Gaze Contingent Control for Optical Micromanipulation

Maria Grammatikopoulou, and Guang-Zhong Yang, *Fellow, IEEE*

Abstract—Optical Tweezers (OT) have the advantage of non-contact interaction with target objects such as cells, overcoming the pitfall of obstructive adhesion forces which are present in contact micromanipulation. It is also feasible to manipulate a number of small microparts simultaneously or 3D structures by using multiple laser traps. These capabilities give rise to the potential to develop a human-robot interface to facilitate microassembly tasks. This paper presents a gaze contingent control framework and a method for 3D orientation estimation for optical micromanipulation. The proposed strategy aims to use OT as an interactive microassembly platform. The framework comprises I) a strategy to recognize the operator's intentions in order to interactively place and reconfigure the optical traps using the operator's eye fixation point, II) haptic constraints generated from the user's eye gaze to assist positioning of the assembled microparts and III) a method for 3D orientation estimation. The performance of the proposed framework is assessed through a set of experiments comparing it to the standard OT user interface. Three-dimensional manipulation and orientation estimation of a non-spherical microstructure are also performed.

I. INTRODUCTION

OT can realize repeatable grasping sequences owing to the absence of physical contact with the object. Since the trapping force is applied by a tightly focused laser beam through light interaction with the matter, the object is released immediately when the optical trap is disabled. Hence, OT can be used as a micromanipulation platform for assembling parts with sizes up to a few tenths of microns into larger structures. Another advantage of the OT is that they allow multiple object manipulation by generating a pattern of optical traps. This can be realized either by high frequency alteration of the laser position in conventional OT or by Spatial Light Modulation (SLM) in Holographic Optical Tweezers (HOT). Such considerations give rise to the development of an intuitive interface for micromanipulation exploiting the OT capabilities. Current control interfaces of available micromanipulator systems allow user input through 2D mouse position or gamepads. Consequently, bimanual task execution is challenging and requires increased user effort due to the limitations of the input devices.

Chen *et al.* have proposed a potential-field based controller in order to move a group of microparticles automatically to a desired position while preserving a fixed laser spot configuration [1]. Multiple groups of objects are manipulated by controlling both the laser source and the piezo stage for transportation within and outside of the image Field of View (FoV) [2]. Multiple laser spot manipulation is

Fig. 1. The system hardware components: The haptic devices, the eye tracking device, the computer screen displaying the live camera feed and the OT mounted on an anti-vibration optical table (from left to right).

also employed for automatic single object grasping using microbead gripping formations [3] and for manual microtool manipulation [4].

Hand-eye coordination can be facilitated by using the concept of perceptual docking [5]. This concept incorporates human sensory information, such as the eye gaze, to the human-machine interface to assist the operator in task execution. Such interfaces have been developed for large scale surgical robotic systems. The eye fixation point of the operator is processed for haptic constraint generation in the work presented in [6]. Microrobotic systems could benefit from similar implementations that assist hand-eye coordination in multiple or 3D object manipulation.

The incorporation of a haptic guidance interface in the microscale could also be beneficial since it enhances ergonomics for ease of manipulation. Previous implementations in OT systems assist the user through haptic guidance along an obstacle-free path and by limiting the user movements to ensure that the particles do not escape from the optical trap [7].

Recovering the pose of the microstructure in 3D is also essential for robotic control as well as for the operator's perception during object manipulation. By estimating the 3D pose, closed-loop control and cooperative control techniques can be developed. Therefore, it is necessary to establish a mapping between the optical trap arrangement and the corresponding object pose for kinematic modelling. Berk Yesin *et al.* proposed a CAD-based multi-camera method using edge detection to recover the microstructure's 3D pose for visually

guided microassembly [8]. A 2D to 3D registration method for guidance in MEMS microassembly is presented in [9] using feature matching between the identified 2D tracked features and the corresponding 3D features from the CAD model. Shape tracking has also been proposed in order to provide stable haptic feedback [10].

The main contributions of this work include: I) How object trapping and transportation are realized when taking into account the user's intentions during manipulation. This is performed transparently by processing the natural gaze of the user. In particular, the gaze information is combined with visual servoing to assist trapping and pattern reconfiguration without interrupting the executed task or compromising the number of objects to be manipulated. II) The operator can activate haptic guidance generated by the eye fixation point to position the assembled parts. III) 3D orientation estimation is performed in non-spherical object manipulation to relate the respective trapping configuration to the achieved object pose. This is one of the first reported attempts of perceptual docking and pose estimation in OT applications.

The system aims to assist the operator seamlessly in hand-eye coordination. The effectiveness of the developed control scheme was evaluated through two sets of experiments; in multiple object manipulation using 2D visual servoing and in 3D structure manipulation. In the first set of experiments, the gaze controller is compared to the standard OT user interface. The results of these experiments demonstrate that the execution time and the manipulation flexibility are improved when using the gaze controller. The second experiment shows how 3D control can be performed using gaze assistance while the orientation of the object is also recovered.

II. SYSTEM OVERVIEW

The operator controls the laser traps via two haptic devices. The microscope-camera feed is displayed to the user with the optical trap positions overlaid on the image. The software controller operates in different states according to the phase of the task. Seamless state switching is realized based on the operator's intention and human sensory input. The user initially controls one laser point with each of the haptic devices and can add more laser traps on demand during the manipulation task. Hence, the operator can manipulate multiple spots with each one of the haptic devices. The laser positions are added according to the eye gaze input and the visual servoing either directly on the identified objects or on the eye fixation point on the respective frame. The positions of the added optical traps can be later adjusted using the eye tracking data. The user can selectively activate a guidance force that helps in microparts positioning.

III. METHODS

A. Gaze contingent controller

The aim of the presented control scheme is to enhance the operator's dexterity in tasks that require handling of multiple microparts or complex structures in a dynamic environment.

Fig. 2. The camera feed as viewed by the user (left). The detected objects (red) and object template (right).

1) *Eye gaze data processing:* The operator's eye gaze is tracked by the device using infra-red video with a measurement rate of 60 Hz. Eye gaze processing is performed continuously in a separate thread regardless of gaze data utilization. A buffer of previous values is kept while the noisy values are discarded. Noisy values include measurements that correspond to the operator blinking or detection of only one eye. The 2D eye gaze is filtered with a low pass filter. The camera provides an 648 x 488 image with each pixel corresponding to 0.074 μm, therefore the visible workspace is 48 x 36 μm. The view covers partially the computer screen and the Graphical User Interface. Hence, only the filtered data points which are within this region of the screen are used in the software controller. In case the user eye gaze is outside of this region, the latest position within the camera image which is kept in the buffer is used in the calculations. Consequently, the operator does not need to continuously look at the camera feed.

2) *Intention recognition and gaze control:* The software controller has three different states: I) Trap pattern generation for object trapping; II) optical trap repositioning and III) haptic guidance for object placement. A new optical trap is added or re-positioned only after user request allowing the primary control to the operator at all times. The processed eye gaze information is used to determine the intentions of the operator and can be identified by addressing the following three questions.

If the user is exploring or has focused on a specific region of the workspace: This user activity is determined by calculating the Candidate Target Position (CTP) which gives an indication of the respective region of interest. To obtain the CTP, the standard deviation of the buffer filtered gaze positions is calculated. The mean value of the sample is set as the CTP if the standard deviation is sufficiently small. The threshold varies according to the desired precision. A threshold of 20 pixels that corresponds to eye gaze being focused at a region with radius of 1.48 μm of the actual workspace was used in the presented setup. This threshold was selected experimentally because it allows consistent targeting of the microspheres. Meanwhile, it is not obstructive to the natural gaze of the user since it does not require the gaze to be focused at a very restricted workspace. It is also worth noting that the standard deviation of the eye gaze data depends on the velocity of the object that the user is trying to follow. For instance, for small particles where the Brownian motion is considerable, the standard deviation

will be larger than in the case of a more static object. Hence, the tolerance threshold is adjusted accordingly. The number of past fixation points also determines the update rate of the CTP. In this implementation, a buffer of 50 previous positions was used because it was compliant with the natural gaze in most experiments and was sufficient to identify the regions of interest. In particular, 50 previous positions with an update rate of 60 Hz corresponds to identifying the region of interest within a time period of 0.83 s. Hence, this update rate can provide a consistent position when the user has focused on a specific region for more than 0.83 s allowing to follow both the exploration phase and the target selection phase. However, for different users and desired precision, the parameter can be adjusted accordingly.

Where the new optical trap should be placed: The new optical traps are added on demand using the eye tracking data and the output of the visual servoing algorithm. The trap can be either placed directly on the detected object which is closest to the operator's fixation point or on the last identified CTP in case there is no object in the area around the CTP. In the first case, the user does not need to physically move the trap to the desired object. It also facilitates more consistent object handling when larger 3D structures are manipulated in terms of adjusting the object position and orientation. Generating a trap on a CTP aims to compensate for visual servoing inconsistencies when the objects are not detected and also add a trap at a different position on the object rather than its contour centroid. Once the laser trap is added, it maintains its relative position to the previously added laser traps.

In order to assist the user in translating the points simultaneously in a more intuitive manner, the centre of the convex hull of the generated laser trap pattern is translated by the haptic device. Let $\mathbf{p}_{ki}, i = 1, \dots, k$ and $k = L, R$ be a trap that belongs to the left ($k = L$) and right ($k = R$) group of points respectively where N_k the total number of traps of the k_{th} pattern. By the terms left and right group we refer to the number of points that are manipulated by the left and right haptic device respectively. Therefore, the centre $\mathbf{c}_k$ of the convex hull C_k of each pattern is given by:

$$\mathbf{c}_k = \frac{\sum \mathbf{x}_i}{N_{C_k}}, \quad \mathbf{x}_i \in C_k, i = 1, \dots, N_{C_k} \quad (1)$$

$$C_k = \left\{ \sum_{j=1}^{N_k} \alpha_j \mathbf{p}_{kj} \mid (\forall i : \alpha_j \geq 0) \wedge \sum_{j=1}^{N_k} \alpha_j = 1 \right\} \quad (2)$$

with $\mathbf{p}_{ki}, \mathbf{c}_k, \mathbf{x}_i \in \mathbb{R}^2$.

How to rearrange the optical trap configuration while maintaining the grasp by the remaining traps: During object transportation, the relative position among the traps of a pattern remains fixed. However, the position of an individual trap may be adjusted in order to rearrange the trapped objects or to change the pose of a 3D object. The trap reconfiguration is realized as follows; The desired laser trap is selected as the closest trap to the CTP and cannot be deselected until the trap reconfiguration is terminated. This serves the purpose of allowing the user to explore the adjacent region and not

Fig. 3. CAD model of the microstructure used in the second experiment

Fig. 4. Registered projection of CAD model (left) and initial pose with respect to global frame (right)

requiring the gaze to continuously follow the selected trap.

Since the operator is manipulating a set of traps with the haptic devices, the motion of one of the haptic devices has to be decoupled from this pattern and coupled to the trap to be repositioned. Hence, during the optical trap reconfiguration, the pattern that belongs to the respective haptic device maintains its position and the second haptic device controls the target trap. At the end of the reconfiguration state, the second haptic device is coupled again with the respective laser point pattern.

3) Gaze Contingent Guidance Constraints: In order to assist the user in placing the assembled microparts with respect to the pattern's centre, a guidance force is activated on demand to guide the operator to the desired position. The target position $\mathbf{p}_d$ is defined as the CTP and is not updated until this position is reached. This is to avoid instabilities in the force profile caused by rapid switching when target positions are close. The guidance force, as defined in [11], is given by:

$$\mathbf{F}_p = \begin{cases} 0 & \text{if } d \geq d_{max} \\ f_A \mathbf{n} & \text{if } d_{max} > d > d_{high} \end{cases} \quad (3)$$

where

$$f_A = k_0 \|\mathbf{p}_t\|^2 (\|\mathbf{p}_t\| - d_{max})^2 e^{k_1 \|\mathbf{p}_t\|} \quad (4)$$

$$\mathbf{p}_t = \mathbf{p}_d - \mathbf{c}_k \quad (5)$$

$$\mathbf{n} = \frac{\mathbf{p}_t}{\|\mathbf{p}_t\|} \quad (6)$$

with k_0, k_1 two scalar values, $\mathbf{p}_d$ the target position, d_{max} the distance between $\mathbf{c}_k$ and the target $\mathbf{p}_d$ at which the guidance force starts to increase, d_{high} the distance between

$\mathbf{c}_k$ and the target $\mathbf{p}_d$ at which the maximum force is rendered and $\mathbf{n}$ the direction towards the target.

B. 2D visual servoing

1) *Template matching for spherical object tracking:* Real-time image processing is performed in order to obtain the 2D position of the microstructures within the visible workspace. For spherical objects, a template matching algorithm is used for object detection since they maintain a relatively constant appearance in all orientations. The user can define the template according to the object that needs to be tracked. In the performed experiments the processing time varies from 1 to 3 ms. In Fig 2, the live camera feed is shown as displayed to user (Fig. 2 (left)). The detected microspheres (size of $1 \mu\text{m}$) are shown in Fig. 2 (right). The minimum matching score is set accordingly in order to identify objects which are moving sufficiently close to the focus plane. However, these features can be adjusted by defining an alternative matching template and similarity score. This method of detection is more robust to noise in the case of small spherical objects where Brownian motion is evident, although it is not particularly effective for overlapping objects. An alternative tracking strategy is implemented for larger non-spherical objects which have different appearance depending on their orientation in the 3D space.

2) *Non-spherical object tracking:* The object boundary is obtained by the following sequence. First, canny edge detection is performed followed by discarding all objects with a small number of pixels in order to reject noise. Then, the boundary is thickened so that a closed boundary is obtained. For convex-shaped objects, the convex hull of the edges is considered as the boundary to ensure closed contours for increased detection robustness. This is sufficient for applications where a closed boundary is more important than an accurate contour around the object. Contour filling and image segmentation is performed to locate the object of interest. The filled contour is used for template matching in 3D orientation estimation, as described in the next section.

C. 3D orientation estimation

1) *Optimization-based initial pose estimation and registration using correspondences:* Consider as $\mathbf{P}_{Mi} \in \mathbb{R}^3, i = 1, \dots, N$ the 3D model point cloud and $\mathbf{P}_{ci} \in \mathbb{R}^2, i = 1, \dots, M$ the boundary points as detected on the image frame. Let $\{O\}$ be the co-ordinate frame attached to the centre of mass of the micro-gear, $\{G\}$ the global co-ordinate frame, $R_o = [\mathbf{x}_o \ \mathbf{y}_o \ \mathbf{z}_o] \in \mathbb{R}^{3 \times 3}, \mathbf{t}_o \in \mathbb{R}^3$ the orientation matrix and translation vector of the gear with respect to $\{G\}$. The object is orthographically projected on the image plane:

$$\mathbf{P}_{ci} = K P (R_o \mathbf{P}_{Mi} + \mathbf{t}_o) \quad (7)$$

where $K = \text{diag}\{p_x, p_y\} \in \mathbb{R}^{2 \times 2}$, p_x, p_y the intrinsic parameters and

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (8)$$

the projection matrix of the 3D object on the image plane. The object boundary is detected as described in Section III-

Fig. 5. (a) Initial pose (top left) with matched template (middle left) and corresponding 3D orientation (bottom left) and (b) final pose (top right) with matched template (middle right) and corresponding 3D orientation (bottom right)

B.2. The detected boundary at the initial position is registered with the projection of the 3D model points using correspondences which are selected manually by the user. The manually chosen 3D points are projected on the image plane. After registering the detected boundary with the projection of the CAD model on the image frame, the initial rotation and translation of the object is estimated by the following minimization problem

$$\hat{R}_o, \hat{\mathbf{t}}_o = \arg \min \| \mathbf{P}_{ci} - K P (R_o \mathbf{P}_{Mi} + \mathbf{t}_o) \|^2 \quad (9)$$

$$\text{s.t } R_o^T R_o = I_{3 \times 3}$$

where $\hat{R}_o, \hat{\mathbf{t}}_o$ the optimal rotation matrix and translation vector respectively and $I_{3 \times 3}$ the 3×3 identity matrix.

2) *Orientation estimation with online template generation:* The previous minimization approach needs robust tracking of the corresponding points at each subsequent frame. It also requires that the corresponding features are always visible. Therefore, template matching is employed in this stage using online generated templates derived relatively to the previously estimated object orientation. The templates are generated from a series of all possible orientation changes, i.e all possible rotations around the axis of the global frame, by a small angle step $\Delta\theta$, including the last estimated orientation. The possible micro-gear orientations are projected on the image plane to create the library of the templates. The number of generated templates is constant for

Fig. 6. The operator view during the performed experiments (left). The processed image feed displaying the detected microspheres (red) and the candidate target position as calculated from the operator's gaze (right)

every iteration. The similarity measure which is used for the matching is the normalized cross-correlation which is equal to

$$\gamma = \frac{\sum_{x,y} (I(x,y) - \bar{I}_{u,v})(t_m(x-u, y-v) - \bar{t}_m)}{\left(\sum_{x,y} (I(x,y) - \bar{I}_{u,v})^2 \sum_{x,y} (t_m(x-u, y-v) - \bar{t}_m)^2 \right)^{0.5}} \quad (10)$$

where I the processed current image, $\bar{t}_m$ the mean of the template and $\bar{I}_{u,v}$ the mean of the region within the template.

D. Experimental Setup

1) *Hardware components*: The system consists of a master and a slave component (Fig. 1). The master component includes two haptic devices (Geomagic, USA) and an eye tracking device (Tobii Technologies AB, Sweden) that tracks the operator's eye gaze. The haptic devices control the position of the laser traps while the data from the eye tracker is used to determine the user's intentions and define the operating states of the software controller as described in Section III-A.2. The slave part is the Acousto-Optic Deflector (AOD) OT (Elliot Scientific, UK). The OT system consists of a 1070 nm fiber laser with variable output power (Ytterbium Fiber Laser, IPG Photonics, USA), an optical microscope with an oil immersion objective lens (Nikon Ti) and a high frame rate CCD camera (Basler AG, Germany). The laser traps are placed within the workspace by direct laser actuation while multiple traps are generated through time multiplexing.

2) *Micro-gear fabrication*: The object used in the conducted experiment is a gear-shaped structure with radius of 8 μm and height of 9.6 μm. The CAD model with the respective dimensions are shown in Fig. 3. The structure was printed on the glass substrate using the Nanoscribe 3D printer (Nanoscribe, Germany) and IP-L photoresist (Nanoscribe, Germany). A set of mechanical micromanipulators (Imina Technologies, Switzerland) was used to detach the printed objects from the glass surface. The detached objects were placed in a solution of deionized water.

IV. EXPERIMENTAL RESULTS

Two sets of experiments are presented demonstrating the functionality of the proposed framework. The first set of experiments compares the performance of the gaze contingent

Fig. 7. Time - first assembly [s]

Fig. 8. Time - second assembly [s]

Fig. 9. Total time for task [s]

Fig. 10. Distance over time [μm s]

TABLE I

MEAN VALUES AND P-VALUES OF METRICS FOR MICROASSEMBLY USER

		EXPERIMENT		
METRIC		GAZE	MANUAL	P-VALUE
First assembly	[s]	49.8	67.7	< 0.001
Second assembly	[s]	13.8	63.3	< 0.001
Total time	[s]	63.6	131	< 0.001
Distance over time	[mm s]	4.19	4.75	0.118

controller to the standard OT control interface. The second experiment demonstrates non-spherical object manipulation using the gaze contingent framework in conjunction with 3D orientation estimation to recover the orientation of the microstructure.

A. Gaze contingent controller performance assessment in microassembly with 2D visual servoing

1) *Task description*: The effectiveness of the gaze control scheme was evaluated through a set of user experiments comparing the execution time, the manipulation flexibility and repeatability of the task when the gaze controller and the standard user interface for the OT are used. The standard user interface uses the mouse as an input device and does not allow the formation and transportation of an arbitrary object configuration. The task was to assemble 7 microspheres in a specific configuration and then move the assembly at a different position. The desired assembly positions were overlaid on the displayed image (Fig. 6). The purpose of the task was to demonstrate how multiple small sized microstructures can be assembled in an arbitrary configuration using the gaze controller. The laser power was set to 200mW. The task was repeated 10 times with the gaze contingent control scheme and 10 times with the mouse interface. The user was familiar with both modes of operation prior to performing the experiment. The 20 experiments were performed interchangeably; once with the gaze controller and once with the mouse control so that the two consequent attempts are performed under comparable experimental conditions. The guidance force was activated

Fig. 11. (a) Initial orientation of micro-gear with respect to global frame as viewed from the XY plane (b) Final orientation of micro-gear with respect to global frame as viewed from the XY plane (c) Trajectory of x-axis of object frame (d) Trajectory of y-axis of object frame (e) Trajectory of z-axis of object frame

at a distance $d_{max} = 6 \mu\text{m}$ from the CTP with maximum force rendered at $d_{high} = 3 \mu\text{m}$. The distances were chosen so that guidance is activated sufficiently close to the CTP. The maximum guidance force was set to 0.8 N to provide soft haptic guidance. Guidance was only activate for the second stage of the experiment.

2) *Results and discussion:* The aim of the repeated experiment is to demonstrate that the execution time of the individual assemblies and overall task as well as the total travelled distance of the laser traps are consistently reduced when using the eye tracking interface. Therefore, the metrics that are used for performance assessment are the time to perform the first assembly, the time for the second assembly (reposition the assembly or repeat the assembly at the desired position), the total execution time of the task and the cumulative distance over time, i.e. the sum of all integrals of the distance covered by each laser trap with respect to time. The means and p-values of the metrics are given in Table I. The p-values were calculated using the Two-sample t -test. The results demonstrate that the execution time is significantly reduced in both subtasks and overall task (Fig. 7 - 9) while the distance over time metric also shows reduced travelled distance throughout the repeated experiments (Fig. 10). The reduced travelled distance suggests that placing the trap directly on the desired object improves the manipulation efficiency compared to the manual case where the trap is placed arbitrarily and has to be physically moved towards the target object. In addition, the synergy of eye tracking and visual servoing compensates for inconsistent object tracking due to the Brownian motion since at that case the trap will be place on the CTP. Similarly, visual servoing allows the user to interact seamlessly since the optical trap is positioned on the tracked object which is closer to the CTP, disregarding the need of unnatural user interaction with the system.

B. Manipulation of a micro-gear and 3D orientation estimation

1) *Task description:* In the second experiment, a micro-gear is manipulated using the gaze contingent controller. The laser power was set to 400mW during the experiment.

Fig. 12. Rotation angles of the z,y and x axis of the object frame (θ_z , θ_y and θ_x) with respect to global frame

An increase of the laser power is required compared to the previous experiment due to the larger dimensions of the microstructure. The task was to demonstrate how 3D rotation can be performed using the gaze controller while the achieved pose is estimated for each trapping configuration. The proposed orientation estimation algorithm complements the gaze controller by assisting the user's perception in understanding the force dynamics and kinematic properties of the object. It can be seen that the relative trap positioning affects the object orientation due to its non-uniform structure. These orientation changes are not evident when manipulating spherical objects hence it is not necessary to recover the sphere's orientation.

2) *Results and discussion:* The results show the estimated orientation when the microstructure is rotating simultaneously about all three axis. The total time that was required for the object to rotate was 16.6 s. The initial orientation $\hat{R}_0$

and translation $\hat{\mathbf{t}}_o$ given by (10) are equal to

$$\hat{R}_o = \begin{bmatrix} -0.9849 & -0.1377 & 0.1055 \\ 0.1554 & -0.9706 & 0.1838 \\ 0.0771 & 0.1975 & 0.9773 \end{bmatrix}, \hat{\mathbf{t}}_o = \begin{bmatrix} 0.4183 \\ 0.4051 \\ 0 \end{bmatrix} \quad (11)$$

as can be seen in Fig. 11 (a). The trajectories of the placed laser traps for each recorded sample are depicted in Fig. 13 and result to the estimated angles defining the object's orientation as shown in Fig. 12. The difference between the number of samples in the two set of data is attributed to the fact that the image feed has a slower update rate than the master control loop.

The initial and final orientation of the micro-gear as well as the recovered orientation trajectory are shown in Fig. 11. Fig. 11 (c)-(e) depict the individual trajectories of the x,y and z axis of the object frame during the performed experiment. The object coordinate frame origin is attached to centre of mass of the object. The consequent orientations depicted in this figure are downsampled for illustration purposes. The proposed algorithm does not depend on feature tracking and the shape of the structure. It also provides a good estimate in case the 3D-printed object is different from the CAD model (Fig. 5).

V. CONCLUSIONS

In this paper, a gaze contingent control framework for optical micromanipulation is proposed. The developed controller allows the operator to add and reposition the optical traps interactively using the eye gaze combined with real time 2D visual servoing. The synergy of the eye gaze and visual servoing aims to identify the objects of interest and to compensate for tracking inconsistencies. Haptic guidance is provided to the user to achieve more consistent object transportation. A method for 3D orientation estimation is also presented. The capabilities of the proposed design are assessed through two sets of experiments; one in multiple microsphere 2D manipulation and a second one in 3D object manipulation. The user experiments for the first task shows that task execution is faster and hand-eye coordination is improved in comparison to the standard control OT mode. It is shown that the proposed orientation recovery algorithm provides a good estimate of the orientation of the object despite the difference between the 3D-printed object and the CAD model. Future work includes the extension of the gaze control to provide more coordinated trap positioning and real time 6 degree-of-freedom pose estimation. These components will allow the derivation of an algorithm that provides kinematic modelling for arbitrary shaped structures and identify the impact of different trapping configurations for improved manipulation effectiveness.

ACKNOWLEDGMENTS

The authors would like to thank Konrad Leibrandt for his contribution to the software implementation, Maura Power for her input on the experimental part and Dr. Stamatia Giannarou for her contribution to the image processing components in this work.

Fig. 13. Optical trap trajectories during object manipulation

REFERENCES

- [1] H. Chen and D. Sun, "Moving Groups of Microparticles Into Array With a RobotTweezers Manipulation System," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1069–1080, 2012.
- [2] R. Haghighi and C. C. Cheah, "Optical Manipulation of Multiple Groups of Microparticles using Robotic Tweezers Optical Manipulation of Multiple Groups of Microparticles using," *IEEE Transactions on Robotics*, vol. 32, no. April, pp. 275–285, 2016.
- [3] S. Chowdhury, P. Svec, C. Wang, W. Losert, and S. K. Gupta, "Gripper synthesis for indirect manipulation of cells using holographic optical tweezers," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2749–2754, 2012.
- [4] F. Arai, T. Endo, R. Yamuchi, and T. Fukuda, "3D 6DOF manipulation of micro-object using laser trapped microtool," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE, 2006, pp. 1390–1395.
- [5] G. Z. Yang, G. P. Mylonas, K. W. Kwok, and A. Chung, "Perceptual docking for robotic control," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5128 LNCS, pp. 21–30, 2008.
- [6] G. P. Mylonas, K. W. Kwok, A. Darzi, and G. Z. Yang, "Gaze-contingent motor channelling and haptic constraints for minimally invasive robotic surgery," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5242 LNCS, pp. 676–683, 2012.
- [7] C. Basdogan, A. Kiraz, I. Bukusoglu, A. Varol, and S. Doganay, "Haptic guidance for improved task performance in steering microparticles with optical tweezers," *Optics Express*, vol. 15, no. 18, p. 11616, 2007.
- [8] K. Berk Yesin and B. Nelson, "Robust CAD model-based visual tracking for 3D microassembly using image space potentials," *2004 IEEE International Conference on Robotics and Automation (ICRA)*, no. April, pp. 1868–1873, 2004.
- [9] B. Tamadazte, E. Marchand, S. Dembele, and N. Le Fort-Piat, "CAD Model-based Tracking and 3D Visual-based Control for MEMS Microassembly," *The International Journal of Robotics Research*, vol. 29, no. 11, pp. 1416–1434, 2010.
- [10] Z. Ni, A. Bolopion, J. Agnus, R. Benosman, and S. Régnier, "Asynchronous event-based visual shape tracking for stable haptic feedback in microrobotics," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1081–1089, 2012.
- [11] K. Leibrandt, C. Bergeles, and G.-Z. Yang, "On-Line Collision-Free Inverse Kinematics with Frictional Active Constraints for Effective Control of Unstable Concentric Tube Robots," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, 2015.