2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Vision-Sensorless Bin-Picking System Using Compliant Fingers with Proximity Sensors

Michihisa Ohara, Keisuke Koyama and Kensuke Harada

***Abstract*— In this study, we propose an exploratory bin-picking system with less installation space than a human worker. The robot adjusts the arm tip position based on quasi-static deformations of the compliant fingers. It also estimates the number of grasped objects by frequency analysis of the dynamic deformations. The proposed system has two advantages because it performs bin-picking using only deformations. The first is that all control can be performed on small on-board computer board (Raspberry Pi 4B). Second, there is no need to place and mount a 3D vision sensor. The control system can also be reconfigured without any learning time when shape of the picked objects changes. In bin-picking experiments, we confirmed that our system achieved a picking success rate of over 85 % and a bin picking tact time within 30 s for two types of metal bolts.**

I. INTRODUCTION

Compact bin-picking robots are required in a cell production system. The arm itself, a 3D vision sensor, and a control PC must be placed in the same space as a human worker. However, the current method requires a 3D vision sensor and a PC with GPU, it requires larger installation space than a human worker. This study proposes a compact bin-picking system that does not use a 3D vision sensor and a PC with GPU. Our method uses simple control equations based on the deformation of the robot gripper fingertip to adjust the grasping positions, and estimate the number of objects grasped. The controller is only a Raspberry Pi 4B, which allows bin-picking with less installation space than a human worker.

Many currently existing bin-picking robots lack the sense of force and tactile sensing in their fingers, and they often rely on high-precision recognition of the image taken from a 3D vision sensor for picking tasks [1]–[3]. The combination of machine learning and the 3D vision sensor makes it possible to accurately grasp a target position of a object. However, there is a problem that the system layout size increases. The 3D vision sensor is often fixed to a tripod or aluminum frames and placed around the robot. Basically, a 3D vision sensor and these frames are needed for each robot, and it is difficult to place several robots closely together [2]. On the other hand, in the case of 3D vision sensors are mounted on the wrist of the robot arm, there is no need to mount a tripod or aluminum frame around the robot [1], [3]. However, the risk of the 3D vision sensor colliding with a shelf increases when inserting the arm tip into a narrow space such as inside a shelf. Furthermore, to quickly process the high-resolution depth information from 3D vision sensors and to run large-scale mashine learning models, a high-performance PC with GPU is necessary. This requirement also limits the number of robots that can be placed. Therefore, it is crucial to develop compact bin-picking systems that do not rely on 3D vision sensors or machine learning.

Michihisa Ohara, Keisuke Koyama and Kensuke Harada are with the Graduate School of Engineering Science, Osaka University, Osaka 560-0043, Japan oohara@hlab.sys.es.osaka-u.ac.jp, koyama@sys.es.osaka-u.ac.jp, harada@sys.es.osaka-u.ac.jp

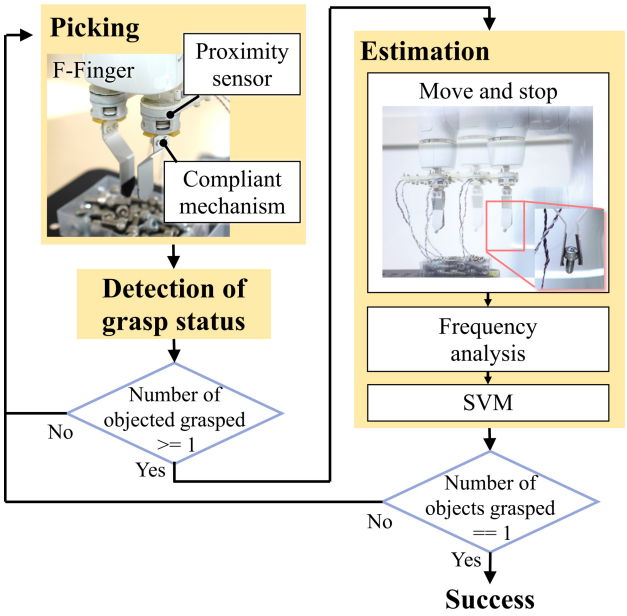


Fig. 1. Overview of the Proposed System. Picking: Vision-sensorless exploratory picking is performed using an F-Finger. Grasp Determination: Grasp success is determined based on the amount of deformation at the fingertips. Estimation: The number of grasped objects is estimated using support-vector machines (SVM) based on the frequency characteristics of in-motion vibrations generated during move and stop.

In this study, we propose an exploratory bin-picking system without using a 3D vision sensor (Fig. 1). The system consists of a compliant fingers and short-range distance and tilt angles sensors (proximity sensors) that measure the deformation of the mechanism. We refer to this finger unit as the Floating-Finger (F-Finger). By using the quasi-static deformation of F-Finger, the position of the robotic hand is adjusted during bin-picking. Dynamic deformation data is used to estimate the number of objects grasped. Since feedback control is performed based on only deformations of the fingers, the control equations can be implemented on a small onboard computer, such as Raspberry Pi 4B. In addition, as no 3D vision sensors are required in the

979-8-3315-3161-4/25/$31.00 ©2025 IEEE

environment or on the wrist of robot arm, this system has the advantage that it can be placed in a smaller space than a human worker.

II. RELATED WORKS

A. Bin-picking using 3D vision sensor or tactile sensor

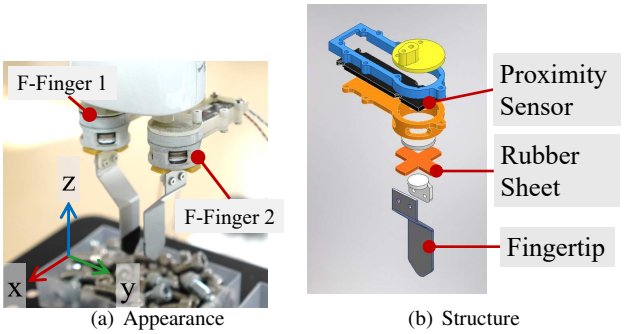
For example, some approach to industrial bin picking estimates the grasp position and orientation from depth images to pick up a target object while avoiding collisions [1], [3]. Additionally, another approach involves grasping one or more than one target objects at once, transferring them to a stage, and isolating them [2], [4]. These methods enable high-precision picking of objects that were previously challenging for robots, such as bolts around one centimeter in length or objects easily get entangled. However, since these methods rely on 3D vision sensors, they require the installation of 3D vision sensors on the robot's wrist or the setup of scaffolding in the environment to mount 3D vision sensors. Moreover, the estimation of grasp position and orientation utilizes machine learning, which requires a computer equipped with GPUs.

As an approach that doesn't rely on 3D vision sensors, several methods use tactile sensors [5], [6]. These methods employ fingers equipped with tactile sensors on their fingertips and isolate target objects by performing in-hand manipulation based on tactile information. However, these methods use a vision-based tactile sensor and machine learning and require a large PC with a GPU. Thus, many conventional bin-picking systems demand significant space for the installation of 3D vision sensors and computers. To address this issue, we propose a bin-picking system operated by a small computer board, such as a Raspberry Pi 4B.

B. Estimation of the number of grasped objects

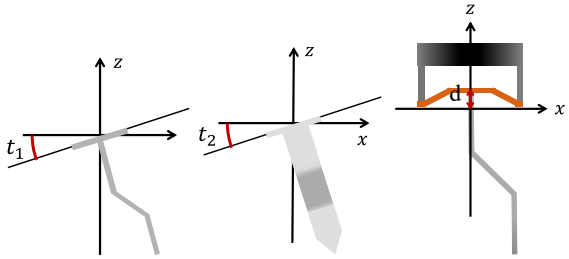
There are a few methods for estimating the number of small objects grasped (such as bolts approximately 1 cm in length). For example, Ishige et al. [5] proposed a method using a gripper equipped with tendon-driven flexible fingers and tactile sensors on the fingertips. In this method, after grasping the target objects, the gripper stabilizes the grasp posture by performing a sliding motion between the fingertip surface. The number of objects grasped is then estimated based on tactile information. This method achieves nearly 90 % accuracy in estimating the number of objects grasped. However, these methods require additional actuators on the fingertips.

On the other hand, several methods exist for estimating the physical properties (e.g., mass) of grasped objects larger than the gripper, such as bottles containing liquid. For instance, a method known as dynamic tactile sensing estimates the objects' physical properties by measuring vibrations generated when the object is shaken, and the vibrations are measured by using tactile sensors on the fingertip [7], [8]. This implies that utilizing the response generated by applying vibrations to the gripper is effective to estimate the number of objects grasped based on their mass. In this study, we propose a method for estimating the number of objects grasped based on vibrations during transport. This approach adopts a gripper with a compliant mechanism at the base of its fingers, enabling implementation without additional actuators on the fingertips.



(a) Appearance (b) Structure

Fig. 2. Overview of F-Finger. A proximity sensor is installed at the base of the compliant finger mechanism using a compliant sheet.



(a) t_1 (b) t_2 (c) d

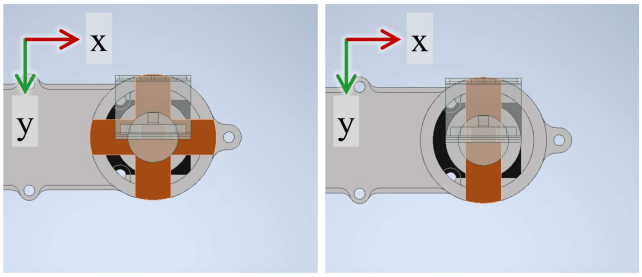
Fig. 3. Three outputs of the proximity sensor. We define t_1, t_2 and d output from F-Finger1 as t_{11}, t_{12} and d_1. We define the outputs from F-Figner2 in the same way.

III. PROPOSAL METHOD

A. Floating-Finger

In this study, we propose a customizable compliant finger mechanism (F-Finger: Floating-Finger) to realize both arm tip positioning and grasp count estimation based on vibration measurement. Overview and structures are shown in Fig. 2. The coordinates in Fig. 2(a) indicate the coordinate system fixed to the hand. F-Finger is equipped with proximity sensor at the base of the finger, a compliant sheet with a reflector surface for proximity sensing attached beneath it, and a fingertip is connected below the compliant sheet. The sensor measure the tilt of the reflector around x and y axis (t_1, t_2) ,as well as the amount of depression in the z direction (d) at the reflector (Fig. 3).

When the fingertip of the F-Finger contacts an object, deformations in the compression and torsion directions occur at the base of the fingertip. At this moment, twisting occurs on the surface of the compliant sheet, generating a restoring force that returns the sheet to the original intermediate position. This restoring force enables the grasping of an object and maintains contact between the fingertip and the object. In addition, using proximity sensors, the deformation of the fingertips is measured when they come into contact with small randomly piled objects, enabling feedback control of the hand.



(a) Cross-shape (b) Bar-shape

Fig. 4. Shape of compliant sheet

For the proximity sensor, a customized version of a commercially available, compact, and thin distance and angle measuring module, TK-01 [1] is used for research purposes. The TK-01 processes infrared reflectance intensity information from multiple elements using its proprietary high-speed, high-precision machine learning model, enabling high-resolution sensing of distances and angles. The measurement range for tilt is $-4.0\,^\circ$ to $4.0\,^\circ$, with a resolution of $0.1\,^\circ$, and for compression amount, the measurement range is $0.15\,\mathrm{mm}$ to $2.0\,\mathrm{mm}$, with a resolution of $0.01\,\mathrm{mm}$. The measurement cycle duration is $5\,\mathrm{ms}$. A machine learning model is built into a small board ($55\,\mathrm{mm} \times 60\,\mathrm{mm}$), so it can be used without a PC equipped with GPU.

The compliant sheet is made of a commercially rubber sheet[2]. The fingertips are made of aluminum plates and the other parts are made of PLA made with a 3D printer. Therefore, the characteristics of the compliant-finger mechanism can be easily altered by changing the thickness or shape of the compliant sheet. In this study, to maintain a sufficient grasping force (restoring force around the x-axis in Fig. 2(a)), we used a sheet thickness of $3.0\,\mathrm{mm}$ and two shapes: cross and bar (Fig. 4).

B. Bin-Picking process overview

This section describes the bin-picking process for the small objects. The process is divided into three steps: 1) Picking, 2) Detection of the grasp status, and 3) Estimation of the number of grasped. Each step is interconnected, as illustrated in Fig. 1. First, the starting point for picking is randomly determined for bins of known positions and sizes. In the "Picking" step, grasping of small objects is performed using a five-degree-of-freedom positioning control with an F-Finger. Through "Detection of grasp status" and "Estimation of the number of objects grasped", high-precision bin picking, which is the task of grasping only one object, is achieved. The subsequent sections discuss these methods in detail.

C. 5DoF Control

In feedback control, the position and orientation of the hand are adjusted to make the fingertip deformations symmetrical [9]. The controlled degrees of freedom are the deformations along the x-, y-, and z-axis (D_x, D_y, and D_z) and the rotations around the x- and z- axes (R_x and R_z), as shown in Fig. 5. The position and orientation of the hand are controlled based on the command values generated using Eqs. 1 to 5.

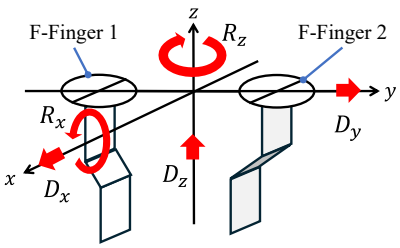


Fig. 5. The coordinate system fixed to the hand and the degrees of freedom controlled during picking.

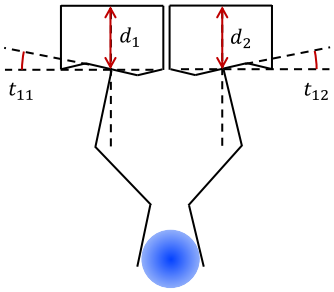


Fig. 6. Detection of grasp status. Using the parameters of the figure, we detect whether the object is grasped.

$$\Delta D_x = K_x(t_{21} + t_{22}) \quad (1)$$
$$\Delta D_y = K_y(t_{11} + t_{12}) \quad (2)$$
$$\Delta D_z = K_z(d_1 + d_2 - d_{ref}) \quad (3)$$
$$\Delta R_x = K_{rx}(d_1 - d_2) \quad (4)$$
$$\Delta R_z = K_{rz}(t_{21} - t_{22}) \quad (5)$$

where t_{11} and t_{12} represent the proximity sensor measurements t_1 for F-Finger1 and F-Finger2, respectively. Similarly, t_{21} and t_{22} represent proximity sensor measurements t_2, and d_1 and d_2 represent proximity sensor measurements d. The effects of each feedback type on the picking operation are discussed below.

1) Feedback of D_x, R_z: These feedback mechanisms aim to eliminate the tilt of each finger around the y-axis. Eliminating this tilt allows the gripper to grasp objects symmetrically. By grasping the object from a symmetrical position, the restoring forces exerted by the F-Fingers after lifting the object prevent it from rotating. Consequently, this enhances the picking accuracy.

2) Feedback of D_z: In this system, the target object is approached without using a 3D vision sensor. In addition, the heights at which the objects are stacked are unknown. Therefore, by setting a target value for the z-axis deformation D_z, the system performs a downward approach.

[1]Thinker Inc.: TK-01:PRODUCT. Accessed on 24.06.2024. URL: https://www.thinker-robotics.co.jp/

[2]Akitsu: amegomu sheet, shore hardness:40, Accessed on 23.11.2024. URL: https://www.monotaro.com/p/0000/2327/

Furthermore, by setting an appropriate speed for the downward approach, the feedback prevents shocks and loads on the hand caused by collisions with cluttered environment.

3) Feedback of R_x, D_y: To perform picking after recognizing contact in a clutteredenvironment, the fingertips must be inserted into the gaps between the randomly piled objects. To achieve this, feedback controls for rotation around the x-axis (R_x) and deformation along the y-axis (D_y) are used in conjunction with the previously mentioned D_z feedback. This allows for adjustments in the hand position such that, for example, if only one finger is being pressed against an object, the system can adjust the hand position to insert the finger into the gap surrounding that object.

D. Detection of grasp status

The object is grasped based on the deformation of the fingertips. First, as illustrated in Fig. 6, the $l2_{norm}$ value defined in Eq.6 is calculated while the object is being grasped. If $l2_{norm}$ exceeds a predefined threshold, the object is determined to be grasped.

$$l2_{norm} = \sqrt{t_{11}^2 + d_1^2 + t_{12}^2 + d_2^2} \tag{6}$$

To prevent incorrect judgments, the grasp status is detected twice: immediately after grasping the object and immediately before entering the move-and-stop phase. If an object is determined to be grasped in either instance, the process proceeds to the next step.

E. Estimation of the number of grasped objects

This section describes the motion and methods for estimating the number of grasped objects, including data preprocessing and estimation methods.

1) Proposal Motion: To estimate the number of grasped objects using vibrations during transport, the following criteria must be satisfied: 1) The action must be executed during transport. 2) the risk of the grasped object slipping should be minimized, and 3) the action must generate vibrations detectable by proximity sensors. To address these requirements, we propose the "Move and Stop" method, where, after grasping the object, the hand is accelerated to the desired speed in a certain derection, moves at a constant velocity, and then stops (see Fig. 7). To minimize the risk of the object slipping owing to vibrations in the grasping direction, acceleration and deceleration are performed in a direction orthogonal to the grasping direction (x-axis in Fig. 2(a)). This method can be executed during transport and facilitates rapid estimation of the number of grasped objects.

Significant in-motion vibrations occur during acceleration and deceleration. Consequently, the proximity-sensor values (y-axis tilt in Fig. 2(a)) are expected to change, as illustrated in Fig. 7. In Fig. 7, the red line represents the sensor values, and the black line indicates the hand displacement. By analyzing these sensor value changes, the number of grasped objects can be estimated.

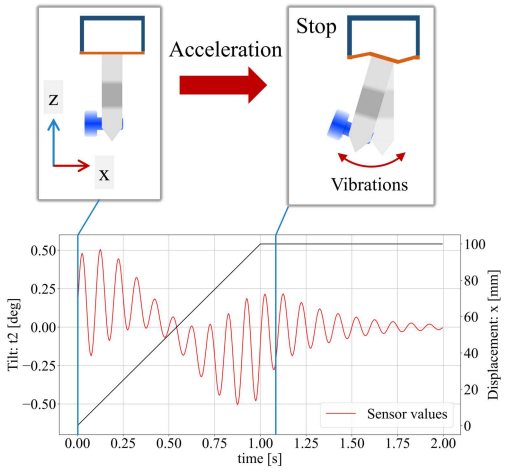


Fig. 7. The image of the motion in the proposal method (top) and the transition of sensor values during the motion (bottom). The red line in the graph represents the sensor value, and the black line represents the displacement of the fingertip.

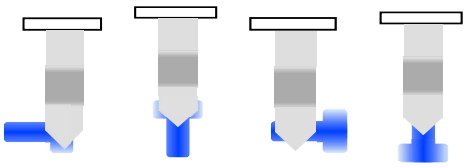


Fig. 8. Grasping pose of objects. From left to right: pose 1: screw head grasped and bolt horizontal to the ground; pose 2: screw head grasped and bolt vertical to the ground; pose 3: screw part grasped and bolt horizontal to the ground; pose 4: screw part grasped and bolt vertical to the ground.

2) Prepossessing for Sensor Values: A Fast Fourier Transform (FFT) is applied to the sensor values recorded during the move-and-stop frequency analysis. The peak positions and magnitudes of the frequency spectrum are expected to vary depending on the weight of the grasped objects, which correlates with the number of objects. Therefore, the frequency spectrum of the in-motion vibrations is used as a feature for estimating the number of grasped objects. However, the specific frequency rangein which the differences owing to the number of grasped objects become evident is unknown. Therefore, the frequency range used as a feature is experimentally determined.

3) Collection Datasets: To estimate the number of objects grasped using the aforementioned data, we collected a dataset consisting of three classes based on the number of objects grasped (0, 1, and 2 objects). Data are collected from 20 trials in each class. In this study, we used two types of metal bolts, M3 and M6, as the small objects. The bolt specifications are listed in Table I.

When grasping a single bolt, the center of gravity of the hand varies depending on the grasping pose. Therefore, we collected a dataset using four grasping poses as shown in Fig. 8. For each grasping pose, five data samples were collected, resulting in 20 samples for the class the represented one grasped bolt.

4) Classification: Based on the aforementioned dataset, we use an SVM to estimate the number of grasped objects.

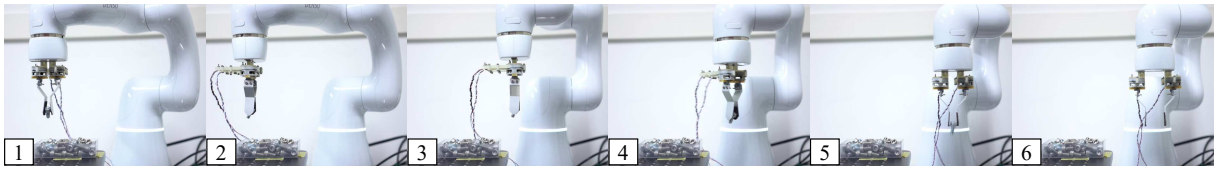
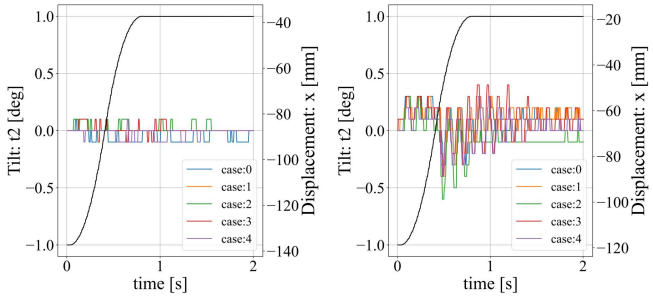


Fig. 9. Snapshot during move and stop (every 0.9 s)

TABLE I
PARAMETERS OF TARGET OBJECTS

	M3	M6
Screw diameter [mm]	3	6
Head diameter [mm]	5	10
Length [mm]	22	18
Weight [g]	1.0	4.8



(a) Cross pattern (b) Bar pattern

Fig. 10. Comparison of in-motion vibrations based on the shape of the compliant sheet. The left vertical axis represents the tilt t_2 , while the right vertical axis represents the displacement x. The black line indicates the displacement of the hand, and the other lines represent the sensor values for each trial.

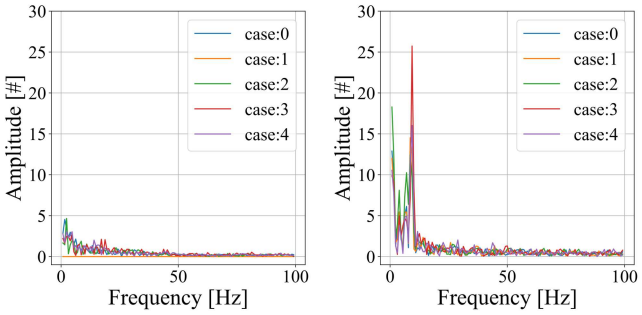
The features used for the estimation are extracted from the frequency spectra of the in-motion vibrations. The specific range of the frequency spectrum to be used as features is determined experimentally, as discussed in Section III-E.2.

Using these features as inputs, we estimated the number of grasped objects using an SVM with an RBF kernel. We used 75 % of the dataset as training data to construct the estimator.

IV. EXPERIMENT

In this section, we describe the experiments conducted to construct the estimator and verify its estimation performance, as well as the results obtained. We conducted experiments using a collaborative robot Cobotta. The proposed action was implemented by accelerating Cobotta's hand to $240\,\mathrm{mm/s}$, performing a $100\,\mathrm{mm}$ linear movement, and then stopping. Snapshots of picking, the proposed motion and placing the object in an another bin taken every $0.9\,\mathrm{s}$ are shown in Fig. 9.

These experiments are conducted using a Raspberry Pi 4B as the PC. Since both the picking and estimation of the number of objects grasped do not require large-scale processing for machine learning, the system can be implemented even on such a compact PC.



(a) Cross pattern (b) Bar pattern

Fig. 11. Comparison of frequency characteristics based on the shape of the compliant sheet. The vertical axis represents the spectral intensity, and the horizontal axis represents the frequency. Since the measurement period of the proximity sensor is 200 Hz, frequencies up to 100 Hz are displayed.

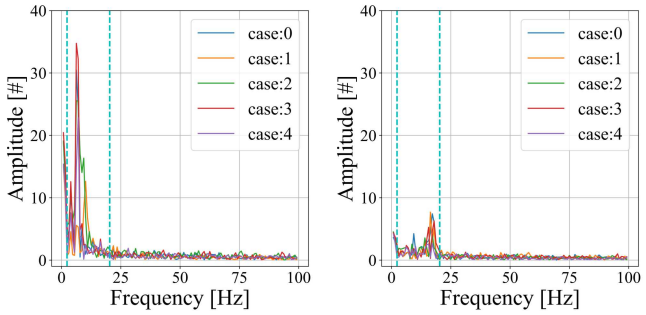
A. Comparison of Vibration by Mechanism

First, we experimentally investigated an F-Finger suitable for the estimation. The mechanical characteristics of the F-Finger can be altered by changing the shape of the compliant sheet. Consequently, the magnitude of vibrations during operation varies depending on the shape of the compliant sheet. To estimate the number of objects grasped from in motion vibrations, a mechanism that maintains a sufficient grasping force for small objects while exhibiting significant vibrations during the operation is necessary.

In this study, we used the two F-Fingers described in Section III-A (Fig. 4). Based on the sensor values recorded during the operation of each F-Finger, we selected a mechanism that was more suitable for estimating the number of grasped objects. The target object was a single M6 bolt, as described in Section III-E.3.

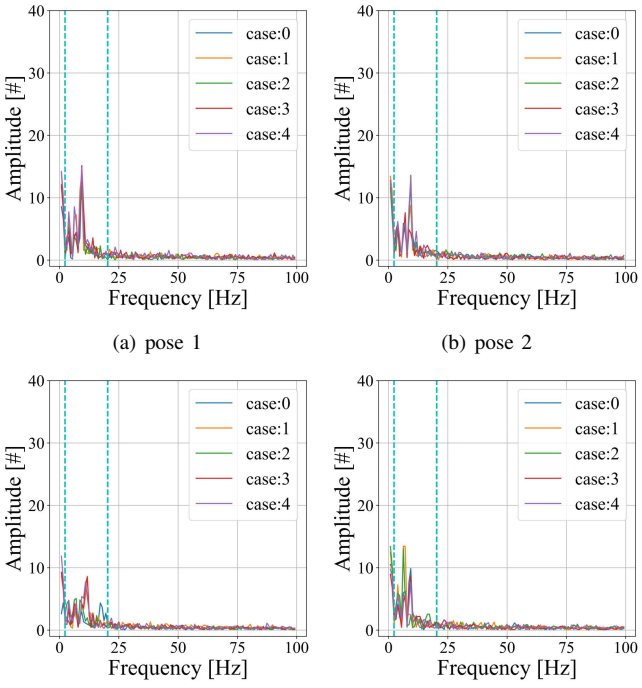
The transitions in the sensor values during the five operations for each mechanism are shown in Fig. 10, and the FFT results are shown in Fig. 11. In Fig. 10, the black line represents hand displacement. In Fig. 10, the horizontal axis denotes the time, the right vertical axis indicates the sensor values (t_2 [°]), and the left vertical axis represents the x-coordinate displacement of the hand. In Fig. 11, the horizontal axis represents the frequency, and the vertical axis indicates the intensity of the frequency spectrum.

In the case of the cross-shaped compliant sheet, minimal vibration was observed compared to the bar-shaped sheet, and the intensity of the frequency spectrum was also lower. From these results, it can be inferred that an F-Finger with a bar-shaped compliant sheet is more suitable for estimating the number of grasped objects. Therefore, subsequent experiments used a bar-shaped compliant sheet.



(a) two objects (b) zero object

Fig. 12. Frequency characteristics for zero and two grasps. The object is M6, and the blue line in the graph shows the range of features to be extracted.



(a) pose 1 (b) pose 2

(c) pose 3 (d) pose 4

Fig. 13. Frequency characteristics of each grasping pose for a single grasp. The object is M6, and the blue line in the graph shows the range of features to be extracted.

B. Comparison of Frequency Characteristics by the Number of Grasps and Pose

In addition to the four grasping poses considered in Fig. 8, we included cases of zero and two objects grasped, resulting in six scenarios. We performed the motion shown in Fig. 9 and measured the vibrations during the motion for each case. This section compares the frequency characteristics of the vibrations for each scenario and discusses the features used for the estimator.

The frequency characteristics of each grasping pose with the M6 bolt are shown in Fig. 12. Additionally, Fig. 13 shows the frequency characteristics for different numbers of grasped objects. The results for each case are based on five trials. These result show that the frequency characteristics differ according to the number of objects grasped. Even among the cases where a single object was grasped, there was a distinct difference in the frequency characteristics between pose 4 and the other grasping poses.

Additionally, it was confirmed that differences in the frequency characteristics occur in the low-frequency range. Therefore, the frequency spectrum within the range indicated by the blue dashed lines in the figures ($2 - 20$ Hz) was used as features for estimation.

C. Estimation Performance Using Test Data

Using test data from the aforementioned dataset, we evaluated the estimation performance. The resulting confusion matrix is shown in Fig. 14.

It was found that the M6 bolts could be estimated with an accuracy of 93 %. Similarly, for the M3 bolts, data were collected, and an estimator was constructed using the same method as for the M6 bolts. The estimation accuracy for the M3 bolts was 87 %. The lower estimation accuracy for M3 bolts compared to that for M6 bolts is likely due to the smaller weight per unit of M3 bolts, which results in smaller differences in weight depending on the number of bolts.

D. Full Pipeline

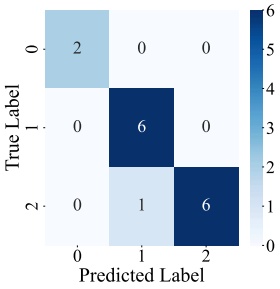
We implemented the picking process shown in Fig. 1 and conducted a fullpipeline experiment. The process is divided into three stages: “Picking,” “Detection of grasp status,” and “Estimation.” In the “Picking” stage, vision-sensorless picking, as described in Section III-C, is performed. After grasping the target object, the process proceeds to “Detection of grasp status.” Here, it is determined whether the object has been grasped, that is, whether the number of grasped objects is one or more. If the number is one or more, the process proceeds to the “Estimation” stage; otherwise, it returns to “Picking.” In the “Estimation” stage, the number of grasped objects is estimated using “Move-and-Stop” as described in Section III-E. If the estimation result indicates that one object has been grasped, the picking is considered successful. Otherwise, the process returns to “Picking.”

Two types of bolts, listed in Table I, were used to grasp the objects. For each type, 200 picking trials were conducted using only the F-Finger and 200 trials using the fullpipeline, and their accuracies are compared. Additionally, the estimation accuracy within the 200 trials conducted with the fullpipeline was investigated.

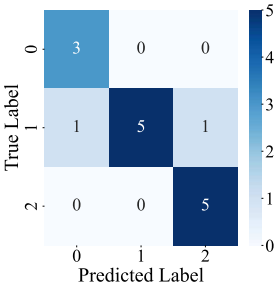
The results are listed in Table II, and the estimation accuracy during the experiment is shown in Fig. 15. For the M6 bolts, incorporating the grasp count determination resulted in a 100 % success rate for grasping. The estimation accuracy within fullpipeline was 87.5 %. However, for the M3 bolts, the grasping success rate only improved by 2 %. The estimation accuracy within the fullpipeline for M3 bolts was approximately 59.9 %, which is lower than that for the M6 bolts. Additionally, in both cases, the time required for determination and re-grasping increased the conveyance time compared to using only 5DoF control.

TABLE II
RESULT: FULL PIPELINE

Object		Mean time for task [s]	Task success rate [%]	The number of grasped [–]		
				1	0	multiple
M3	5DoF control only	14.6	84.5	169	4	27
	5DoF control + Estimation	27.8	86.5	173	2	25
M6	5DoF control only	14.2	97.0	194	4	2
	5DoF control + Estimation	21.6	100	200	0	0

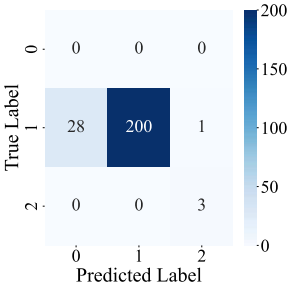


(a) M6

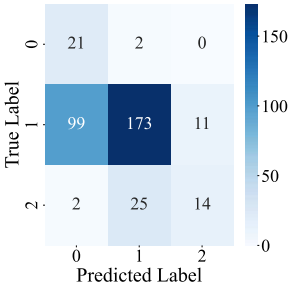


(b) M3

Fig. 14. Estimation results of the number of grasped objects using test data (Label: the number of objects grasped)



(a) M6



(b) M3

Fig. 15. Estimation results of the fullpipeline (Label: the number of objects grasped)

V. DISCUSSION

From Section IV-D, our proposed system demonstrates a task success rate exceeding 85 % with only 5-DoF control. This indicates that our control method (refer to Section III-C) is effective for bin-picking of small objects. Additionally, the estimation of the number of objects grasped improved the task success rate for all target objects. The main reason is that the system was able to estimate the number of objects grasped using a small-scale dataset with only 20 trials per class, and the dataset described in Section III-E.3 effectively covered most grasping pose observed during the picking.

On the other hand, compared to M6 bolts, the task success rate for M3 bolts was lower. This is likely because the same control parameters used for M6 bolts (refer to Section III-C) were applied when picking M3 bolts. Adjusting the control parameters to better suit the target object may improve the task success rate. As a limitation, the accuracy to estimate the number of objects grasped may decrease when it is challenging to grasp near the center of gravity of the object. The frequency characteristics of in-motion vibration depend not only on the weight but also on a composite center of gravity position of fingertip part and object. Consequently, the method is less applicable to objects such as L-shaped or long rod-shaped objects. This likely explains the decrease in estimation accuracy for M3 bolts.

VI. CONCLUSIONS

This paper proposed a compact bin-picking system that occupies less space than a human worker. We achieved such system by using the quasi-static deformation and dynamic deformation, which are measured by F-Finger. By using the feedback from the quasi-static deformation when the fingertip contacts the environment, we achieved the vision-sensorless bin-picking. By analyzing the dynamic deformation of the fingertip caused by in-motion vibration during arm tip position movement, the number of objects grasped was estimated . Every processes can be executed on small on-board computer board (Raspberry Pi 4B).

By using proposal system, a task success rate of more than 85 % and a bin-picking tatc time within 30 s were achieved for small metal bolts. Specifically, for M6 bolts, the grasp count was estimated with 87.5 % accuracy, achieving a 100 % task success rate. Future challenges include adapting this system to objects other than metal bolts, aiming to construct a more generic system.

REFERENCES

[1] Zhao, C., *et al.*: Learning to Pick by Digging: Data-Driven Dig-Grasping for Bin Picking from Clutter, 2022 International Conference on Robotics and Automation (ICRA), pp. 749–754, 2022.

[2] Zhang, X., *et al.*: Learning to dexterously pick or separate tangled-prone objects for industrial bin picking, IEEE Robotics and Automation Letters, 2023.

[3] Domae, Y., *et al.*: Fast graspability evaluation on single depth maps for bin picking with general grippers, 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 1997–2004, 2014.

[4] Cao, H., *et al.*: Two-Stage Grasping: A New Bin Picking Framework for Small Objects, 2023 IEEE International Conference on Robotics and Automation(ICRA), pp. 2584–2590, 2023.

[5] Ishige, M., *et al.*: Blind Bin Picking of Small Screws Through In-finger Manipulation With Compliant Robotic Fingers, 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 9337–9344, 2020.

[6] Do, W. K., *et al.*: Inter-finger Small Object Manipulation with Dense-Tact Optical Tactile Sensor, IEEE Robotics and Automation Letters, 2023.

[7] H. P. Saal, *et al.*: Active estimation of object dynamics parameters with tactile sensors, IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS), pp. 916-92, 2010.

[8] Huang, H. J, *et al.*: Understanding Dynamic Tactile Sensing for Liquid Property Estimation, arXiv: 2205.08771, 2022.

[9] Koyama, K., *et al.*:Integrated control of a multiple-degree-of-freedom hand and arm using a reactive architecture based on high-speed proximity sensing, The International Journal of Robotics Research, Vol. 38, No. 14, pp. 1717–1750, 2019.