Vision-based closed-loop control of mobile microrobots for micro handling tasks

Axel Buerkle*[a], Ferdinand Schmoeckel*[a], Matthias Kiefer*[a],
Bala P. Amavasai**[b], Fabio Caparrelli**[b], Arul N. Selvan**[b] and Jon R. Travis**[b]

[a]Institute for Process Control and Robotics, Universität Karlsruhe (TH), Germany
[b]Electronics Research Group, School of Engineering, Sheffield Hallam University, UK

ABSTRACT

As part of a European Union ESPRIT funded research project a flexible microrobot system has been developed which can operate under an optical microscope as well as in the chamber of a scanning electron microscope (SEM). The system is highly flexible and configurable and uses a wide range of sensors in a closed-loop control strategy. This paper presents an overview of the vision system and its architecture for vision-controlled micro-manipulation. The range of different applications, *e.g.* assembly of hybrid microsystems, handling of biological cells and manipulation tasks inside an SEM, imposes great demands on the vision system. Fast and reliable object recognition algorithms have been developed and implemented to provide for two modes of operation: *automated* and *semi-automated* robot control. The vision system has a modular design, comprising modules for object recognition, tracking and depth estimation. Communication between the vision modules and the control system takes place via a shared memory system embedding an object database. This database holds information about the appearance and the location of all known objects. A depth estimation method based on a modified sheet-of-light triangulation method is also described. It is fully integrated in the control loop and can be used both for measuring specific points or scanning a complete field of view. Furthermore, the novel approach of electron beam triangulation in the SEM is described.

Keywords: vision sensors, object recognition, tracking, depth estimation, micromanipulation, microrobots

1. INTRODUCTION

The utilisation of mobile microrobots for the execution of assembly and manipulation tasks in industrial applications is growing rapidly. For instance, in the field of biological manipulation, a flexible miniaturised robot, operating under an optical microscope, can be entrusted to locate, grasp and transport cells from one location to another using miniature tools mounted on its mobile platform. In such systems, characteristics such as non-destructive manipulation, accurate positioning and task repeatability are generally required. Another example can be found in the assembly of small (micron-sized) mechanical parts, operating either under an optical microscope or within the chamber of a scanning electron microscope (SEM).

In the micron range, position sensors are complicated to fabricate and utilise. The MINIMAN robots are not equipped with internal position sensors. To this end, the use of non-contact sensors, such as CCD and CMOS cameras, offers an attractive alternative which has the additional advantage of being highly flexible and adaptable to a wide range of application domains.

This paper presents the work carried out by two of the research groups in the MINIMAN project consortium. It involves the development of a microrobot-based microassembly station which is to be used for manipulation tasks and which employs a vision subsystem to provide positional feedback.

---

* {abuerkle, schmoeckel, kiefer}@ira.uka.de; http://www.miniman-project.com; http://wwwipr.ira.uka.de; Institute for Process Control and Robotics, Universität Karlsruhe (TH), Kaiserstr. 12 (Geb. 40.28), D-76128 Karlsruhe.
** {b.p.amavasai, f.caparrelli, a.n.selvan, j.r.travis}@shu.ac.uk; http://www.shu.ac.uk/mmvl; Electronics Research Group, School of Engineering, Sheffield Hallam University, City Campus, Pond St, Sheffield S1 1WB, UK.

In the next section, an overview of the MINIMAN system is presented together with its main components. Section 3 gives a detailed description of the inter-process communication protocol between the main controller and the vision subsystem including how information channels are created and how robot tasks are initiated and executed by the controller. Section 4 presents a full account of the 2-D vision system comprising the automatic focussing system (for optical microscope operations), the object recognition module and the tracking module. In section 5, a purpose-built laser system is described to accomplish object depth estimation, based on the sheet of light triangulation principle, for use under both the optical microscope and – using the electron beam instead of a laser - inside the chamber of an SEM. Section 6 describes the main features of the robot controller and final conclusions are drawn.

2. SYSTEM OVERVIEW

MINIMAN is a highly flexible and configurable microrobotic system which uses a wide range of sensors in a closed-loop control strategy. The motion drive of the robot is based on piezo-actuators allowing them to cover long distances with a maximum speed of 30 mm/s. The manipulating unit is also driven by piezo-actuators and can be easily replaced to suit the application in hand. A positional accuracy of down to 20 nm can be achieved.

The following tasks have been accomplished in teleoperated mode to verify the capabilities of the MINIMAN microrobots.

1. Assembly of micro-mechanical parts;
2. Handling (grasping, transporting and releasing) of biological cells in aqueous solution under an optical microscope;
3. Handling (grasping, transporting and releasing) of small particles (a few μm) inside the chamber of an SEM.

Figure 1 shows the configuration of the hardware and software system that controls one robot. The hardware system can control the microscope directly for automatic focussing. Communication between the microrobot controller and the vision subsystem takes place via a shared memory segment.

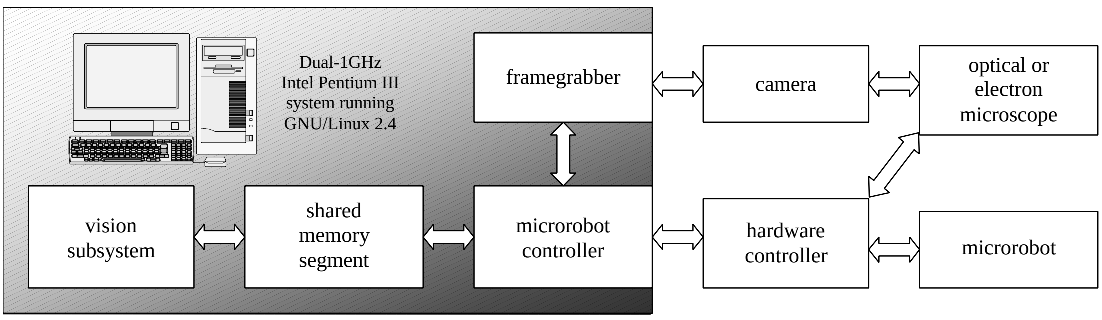


Figure 1: System overview and hardware/software configuration of the microrobot platform.

The vision subsystem which is embedded in the main system controller, plays an important role in providing the necessary information needed by the controller to locate and track both the objects to be handled and the micro-grippers attached to the robot platform. A graphical user interface provides the mechanism for the operator to select the task, point to the desired part to be handled (*click-to-grasp*) and start the movement of the robot. Automatic tracking of the robot position and of its micro-grippers is executed and no other user input is required, except for special cases where an unexpected situation is encountered (*recovery mode*).

3. COMMUNICATION INTERFACE AND OBJECT DATA-BASE

To ensure co-ordinated operation between the vision subsystem and the microrobot controller, an interprocess communication (IPC) system has been implemented. IPC takes place through shared memory using semaphore control. This allows for bi-directional flow of information data. Figure 2 shows a block diagram of the IPC system. To install a communication channel, either the vision subsystem or the microrobot controller is first initiated and upon the successful creation of the IPC link, set into wait mode. Operations may only take place once both the vision subsystem and the microrobot controller are launched. Three shared memory segments are allocated, one for the predefined object database, one for the scene images and one for objects located in the scene.

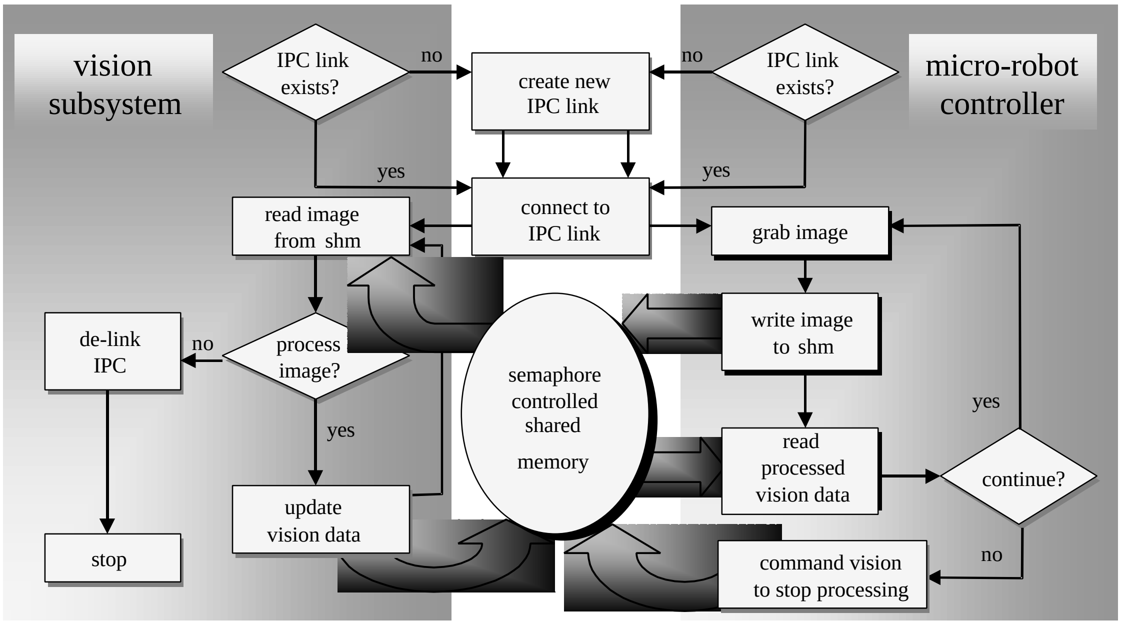


Figure 2: The shared memory controller. The microrobot controller and vision subsystems communicate via a shared memory segment (shm).

The information shared through this IPC mechanism includes:

1. the object database defined by the vision subsystem and shared with the microrobot controller
2. the scene images streaming in from the framegrabber and shared by the microrobot controller
3. the updating of object location shared by the vision subsystem

The vision system writes the predefined object database into the shared memory which holds information relating to models of the objects that could be present in the scene. The models are represented by (x,y) co-ordinate pairs of the vertices of the polygon which describes the shape of the object. Each model is also paired with a set of gripping co-ordinates that can be used by the microrobot controller to identify the gripping points of the object. Both the left and right jaw of the microrobot gripper are categorised separately as special objects with the gripping point of the gripper specified. Figure 3 shows an outline of the object database.

object index	Vertices of polygon surrounding each object	z-coord (depth)	Gripping Coordinates pairs
$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$(x_{i,1},y_{i,1})$ $(x_{I,2}, y_{i,2})$ $\vdots$ $(x_{i,n}, y_{i,n})$	depth information entered by laser system	$(x_{i,a},y_{i,a})$ $(x_{i,b},y_{i,b})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Special objects i=0 left gripper i=1 right gripper		$\vdots$	Special objects: gripper $x_{0,a} = x_{0,b}$ and $y_{0,a} = y_{0,b}$ $x_{1,a} = x_{1,b}$ and $y_{1,a} = y_{1,b}$

Figure 3: Structure of the object database stored in shared memory.

The closed-loop system for acquiring and processing scene images operates as follows. The scene images are captured by the microrobot controller via a framegrabber and are written to shared memory so that they may be accessed by the vision subsystem. The microrobot controller then instructs the vision subsystem to process the scene images. The vision subsystem then feeds back information regarding object location in the scene via shared memory to the microrobot controller. The vision subsystem then waits for the next scene image to be written into shared memory by the controller.

The use of a shared memory system simplifies distributed development since the vision subsystem and the microrobot controller may be developed independently. Furthermore, in order to increase performance, a distributed shared memory system may be used allowing multiple computers to be employed for different tasks.

4. THE 2D VISION SYSTEM

4.1 Focus estimation and auto-focusing

A major problem when dealing with images obtained through an optical microscope is the limited depth of field. Unlike a scanning electron microscopes which has a much wider depth of field, focus control is required here. Object recognition and tracking requires image information to be clear and sharp since image entropy is at maximum for a perfectly focussed image.

In order to assess the focus quality of an image, a metric or criterion of sharpness is required. When comparing a focussed image with a defocused one, it is observed that the defocused image lacks high-frequencies. Krotkov[1] evaluated and compared various methods that respond to high-frequency content in the image, such as a Fourier transform methods, gradient magnitude maximisation, high-pass filtering, measurement of histogram entropy, histogram local variations and grey-level variance. Although Fourier transform methods are very promising, it was discounted due to its high computational complexity. A comparison of the other criterions showed that the gradient magnitude maximisation method, the Tenengrad criterion (named after Tenenbaum and Schlag), was the most robust and functionally accurate. Pappas[2] summarizes the advantages of the Tenengrad criterion as follows:

1. the measured sharpness varies monotonously as the object is moved towards the position of best focus (no local maxima, single global maximum)
2. the criterion shows a strong and sharp response at the position of best focus
3. the criterion is robust to low signal to noise ratio
4. it is easy to implement and has linear complexity

This criterion, as well as most other sharpness criterions, is sensitive to noise induced by image acquisition as well as variations in illumination. The Tenengrad value of an image I is calculated from the gradient $\nabla I(x,y)$ at each pixel (x,y), where the partial derivatives are obtained by a high-pass filter, *e.g.* the Sobel operator, with the convolution kernels i_x and i_y. The gradient magnitude results to

$$S(x,y) = \sqrt{(i_x * I(x,y))^2 + (i_y * I(x,y))^2} \qquad (4.1)$$

and the final criterion for a given focus position *z* is defined as

$$C(z) = \sum_x \sum_y S(x,y)^2 , \qquad \text{for } S(x,y) > T, \qquad (4.2)$$

where T is a threshold, *e.g.* 75% of the maximum value at the last focusing distance (as suggested by Tenenbaum).

With this criterion an auto-focus strategy has been implemented. The focus setting, *i.e.* the z-position of the microscope stage, is adjusted until the sharpness function reaches a maximum. To reduce the number of evaluation steps a search strategy for maximising the sharpness function is required. Krotkov[1] describes the Fibonacci search technique which is optimal for unimodal functions. When considering only a small region of interest (ROI) of the image, the unimodality of the Tenengrad criterion function can be guaranteed. But when *a priori* information about the object locations in the scene is not available the whole image should be focussed. A typical scenario is when the micro-gripper is located above an object with the depth of field being smaller than the height distance between gripper and object. In this case it is not possible to have both components in focus. The corresponding Tenengrad function for the whole image will not be unimodal, the Fibonacci search converges to one of the two maxima, *i.e.* one of the two objects will be in focus. Such situations can be resolved by user interaction. The operator identifies a rectangular area which is to be in focus.

The sharpness criterion can also be used to estimate the depth of an object (*depth from focus*). Although such an approach is not suitable for on-line measurements, it can be useful to control the microrobot's gripper when there is no possibility of adjusting focus during a micro-handling operation, *e.g.* in a setup with a CCD or CMOS camera equipped with a macro objective having fixed focus on the objects to be handled. The gripper can be navigated towards the height of the objects without measuring the exact gripper height. While slowly moving downwards, the sharpness criterion of the gripper ROI is continuously calculated until it reaches its maximum.

Intensive tests showed that the implemented auto-focus algorithm performs extremely well. In general, ten iterations are sufficient to scan the whole z-range of the microscope stage (25 mm) and bring the desired area in focus. This can even be improved by restricting the search area or using *a priori* information.

4.2 Object recognition

A number of different feature-based techniques for 2-D object recognition have been implemented within the vision system. Arbitrary shapes can be described and represented by selecting the appropriate features to extract from a scene. Object recognition strategies based on image features typically comprises two distinct stage, firstly features are extracted from the scene and secondly a search is performed to find groups of features that are consistent with the stored models in an object database.

One of the two object recognition schemes implemented makes use of pairwise geometric histograms (PGH)[3]. PGH is a robust, statistically based method that allows scene image features to be classified according to known model features. The method can handle both occlusion and clutter and is well suited to many 2-D recognition tasks.

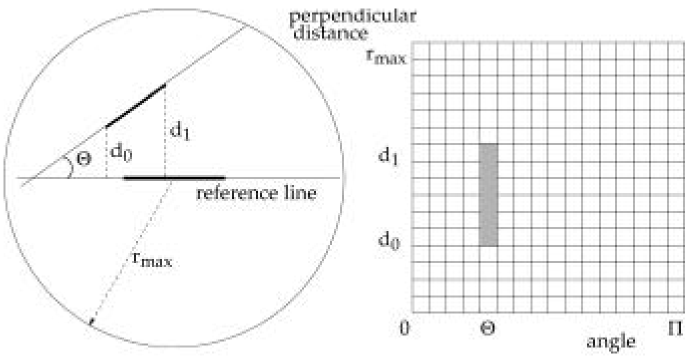


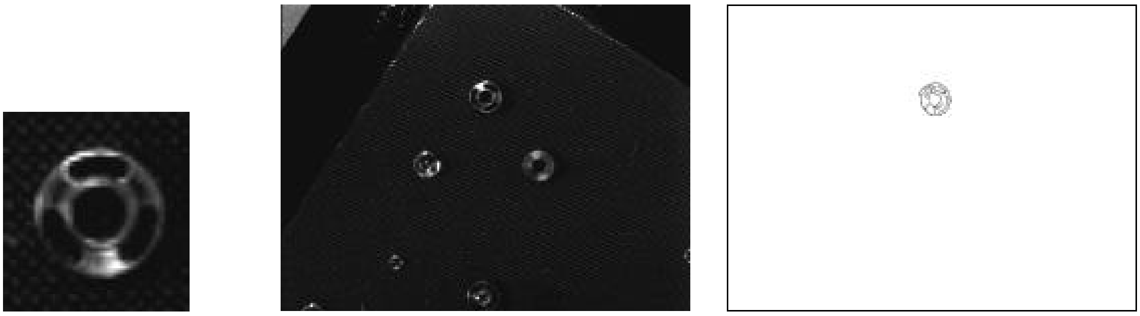
Figure 4: Construction of a PGH for a single comparison between two line segments.

A PGH is essentially a 2-D histogram that accumulates measurements of distance and orientation between each given line segment and every other line segment that is within a specified distance from it. In order to construct PGHs, edges in an image are first detected using the Canny[4] edge algorithm and then approximated using a variant of the Ballard's[5] straight line recursive-split approximation algorithm. Figure 4 shows an example of how a PGH is constructed. By applying this procedure to all the line segments forming the object shape, it is possible to represent the entire object using a distinct set of model histograms. In general, different shapes are mapped to very different sets of histograms, producing a well-founded criterion for recognition.

When an image is acquired by the framegrabber, PGHs are generated as described above and these are used to construct scene histograms. Scene histograms are matched against the model histograms using the Bhattacharyya metric[6] and a winner-take-all layer. Object classification is validated by finding consistent labelling within the scene image using a probabilistic generalised Hough transform. This stage also determines the location and orientation of one or more object models in the scene. The PGH algorithm is computationally expensive and therefore it is only used in specific situations within the context of micro-assembly. However, various optimisation schemes, including parallelisation may be used.

The PGH algorithm is used for a microassembly task. The task is to recognise and assemble components that made up a micro-lens system (2 mm diameter). The model of the lens-holder is constructed from a template image which is shown in Figure 5(a). The image of a captured scene is shown in Figure 5(b) and the recognised lens-holder is shown in Figure

5(c). The lens components are lying on a highly textural gel pack, and in order to reduce image features and to increase the speed of the PGH algorithm, the lighting conditions and the aperture of the camera have been suitably adjusted. If there are too many features in the scene, the search time of the PGH algorithm will increase significantly and if there are too few features then the PGH algorithm will fail to recognise the objects. A compromise has therefore to be found.



(a) magnified view of the lens holder from which the PGH model is created

(b) incoming scene image with objects arbitrarily rotated and translated

(c) lens holder recognised using the PGH algorithm

Figure 5: Example application of the PGH algorithm. Here the objective is to recognise and localise the lens-holder in the scene.

Another handling task consists of grasping and transporting cells in an aqueous solution using a micro-pipette. In this case, the task of the vision system is to recognise the boundaries of the cells, which have non-rigid shapes, making it unsuitable for the PGH technique. Hence the second recognition scheme implemented here makes use of active contours[7].

The task commences by the operator identifying the cell of interest, *e.g.* by a mouse click. An active contour, larger than the size of the cell, is placed around the point selected (Figure 6) and energy minimisation is then performed on the active contour to minimise it to the boundaries of the cell. The objective function to be minimised is made up of three parts, one of which is the strength of the edge boundaries. By setting this as a pre-calibrated value, the active contour can be controlled such that it will not minimise past the cell boundary.

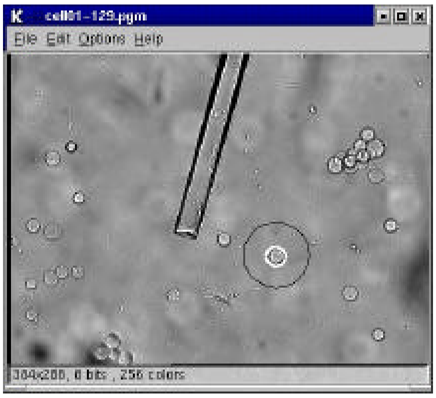


Figure 6: Demonstration showing the active contour module operating in *click-to-grasp* mode. The initial contour is shown as a black circle and the final contour as a white circle encompassing the cell.

4.3 Tracking

The development of a real-time tracking algorithm for the control system has involved the analysis and trial implementation of various methods. It has been found that simple correlation based methods working directly on grey-level images were the most appropriate to meet the real-time performance requirements of the system. It was further found that when objects contain weak features, feature-based tracking methods often fail, for example the Kanade-Lucas-Tomasi tracker was unable to track cells with weak features.

The method for tracking moving objects makes use of normalised sum-squared-difference (SSD) correlation[8]. A particular condition that has to be maintained with this method is that image intensities across the image sequences must be kept constant and consistent with the tracking template. In practise this is easily met due to:

1. the controlled operating environment which allows for lighting conditions to be controlled
2. the use of a scheme to update templates during tracking of objects.

Normalised SSD correlation is performed according to the following equation:

$$C_{x,y}(d_x,d_y)=\frac{\sum_{(a,b)\in\Re}\left(I_t(a,b)-I_f\left(x+a+d_x,y+b+d_y\right)\right)^2}{\sqrt{\sum_{(a,b)\in\Re}I_t(a,b)^2\sum_{(a,b)\in\Re}I_f\left(x+a+d_x,y+b+d_y\right)^2}} \tag{4.3}$$

where I_t and I_f are the pixel intensity points of the template and incoming frame respectively, $\Re$ is the area within the polygonal template and $d_x \in [-m,m]$ and $d_y \in [-n,n]$ specifies the search region. The objective is to minimise the correlation error such that the best match is produced by $\arg\min_{d_x,d_y} C_{x,y}\left(d_x,d_y\right)$. This is based on the fact that the minimum value of the correlation error gives the best position estimate of the template in the matching region. A common method for obtaining subpixel tracking is by oversampling the template and the scene image down to the required accuracy but this is computationally very expensive. Alternatively, an estimate can be obtained by interpolating between the distances having the lowest correlation errors.

In order to improve tracking, the template update scheme mentioned previously can work in two modes. The first mode requires the tracking template to be updated repeatedly after a fixed number of frames. In the second mode the template is updated only after the correlation value falls below a particular threshold - this removes the need for predicting a template update rate. When the correlation value drops below the threshold, one of the following situations may have occurred: (1) the lighting conditions of the scene have changed, (2) the background of the object template has changed or (3) the object has rotated slightly. Because of these, application of the template update scheme is necessary in order to obtain accurate tracking. An example showing the tracking of a micropipette tip using normalised correlation is given in Figure 7. It is worth noting that the implementation is such that arbitrary polygonal shapes may be tracked, *i.e.* it is not limited to rectangular areas.

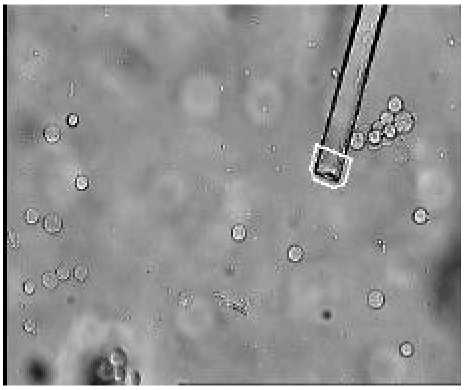


Figure 7: Tracking of pipette tip using normalized correlation. The white polygon shows the area that is being tracked.

5. THE 3D VISION SYSTEM

5.1 Depth recovery from the light microscope

Since the microscope image provides only two-dimensional position information, the control of the robots in three dimensions requires an additional sensor for depth measurements. Although there exist methods to directly extract depth information from standard optical microscopes, these methods appear to be unsuitable for continuous monitoring of assembly operations. These methods either require a series of images captured at different focus levels, which makes them inapplicable in real-time, or they measure the fuzziness of blurred structures, typically edges[9], which requires the presence of edged object contours. Instead, a method for gaining microscopic depth information by sheet-of-light triangulation has been introduced[10]. The proposed system combines (1) ease of integration into an existing microscope system, (2) fast and robust depth measurement and (3) affordability.

The measuring system consists of a line laser mounted on a micro positioning table which in turn is fixed to the microscope (Figure 8, left). The laser can be moved vertically in 100 nm steps allowing the laser line to be positioned on arbitrary objects in the scene. The measuring principle is derived from the sheet-of-light triangulation method. In the general approach, the intersection of the laser sheet-of-light with the projection ray of the microscope image point to be measured is calculated. This requires the exact laser position and projection angle as well as the projection parameters of the optical sensor system formed by microscope and camera, to be known. This causes specification inaccuracies to be accumulated. However the system presented here makes use of the robot's planar working surface. The line projected on the ground serves as a reference line. An object in the range of the laser sheet generates a displacement of the line. This offset directly corresponds to the object's height, described by the equation

$$? h = ? y / \tan f \qquad (5.1)$$

where f is the angle between the sheet of light and optical axis of the camera system. For instance, to measure the vertical alignment of the robot's gripper, a short horizontal segment of the laser line that lies on the gripper is chosen. ? y is the y-displacement between the chosen segment (herewith referred to as object line) and reference line. In contrast to the abovementioned general approach, the relative height ? h of an object only depends on two parameters, ? y and f .

Since a precise micro measuring table is used to position the laser, it is not necessary to have the reference and object line in view at once. The location of the reference line is determined only once, *i.e.* by storing a pair of parameters h_1 (laser table position) and y_1 (reference line position in the camera image). The laser is then moved to position h_2 where its light intersects the object at position y_2 in the camera image. The movement of the laser of $h' = h_2 - h_1$ results in a displacement of the reference line of $y' = h' * \tan f = (h_2 - h_1) * \tan f$. The height ? h of the object line is determined in the following way (see Figure 8, right):

$$\begin{aligned} ? h &= ? y / \tan f \quad \text{with } ? y = y_1 + y' - y_2 \\ &= (y_1 + y' - y_2) / \tan f \qquad (5.2) \\ &= (y_1 - y_2) / \tan f + h' \end{aligned}$$

This effectively means that in order to obtain the object height ? h, the laser displacement h' has just to be added.

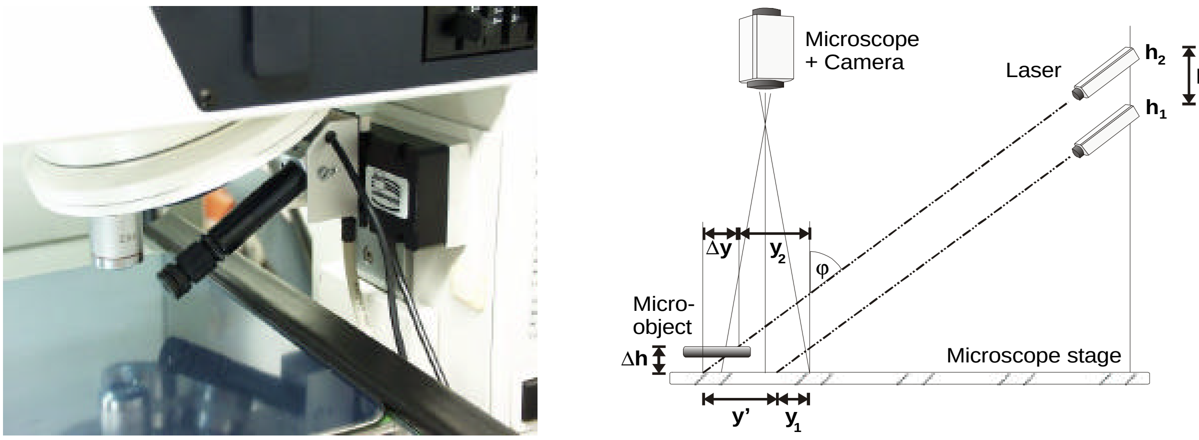


Figure. 8: Laser depth measuring system integrated into the optical microscope (left), and measuring principle (right)

Calibration of the system is performed in two steps: first, the scaling parameter of the optical system, *i.e.* microscope plus camera, is done semi-automatically using a stage micrometer and some image processing[11]. Second, the angle between the laser sheet of light and optical axis of the camera system is calculated automatically according to eq. 5.1

after moving the laser table by ? h and determining the corresponding ? y. Several measurements are carried out and averaged.

A fuzzy logic based image processing algorithm has been implemented to segment the line or line segments in the microscope image. Standard segmentation methods proved unsuitable because they do not cope with the laser line's speckled boundary or pseudo line segments introduced by reflections.
In a pre-processing step, the RGB image is converted into a grey-level image by subtracting the green channel component from the red one. This significantly reduces the white component from the microscope illumination while leaving the red laser line almost intact. In the next step, all possible vertical line segments in each image column are assigned membership values according to brightness, length and number of adjacent segments. After de-fuzzification all segments with a membership value above a threshold are considered to belong to the laser line. For each pixel of the skeleton of these segments a height value can be calculated using the method explained above.

Various experiments with different objects, *e.g.* grippers of different shape and material, coins and screws, showed that the measuring system is very robust. Fuzzy segmentation only fails at the occurrence of heavy reflections. This could be improved by automatic adaptation of the segmentation parameters. The accuracy of the height measurement depends on the microscope magnification and is slightly better than the lateral image resolution. The measurement time (excluding image acquisition) varies from 10 to 40 µs. A sampling rate of ten measurements per second (including image acquisition and table motion) is achievable.

An additional feature of the depth recovery system is the possibility of measuring the complete profile of a scene. Figure 9 (left) shows a section of a coin which has been scanned in steps of 25 µm. A complete depth map is obtained by interpolating between the depth values. Figure 9 (right), shows a 3-D reconstruction of the scan, generated from the depth map using the visualisation tool *vtk*[12].

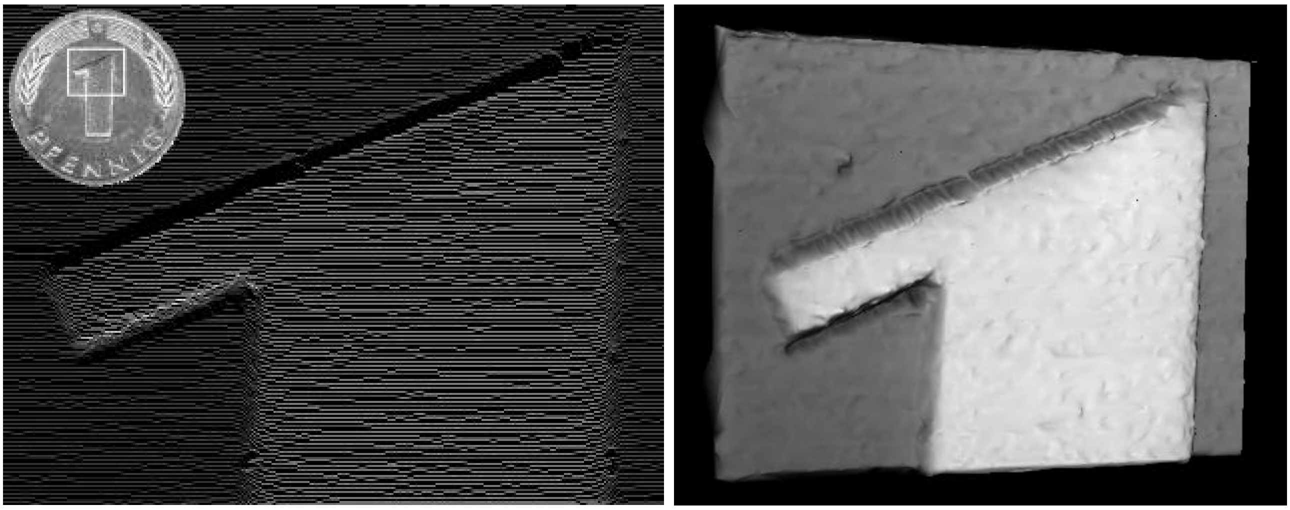


Figure 9: Profile of a coin (left), and corresponding surface model generated with *vtk* (right)

5.2 Electron beam triangulation within the SEM
Depth information is also required for applications within the SEM. One method is to install a second electron gun for lateral SEM images, but this is prohibitively expensive. The same applies for stereo SEMs. Moreover in the case of stereo, the correspondence problem which links image features in a stereo pair is difficult to perform due to the generally noisy images obtained from SEMs.

A system similar to the laser triangulation under the optical microscope can be established inside the SEM. Instead of a laser the electron beam is used to induce a line. The digitally controlled positioning of the electron beam is very fast and flexible. For electron beam triangulation, a miniaturised light microscope is mounted inside the vacuum chamber. It provides images of the luminescent points of the electron beam. This principle is explained in Figure 10. As the positions of the miniature microscope and the electron beam are known, the height of the electron beam's point can be calculated by triangulation from its image on the CCD chip[13].

In order to produce a sufficiently bright spot, the surface to be measured must be coated with cathodoluminescent material (scintillator). In order to allow for the measuring of the whole robot configuration, scintillator material was attached to the gripper tips in form of a Z-pattern. If the electron beam scans a line across this pattern, up to six bright spots are visible and are picked up by the miniature microscope (Figure 10). The co-ordinates of these spots in the microscope image are determined by simple image processing algorithms. By triangulating the heights of these points are calculated. Their distances from each other provide information about the position of the Z patterns and herewith almost the total configuration of the grippers including the gap between the gripper tips. In order to obtain the remaining degree of freedom, which is the rotation around the axis determined by the points, a second line can be scanned by the electron beam. Due to the redundancy in the measurements, the accuracy of particular parameters can be further enhanced. However, the position of the Z-patterns must be known with a certain degree of accuracy, in order to maximise the distance between the two lines.

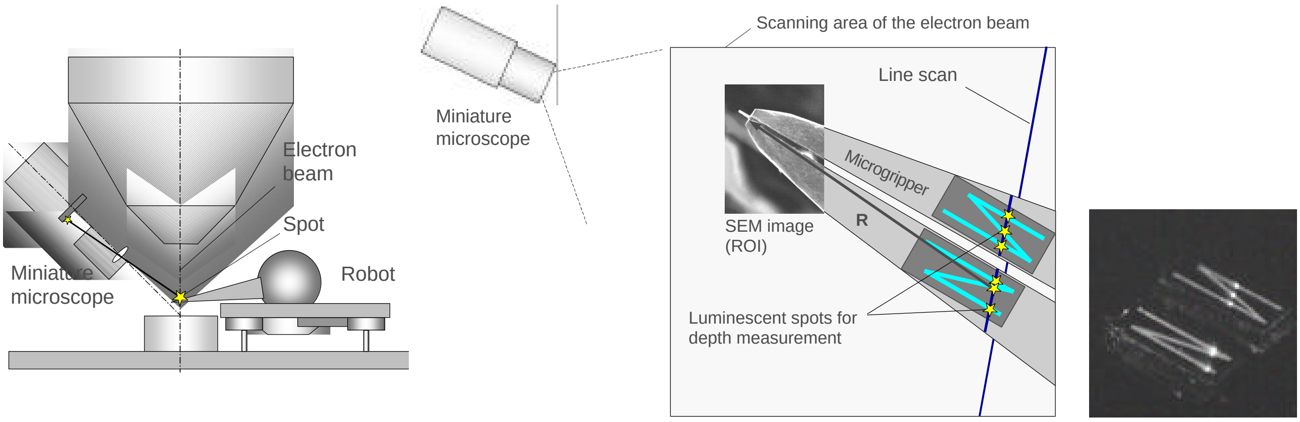


Figure 10: Sensor principle: global side view (left), top view (middle) and section of the miniature microscope image (right)

As a first experiment, silicon chips with 2 mm microstructured grooves are glued to the grippers. These grooves had been filled with P47 scintillator powder. First, the position of the grooves relative to the gripper must be measured as accurately as possible, *e.g.* using the SEM. This measurement can be avoided if the grooves are integrated in a microstructured gripper. In this case, further smaller Z-patters could be integrated closer to the gripper tips. By interchanging the zoom objective for the miniature microscope, different resolutions and working ranges can be selected.

Calibration of the system requires the determination of the SEM image and the miniature microscope parameters. The latter are the 11 parameters of Tsai's camera model[14]. They are obtained by using a small calibration grid[13]. Since the SEM image is formed as a central projection, the electron beam seems to come from the so-called virtual pivot point. To be able to determine the exact position of the electron beam, this point must be known. An object – preferably a very thin wire – is mounted in the centre of the SEM image and in the height h above the present object plane. If it is moved horizontally by the length B, the SEM image shows a shift of the length A. This length is measured in the SEM image. Using the intercept theorems, the height of the pivot point can be calculated[13].

6. THE CONTROL SYSTEM

The microrobot control system has two main task: global and local position control. Global control employs information from the global vision system, *i.e.* a CCD camera supervising the microrobots' workspace of about 250 x 200 mm² in size. Each robot is equipped with a unique configuration of four infrared LEDs that are tracked by the global camera[15]. The local sensor system is either a light optical microscope, a camera with macro objective or a scanning electron microscope.

There are several problems to be considered when designing a closed-loop position controller for a complex system such as a mobile microrobot:

1. comparably large sampling times due to the complex image processing (especially for local control)
2. the robot's behaviour and systematic disturbances are extremely complex to model
3. the number of manipulated variables is larger than the number of measurable parameters

The platform configuration of the MINIMAN robots, which stand on three legs, consists of 9 parameters (direction vectors and speed values for each leg). However, the sensor system can only record three parameters (platform x and y position and orientation), *i.e.* deviations in the behaviour of the legs are not known to the closed-loop control because the oscillation of the legs is not measured. This leads to systematic disturbances for the resulting parameters, *i.e.* position and orientation of the platform's reference point – its centre of gravity. The manipulated variable of the control loop is broken down into the necessary parameters for the three legs. Each leg has its own coordinate system, so that the required leg movements can be calculated by simple matrix and vector operations. Here, another problem arises. The relation between oscillation frequency and speed of the legs is non-linear and thus difficult to describe mathematically. A speed table, which is generated offline for each robot allows the controller to choose the right oscillation frequency for a given speed.

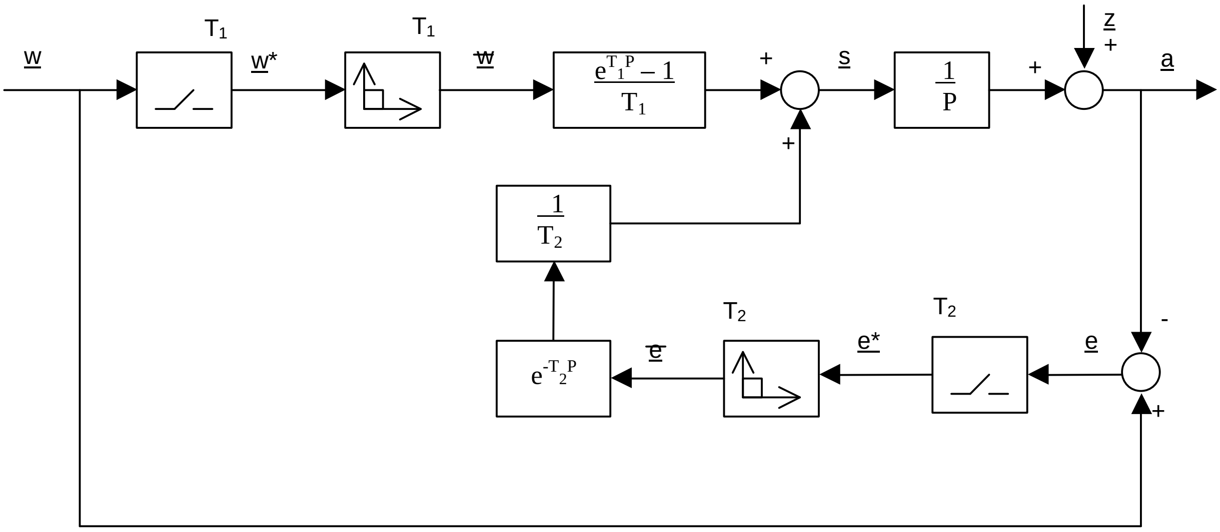


Figure 11: Structure of the implemented closed-loop controller for platform positioning

Figure 11 shows the structure of the implemented controller for platform positioning with the disturbances $\underline{z}$ and the controlled variable $\underline{w} = (w_x, w_y, w_\omega)^T$, where w_x, w_y and w_ω are the desired values of the robot's x and y position and orientation. The controller has been found to perform well in response to pulse-like disturbances. However, constant disturbances cannot be compensated for completely.

7. DISCUSSION AND CONCLUSION

In this paper a novel approach to applying computer vision to provide control for a microrobot has been described. Various issues pertaining to the requirements of the vision subsystem have been described and demonstrated. It has been found that many known vision algorithms are adaptable to the micro level if the operating environment is taken into account. However a number of outstanding issues still exist. Furthermore, the widely-used method of laser triangulation has been adapted to gain depth information from an optical microscope. Inside the SEM, depth is estimated by the novel approach of electron beam triangulation.

In section 4.2, the creation of the predefined object database holding information relating to models of the objects in the scene by the vision subsystem is a time consuming operation. This is because the creation of the PGH models involves the computationally expensive task of building the line segments specifying the object using the iterative process of edge detection and line linking. Parallelisation methods are being developed to speed up the process.

At present, the accuracy of the depth measurement inside the SEM is limited to about 30 µm, mainly by the employed miniature microscope. Using a higher resolving camera and a smaller field of view, an accuracy of a few microns will be reached. Beyond this, the depth of focus of the light microscopy becomes the limiting factor. In comparison to the possible resolution of an SEM this still seems to be low. However, it is sufficient for many tasks in the micro- and millimetre range. The measurement time (excluding image acquisition) varies from 20 to 30 µs.

With regards to the controller described in section 6, in the current implementation the maximum deviation between the desired and the actual trajectory ranges from 1 to 2 mm. Current work is concentrated on improving the control system. Concepts like dead beat response are expected to improve the controller significantly. Furthermore, shorter sampling times (T_2 in Figure 11) are aspired through speeding up the image processing, thus improving the response time on disturbances and reducing the total trajectory deviation. In addition, the speed tables, which are currently only estimated, will be generated more accurately. A path planning module is being developed and will provide trajectory planning that can be fed directly into the controller. Finally, the control structure for platform control will also be employed to control the microrobot's gripper.

ACKNOWLEDGEMENT

This research work is currently being carried out at the Institute for Process Control and Robotics, Computer Science Department, University of Karlsruhe and the Electronics Research Group, School of Engineering, Sheffield Hallam University. The research work is being supported by the European Union (ESPRIT Project "MINIMAN", Grant No. 33915). We would like to thank Jan Wedekind for his contribution.

REFERENCES

1. E. P. Krotkov: "Active computer vision by cooperative focus and stereo", Springer, 1989.
2. I. P. Pappas: "Adaptive Visual Control for Microrobots", dissertation, ETH Zürich, 1998.
3. A.C. Evans, N.A. Thacker and J.E.W. Mayhew, "The use of geometric histograms for model based object recognition", *Proceedings of the 4th British Machine Vision Conference (BMVC)*, 1993.
4. J.Canny, "A computational approach to edge detection", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp.679-698,1986.
5. D.H.Ballard, "Strip trees: a hierarchical representation for curves", *Communications of the ACM*, vol. 24, no. 5, pp.310-321, 1981.
6. F.Aherne, N.A.Thacker and P.I.Rockett, "The Bhattacharyya metric as an absolute similarity measure for frequency coded data", *Kybernetika*, vol. 32, no. 4, pp. 1-7, 1997.
7. D.J. Williams and M. Shah, "A fast algorithm for active contours", *CVGIP: Image Understanding*, **55**, 1992.
8. A. Fusiello, V.Roberto and E.Trucco, "Efficient stereo with multi-windowing", *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, Puerto Rico, 1997.
9. A. Sulzmann, P. Boillat, and J. Jacot: "New developments in 3D Computer Vision for microassembly", *Proc. of SPIE Int. Symposium on Intelligent Systems & Advanced Manufacturing*, Vol. 3519, 1998.
10. S. Fatikow, A. Buerkle, and J. Seyfried: "Automatic Control System of a Microrobot-Based Microassembly Station Using Computer Vision", *SPIE's International Symposium on Intelligent Systems & Advanced Manufacturing, Conference on Microrobotics and Microassembly*, Boston, Massachusetts, 1999.
11. A. Buerkle and S. Fatikow: "Laser measuring system for a flexible microrobot-based micromanipulation station", *IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, Takamatsu, Japan, 2000.
12. vtk, The Visualization Toolkit, http://www.kitware.com/vtk.html
13. F. Schmoeckel, H. Wörn, M. Kiefer: "The Scanning Electron Microscope as Sensor System for Mobile Micro-robots", *8th IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA 2001)*, 2001.
14. R.Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and Automation*, No 4, 1987.
15. A. Buerkle and S. Fatikow: "Computer Vision Based Control System of a Piezoelectric Microrobot", *Proc. of the Int. Conf. on Computational Intelligence for Modelling, Control and Automation (CIMCA'99)*, Vienna, 1999.