Stable Haptic Feedback based on a Dynamic Vision Sensor for Microrobotics

Aude Bolopion1∗, Zhenjiang Ni2∗, Joël Agnus1, Ryad Benosman3 and Stéphane Régnier2

Abstract—This work presents a stable vision based haptic feedback for micromanipulation using both an asynchronous Address Event Representation (AER) silicon retina and a conventional frame-based camera. At this scale, most of the grippers used to manipulate objects lack of force sensing. High frequency vision detection thus provides a sound solution to get information about the position of the object and the tool to provide virtual haptic guides. Artificial retinas present high update rates, which enables to address one of the major challenge of haptic feedback teleoperation systems, namely stability. However static objects are not detected. The haptic feedback is thus based on an asynchronous silicon retina to provide a high update rate of moving objects and a frame-based camera to retrieve the position of the target object. This approach is validated by pick-and-place of microspheres (diameter: around 50 micrometers) using a piezoelectric microgripper. The displacement of the tool, as well as the opening and closing of the gripper are controlled by the haptic device. Haptic virtual guides are transmitted to users to assist them in the different steps of the pick-and-place task: a virtual stiffness ensures the correct alignment of the tool with respect to the object, a repulsive haptic force enables users to monitor the gripping step, and operators are assisted while picking and placing the object.

I. INTRODUCTION

Micromanipulation has received an increasing interest in the last few years, since the ability to handle objects with a size below 100 micrometers will open a wide range of applications, from electronics to biological fields. Several manipulation techniques have been developed, based on rolling, pushing-pulling or picking-and-placing techniques [1], [2], [3]. However many issues must still be overcome to get intuitive manipulation platforms. In particular, only skilled operators can perform manual micromanipulation due to the scale difference, the fragility of both the tools and the objects, the complexity of the force fields and the high sensitivity of the systems to environmental conditions. One solution is to perform automated tasks. Good results have been obtained (i.e. high throughput and accurate positioning) on objects with a size of hundreds of micrometers [4]. This manipulation mode fits the needs of repeated tasks on large number of objects. However, operator’s knowledge is not exploited, users cannot interact while the automated task is being performed and they do not receive enough feedback to learn from the manipulation. The use of assistance for manipulation appears as a promising alternative, which benefits from the interactivity of manual manipulation, while facilitating it. Haptic feedback is one of the solution to intuitively transmit information to operators, as it provides indications that users are used to receive when they perform manipulations at macroscales [5], [6].

Several teleoperated micro- and nanoscale tasks are reported in the literature, from feeling the substrate topology [7] to inducing a modification of the sample through indentations (e.g. direct patterning on a substrate [8]). Tasks involving pushing/pulling or cutting objects are also presented [9], [10], [11]. In addition to transmitting micro- and nanoscale interactions, haptic feedback is used as an enhancement for user assistance, with virtual guides for pushing and pick-and-place by adhesion tasks [12]. In this case, haptic feedback is used to keep the user’s motion on a specified path. However these examples are mostly limited to proof of concepts, and none of them can be extended to industrial applications.

To get an efficient and versatile teleoperation system for microscale applications a major issue must be faced: the lack of position and force feedback [13]. Sensors have been developed [14], [15], but their integration into the dedicated tools induces an important increase in the complexity and the cost of their fabrication. In particular, even if some microgrippers offer sensing capabilities at the expense of complex designs [16], [17] most of them still lack force measurements [18], [19].

To avoid the complexity of sensors’ integration, vision is a promising solution [20], [21]. However, to ensure a stable haptic coupling, a frequency of 1 kHz is needed for the sensor output [22]. This is even more critical in case of micromanipulation since the dynamic of the objects is important due to their low inertia. Conventional frame-based cameras’ lack of dynamic information and their redundancies set an important limit to potential micromanipulations. Event-based computer vision based on Address Event Representation (AER) provides a sound solution to high-speed vision problems [23]. This newly developed discipline is motivated by mimicking biological visual systems [24]. The Dynamic Vision Sensor (DVS) silicon retina reacts to changes of contrast that are converted in a stream of asynchronous time-stamped events [25]. The reduction of redundant information makes this technique promising for high-speed tracking.

This paper presents a haptic feedback teleoperation system for microscale applications based on visual sensing obtained from an asynchronous Address Event Representation (AER)
silicon retina and a conventional frame-based camera. The Dynamic Vision Sensor enables to record high frequency phenomena, and to provide feedback at a high sampling frequency, but it gives no indication about immobile objects. The position of static objects are provided by a conventional frame-based camera. To validate this approach, a teleoperated pick-and-place task of spheres (diameter around 50 µm) using a piezoelectric gripper is presented. Haptic feedback directly estimated from the output of both the event-based retina and the conventional frame-based camera is provided to assist users during the manipulation. The first pick-and-place manipulation with 3D haptic feedback using a microgripper is successfully achieved. This work is based on several previous studies on the stability of haptic coupling schemes for applications at microscales [26], on the definition of virtual guides for pick-and-place operation of microspheres based on two atomic force microscopy cantilevers [27], on the use of visual feedback coming from a scanning electron microscope for teleoperation [28] and on the use of dynamic vision sensors [29]. However, it is the first time, to our knowledge, that dynamic vision sensors are used to provide stable haptic feedback to assist users while performing a pick-and-place operation. This approach will enable complex teleoperated assembly tasks.

This paper is organized as follows. The experimental setup is presented in Section II. Section III details the teleoperation system. The proposed approach is validated by experimental results presented in Section IV. Finally, Section V concludes the paper.

II. Setup

A. Micromanipulation tool

The micromanipulation end-effector developed in FEMTO-ST Institute [30] is a microgripper with tightened fingers that proves to be efficient and intuitive for microhandling problems (Fig. 1). This microgripper, named MMOC (Microprehensible Microrobot On Chip), has 4 degrees of freedom and is able to grasp, hold and release submillimetric-sized objects up to several tens of micrometers. The Fig. 1(d) and 1(e) show respectively the manipulation of a 20 µm silicate grain and 100 µm glass sphere.

Each finger is actuated by a piezoelectric cantilever with local electrodes, called duo-bimorph (Fig. 2). As mentioned on the cross section view, one finger has four electrodes referred to a central ground and two voltages ($V_1 = V_z - V_y$ and $V_2 = V_z + V_y$) are necessary to impose the deflections $y$ and $z$ based on the deflections of the piezoelectric beam. This configuration offers number of capabilities: open-and-close motion as well as up-and-down motion. Indeed, each finger is able to move independently from the others in two orthogonal directions (right picture of Fig. 2). The up-and down motion can be useful for fine motion, for release strategies of objects by crossing the fingers or for insertion of microparts. It is also particularly convenient to align the finger-tips in case of misalignment after microfabrication and assembly process of the microgripper.

Specific silicon end-effectors able to grasp objects from 10 to 100 micrometers have been made (Fig. 1(a) and 1(b)). They are designed to minimize the sticking effects between the end-effector and the objects, to facilitate the release step. They are microfabricated on SOI wafer with a thickness of 10 µm, mainly using DRIE process [31]. The gap and the alignment of the tips are assured by breakable-parts. These parts are removed after the pair of fingers have been glued on the spatulas of the actuators (Fig. 1(a) and 1(b)).

Such a microgripper presents a typical stroke of open/close motion and up/down motion of respectively 320 and 200 micrometers at the end of the finger tips for ±100V and a blocking force of respectively 80 and 30 mN at the end of the actuator for 100 V. The gripper is controlled by commands sent from a personal computer to a high voltage interface (four channels of ±150V) via RS232 link.
B. Vision system

As shown in Fig. 3, the observed scene is monitored by two optical sensors, that record the same view. The light beam is divided into two optical paths, and redirected to an asynchronous event-based silicon retina (DVS, prototype from ETHZ) and a conventional frame-based camera (GigE vision, Basler). The scene recorded by both sensors is magnified by a 10× objective (Olympus).

The DVS used in this work is a silicon retina with 128 × 128 pixels [25]. It mimics the biological retinas and outputs compressed digital data in the form of events, removing redundancy and increasing dynamic range compared with conventional imagers. A review of the existing similar sensors are surveyed in [24]. The DVS’s output consists of asynchronous address-events that signal scene luminance changes. Each pixel is independent and detects changes in log intensity larger than a threshold since the last event is emitted. When the change exceeds a set threshold, a +1 or −1 event is generated by the pixel depending on whether the log intensity is increased or decreased. The timing of events can be conveyed with a very accurate temporal resolution of 1 µs. The events are then transmitted to a computer using a standard USB connection. The advantage of this sensor over conventional clocked cameras is that only moving objects produce data, thus increases significantly the processing speed. The reactivity and the stability of the haptic system can therefore be enhanced.

C. Teleoperation system

The gripper is mounted on a 3-axis motorized micromanipulator (MP285, Sutter Instrument)1 to allow an accurate positioning with respect to the substrate. The manipulator used relies on stepper motors with a step size of 0.040 µm. It is a cable-driven system with cross roller bearings, and it has a sub micrometer resolution and a travel range of 25 mm. The manipulator was originally controlled through a serial port; however, to increase the communication frequency, a joystick is emulated by programming the manipulator’s parallel port using a PCI6259 National Instrument acquisition card.

Both the positioning of the gripper and its actuation (opening and closing operations) are controlled through an Omega haptic device2. This device has 3 degrees of freedom for both displacement and force feedback. Haptic feedback is computed based on the vision detection. Forces higher than 5 N are saturated to avoid any damage to the interface.

A single PC (Intel Xeon core, 2.93 GHz) operating under Windows 7 runs the threads corresponding to the gripper, the micromanipulator, the vision detection, and the haptic feedback.

III. HAPTIC FEEDBACK TELEOPERATED SYSTEM

A. Manipulation modes

The coupling between the haptic device and the microgripper is represented in Fig. 3. Using the haptic interface, the user controls the displacement of the micromanipulator as well as the opening and the closing of the gripper. Scaling factors αd and αc are introduced to convert the position of the haptic handle to the variables used to control the positioning and the actuation of the gripper.

![](image)

Fig. 3. Haptic coupling scheme. The user sets the position of the gripper and the actuation of the gripper’s fingers by controlling the position of the haptic interface $R_h$. Scaling factors $a_d$ and $a_c$ convert the position of the haptic handle to commands to control the gripper. The operator receives haptic feedback through the device. The haptic force is based on the distance between the gripper and the sphere, determined from vision algorithms.

To ensure ease of manipulation, different modes have been defined:

- planar displacement: the operator controls the displacement of the gripper in a plane parallel to the substrate, the $(x, y)$ plane,
- vertical displacement: the operator controls the displacement of the gripper along the vertical direction. It enables to lift-off the object, and to place it down,
- gripper control: the operator controls both the opening and closing of the gripper and its position along the $y$-axis to align the gripper with the middle line of the object.

The operator selects the appropriate manipulation mode on a graphical user interface developed in C++.

B. Determination of the relative position of the object and the tool

The haptic feedback is based on the distances between the gripper and the object. To determine these distances, both the DVS silicon retina and the frame-based camera are used.

The DVS sensor is installed to provide a fast vision detection. To track the gripper, the most recent events within a time period $\Delta t$ (called “active events” hereafter) are continuously stored in memory. Following the Event based Iterative Closest Point (EICP) algorithm, a model composed of points of the contour of the microgripper is required. An active event is matched with a model point by computing the minimal Euclidean distance between the event’s position and all the points of the model. After the matching step, the rigid body transformation composed of rotations and translations is estimated by minimizing a mean square cost function. Readers interested in the algorithm details can refer to [32]. In our manipulation, the active events’ decay time $\Delta t$ is tuned around $10 – 15$ ms. The processing is event driven, and its
update frequency has a mean value of 4 kHz. The algorithm is implemented in Java under JAER open-source project [33].

The frame-based camera serves as a complement to DVS silicon retina solely for the static object detection. The focal planes of both the DVS (128 × 128 pixels) and the frame camera (659 × 494 pixels) are related by a homography transform as both observe the same 2D plane [34]. The homography is estimated off-line by extracting from both sensors’ the coordinates of six corner points of the gripper fingers and linking them to the actual metric of the gripper’s points in the scene (see Fig. 4) [35]. During the manipulation, the circle corresponding to the sphere to be manipulated is detected using a Hough transform through conventional camera output. The distance between the gripper fingers $\delta_2$ is estimated from DVS output. By combining the processing of the two sensors, if an object is detected between the two fingers, the relative finger-object distance on the left and the right sides $\delta_{l2}$, $\delta_{r2}$ and the distance between the center of the sphere and the gripper $\delta_l$ (Fig. 4) is calculated instead. These various distances are used to compute the haptic feedback.

C. Haptic feedback

To increase the success rate of the pick-and-place operation two criteria should be met: the sphere should be grasped on its middle line, and the grasping force should be enough to lift the sphere but controlled to avoid any damages to the object. The haptic feedback must assist the user for these two operations. The operator receives a force feedback $F_h = \begin{bmatrix} F_{oc} & 0 \\ 0 & F_y \end{bmatrix}$ based on the distance between the gripper and the sphere $\delta_l$, $\delta_{l2}$, $\delta_{r2}$ and $\delta_l$ determined from vision. The haptic force $F_y$ is transmitted on the vertical axis of the haptic device to ensure a correspondence between the visual feedback (on the screen the gripper is moving in the vertical plane) and the haptic feedback.

To help the user align the gripper with respect to the middle line of the sphere, a haptic force corresponding to a spring of stiffness $k$ between the position of the gripper and the sphere is provided:

$$F_y = -k\delta_y$$

(1)

where $\delta_y$ is the distance between the center of the gripper and the center of sphere along the y-axis (see Fig. 4).

A haptic feedback $F_{oc}$ is provided so that the user can monitor the grasping force. Contrary to what is commonly presented in the literature [36], we are not interested here in computing the exact efforts applied on the object but only in deriving information to assist the user while performing a given task. The calibration process, which enables to relate the tool deformations to the applied force, is thus unnecessary. While closing the gripper, the user has to counteract a haptic force $F_{oc}$:

$$F_{oc} = \begin{cases} F_{max}e^{-\frac{\delta_{l2}^2}{\alpha^2}} & \text{if not in the contact zone} \\ F_{contact} & \text{if in the contact zone} \end{cases}$$

(2)

where $\delta_{l2}'$ is the free space between the two gripper’s fingers. If the sphere is situated between the fingers $\delta_{l2}' = \delta_l + \delta_{l2}'$, which corresponds to the sum of the distances between each of the fingers and the sphere; otherwise, $\delta_{l2}' = \delta_l$. $F_{max}$ is the maximum force that can be transmitted to the user when the gripper is close to the sphere but has not entered the contact zone yet. $\alpha$ is a constant chosen to tune the decrease of the haptic force as the distance between the two fingers increases. $F_{contact}$ is the force sent while the gripper is grasping the sphere. The step between $F_{max}$ and $F_{contact}$ must be high enough to indicate clearly the contact between the sphere and the gripper. The contact zone is reached if $\delta_{l2}'$ and $\delta_{r2}'$ are less than a given distance $\delta_l = 3$ µm (which corresponds to 6% of the sphere diameter). The gripper will then reach the non contact zone if $\delta_{l2}'$ and $\delta_{r2}'$ are greater than $\delta_l = 7.5$ µm. This hysteresis avoids undesirable frequent transitions between contact and non-contact modes because of noises and tracking errors. The hysteresis values $\delta_l$ and $\delta_r$ are chosen according to our experience to achieve a comfortable user sensation. The force step that is sent to the user when contact is detected is filtered to avoid brutal force changes. Even if the user does not receive the maximum force feedback at the instant of contact, he/she can distinctly feel the increase in the force, and infers that contact happened.

IV. EXPERIMENTAL RESULTS

Experiments are performed to validate the integration of event-based vision sensor for stable haptic feedback. The microspheres are glass beads of 30 ~ 50 micrometers of diameter from Polysciences, Inc. A Gel-Pack substrate has been used to provide enough adhesion to prevent the sphere from sticking to the gripper. The experiment consists in the following tasks: positioning the gripper with respect to the sphere (in plane displacements), grasping it, taking it off, moving it, placing it down, and releasing it.

3http://www.polysciences.com/
When the sphere is at the correct location, the user can place it down. The last step is to release the sphere (see Fig. 5(b)). The plot starts at $t = 44$ s approximately since it took around $30$ s to the user to lift the object, move it, and place it down. This duration depends on the distance between the starting and final positions, as well as on the user expertise with haptic devices. At the beginning of the operation, as the gripper contacts the sphere, the operator can feel a constant force ($5$ N) that assist him/her to open the gripper. The user releases the sphere at time $45.3$ s. It can be noted that the curve does not reach zero as the force is still helping the user to open the gripper and avoids unexpected closing. Since the gripper and the sphere are aligned during releasing, $F_y$ is around zero.

During all the grasping-releasing and pick-and-place operation, the user receives haptic feedback based on vision detection that helps him/her to achieve comfortable and easy sphere manipulation. The dynamic vision sensor provides a high-frequency feedback that enables to provide a stable haptic system. Users successfully performed a 3D teleoperated manipulation on micron-sized objects. This system will surely benefit teleoperated or automated microassembly and opens new perspectives for complex micromanipulations. In the future a vertical haptic feedback should be provided to assist the user while positioning the gripper along the vertical axis during the grasping and the placing steps. To determine the distance between the gripper and the substrate the focusing and defocusing information might be used.

V. Conclusion
To overcome the lack of sensing capabilities at microscale, a vision-based system is presented. To ensure a stable haptic feedback, the frequency of the vision detection must be higher than $1$ kHz at this scale as low inertia induces high dynamic phenomena. This is ensured by the output of a DVS
sensor that enables to track the tool at more than 4 kHz. This feedback is combined with the output of a classical frame-based camera, used to derive information about static parts of the scene, and in particular the position of the object that must be manipulated. This approach is validated on a pick-and-place experiment of glass spheres with a diameter of around 50 µm using a piezoelectric gripper. Haptic feedback is provided to assist users while aligning the gripper with respect to the tool, grasping it, picking and placing it and releasing it. This work will enable to provide stable haptic feedback to perform complex assembly tasks even with sensor-deprived systems.

Future works include extending this approach to other applications, involving different objects or tools. Automated tasks could also benefit from this high frequency vision detection.

VI. ACKNOWLEDGMENT

This work was supported by the French National Agency of Research, through the NANOROL project (contract ANR-07-ROBO-0003) and the NANOROBUST project (contract ANR 2011 NANO 006).

REFERENCES


