

A large wooden hand rest is used to support hand weight during calibration (Figure 3B). The wide base of the hand rest provides stability to the structure and allows for the addition of the unbalanced weight of the aluminum insert and finger restraint. The finger restraint has an L-shaped base that conforms to the tree shape of the wooden hand rest, capable of restraining a finger on either hand. The constraint prevents unwanted finger movement that may cause disturbances and noise in the collected data.

B. Motion and Force LfD Station

Both position and force are measured with two Point Grey Flea video cameras with 16 mm lenses, attached on two parallel half circle rails as shown in Figure 4A. The number of the cameras and the selection of the lenses on those cameras are decided based on the tasks. For a pick-and-place task using two fingers in a confined space, two cameras with relative narrow lenses are sufficient. More cameras would certainly provide better coverage for larger motions.



Fig. 4. (A) Grasping force demonstration platform; (B) The robot grasping and manipulation station.

The two cameras are configured to capture the images of the fingernail and surrounding skin areas of thumb and index finger along with the fingertip motion during a graspingplacing process. Software was developed to capture the fingertips of both index finger and thumb from both cameras simultaneously at 30 frame-per-second (fps). The fingertips are tracked and segmented from the background and then the calibrated force estimation model is used to estimate the force of both fingers during the tasks. The estimated force then combined with the fingertip motion to provide the training data.

C. Robot Grasping and Manipulation Station

As shown in Figure 4(B), the learning results from the demonstration are applied to a 6-DOF Fanuc L200IC robotic arm and a BarrettHand. Two ATI Nano 17 force sensors are embedded in two fingertips of the BarrettHand to measure fingertip force for force control. A Point Grey Firefly MV video camera is mounted on the wrist of the Fanuc robotic arm to provide visual feedback. Computer software is developed to control both the Fanuc robotic arm and the BarrettHand with the feedback from the Point Grey camera and the Nano 17 force sensors in real-time. The BarrettHand has a control loop running at 50 Hz and the Fanuc arm has control loop running at 100 Hz. The low level controller Fanuc R30A runs the inverse kinematics and motor control at 1000 Hz.

In software, a visual servoing controller uses the visual feedback from the Point Grey camera to guide the robot arm to the right grasping position. The Fanuc R30A controller then moves the robot arm to the right position related to the object. After the robot arm reaches the grasping position, a force input is generated from the learning from demonstration to control the BarrettHand to apply proper force on the object with the feedback from the force sensor.

IV. EXPERIMENTAL RESULTS

We have designed an experiment to verify the proposed grasping force learning from demonstration framework. For a pick-and-place task with two fingers, only the normal pinch force is controlled for grasping. Therefore we only measured and learned normal grasping force and compare the results with the readings from two thin-film force sensors. First a volunteer used the calibration station to calibrate his index finger and thumb with a designed normal force trajectory in the range of 0 to 8 N [9]. Three sets of training data were taken. One original image is shown in Figure 5A. Its fingertip image after segmentation and normalization is shown in Figure 5B.

A. Fingertip Force Measurement



Fig. 5. (A) The fingertip image after segmentation and normalization; (B) Original image taken with the camera at the calibration station.

To verify the calibration result, we used two sets of training data to train the GLS model in Section 1, and then verify it with the third data set. Figure 6 shows the verification result of the calibration. The root-mean-square (RMS) error is 0.3012 N; that is consistent with our previous finding [18].



Fig. 6. The estimation result of the GLS model on a new verification data set. (A) The verification result for an index finger; (B) the verification result for a thumb.

To compare the result of the force measurement from our imaging approach to the embedded sensor approach, we placed two FlexiForce force sensor on one object for a pickand-place task. FlexiForce is thin film force sensor which measures one dimensional pressure. The other types of force sensors are able to measure higher dimensional and more accurate forces, but are much larger in size. Applying the trained GLS model, the force during a pick-and-place task is estimated and displayed in Figure 7. For comparison, the force measured with the FlexiForce thin film force sensor is displayed in Figure 8, showing a large noise of the force measurement by FlexiForce sensors. According to Park (1999) [12], the repeatability of the FlexiForce senor ranges from 75% to 91% and it is worst for lowest forces.

Compared to regular force sensors, the imaging approach to measure force is not intrusive, though accurate enough for the applications of force measurement and grasp studies.



Fig. 7. Estimated force during a pick-and-place process from the images of the fingernail and surround skin. (A)The estimated force on the index finger; (B) The estimated force on the thumb.



Fig. 8. The measured force by the FlexiForce sensor. (A) the measured force on the index finger; (B) The measured force on the thumb.

B. Motion and Force LfD Model

Several objects are used to verify the grasping force learning from demonstration framework. For example, Figure 9 shows the demonstration grasping force reading from index finger of picking and placing a red pepper and position along the vertical direction of the end-effector. Most attentions are paid to the normal force and height because they are highly related to each other in the pick-and-place task. The dataset of force and motion during 3 trials of pick-and-place demonstration is modeled as in Section II-B. The GMM model is shown in Figure 10. The number of components is selected to be 4 so that the components naturally represent the mental intentions of the human user - grasp, lift, place and release.



Fig. 9. The demonstrated dataset.



Fig. 10. The GMM model result.

C. Apply Learning Result



Fig. 11. Generated trajectory with GMR.

The generated GMR trajectories (Figure 11) are input to the controller as control signals. Figure 12 is the execution result, showing the actual force and motion trajectories the robot applied. At the beginning, the robot increases the contact force with the object for holding the object, while the position remained before a certain force is achieved. Then the robot picks up the object away from the table and then place back the object, while the controller controls the robot to keep holding the object using a certain force. When the object is placed back to the table, the robot releases the object.



Fig. 12. The normal force and height results recorded from the robot execution

V. CONCLUSIONS

This paper presents a novel motion and force learning framework that allows a robot to learn fingertip force for a grasping and manipulation process from a human teacher. In addition to demonstrated motions, the fingertip force is measured with our previously developed image-based force estimation approach, which provides us important additional dimensions for the robot. A demonstration station has been setup with two video cameras to capture the forces on the fingertips of an index finger and a thumb. The Camera system is beneficial to grasping demonstration compared with force sensors, due to its nonintrusive yet accurate attributes. However, the camera system has its limitations, that is, the demonstrated motions have to be in range of the cameras. This limitation can be solved by adding more cameras.

The measured fingertip forces of two opposite fingers are modeled with GMM based machine learning approach. The learned force distributions are then used to generate fingertip force trajectories with a GMR approach. Instead of defining a certain grasping force value, force trajectories are used to control the robotic fingers in contact with an object to carry out a learned grasping and manipulation task.

The proposed framework and the force learning concept have been verified with several grasping tasks. The estimated force with our image-based fingertip force estimation approach appears to produce more consistent measurement than the FlexiForce thin film force sensor. The learning results allow the robot to apply proper force on the objects. The learned force are controlled using Hybrid motion and control so that both motion and force can be simultaneously controlled in separate space. In this study, we only recorded normal force on an index finger and a thumb. Human fingers are capable of controlling and exerting more complicated forces in higher dimensions.

Grasping force synergy studied in neuroscience can also be learned from demonstration and be taught to robots for various tasks. In the future, we plan to expand our framework to include shear force and torque. We will integrate the current setup with a 5DT dataglove and a motion tracking system to capture more detailed hand and arm motion. More fingers will be tracked and monitored to measure fingertip forces on them. We will also use our framework to study and learn grasping force and motion for more complicated manipulation tasks. Incorporated with motion and force, this framework would provide a way to learn the strategy of how human interact with the environment during manipulation.

VI. ACKNOWLEDGMENTS

This research was supported by a University of South Florida Seed Grant.

References

- A. Billard, S. Calinon, R. Dillmann, and S. Schaal, Robot Programming by Demonstration, *Handbook of robotics*, Springer, 2007.
- [2] J.D. Brown, J. Rosen, L. Chang, M. Sinanan, and B. Hannaford, "Quantifying Surgeon Grasping Mechanics in Laparoscopy Using the Blue DRAGON System", *Studies in Health Technology and Informatics - Medicine Meets Virtual Reality*, vol. 98, pp. 34-36, IOS Press, January 2004
- [3] S. Calinon and A. Billard, Incremental learning of gestures by imitation in a humanoid robot, Proceedings of the ACM/IEEE international conference on Human-robot interaction, 255-262,2007
- [4] D. Cohn, Z. Ghahramani, and M. Jordan. Active learning with statistical models. Artficial Intelligence Research, 4:129145,1996.
- [5] A. Dempster and N. L. D. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society B*, 39(1):1-38, 1977.
- [6] J.R. Flannagan, M.K.O. Burstedt, and R.S. Johansson, Control of Fingertip Forces in Multidigit Manipulation, *Journal of Neurophysiol*, 81, 1706-1717, 1999
- [7] L. Fogassi, V. Gallese, G. Buccino, L. Craighero, L. Fadiga, and G. Rizzolatti, Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study, *Brain*, vol. 124, no. 3, pp. 571586, Mar. 2001.
- [8] P. Jenmalm, S. Dahlstede, and R.S. Johansson, Fingertip Forces and Grasp Kinematics in Human Dexterous Manipulation, *Journal of Neurophysiol*, 84, 29842997, 2000
- [9] Y. Lin and Y. Sun, "5-D Force Control System for Fingernail Imaging Calibration," Proc. IEEE Intl. Conf. Robotics and Automation, pp. 1374-1379, 2011
- [10] M. Lopes, J. Santos-Victor, Visual learning by imitation with motor representations, *IEEE Transactions on Systems, Man, and Cybernetics*, *Part B*, 35 (3), 438-449, 2005.
- [11] M.N. McDonnell, M.C. Ridding, S.C. Flavel and T.S. Miles, Effect of human grip strategy on force control in precision tasks, *Exp. Brain Res.*, 161, 368-373, 2005
- [12] S. K. Park, The development of an exposure measurement system for assessing risk of work-related musculoskeletal disorders, Ph.D. thesis, University Park, PA: The Pennsylvania State University, 1999.
- [13] R. Platt, Learning Grasp Strategies Composed of Contact Relative Motions, HUMANOIDS, 2007, 49-56
- [14] M. A. Smith and J.F. Soechting, Modulation of Grasping Forces During Object Transport, J Neurophysiol, 93, 137-145, 2005
- [15] Y. Sun, J.M. Hollerbach, and S.A. Mascaro, "Measuring fingertip forces by imaging the fingernail," Proc 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 125-131, 2006.
- [16] Y. Sun, J.M. Hollerbach, and S.A. Mascaro, "Dynamic features and prediction model for imaging the fingernail to measure fingertip forces," Proc IEEE Intl. Conf. Robotics and Automation, pp. 2813-2818, 2006.
- [17] Y. Sun, J.M. Hollerbach, and S.A. Mascaro, "EigenNail for finger force direction recognition," Proc. IEEE Intl. Conf. Robotics and Automation, pp. 3251-3256, 2007.
- [18] Y. Sun, J. M. Hollerbach and S. A. Mascaro, Predicting Fingertip Forces by Imaging Coloration Changes in the Fingernail and Surrounding Skin, *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 10, pp. 2363-2371, 2008.
- [19] R.M. Voyles, P.K. Khosla, A multi-agent system for programming robots by human demonstration, *Integrated Computer-Aided Engineering*, 8(1), 59-67, 2001.