Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Grip Force Perception Based on dAENN for Minimally Invasive Surgery Robot*

Yongchen Guo, Bo Pan, Yili Fu
State Key Laboratory of Robotics and System
Harbin Institute of Technology
Harbin, Heilongjiang,China
panbo4034@163.com

Max Q.-H. Meng
Department of Electronic Engineering
Chinese University of Hong Kong
Hong Kong, China
qhmeng@ee.cuhk.edu.hk

***Abstract*— Although robot assisted minimally invasive surgery brings the gospel to patients, force perception is gone. Among all the contact forces during surgery, the instrument grip force plays the most important role. In this paper, a grip force perception method based on denoising AutoEncoder Neural Network (dAENN) is proposed. The method utilizes sensor data including encoder readings and motor current over a time window as the input of dAENN for sufficient information. An Artificial Neural Network (ANN) is then introduced as a machine learning tool to learn the nonlinear mapping between the compact features and grip force labels. Feature extraction is first introduced into grip force perception problem in this paper. Experiment results shows adequate expressive capability of the extracted coding as well as the superior grip force perception performance over several popular data-based methods under the same dataset.**

***Index Terms*— grip force estimation, dAENN, surgical instrument, minimally invasive surgery.**

I. Introduction

Minimally invasive surgery (MIS) is recently improved by surgical robots which extend the surgeon's accessibility and accuracy of movement, restrain the hands' tremor, and relieve the fatigue [1]. However, few commercial surgical robot systems can provide surgeon with the haptic feedback which can improve the surgeon's surgical experience and reduce the misoperation [2], [3], [4]. Accurate haptic perception is the first step for haptic feedback. Among different types of haptic perceptions, grip force perception improves the control performance of teleoperation significantly [5].

Due to limited space, the surgical instruments are usually design compactly. In other words, the integration of extra force sensors should be considered at the beginning of the whole design. In this way, there are several serious challenges to perceive grip force clinically such as limited space, sterilization problem, etc [6], [7], which are the reasons why grip force perception is still not used widespread.

This paper proposes a novel grip force perception method based on denoising AutoEncoder Neural Network (dAENN). Only the actuating motor current and encoder readings are used as the learning inputs. The contributions are as follows: 1) This paper for the first time conducts feature extraction on grip force perception problem. 2) dAENN is introduced to extract features as a preliminary trial, and it reduces the reconstruction error by 14.76% compared with the classic AENN. 3) The experiment results show superior performance over several state-of-the-art methods under the same dataset.

II. Related Work

Up to now, there are three types of schemes to perceive grip force on the instrument tip: sensor-based ones, model-based ones, and data-based ones.

A. Sensor-based Grip Force Perception

Mounting additional sensors on the instrument is a straightforward way. Uikyum Kim designed a force sensor integrated surgical instrument [8]. It can directly sense the normal and shear forces at surgical instrument tips with a Root Mean Squared Error (RMSE) around 0.1 N. Man Bok Hong and Yung-Ho Jo designed a novel 2 degree-of-freedom (DOF) compliant forceps [9]. They considered the forceps itself as a force sensor, and eventually achieved an RMSE of 0.1 N. Although these force sensors can sense the grip force accurately, sterilization prevents them from widespread usage. Sterilization at the instrument tip can cause damage to the force sensors, making the surgical instrument not reusable. Chao He mounted a force sensor on the pulley to sense the cable tension, and calculated the grip force with a Maximum Error (ME) of 0.4 N [10]. This method solved the instrument sterilization problem because the sensors are placed away from the sterilization area, although it has a relatively higher error level. However, a surgical instrument must be abandoned after being used for several times, and adding extra sensors on it can increase the surgery costs.

*This work is partially supported by National Natural Science Foundation of China (Grant No. 61803341), Stake Key Laboratory of Robotics and Systems (Grant/Award No. SKLRS201601C), and Heilongjiang Postdoctoral Scientific Research Foundation (LBH-Q17070, LHB-Q17068). The corresponding author is Bo Pan (panbo4034@163.com).

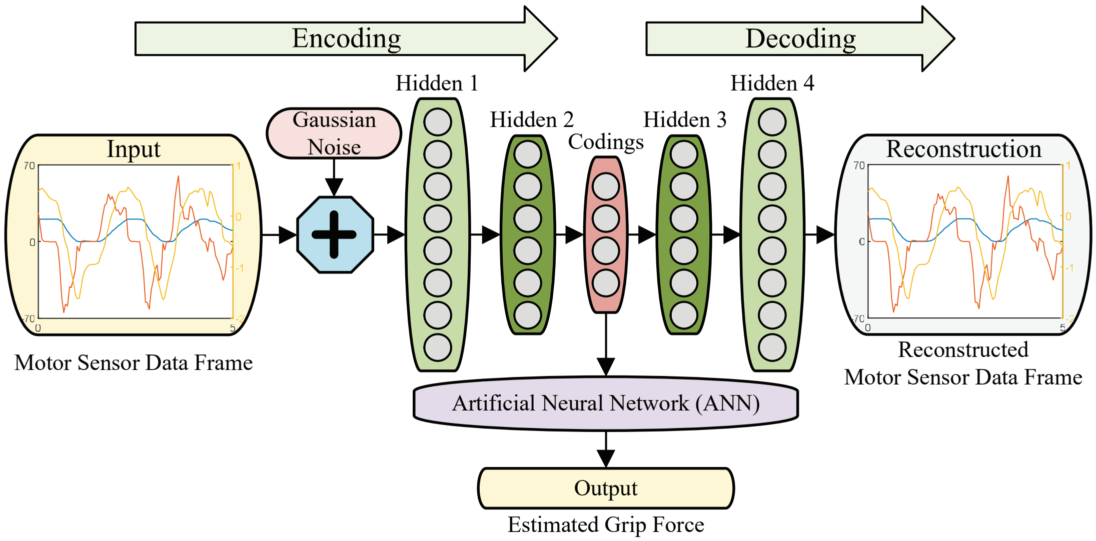


Fig. 1. Grip Force Estimation Method Based on dAENN

Therefore, there are still many obstacles to the widespread use of real sensors.

B. Model-based Grip Force Perception

Model-based methods depend on the dynamic modeling of surgical instrument. A fuzzy-logic-based Sliding Mode Control with Sliding Perturbation Observers (SMCSPO) was developed by Nahian Rahman and Min Cheol Lee, which is based on the dynamic model to estimate the grip force (considered as a perturbation) [11], [12]. A model-based estimator was developed by Yangming Li using Unscented Kalman Filter (UKF) and a bounding filter [13]. Furthermore, a model-based estimator estimating cable tension was developed by Mohammad Haghighipanah [14], and the grip force was calculated by the estimated cable tension. The performance of these model-based methods mainly depends on the accuracy of the dynamic model, but the nonlinearity in cable-driven system is very difficult to model. So the accuracy of model-based methods up to now is not satisfying enough. Although [15] proposed two simplified quasi-static models to model frictions and compliance of the cables and achieved a sufficiently close estimation of the output torque, the complexity of the model prevents it from online identification. Therefore, model-based methods have limitations.

C. Data-based Grip Force Perception

Recently, data-based methods emerge and have the potential to solve the grip force estimation problem. A data-based scheme was put forward by Yangming Li [16]. This method utilized pre-measured dataset including motor current and encoder readings to learn the sophisticated model of the surgical instrument. It illuminates the potential advantages of data-based methods, and Trevor K. Stephens tested several machine learning algorithms and found out that a single-layer neural network with 30 nodes in the hidden layer has the best performance [17]. Although with improvements, these methods neglect the feature extraction but using raw sensor data as the inputs of machine learning algorithms. It is necessary to extract features considering the machine learning algorithms they use only has little capability of feature extraction.

III. Grip Force Perception based on dAENN

The performance of data-based grip force perception methods surpasses the traditional model-based ones significantly. That is why the co-authors pay much attention on this kind of methodology. Theoretically, machine learning techniques substitutes the dynamic model to calculate the instrument grip force. Hence, whether the learning algorithm can learn the true dynamic model from the training set is the core issue.

However, data-based methods up to now only use raw sensor data as the input of a machine learning algorithm with little feature extraction capability. This paper attempts to pre-extract features from sensor data over a time window frame. The compact features then becomes the input of a typical machine learning algorithm.

A novel grip force estimation method based on denoising AutoEncoder Neural Network (dAENN) is proposed in this paper. Fig. 1 demonstrates the entire process of the method. The input data comprise position, velocity, and current of the actuating motor. Added with Gaussian noise, the input data is encoded into compact codings which can be reconstructed using the decoder. The decoder is actually an inverse operation of the encoder. Root Mean Squared Error (RMSE) between the reconstructed sensor data and the input with noise is chosen to be the loss function of dAENN. After

TABLE I

NETWORK STRUCTURES

dAENN		ANN	
Layer	Number of Neurons	Layer	Number of Neurons
h1	20	h1	30
h2	15	h2	15
c	9	output	1
h3	15		
h4	20		

training dAENN, the codings will be input into an Artificial Neural Network (ANN) whose output is trained to be the estimated grip force. In order to cover enough information, the input of dAENN is not only the current time sensor data but also several history data as is depicted in Fig. 1. That is to say, sensor data over a time window frame is the input of our method. As a preliminary attempt, the sample time is 50 ms and the time window is 500 ms. The performance when sample time and time window can vary will be studied in the future.

A. Sensor Data Preprocessing

In order to speed up the network optimization process and improve accuracy, normalization is conducted before training. In this work, a standard Max-Min normalization is used because it is not necessary to use distance to measure similarity. The normalization is as follows

$$x_N^i = \frac{x_U^i - v^i}{u^i - v^i} \tag{1}$$

where $i = 1, 2, 3$ represents position, velocity, and current of the actuating motor respectively. x_N^i is the normalized sensor data, and x_U^i is the unnormalized sensor data. v^i is the lower normalization limit, and u^i is the upper normalization limit. v^i and u^i are chosen to be a little bit larger than the range of training sensor data so that the efficient normalization is guaranteed in the testing set. In this way, x_N^i can be normalized into [0, 1] which is suitable for dAENN.

B. Feature Extraction & Grip Force Perception

The proposed method is divided into two phases: Coding Phase (CP) and Perception Phase (PP). In CP, normalized sensor data x_N^i is added with gaussian noise and becomes a partially destroyed version $\tilde{x}_N^i = x_N^i + \zeta_G^i$. This process can force the Auto Encoder to learn the more robust compact coding of the learning inputs. In this paper, the magnitude of ζ_G^i is chosen to be 0.01 after normalization. After that, we can obtain the 1st hidden layer using the following formula

$$\boldsymbol{h_1} = s\left(\mathrm{W}_1 \boldsymbol{\tilde{x}_N} + \mathrm{b}_1\right) \tag{2}$$

where $\boldsymbol{h_1}$ is the vector of the 1st hidden layer, W_1 and b_1 are the weights and bias respectively. $s(\cdot)$ is sigmoid activate function. Sigmoid function is chosen because it is convenient to find the derivative. The calculation in other layers are similar to Eq. (2). Due to the decoder is the opposite operation of the encoder, the weight matrix satisfies $\mathrm{W}_j = {\mathrm{W}_{7-j}}^{\mathrm{T}}$, $j = 1, ..., 6$. For simplicity, the encoding and decoding processes can be represented as

$$\boldsymbol{c} = f_\theta\left(\boldsymbol{\tilde{x}_N}\right), \theta = \{\mathrm{W}_j, \mathrm{b}_j\}\ (j = 1, 2, 3) \tag{3}$$

$$\boldsymbol{z} = g_{\theta'}\left(\boldsymbol{c}\right), \theta' = \{\mathrm{W}_j, \mathrm{b}_j\}\ (j = 4, 5, 6) \tag{4}$$

where $\boldsymbol{c}$ is the extracted coding, and $\boldsymbol{z}$ is the reconstructed sensor data. The square error between $\boldsymbol{\tilde{x}_N}$ and z is the loss function to measure the difference between noised input and reconstructed sensor data

$$L_{dAENN}\left(\boldsymbol{\tilde{x}_N}, \boldsymbol{z}\right) = \left\|\boldsymbol{\tilde{x}_N} - \boldsymbol{z}\right\|^2 \tag{5}$$

In PP, the already trained coding $\boldsymbol{c}$ is put into an ANN whose output is the perceived instrument grip force $\hat{F}_g$. There are in total 3 layers in the ANN. The first two layers employ sigmoid function as the activate function, while the third layer (output layer) has no activate function because grip force is not limited in [0, 1].

The loss function of the ANN is the square error between the ground truth F_g and the perceived value $\hat{F}_g$

$$L_{ANN}\left(\boldsymbol{\hat{F}_g}, \boldsymbol{F_g}\right) = \left\|\boldsymbol{\hat{F}_g} - \boldsymbol{F_g}\right\|^2 \tag{6}$$

The structure of the dAENN and ANN is listed in Table I. The hyper-parameters are determined by trial and error.

Our dAENN and ANN use Adam optimizer to optimize the parameters, and the learning rate is set to be 0.005.

IV. EXPERIMENT AND DISCUSSION

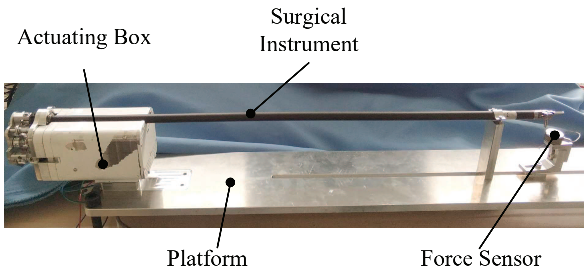
A. Hardware Introduction

Fig. 2 shows the experiment system we establish. Motors and drivers in the actuating box drive the four degree-of-freedoms of the surgical instrument. The surgical instrument of HuaQue III Minimally Invasive Surgery Robot System is mounted on the platform. The master hand gripper shown in Fig. 2b is manipulated mannually to control the instrument gripper when collecting datasets.

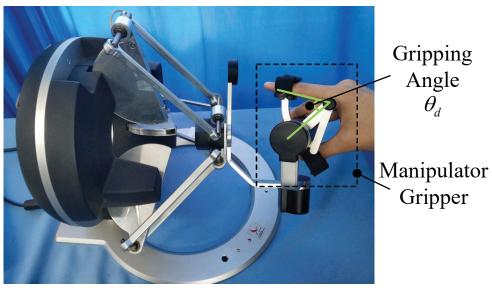
A discrete PID controller is utilized to let the instrument gripper follow the master hand. The controller is as follows

$$u\left(kT\right) = K_p\left(e\left(kT\right) + \frac{T}{T_i}\sum_{j=0}^{k} e\left(jT\right) + \frac{T_d}{T}\left(e\left(kT\right) - e\left(kT - T\right)\right)\right) \tag{7}$$

where $u\left(kT\right)$ is the control law which is the desired velocity $\dot{\theta}_d$ in our case, $e = \theta_d - \theta$ is the position error of instrument gripper, and T is the control cycle. In our case, control cycle T is 10 ms, while the sample time of sensor data is 50 ms in order to include a larger period of time.



(a)



(b)

Fig. 2. 3D model of the experiment platform: (a) Overall view of the platform. (b) Detailed transmission mechanism of the platform.

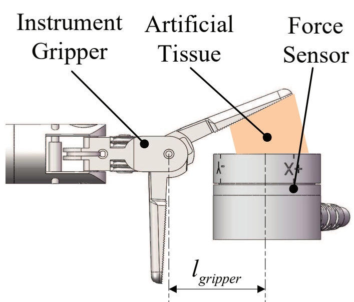


Fig. 3. Grip Force Estimation Method Based on dAENN

Near the instrument tip, a one-dimension force sensor is fixed on the platform. Fig. 3 demonstrates the way we collect datasets. An artificial tissue is gripped between the instrument gripper and the force sensor. When collecting grip force, $l_{gripper}$ remains the same. EPE foam wedge is employed as the artificial tissue in this work. It is hypothesized that the gripping action is a quasi-static process because surgeons usually grip objects slowly in surgery.

There are in total three datasets we collect, training dataset, validation dataset, and testing dataset which includes 5392, 1637, and 1120 samples respectively. Training dataset is to train the dANEE and ANN, validation dataset is to help determine network structure, and testing dataset is to verify the performance of our method. When collecting datasets, the master gripper is operated randomly including zero-load case and gripping case. The reading of force sensor is traversed from the lowest to the highest in gripping case, while in zero-load case the opening angle is varied.

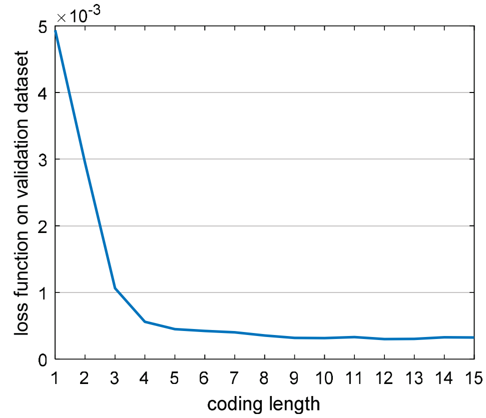


Fig. 4. Relationship Between Loss Function and Coding Length

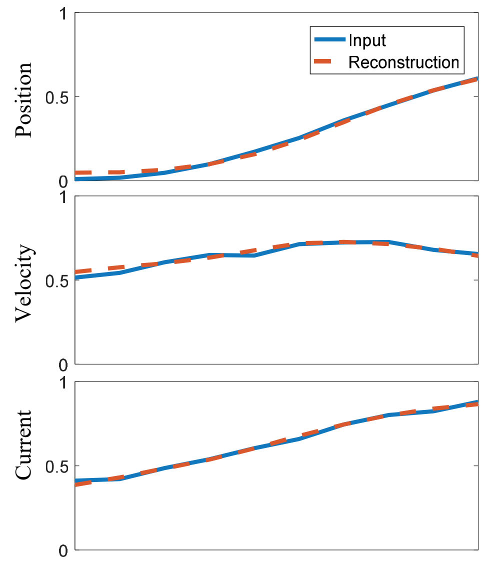


Fig. 5. Sensor Data Reconstruction Result

B. Feature Extraction Results

The dAENN aims to extract compact coding from the raw sensor data frame. In order to decide the coding capacity, we vary the length of the coding vector from 1 to 15, and observe the loss value on the validation dataset. Fig. 4 depicts the relationship between loss function and coding length. Obviously, the loss value descends considerably when the coding length grows. Once the length grows up to 4, the loss function begins to flatten, and after growing up to 9, loss value tends to be stable. This reveals that at least 4 features are needed to represent raw sensor data roughly, and at least 9 features are required for better representation. To avoid redundancy, the coding length is chosen to be 9 as a compact feature set.

The comparison between raw sensor data frame and the reconstructed one is demonstrated in Fig. 5. The reconstructed data trend is consistent with the original data as a whole, and the square error on validation set is 3.1765×10^{-4} , which indicates that the extracted compact features manifest sufficient expressive power. To reveal the contribution of the gaussian noise we add, a classic AENN with the same structure is established, and the square error on validation set is 3.7267×10^{-4} which is higher than dAENN. That is to say, the use of dAENN reduces the reconstruction error by 14.76%.

C. Grip Force Perception Results

After the compact coding is obtained, it is sent to the ANN to perceive grip force. A fragment of perception result is shown in Fig. 6. Overall, the perceived grip force matches the true grip force well although with fluctuation. The mean average error (MAE), max error (ME), and standard deviation (SD) on the testing dataset is calculate to quantify the performance

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|\bar{F}_P^i - F_{GT}^i\right| \tag{8}$$

$$\mathrm{ME} = \max\left(\left|\bar{F}_P^i - F_{GT}^i\right|\right) \tag{9}$$

$$\mathrm{SD} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(E^i - \bar{E}\right)^2} \tag{10}$$

where m is the sample amount, $\bar{F}_P$ is the perceived mean value of grip force, F_{GT} is the ground truth, $E^i = \left(\left|\bar{F}_P^i - F_{GT}^i\right|\right)$ is the perception error, and $\bar{E} = \mathrm{mean}\,(E)$ is the mean perception error. The quantified performance is summarized in Table II.

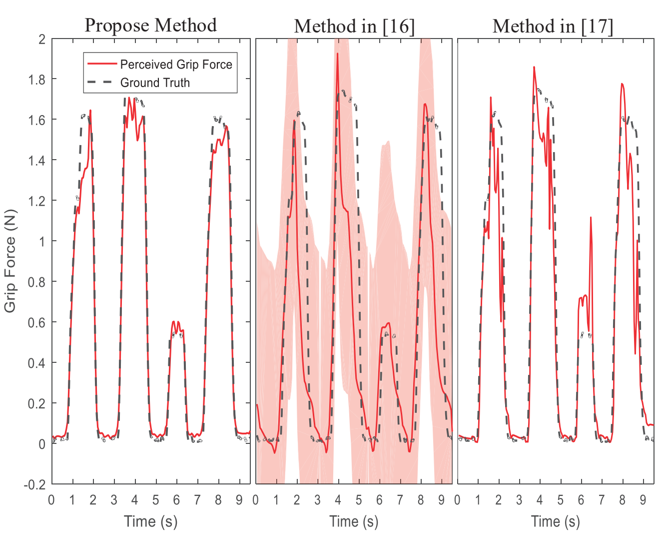


Fig. 6. Grip Force Perception Results

TABLE II
COMPARISONS WITH STATE-OF-THE-ART METHODS

	Proposed method	Data-based Method [16]	Data-based Method [17]	Model-based Method [13]
MAE (N)	0.0641	0.2719	0.1418	0.1882
ME (N)	0.3107	1.3478	1.1112	1.0589
SD (N)	0.0852	0.4169	0.2301	0.1991

In order to compare our method with the state-of-the-art data-based methods, we reproduce the methods in [16] and [17] on our datasets because datasets can influence the performance significantly. For instance, the difference in maximum grip force, number of included gripping period, and the proportion of zero-load case in testing set will all make the perception results inconsistent. The visualized perception results are shown in Fig. 6, and the quantified accuracy can be found in Table II. Datasets only have little influence on model-based methods, so the results in [13] are directly cited in Table II.

D. Discussion

According to Fig. 6 and Table II, the performance of our proposed grip force perception method in our datasets surpasses the popular methods. In particular, the performance of [16] and [17] obviously drops in the second half of the gripping period as is depicted in Fig. 6. It is because the cable-pulley mechanism of surgical instrument has hysteresis which influences the perception performance. While the performance of our method is stable during the entire gripping process, and this thanks to the feature extraction before training.

Note that, our reproduced results in [16] is not as good as what was reported. It is because the maximum perception value of [16] is only around 1 N, while the maximum grip force is around 1.8 N in our datasets. Besides, in our testing dataset, the proportion of the hysteresis phase is larger than what in [16]. These two reasons are why our produced performance of [16] is worse than it reported. We consider that the reasons for [17] are similar although the datasets of it were not described clearly.

V. CONCLUSION

In this paper, a novel grip force perception method for surgical instrument based on dAENN is proposed. Feature extraction technique is introduced for the first time. dAENN reduces the reconstruction error by 14.76% compared with classic AENN. Experiment results reveal that the extracted coding manifest sufficient expressive power, and our method outperforms several state-of-the-art data-based and model-based methods with lower errors under our datasets.

In the future, the performance when sample time and time window can vary will be studied. Moreover, hyper-parameters including network structures and gaussian noise magnitude should also be chosen more reasonably. Last but not least, more appropriate feature extraction strategy will be investigated in order to further improve the grip force perception performance.

ACKNOWLEDGMENT

This project is sponsored by National Natural Science Foundation of China (Grant No. 61803341), Stake Key Laboratory of Robotics and Systems (Grant/Award No. SKLRS201601C), and Heilongjiang Postdoctoral Scientific Research Foundation (LBH-Q17070, LHB-Q17068).

REFERENCES

[1] B. Siciliano and O. Khatib, *Springer handbook of robotics*. Springer, 2016.

[2] C.-H. King, M. O. Culjat, M. L. Franco, C. E. Lewis, E. P. Dutson, W. S. Grundfest, and J. W. Bisley, "Tactile feedback induces reduced grasping force in robot-assisted surgery," *IEEE transactions on haptics*, vol. 2, no. 2, pp. 103–110, 2009.

[3] J. C. Gwilliam, M. Mahvash, B. Vagvolgyi, A. Vacharat, D. D. Yuh, and A. M. Okamura, "Effects of haptic and graphical force feedback on teleoperated palpation," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 677–682.

[4] C. E. Reiley, T. Akinbiyi, D. Burschka, D. C. Chang, A. M. Okamura, and D. D. Yuh, "Effects of visual force feedback on robot-assisted surgical task performance," *The Journal of thoracic and cardiovascular surgery*, vol. 135, no. 1, pp. 196–202, 2008.

[5] R. P. Khurshid, N. T. Fitter, E. A. Fedalei, and K. J. Kuchenbecker, "Effects of grip-force, contact, and acceleration feedback on a teleoperated pick-and-place task," *IEEE transactions on haptics*, vol. 10, no. 1, pp. 40–53, 2017.

[6] P. Puangmali, K. Althoefer, L. D. Seneviratne, D. Murphy, and P. Dasgupta, "State-of-the-art in force and tactile sensing for minimally invasive surgery," *IEEE Sensors Journal*, vol. 8, no. 4, pp. 371–381, 2008.

[7] Z.-t. WANG, L.-t. WANG, and J.-m. LEE, "Overview on force sensing techniques in robot-assisted minimally invasive laparoscopic surgery," *DEStech Transactions on Computer Science and Engineering*, no. aita, 2017.

[8] U. Kim, D.-H. Lee, W. J. Yoon, B. Hannaford, and H. R. Choi, "Force sensor integrated surgical forceps for minimally invasive robotic surgery," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1214–1224, 2015.

[9] M. B. Hong and Y.-H. Jo, "Design and evaluation of 2-dof compliant forceps with force-sensing capability for minimally invasive robot surgery," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 932–941, 2012.

[10] C. He, S. Wang, H. Sang, J. Li, and L. Zhang, "Force sensing of multiple-dof cable-driven instruments for minimally invasive robotic surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 10, no. 3, pp. 314–324, 2014.

[11] N. Rahman and M. C. Lee, "Actual reaction force separation method of surgical tool by fuzzy logic based smcspo," *International Journal of Control, Automation and Systems*, vol. 13, no. 2, pp. 379–389, 2015.

[12] S. M. Yoon, M. C. Lee, and C. Y. Kim, "Sliding perturbation observer based reaction force estimation method of surgical robot instrument for haptic realization," *International Journal of Humanoid Robotics*, vol. 12, no. 02, p. 1550013, 2015.

[13] Y. Li, M. Miyasaka, M. Haghighipanah, L. Cheng, and B. Hannaford, "Dynamic modeling of cable driven elongated surgical instruments for sensorless grip force estimation," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 4128–4134.

[14] M. Haghighipanah, M. Miyasaka, and B. Hannaford, "Utilizing elasticity of cable-driven surgical robot to estimate cable tension and external force," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1593–1600, 2017.

[15] F. Anooshahpour, I. G. Polushin, and R. V. Patel, "Quasi-static modeling of the da vinci instrument," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 1308–1313.

[16] Y. Li and B. Hannaford, "Gaussian process regression for sensorless grip force estimation of cable-driven elongated surgical instruments," *IEEE robotics and automation letters*, vol. 2, no. 3, pp. 1312–1319, 2017.

[17] T. K. Stephens, J. J. ONeill, N. J. Kong, M. V. Mazzeo, J. E. Norfleet, R. M. Sweet, and T. M. Kowalewski, "Conditions for reliable grip force and jaw angle estimation of da vinci surgical tools," *International journal of computer assisted radiology and surgery*, pp. 1–11, 2018.