Proceedings of the 6th IFAC Symposium on Mechatronic Systems, Zhejiang University, Hangzhou, China, April 10-12, 2013

IFAC MECH2013-0056

Scale and Rotation Invariant Two View Microgripper Detection that Uses a Planar Pattern

Juha Hirvonen*. Pasi Kallio*

**Micro- and Nanosystems Research Group, Department of Automation Science and Engineering Tampere University of Technology, Tampere FIN-33101, Finland (e-mail: {juha.hirvonen,pasi.kallio}@tut.fi).*

Abstract: In automated grasping of microparts or objects with unknown dimensions and orientations, at least two cameras have to be used to acquire the depth information. In addition to recognition and reconstruction of the real-world coordinates of the target objects, the system has to be able to detect also the real-world coordinates of the microgrippers from the images. This paper presents a scale and rotation invariant microgripper detection method that uses a planar pattern. The method is suitable especially for prototyping systems, whose composition might vary between the experiments. The gripper detection is shown to be accurate enough for challenging micromanipulation tasks of small electronic components and individual paper fibers.

Keywords: Microgripper, computer vision, micromanipulation, grasping, pattern recognition.

1. INTRODUCTION

Knowing the position of the jaws of the microgripper is an essential part of any automated grasping task in a micromanipulation system. Conventionally, a camera and computer vision algorithms are used to guide the gripper towards the target objects. In general microassembly and micromanufacturing tasks, the dimensions of the parts to be manipulated are known. Then, the parts can be detected by using template matching (Pawashe and Sitti 2006, Anis et al. 2007, Zhang et al. 2008) or CAD-models (Tamadazte et al. 2009) and the grippers can be moved to pre-known depths to grasp the parts. Thus, the depth information of the parts is not needed to be acquired with the imaging system. However, when manipulating unknown parts or biological and fibrous objects, whose shape and dimensions may vary drastically, the system needs to be able to measure accurately also the depth of the objects.

At least two views are needed to reconstruct a 3D scene from image data. The camera matrixes describing the connection between the image points and the real world points have to be calculated for this task. Then, a real-world position of any point seen in both of the images can be calculated using the camera matrixes and the point correspondences. (Hartley and Zisserman 2004). Thus, a simplified strategy for automated grasping would be the following: detect the gripper and the target object from both of the images, calculate their real-world positions, move the gripper jaws to the object position, and grasp the object.

The gripper can be detected from images by using template matching (Anis et al. 2007, Bilen and Unel 2008) or certain assumptions of the part of the gripper visible to the camera (Chang and Shiu 2011). These methods require that the gripper orientation remains more or less fixed. However, especially in prototyping environments, the actuator angles, the platform composition and even the camera poses may change between the experiments. This requires updating the templates and the detection rules, which is a time-consuming and error-sensitive process. Also, if the image area is big compared to the size of the gripper, the details of the gripper might not be clear enough in the images for these methods to work accurately.

The shape of the microgripper jaws varies according to the type of objects they are used to manipulate. Often, the jaws of the grippers do not resemble sharp probes but they are rather blocky. This generates its own challenges to the detection. For example, the lower corners of the gripper jaws are not seen from above; yet, the lower corners are the first ones to clash with the platform. Thus, the crucial points of the gripper are not always visible in the images

To overcome the obstacles described above, the method to detect the gripper should be

1. rotation invariant
2. scale invariant
3. utilizing the known geometry of the gripper to detect also points that are not visible in the images.

Since use of planar patterns such as chessboard or dot grid are commonly used in camera calibration (Heikkila and Silven 1997, Zhang 1999, Bouguet 2010), there exist many ready-made methods to detect these patterns in arbitrary position from images accurately (de la Escalera and Armingol 2010, Zhu et al. 2009, Alvarez et al. 2007, Renbo et al. 2008, Bradski and Kaehler 2008). This makes use of these patterns as markers facilitating the actuator detection tempting. If the position of the point on the gripper jaw

related to the pattern is known, the position of the point on the jaw can be calculated after detecting the pattern.

This paper presents a method to detect a certain point on the microgripper jaw accurately utilizing a planar pattern fabricated on a photopaper. The test environment is a microrobotic two view test bench for manipulating paper fibres. The paper is organized as follows: Section 2 describes the test environment, fabrication of the dot grid pattern and the mathematical methods to detect the gripper utilizing the dot grid pattern. The means to measure the gripper position relative to the grid as well as the techniques to use this information for obtaining the gripper position from any image pair are covered. Section 3 discusses the experiments performed to validate the methods for gripper detection. The accuracy of the camera matrixes generated for 3D measurements are assessed, and the actual gripper detection is tested by image-based visual methods and actual grasping tests. The conclusion is drawn in Section 4.

2. MATERIALS AND METHODS

2.1 Test Environment

The test environment is a microrobotic platform for paper fibre analysis. It has been successfully used in mechanical characterization of individual paper fibres (Saketi and Kallio 2011b) and measuring strengths of individual paper fibre bonds (Saketi and Kallio 2011a). The platform has a Stacked Gantry Crane configuration. There are two tailored microgrippers attached to 3-DOF micropositioners. Under the grippers there is an XY-table, on which there is a rotary table with a tailor-made backlight illumination system for fibre analysis. All the actuators were manufactured by SmarAct GmbH, Germany.

The used micropositioners are equipped with position sensors and a control module, which allows accurate positioning with a computer. The position information is sent to and received from the actuators via SCU3DControl API, which is a manufacturer specific library for controlling the micropositioners. More detailed information of the control software is available in (von Essen et al. 2010).

There are two cameras in the system, one on top (Manta G-504b, AVT GmbH, Germany) and one diagonally on side (XCD-U100, Sony, Japan). The cameras are equipped with Navitar 12x motorized zoom and Zoom 7000 optics, respectively (Navitar, USA). In this paper, the image resolutions 1224*1028 and 1600*1200, respectively, were used. There is a tailor-made LED illumination system on top of the actuator. The field of view of the cameras is approximately 2.2cm*3.0cm in maximum. Fig. 1 presents the test environment. As it can be seen in Fig. 1B, the angle of the microgrippers can be adjusted. Different angles are used in different kind of fibre manipulation experiments. Also, the camera positions may change between the tests.

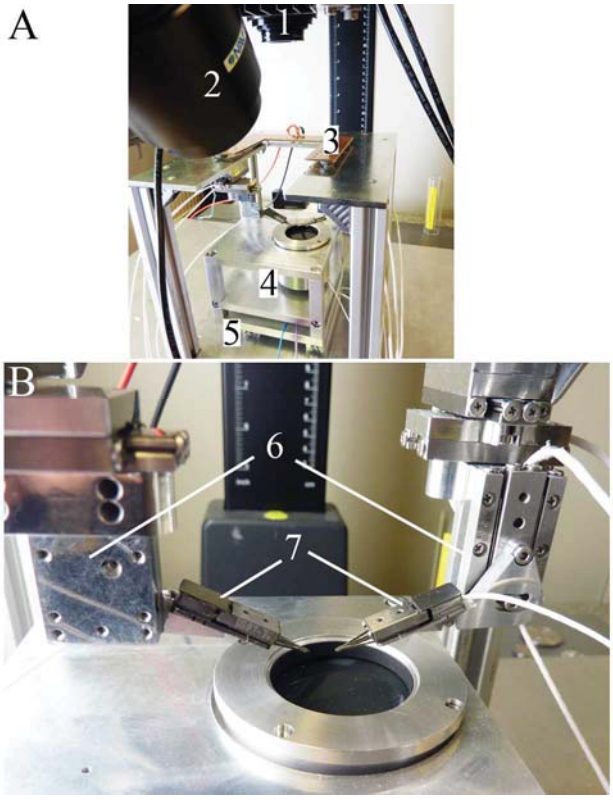

Fig. 1. The test environment. The whole platform depicted in (A) and a close-up of the gripper actuators illustrated in (B). The cameras and optics (1, 2), the LED illumination system (3), the rotary table with the backlight system (4), the XY-table (5), the 3-DOF micropositioners (6) and the microgrippers (7).

In the experiments described in this paper, only the cameras (1 and 2 in Fig. 1), the LED illumination (3 in Fig. 1) and the microgrippers attached to the micropositioners (6 and 7 in Fig. 1) were used.

2.2 Pattern Fabrication

The pattern used in tracking of the microgripper is a 5*4 dot grid with dot diameter of 0.35mm and centre-to-centre distance of 1.0mm. It was exposed on a photopaper using a precise photo mask (Micro Lithography Services Ltd, UK). The pattern was then fixed on a 1mm thick aluminium plate and attached on the microgripper. The plate was used to place the dot grid closer to the gripper jaws to increase the area where the gripper can move so that the grid is still visible for the cameras. The dot grid attached on the gripper is seen in Fig. 2.

The pattern was designed to be small enough to fit the gripper but yet big enough to contain a reasonable number of dots for tracking. The size of the dots makes them easily distinguishable in the images. The asymmetric configuration helps in defining the orientation.

2.3 Notation

In this paper, the term image coordinate relates to a 2D coordinate pair, *[x, y]*, fixed to the columns and the rows of pixels in an image, and the term 3-space relates to 3D coordinates, *[x, y, z]*, in metric units. Unless otherwise stated, all coordinates are given in their homogeneous matrix form. This means that also the scale is given. The image coordinate matrixes are named with a lower case x and the 3-space coordinate matrixes with a capital x. The further naming is given in the subscript. If there are multiple coordinates in the matrix, there is *i* in the subscript. For example, an *N*-point image coordinate matrix $\mathbf{x}_i$ is defined in (1) and a 1-point 3-space coordinate matrix $\mathbf{X}$ is defined in (2).

$$\mathbf{x}_i = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \\ w_1 & w_2 & \dots & w_N \end{bmatrix} \tag{1}$$

$$\mathbf{X} = \begin{bmatrix} X & Y & Z & W \end{bmatrix}^T , \tag{2}$$

w_i and W denote the scale, which is normally set to 1.

2.4 Gripper Measurement

The relative position of the desired point on the gripper jaw to the dot grid has to be measured to utilize the grid in detection of the jaw. For this, we take an image of the gripper with the grid attached in such manner that the camera is in 90° angle to the grid. Then, we detect the dot centres of the grid from the image using the algorithms implemented in OpenCV (Bradski and Kaehler 2008). We want to track the lower corner of the gripper jaw tip, and therefore we select the interest point above the lower corner lying on the same plane as the grid from the image by clicking it with the mouse. These coordinate matrixes for the grid and the interest point, $\mathbf{x}_{G,i}$ and $\mathbf{x}_I$, respectively, are then saved to a file. Fig. 2 shows $\mathbf{x}_{G,i}$ and $\mathbf{x}_I$ plotted on the image of the gripper.

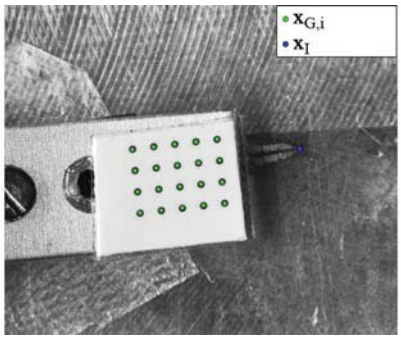

Fig. 2. The detected dot grid centres $\mathbf{x}_{G,i}$ and the chosen interest point $\mathbf{x}_I$ plotted on the image.

With this information, the image coordinate of a point corresponding to the interest point and lying on the same plane as the grid can be detected in any image taken from the gripper. Two images and the camera matrixes for the images are needed to solve the position of the interest point in 3-space. To be able to calculate the position of the lower corner of the gripper jaw, also the distance d_z between the plane the grid is and the lower corner must be known. The distance can be measured or calculated utilizing the information of the actuator dimensions given by the manufacturer.

2.5 Calculating Camera Matrixes

The purpose of the camera matrix P is to link the homogeneous image coordinates to the homogeneous metric coordinates in 3-space as follows:

$$\mathbf{x}_i = P\mathbf{X}_i \tag{3}$$

The camera matrix P is a 3x4 matrix and it can be decomposed to the intrinsic and extrinsic camera parameters by using the formula

$$P = K[R \,|\, \mathbf{t}], \tag{4}$$

where K is defined as the 3x3 calibration matrix comprising the intrinsic parameters as the focal length, the principal point and the pixel skew, and R is the 3x3 rotation matrix, which defines together with the 3x1 translation vector $\mathbf{t}$ the extrinsic parameters, i.e. the orientation of the camera coordinate frame (Hartley and Zisserman 2004).

Often, (4) is utilized in calculating the camera matrix. The calibration matrix is solved by taking multiple images of a calibration grid in different poses some of the methods depicted in (Heikkila and Silven 1997, Zhang 1999, Bouguet 2010). Then, the camera coordinate frame is fixed to one of the poses of the calibration grid. However, this is problematic in robotic systems since the relationship between the camera coordinate frame and the coordinate frame(s) of the actuator(s) of the system is unknown. This can be solved by touching certain points of the grid with the actuator(s) and linking the position data of the actuator with the camera coordinate frame (Fontana et al. 2012). However, this extra step is complicated especially in the microsystem applications since there is no valid depth information until this step is done and the collision between the sensitive actuator and the fragile grid might damage either of them. Furthermore, the calibration and the linking step are both hard to automate in microrobotic systems, and the space constraints may limit the use of the calibration grid in the system.

The camera matrix can also be calculated straight from the definition given in (2) by using known image point –3D point correspondences. As a rule of thumb, at least 28 correspondences should be used for a good estimation though six is enough for minimal solution. For each correspondence $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$ the following relationship can be defined

$$\begin{bmatrix} \mathbf{0}^T & -w_i\mathbf{X}_i^T & y_i\mathbf{X}_i^T \\ w_i\mathbf{X}_i^T & \mathbf{0}^T & -x_i\mathbf{X}_i^T \end{bmatrix} \begin{bmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{bmatrix} = \mathbf{0} , \tag{5}$$

where $\mathbf{P}^{iT}$ is the *i*-th row of *P*. *P* can be solved from (5) by singular value decomposition (SVD). (Hartley and Zisserman 2004).

The advantage of this method is that an actuator with a position sensor can be used in the calibration. The actuator is moved to certain positions, images are taken in each position, a point in the actuator is tracked in the images, and these image coordinates and the 3-space positions given by the sensor are linked. This kind of virtual point strategy has been used also in (Bilen and Unel 2008, Ammi et al. 2009). If the end-effector tip i.e. the gripper jaw corner could be easily detected from the images, the generated camera matrix would already link the tip location in the images to the metric actuator position in 3-space. However, as discussed in Introduction, this is not always feasible.

Here, we utilize the dot grid attached to the gripper. We move the gripper to 32 positions, take an image of the gripper in each position with both of the cameras, and detect the dot grid from each image. We link the image coordinates of the origin of the dot grid with the 3-space positions given by the position sensors of the system, and solve the camera matrixes for the first and the second camera, *P* and *P'*, using (4). Fig. 3 illustrates the procedure.

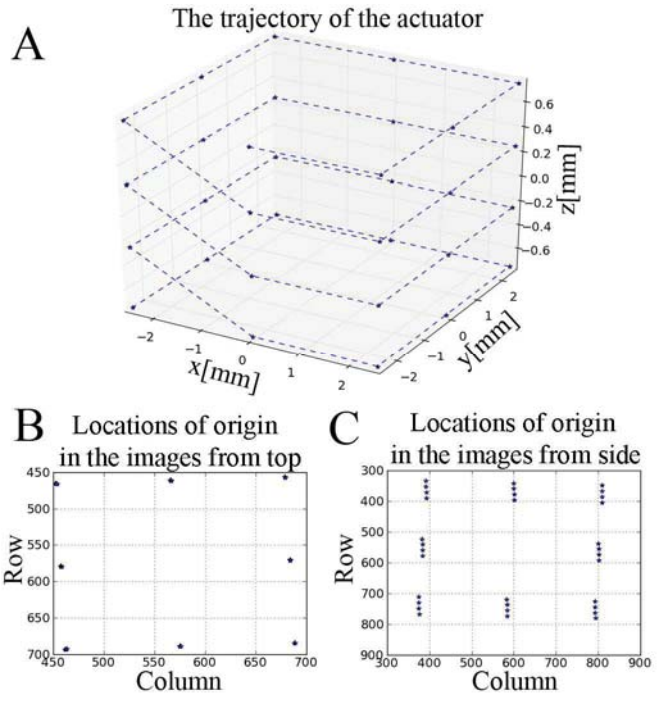

Fig. 3. The trajectory of the gripper (A), the image coordinates of the origin of the dot grid in the top camera images (B) and the corresponding image coordinates in the side camera images (C).

The resulting camera matrixes have now coordinate frames parallel to the coordinate frames of the gripper. Therefore, they can be conveniently used to solve the required movements to reach the target object when the gripper jaw position is known.

2.6 Detecting Gripper Jaw

Let $\mathbf{x}_{grid,i}$ denote the image coordinates of the dot grid centers in the image of the first camera and $\mathbf{x}'_{grid,i}$ in the image of the second camera. Using these image coordinates, the earlier measurement result $\mathbf{x}_{G,i}$ (Section 2.4) and the direct linear transformation (DLT) method (Hartley and Zisserman 2004), we can find homographies *H* and *H'* so that

$$\begin{aligned} \mathbf{x}_{grid,i} &= H\mathbf{x}_{G,i} \\ \mathbf{x}'_{grid,i} &= H'\mathbf{x}_{G,i} \end{aligned}. \tag{6}$$

Now, the image coordinates of the interest points coplanar to the grid $\mathbf{x}_{interest}$ and $\mathbf{x}'_{interest}$ can be solved by using the homographies and $\mathbf{x}_I$ measured in Section 2.4:

$$\begin{aligned} \mathbf{x}_{interest,i} &= H\mathbf{x}_I \\ \mathbf{x}'_{interest,i} &= H'\mathbf{x}_I \end{aligned}. \tag{7}$$

To find the position of the lower corner of the gripper jaw $\mathbf{X}_{ee}$ in 3-space, we first need to calculate the position of the interest point $\mathbf{X}_{interest}$ and then solve the normal to the plane the grid defines. Since we know the camera matrixes *P* and *P'*, $\mathbf{X}_{interest}$ can be determined by solving the following pair of equations:

$$\begin{cases} \mathbf{x}_{interest} = P\mathbf{X}_{interest} \\ \mathbf{x}'_{interest} = P'\mathbf{X}_{interest} \end{cases} \tag{8}$$

by using e.g. the DLT method.

To find the normal to the plane the grid defines, we need to calculate also the positions of the dot grid centres in 3-space, $\mathbf{X}_{grid}$, similarly to (8). By fitting $\mathbf{X}_{grid}$ to the equation of a plane

$$ax + by + cz + d = 0 \tag{9}$$

we get the unit normal $\|n\|$

$$\|n\| = \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2+c^2}} & \frac{b}{\sqrt{a^2+b^2+c^2}} & \frac{c}{\sqrt{a^2+b^2+c^2}} & 0 \end{bmatrix}^T. \tag{10}$$

Now, $\mathbf{X}_{ee}$ can be solved

$$\mathbf{X}_{ee} = \mathbf{X}_{interest} - d_z \|n\|. \tag{11}$$

The point $\mathbf{X}_{ee}$ can be projected back to the images to points $\mathbf{x}_{ee}$ and $\mathbf{x}'_{ee}$ for visual assessments of the detection accuracy utilizing the camera matrixes *P* and *P'* as in (3). Fig. 4 presents the detection procedure by visualizing the points in the images.

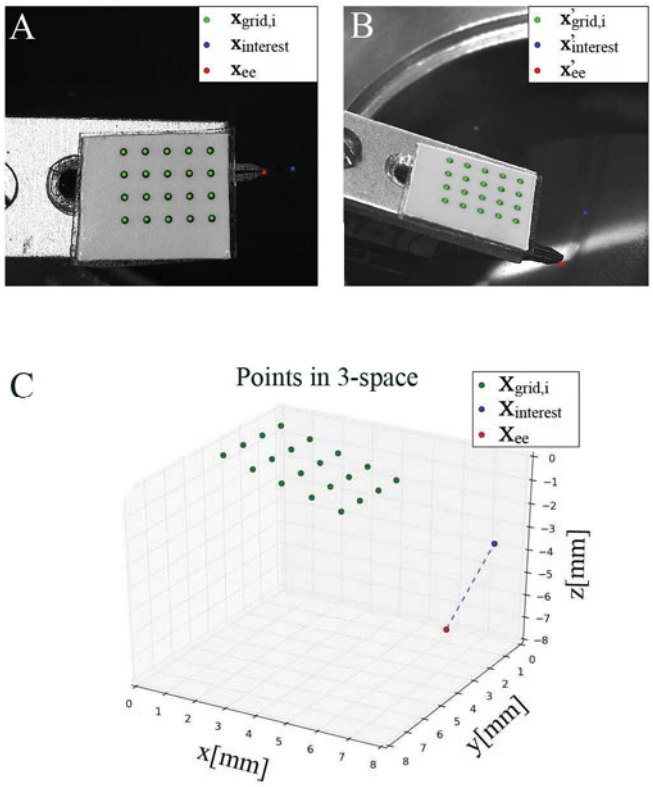

Fig. 4. Detecting the gripper jaw corner from two images by using the dot grid. The first image (A), the second image (B) and the detected points in 3-space (C).

With the algorithm described in this section, we can calculate the position of the gripper jaw from any pair of images in which the dot grid is visible.

2.7 Factors Affecting the Accuracy of the Method

The method is image-based and thus the lens distortion of the optics affects the accuracy of all the steps of the method. According to the manufacturer, the worst case distortion is 0.5%. If the image corners are avoided, this causes at worst a couple of pixels distortion in the images. As a comparison, the diameter of one dot of the grid is approximately 50 pixels in the images used in Section 2.4 and 10 – 15 pixels in the images used in Sections 2.5 and 2.6.

The dot grid detection algorithm is used in all the steps and thus its accuracy affects all the measurements. The algorithm has subpixel accuracy and it is generally claimed to be better than 1/10 pixels.

We assume that the user can select the interest point $\mathbf{x}_i$ in ±1 pixel accuracy in the measurements done in Section 2.4. However, the greater d_z is, the more perspective error occurs in the measurement. In our case d_z was 4.1mm and the perspective error was negligible.

Since the camera matrixes are calculated by using the point correspondences, the accuracy depends on the accuracy of the points used. The image point measurement accuracy depends on the detection method and the lens distortion, which were discussed already in the previous paragraphs. The resolution of the position data read from the sensor of the individual micropositioners is 0.1µm but since the 3-DOF micropositioner is comprised of three linear micropositioners (Fig. 1), also the possible alignment error affects the accuracy of the 3-space points. Since it is difficult to define the overall accuracy of the camera matrixes theoretically, experimental evaluation is performed in Section 3.1.

3. TESTS AND RESULTS

3.1 Evaluation of Camera Matrixes

The accuracy of the camera matrixes affects the calculation of all 3-space coordinates and thus they are assessed with two tests. In the first test of the accuracy of the camera matrixes, we moved the actuator to 144 positions inside the trajectory used in the camera matrix calculation, and we compared the position data given by the sensors with the positions calculated from the image data. Table 1 presents the position error.

Table 1. The position errors inside the camera matrix calculation trajectory.

	Min	Mean	Max
Position error / µm	0.6	3.0	6.7

As seen in the table, the position error is small compared with the volume the gripper moves in and the gripper size.

In the second test, we generated new camera matrixes utilizing (4) and the calibration procedures of OpenCV. The calibration object used was a commercial test target with dots with a diameter of 0.25mm printed with 1.0mm spacing on opal glass (Edmund Optics, USA). We compared the performance of these camera matrixes with the camera matrixes calculated in Section 2.5 by calculating a known 3D structure from the same image pair with the both camera matrix pairs. The 3D structure used was the aforementioned calibration grid lying inclined on the platform. We detected the dots by using OpenCV and calculated their 3-space positions with both camera matrix pairs. It is worth noting that the 3-space coordinates the new camera matrix pair produces are not aligned with the grippers. Fig. 5 shows the grid and the calculated 3-space positions.

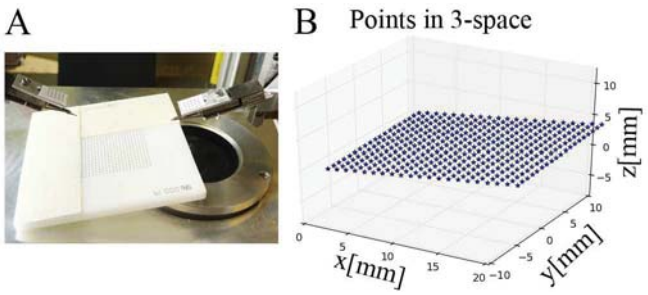

Fig. 5. Tilted calibration grid on the platform (A) and the dot centres of the grid in 3-space (B).

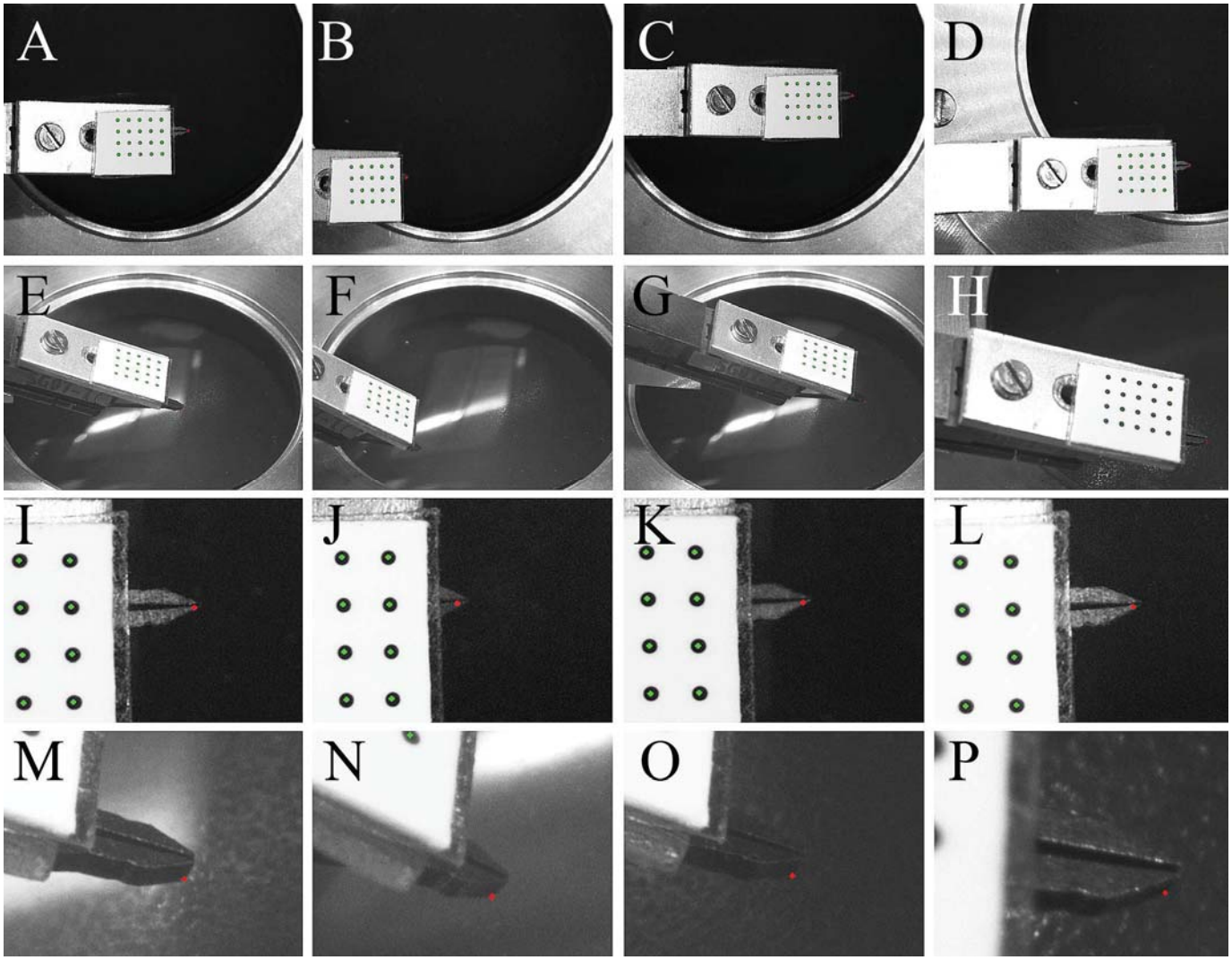

Fig. 6. The calculated gripper jaw position projected back to the images. The images of the first camera (A–D), the images taken with the second camera (E–H), close-ups from the images of the first camera (I–L), and close-ups from the images of the second camera (M–P). The gripper back-projected point is shown in red, the dot grid centres are shown in green.

The average distance between the dots i.e. the dot spacing was calculated from both of the 3-space position matrixes gained. Table 2 shows the results.

Table 2. The reported dot spacing and tolerance of the commercial calibration grid compared with the measured dot spacing and standard deviation by two methods.

	Dot spacing	*Tolerance / mm*
Reported	1.000	±0.002
Camera matrix generation style	*Measured avg. dot spacing / mm*	*Std / mm*
Calibration object	0.993	0.005
Actuator	1.008	0.009

As seen in the table, the actuator-based camera matrix computation gives slightly worse results but the difference is not big. Still, the errors are small compared to the size of the grippers and the field of view. It is probable that the main error source is the assembly error in the 3-DOF micropositioner.

The errors in the camera matrixes calculation utilizing a calibration object are mainly caused by blurring of some of the dots when the grid is tilted to different angles as required in the calibration. This causes error to the detection of the dot centres and thus hinders the following steps.

3.2 Gripper Detection Tests

To test the gripper detection accuracy, we performed five series of 32 movements with the gripper. We took images with both of the cameras at each gripper position. The trajectory in each series resembled the one shown in Fig. 3C, except the starting point was always different. We changed the gripper angle between the series and changed also once the pose of the second camera. We then calculated the position of the lower corner of the gripper jaw from each image pair using the method described in Section 2.6. Since

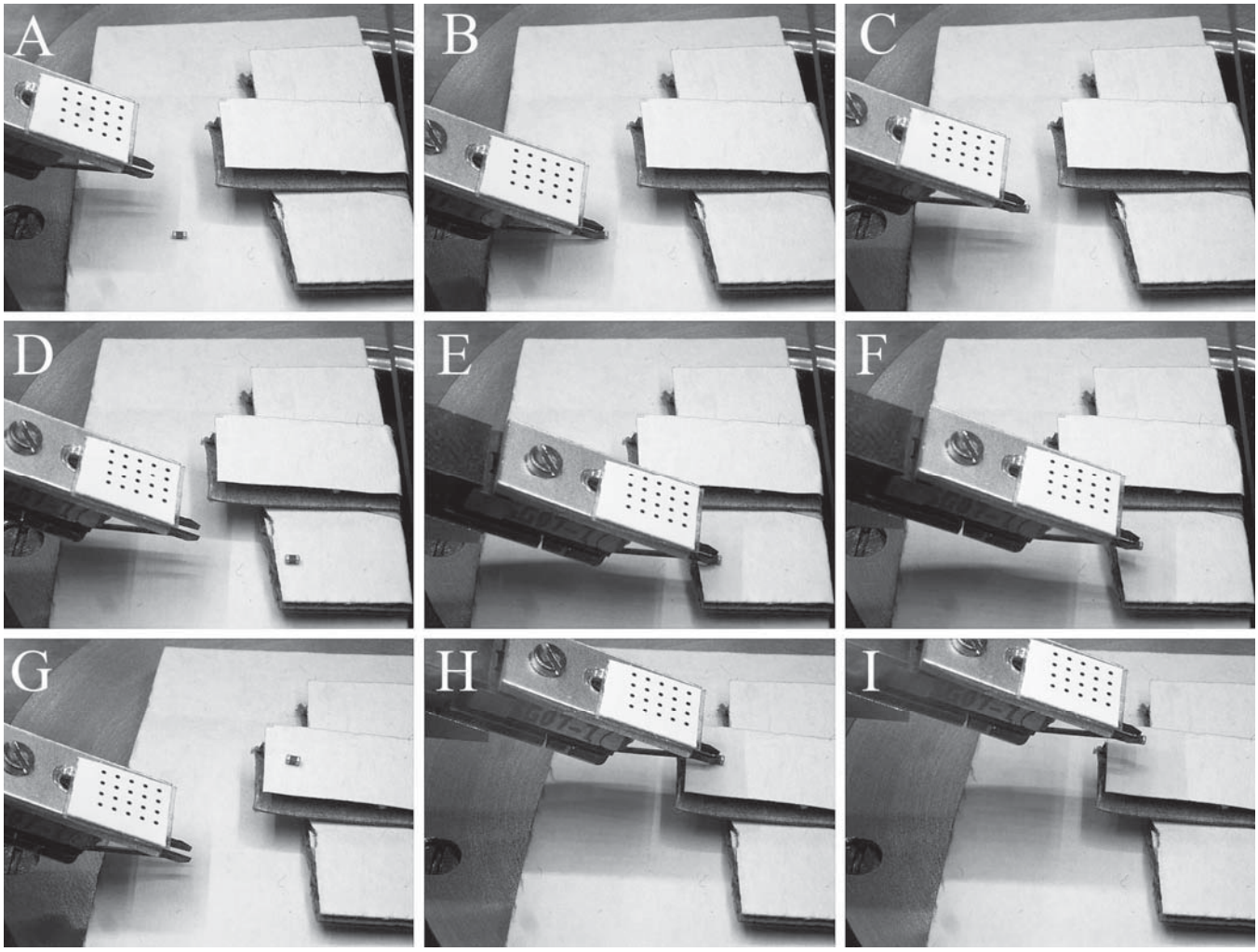

Fig. 8. Grasping the capacitor with the gripper. Initial position (A, D, G), grasp (B, E, H) and lift (C, F, I).

obtaining a reference measurement for the gripper jaw position is extremely challenging, we used the back-projection and visual assessment described in the end of Section 2.6 to provide the qualitative accuracy of the results. Fig. 6 shows the detected positions of the jaw back-projected to four of the images. The detection accuracy does not change as a function of gripper angle or camera pose, as shown.

3.3 Grasping Tests

To test the performance of the gripper jaw position detection in real operation, we performed two kinds of grasping tasks. In the first test set, we manipulated a surface mount capacitor (Taiyo Yuden, China) with case dimensions 0.4mm*0.2mm*0.2mm with the gripper. We placed a rough cardboard surface with different levels in the test bench to create random altitudes. We placed the capacitor to different locations on the surface, moved the gripper to a random position, and took an image with both of the cameras. We selected the approximate centre point of the capacitor from the images manually, detected the gripper automatically, and calculated their 3-space positions. Then, we opened the gripper, moved it the difference of the 3-space positions, and closed it. The capacitor was always between the gripper jaws and the jaws never touched the surface. Grasping tests are illustrated in Fig. 8.

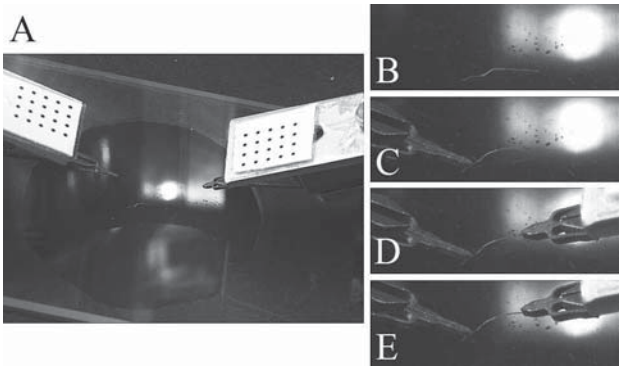

Fig. 7. Grasping a paper fibre with two grippers. Initial positions (A), a close-up of the initial position of the fibre (B), grasping the fibre with the left gripper (C), grasping the fibre with the right gripper (D), lifting the fibre (E).

In the second test set, a paper fibre (pine, average diameter ~30μm) was grasped from its both ends. For this task, we attached 100μm wide gripper tips on the gripper jaws since they are commonly used for grasping fibrous materials (Saketi and Kallio 2011a, Saketi et al. 2010). This required repeating the measurement procedure described Section 2.4 since the gripper dimensions changed. Also, we used two grippers in this manipulation task, and thus the measurement and camera matrix calculation had to be done also for the second gripper.

Since the fibre tends to move and bend when it is grasped, we calculated first the left end 3-space position and grasped the left end with the left gripper, and repeated then the same steps for the right end. Otherwise, the measurement strategy was similar to the case with capacitor. We successfully grasped the paper fibre by utilizing the gripper jaw detection algorithm. Fig. 7 illustrates grasping the fibre.

The purpose of these grasping tests was only to show that the accuracy of the gripper jaw detection is adequate for challenging grasping tasks. Object recognition and grasping point calculation algorithms were outside the scope. Therefore, the operator recognized the target objects and selected the grasping points. Also, the trajectories used were naive axis-by-axis movements.

4. CONCLUSION

This paper presented a scale and rotation invariant method for detecting a microgripper jaw in a two view microrobotic system accurately enough for challenging micrograsping tasks. The method is based on using a planar dot grid attached to the gripper. The position of the microgripper jaw in relation to the grid is measured from an image, and after this the microgripper jaw position can be calculated by detecting the dot grid from an image pair. A ready-made function in OpenCV library was used for the dot grid detection.

The camera matrixes needed for the 3-space measurements were generated by moving a trajectory with the microgripper and linking the data of the position sensor of the gripper and the image-based measurements of the location of the dot grid. The factors affecting the accuracy of the camera matrixes and the detection method in general were discussed. Also, the accuracy of the camera matrixes was validated experimentally.

The gripper detection accuracy was evaluated visually and by performing grasping tests of a tiny surface mount capacitor and a paper fibre. The results showed good accuracy and repeatability. The future experiments include implementing the algorithm to fully automated grasping tasks.

REFERENCES

Alvarez, L., Salgado, A. and Sánchez, J. (2007). Robust detection and ordering of ellipses on a calibration pattern. *Pattern Recognition and Image Analysis,* 17 (4), pp. 508-522.

Ammi, M., Frémont, V. and Ferreira, A. (2009). Automatic camera-based microscope calibration for a telemicromanipulation system using a virtual pattern. *IEEE Transactions on Robotics,* 25 (1), pp. 184-191.

Anis, Y.H., Mills, J.K. and Cleghorn, W.L. (2007). Automated microassembly task execution using vision-based feedback control. *Proceedings of the 2007 International Conference on Information Acquisition, ICIA* 2007, pp. 476-481.

Bilen, H. and Unel, M. (2008). *Micromanipulation using a microassembly workstation with vision and force sensing.*

Bouguet, J.-Y. Camera Calibration Toolbox for Matlab. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html [10/01, 2012].

Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer Vision with the OpenCV Library.* O'Reilly Media. CA, USA.

Chang, R.J. and Shiu, C.C. (2011). Vision-based control of SMA-actuated polymer microgripper with force sensing. *2011 IEEE International Conference on Mechatronics and Automation, ICMA 2011* 2011, pp. 2095-2100.

de la Escalera, A. and Armingol, J.M. (2010). Automatic chessboard detection for intrinsic and extrinsic camera parameter calibration. *Sensors,* 10 (3), pp. 2027-2044.

Fontana, G., Ruggeri, S., Fassi, I. and Legnani, G. (2012). Calibration Strategies for a Micromanipulation Work-cell *The 8th International Workshop on Microfactories, IWMF 2012* 2012.

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision.* Cambridge University Press. West Nyack, NY, USA.

Heikkila, J. and Silven, O. (1997). Four-step camera calibration procedure with implicit image correction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1997, pp. 1106-1112.

Pawashe, C. and Sitti, M. (2006). Two-dimensional vision-based autonomous microparticle manipulation using a nanoprobe. *Journal of Micromechatronics,* 3 (3), pp. 285-306.

Renbo, X., Jibin, Z., Weijun, L. and Jinting, X. (2008). Fully automatic matching of circular markers for camera calibration. *Proceedings of 2008 3rd International Conference on Intelligent System and Knowledge Engineering, ISKE 2008* 2008, pp. 1065-1070.

Saketi, P. and Kallio, P. (2011a). Measuring bond strenghts of individual paper fibers using microrobotics. *Progress in Paper Physics Seminar* 2011a.

Saketi, P. and Kallio, P. (2011b). Microrobotic platform for manipulation and mechanical characterization of individual paper fibers. *Fine Structure of Papermaking Fibres, COST Action E54 "Characterisation of the fine structure and properties of papermaking fibres using new technologies* 2011b, pp. 129-142.

Saketi, P., Treimanis, A., Fardim, P., Ronkanen, P. and Kallio, P. (2010). Microrobotic platform for manipulation and flexibility measurement of individual paper fibers. *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on* 2010, pp. 5762-5767.

Tamadazte, B., Le Fort-Piat, N., Dembélé, S. and Marchand, E. (2009). Microassembly of complex and solid 3D MEMS by 3D vision-based control. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009* 2009, pp. 3284-3289.

von Essen, M., Kuikka, S. and Kallio, P. (2010). Control software for automated microrobotic paper fiber characterization. *Proceedings of ICCIA 2010 - 2010 International Conference on Computer and Information Application* 2010, pp. 378-381.

Zhang, H., Burdet, E., Poo, A.N. and Hutmacher, D.W. (2008). Microassembly fabrication of tissue engineering scaffolds with customized design. *IEEE Transactions on Automation Science and Engineering,* 5 (3), pp. 446-456.

Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientations. *Proceedings of the IEEE International Conference on Computer Vision* 1999, pp. 666-673.

Zhu, W., Ma, C., Xia, L. and Li, X. (2009). A fast and accurate algorithm for chessboard corner detection. *Proceedings of the 2009 2nd International Congress on Image and Signal Processing, CISP'09* 2009.

