Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

IFAC

Automated Nano-Assembly in the SEM I: Challenges in setting up a warehouse

T. Wich[*, **], C. Stolle[*], O. Frick[*], S. Fatikow[*, **]

[*)]: OFFIS e.V., Division MN, Oldenburg, Germany (thomas.wich@informatik.uni-oldenburg.de)
[**]: Division Microrobotics and Control Eng. (AMiR), Univ. of Oldenburg, Oldenburg, Germany

Abstract: This paper describes the implementation of setting up a warehouse for automated nano-assembly, i.e. the automated registration of the parts processed during the automated assembly. Starting with a brief description of the goals of the assembly process, the overall process structure and its challenges with respect to the micro- and nanoscale are explained. The warehouse task is used as an example for describing the task planning, the necessity to minimize the number of subtasks taking further constraints from the experimental setup into account. For controlling such a complex assembly process, an intelligent and flexible control and communication system architecture is necessary, which will be explained in detail. Based on these considerations, the implementation of the warehouse task and its challenges, including the representation of the parts in the control system, object recognition using generalized models and methods for part origin determination will be presented.

1. INTRODUCTION

In contrast to assembly processes on the macro-scale, i.e. with critical part size above the mm-range, manual assembly on the micro- and nano-scale is hardly feasible due to the critical part sizes in the µm respectively sub-µm or nm-scale. A growing need for assembly on the micro- and nano-scale within the last ten years (Clevy et al., 2006), demands for suitable tools to servo (e.g. through teleoperated devices), semi-automate or to completely automate assembly processes.

Assembly processes on the nano-scale can be classified in two categories. The parallel or bottom-up approach utilizes force fields to handle multiple parts at a time (Böhringer et al., 1999), where the advantages of batch processing used in CMOS or more generally in semiconductor processing technology are maintained. For example, electric fields are used to position and orientate parts on a silicon wafer. In contrast, the serial or top-down-approach transforms the conventional assembly principles known from the macro-scale to micro- and nanoscale. Certainly, as the critical part size decreases, more and more attention has to be paid to surface effects and parasitic forces (Fearing, 1995). Assembly on the micro- and nano-scale following the serial approach usually requires a vision sensor, on the one hand to allow the operator to observe the process, On the other hand they are an essential sensor for gathering information about spatial arrangement of the parts (Sievers and Fatikow, 2006) and tools. They can be utilized to infer physical properties of the observed objects, like forces (Wich et al., 2006). Several microscopy technologies can be used as vision sensors in nano-assembly processes, *e.g.* light microscopes, scanning electron microscopes (SEM) and scanning probe microscopes (SPM). A more detailed overview over their respective pros and cons is given in (Wich and Hülsen, 2008). Within this paper, a SEM is used as vision sensor, providing adequate resolution, fast image acquisition and the electron beam can be used as a tool for single tasks in the process (Wich et al., 2006). The purpose of this paper is to present a general introduction to assembly processes in the SEM. Based on the constraints imposed by the geometric scale, specialized methods for automation will be introduced by means of an example process in the next paragraph. In the following section, a detailed planning for the warehouse process will be given, which is used for registering the single parts later used for assembly. In section two, the control and communication system necessary for triggering and evaluating the single tasks will be described. An implementation of the system will be shown in section three, giving a detailed view on automation challenges. In the last section, an outlook towards the next steps in assembly automation on the micro- and nano-scale will be given.

1.1 Description of the Automated Process

Assembly processes are generally composed of different tasks as there are:

- Separation of the parts which should be assembled
- Parts (and tools) handling
- Joining of parts by means of material, force or form closure
- Releasing or detaching (e.g. cutting, breaking etc.)
- Inspection of the (sub-) assembly, e.g. quality assurance

These tasks can be divided into subtasks or primitives (Wich and Hülsen, 2008), which are the atomic operations during the process.

As an example nano automation process we have choosen bonding of carbon nanotubes (CNT) to the point of a scanning tunnelling microscope (STM) tip. The process flowchart is given in Fig. 1.

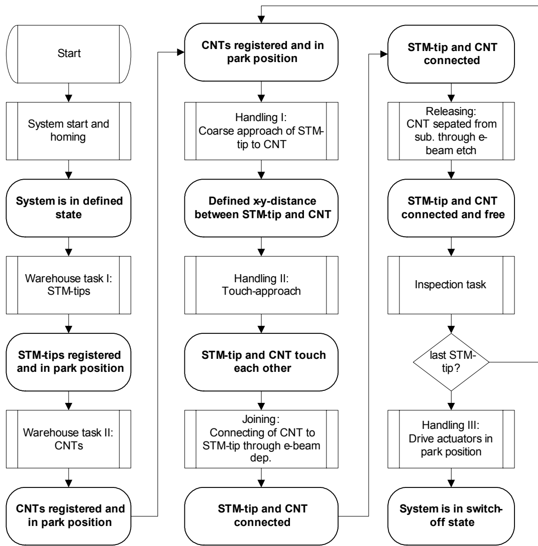

Fig. 1: Process-flowchart for automated bonding of CNTs to STM-tips. The symbols are used on basis of the "Specification and Description language" (SDL)

The CNTs used for this assembly process have a predefined length of approx. 10 µm and a diameter of 200 nm. They are grown by CVD in a predefined matrix on a quadratic silicon wafer, and thus are well suited for automated assembly tasks (Wich and Hülsen, 2008). The STM-tips are manufactured by etching tungsten wires (diameter 0.2 mm, cut to a length of about 10 mm) in NaOH. The tips are then put in a holder with a 2.54 mm raster (cp. Fig 2a). In the whole process no further tools (e.g. grippers) are used. Instead, the parts are directly assembled in order to maximize the reliability by reducing the number of subtasks (Wich and Fatikow, 2007). Therefore, both the wafer with CNTs and the STM-tip holder are mounted on a three-axis linear stage each, opposite to each other.
The core process consists of six tasks. However, before the process can start, the so-called "warehouse" task has to be accomplished, which will be explained in detail in the following sections. The core tasks comprehend the positioning of the parts relative to each other (Handling I), bringing the parts in contact using specialized methods (Handling II), bonding of the CNT to the STM-tip (Joining), cutting the CNT from its substrate (Releasing), checking the stability of the bond connection between CNT and STM-tip (Inspection) and finally moving the assembled part to a predefined position (Handling III).

1.3 Motivation, Goals and Challenges for Setting up the Warehouse

The warehouse-task, used for registration of the single parts, has until now not been discussed in literature for nano-assembly automation. Its relevance in the process chain emerges especially with regard to the micro- and nano-scale, which is based on a couple of reasons:

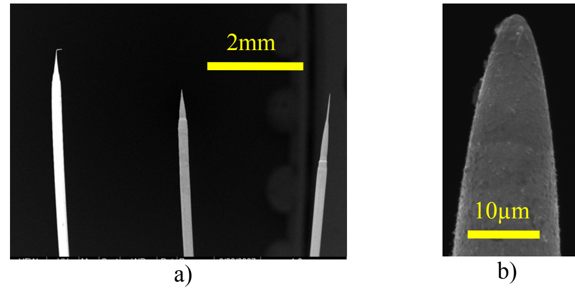

Fig. 2: a) Tungsten made STM-tips at low magnification in their holder on a linear actuator. b) Single STM-tip observed at a higher magnification. The tips tip is referred to as the tool center point (TCP).

- Small position tolerances in relation to the part size on the macro-scale lead to huge tolerances on the nano-scale. Consequently, the exact position of every part has to be measured and recorded individually.
- Parts, which exhibit a high degree of similarities on the macro-scale, can exhibit a high degree of individuality on the nano-scale. Thus, for finding and tracking STM-tips reliably in high magnifications, the visual object recognition must either know the individual shapes of the tips or it must be able to recognize tips by their common shape properties.
- The differentiation between wasted and usable parts is often possible only at high magnifications, where details of part shape and texture are visible and analyzable.
- The parts tool center point (TCP), i.e. the point on a part which is used as the centre of reference for assembly, can for nano-assembly only be defined in a high-magnification observation, because of the limited accuracy in the image recognition (Sievers and Fatikow, 2006, Wich and Hülsen, 2008, Wich et al., 2006).

These arguments reveal that for nano-assembly purposes it is necessary to provide very flexible but robust tasks and subtasks, i.e. precise actuators and sensors on the hardware side as well as a robust and flexible control and image recognition infrastructure.
Until now it has been assumed that the registration and cataloguing is performed in a separate, the actual assembly process preceding, task. This "warehouse" approach has several advantages: Based on the recorded information (*e.g.* position, images in several magnification levels etc.), distinct rules for collision prevention can be derived. Furthermore, the information about the geometric properties of the parts can be used for assembly of selected parts which individual properties fit to one another (*e.g.* "compensatory tolerance assembly"). Alternatively, it is also possible to split the warehouse-task and integrate it separately into the first handling task (i.e. task "Handling I" in Fig. 1) for every

assembly cycle. This “inline”-approach also provides several advantages: Firstly, the amount of data provided for the whole parts database can be reduced tremendously, i.e. one entity per STM-tip respectively CNT can be sufficient. Secondly, the inline-recorded dataset is very fresh compared to datasets recorded in a warehouse task. Thus, negative effects influencing the reliability of the datasets like thermal drifts in the actuators, charging effects etc. affect the assembly process less. For both the “warehouse” and the “inline” approach, the subtask-count is approximately equal, with a slight tendency to fewer tasks for the “inline” approach, depending on the individual organization of the subtasks.

In order to implement either an “inline” or “warehouse” task, a dedicated infrastructure is necessary. Besides the requirements like position or measurement accuracy etc. for the hardware side, the control systems for the actuators have to be flexible with regard to their feedback sensor (e.g. onboard position sensor for low magnifications, image recognition for high magnifications (Wich and Hülsen, 2008). Thus, the data of the various sensors have to be provided to multiple control loops, which requires a communication system with sufficiently low latency and high data-throughput. Compared to macro-scale assembly processes, the number of subtasks is higher on the micro- and nano-scale. In order to allow for flexible and robust implementation of the process plan, a specialized high level control system is necessary. This system will be described in the next section.

2. CONTROL SYSTEM

2.1 Control System Architecture

The architecture for the automated bonding of CNTs is the distributed control architecture for automated nano-handling (DCAAN) (Stolle, 2007). For our purpose the system has been adapted to our hardware requirements.

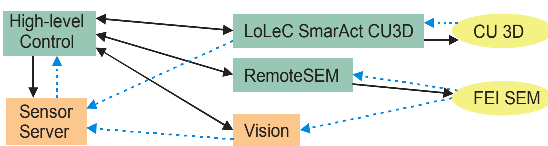

Fig. 3: DCAAN components involved in the automation process with control connections (solid black arrows) and sensor data connections (dotted blue arrows).

The system components (Figure 3) involved in the warehouse task for STM tips are:

- LoLeC SmarAct CU3D – low-level controller (LoLeC) for a cartesian (x,y,z) actuator group. It receives positioning commands from high-level control and executes them closed- or open-loop.
- Vision – image acquisition software that extracts position data from camera / SEM images and sends the position data to SensorServer
- SensorServer – receives sensor data at different update rates from sensor programs. LoLeCs and high-level control can request these measurements.
- RemoteSEM – SEM interface for setting and getting SEM specific parameters like scan field area, brightness, beam shift, etc.
- High-level Control (HiLeC) – processes user input and automation sequences by sending control messages to all other components in the network.

The architecture uses Common Object Request Broker (CORBA) of the Object Management Group (OMG) as network middleware for communication. So every component is a C++ based server. All DCAAN components where set up on one dual core PC for the experiments. However it is possible to run almost all parts on different PCs for better scalability. RemoteSEM is the only exception because it requires a direct link to the SEM.

2.2 Process automation in HiLeC

HiLeC ist the automation control system. For this pupose it is able to process user input (e.g. tele-operation) as well as processing sequences of control commands. The automation of CNT bonding has very high demands on the automation language. Therefore the automation sequence itself is written in C++ as part of HiLeC.

Every DCAAN component specifies its own process primitives as a list of commands, parameters and return types. This specification is read by HiLeC at connection time. The process primitives are called by HiLeC inside the command execution scheduler asynchronously. The components process these calls and return the result and execution status (e.g. successful or failed). In an automation sequence process primitives can be used as synchronous or asynchronous commands. The commands are triggered by inserting command strings into the scheduler, which hides the pure asynchronous network communication.

Tasks or subtasks are defined in HiLeC as C++ functions combining several process primitives (Fig. 4). While process primitives are single DCAAN component commands, tasks can only be applied if the system is in a appropriate state (*i.e.* all preconditions are met). They should provide sufficient error checking such that after execution certain postconditions do hold.

```
sem.GetMagnification(old_mag);
sem.SetMagnification(mag);
vision.EnableModel(model,0,0,0);
stm.StartClosedLoop(x,y,z);
vision.DisableModel(model);
sem.SetMagnification(old_mag);
```

Fig. 4: Example task without error checking code that moves the actuator using visual feedback. As input magnification (mag), the tracking model and the coordinates (x,y,z) are given.

2.3 Entities for parts

In automation sequences like CNT bonding lots of information need to be stored (e.g. tool center point (TCP) positions, vision models, etc.) for later usage. This information is stored directly in HiLeC. The only exception is the model data for the vision tracking system, which is too much data such that it is not acceptable to transmit it through the network on every command. Therefore Vision keeps a local model database. In a future step the data might get stored in a database management system for better accessibility and data consistency. In an offline warehouse process most of the data is collected during the initialization phase ahead of or as first step of the automation sequence. Online algorithms in contrast acquire the data on demand and most often do not store them for later usage. During the offline warehouse task the following data entities are acquired:

- Magnification level
- Tip x,y position using axis internal sensors for every magnification step
- A new tracking model for every magnification step and tip except for the lowest magnification
- The position of the upper most part of the tip in the model (TCP)
- Area of each pixel in meters

This information is stored to be able to address every single tip at every magnification step. Similar data will have to be acquired for the warehouse step II for CNTs.

3. IMPLEMENTATION OF THE WAREHOUSE TASK

3.1 Process plan for the warehouse task

Fig. 5 shows the flowchart for the "Warehouse task I", which is used for finding and registration of the STM-tips. The first important step in this task is the recognition and thus localization of the single STM-tips position ("Localization-step" in fig.5). For recognition of the STM-tips a Region-of-Interest (ROI) is defined in the SEM image through which all the STM-tips are moved with constant velocity. The width of the ROI is chosen so that only one tip at a time is visible to the object recognition. For recognition and tracking of the STM-tips, a generic tip-model is used (cp. Section 4.3). These results are stored in datasets providing the estimated tip positions in actuator coordinates.
In the successive "STM-tip"-loop (Fig. 5), every single tip's characteristics, e.g. position, tip shape and orientation is recorded. Therefore, a second loop is used for repeated magnification and positioning of the tip in the centre of the SEM image. This is referred to as the Zoom-and-Center-step (ZAC-step) (Wich and Hülsen, 2008), which is repeated m times for every tip. In every magnification step, an image of the tip is taken, which is stored in the database as a model for this individual tip at the respective magnification. For the next higher magnification ($m+1$), the m-th image is scaled and stored as initial model.
The warehouse task generates a complete set of the STM-tips characteristics, *i.e.* their positions on the STM-tip holder given actuator coordinates and tracking models for different levels of SEM magnification.

3.2 Object recognition using generalized models

The Vision program uses Normalized Cross-Correlation (NCC) pattern-matching for detecting and tracking objects in the SEM video frames. NCC uses model images to search for object occurrences in an image. Each occurrence is assigned a score (a real value between 0% and 100%), which indicates the degree of similarity between the model and the occurrence. A score of 100% is given if occurrence and model are identical.
Several factors can have a negative influence on the reliability of the NCC-based object recognition: noise in the image, overlapping objects, deformations of the object shapes caused by non-linear geometry of the image acquisition, motion-blur if objects move too fast, vibrations of the tips induced by high-frequent motions of the slip-stick drives and by accelerations. All these factors decrease the score so that the recognition can miss objects it should detect. On the other hand, objects that are no model occurrences but yield a high score can lead to false detections. To prevent wrong detection, optimal score threshold values have to be chosen. Other counter measures include the optimization of SEM scanning parameters like increasing the beam current to reduce noise (see below) and optimizing brightness and contrast.
The shapes of STM-tips are similar enough to use one generic model image for the detection of different tips. The generic model is a high-quality snapshot image of the upper end of an STM-tip that is considered representative and has an undistorted, clean shape. The generic tip model should yield high scores for a great variety of tip shapes. To improve scores for tips that differ greatly from the model, the model and video frames can be blurred before the NCC recognition is performed. Figure 6 shows the measured scores for six tips moved through the image recognition area (ROI) with constant velocity. The model image was taken as a snapshot of STM-tip 5. The diagram shows the scores of all six tips for different degrees of blurring with a Gaussian filter. The jagging of the curves at the rising and falling slopes is caused by distortions of the SEM scanning at the left and right border of the ROI. It is not surprising to see an almost perfect score for model 5 if model and tip are sharp, i.e. not blurred, since both are identical. The scores for tips 1 to 4 are relatively low, especially for tip 3. The situation changes if blurring is applied to the model and/or the video frames. The score for tip 5 decreases slightly because tip and model are no longer identical (the image is blurred with a Gaussian with standard deviation 1, the model with s.d. 4). On the other hand, the scores for tips 1 to 4 are now considerably higher. One reason for this result is that blurring suppresses noise, which, as noted above, is one cause for a bad score. Blurring also makes objects more similar to each other because it eliminates image details. STM-tip images differ mostly in details, so their blurred versions are more similar and therefore yield a higher correlation score. It should be noted that this effect is not always desired in object recognition.

Making objects more similar in the image pre-processing increases the risk of false detections. A balance between improved correct detection and the risk of false detection has to be found. Since in the localization-step STM-tips are the only objects visible, false detections cannot occur. Blurring is therefore risk-free in this task.

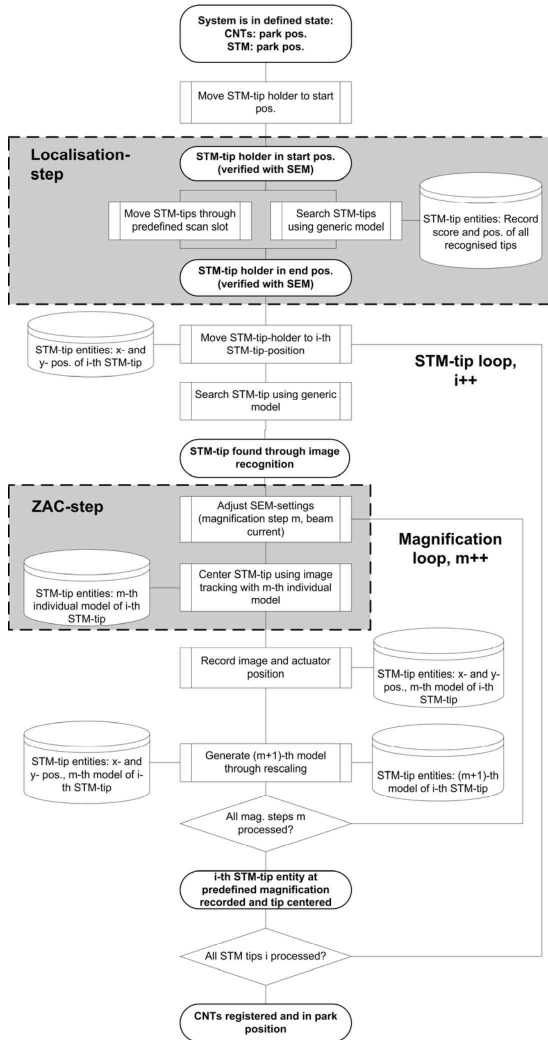

Fig. 5: Process-flowchart for of the warehouse task I, i.e. automated finding and registration of the STM-tips

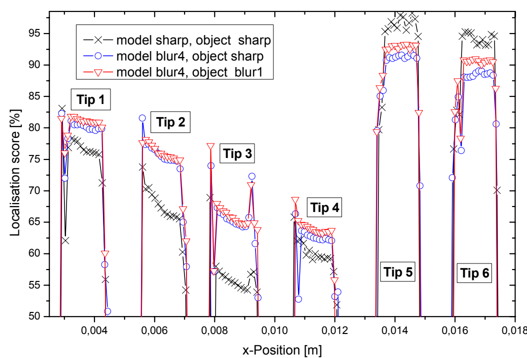

Fig. 6: Measured scores during the tip localization. The single curves are recorded by moving all STM-tips trough the ROI.

In Fig. 7 the average over the measured scores is plotted against the different degrees of blurring. This figure shows that STM-tips can be detected during the localization with NCC pattern recognition and optimized settings for hard- and software. With blurring applied, a score threshold of 60 should allow a robust detection of tips during the localization step. Later on in the assembly process, when every single STM-tip is tracked using its original model recorded during the warehouse task, a threshold score of 80% is applied.

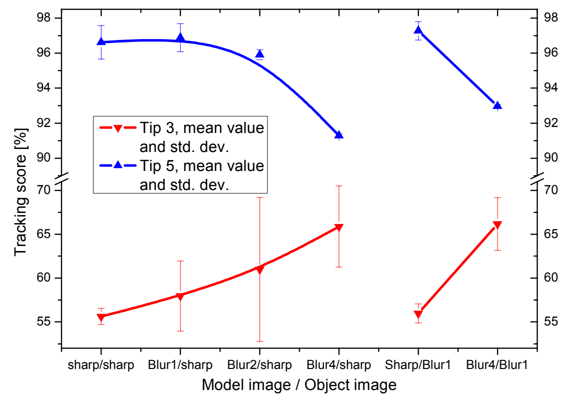

Fig. 7: Comparison of the tracking scores for tip 3 and 5.

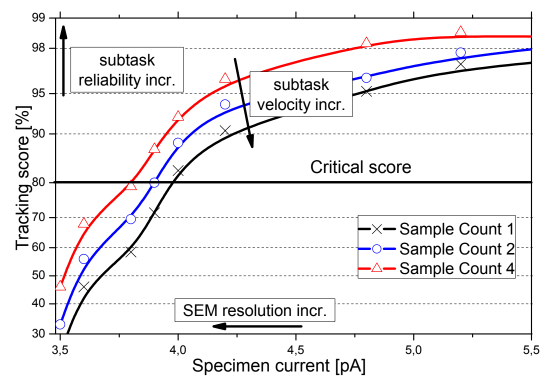

Fig. 8: The object tracking score plotted against the probe current with the sample count, which is proportional to the scan time, as a parameter. The black arrows indicate the trend for relevant process parameters, i.e. SEM resolution, subtask reliability and velocity.

3.3 The relation between SEM-parameters and object recognition and tracking

During many process subtasks, closed loop positioning based on sensor feedback data derived from image recognition and tracking is used, utilizing the SEM as vision sensor (Jähnisch et al., 2005, Sievers et al., 2006). The noise in SEM-images depends on the scan time and the beam current (Reimer, 1998). The duration the beam stays on the position that is mapped to an image pixel is controlled by the sample-count parameter. However, also the resolution of the SEM depends on the beam current, i.e. the resolution increases with decreasing beam current. Image recognition and tracking algorithms depend on a certain image quality for reliable tracking, although special algorithms have been developed for SEM-based recognition and tracking (Sievers and Fatikow, 2006) . The general relationships between SEM-

beam current and its resolution, object tracking score and subtask reliability, sample count and subtask velocity are shown in Fig. 8. Due to the discrepancy between the above requirements, a specialised approach for the warehouse tasks has been chosen. In general, the SEM's beam current is set to meet the least necessary resolution in the whole process, e.g. a probe current of 4 pA. For low sample counts, i.e. fast scan rates, the object tracking score is close to the critical limit of 80%. Thus, the sample count has to be increased, which results in a decreased process velocity, due to the decreased update rate of the "vision" sensor system. In the single ZAC-steps, the specimen current is switched between a high value, e.g. 5.2 pA, for low magnification, and 4.0 pA for higher magnification. For registering the STM-tips' TCP, the sample count is additionally increased to maximize the tracking score.

4. CONCLUSION

For automation of assembly tasks on the nano-scale, the process of bonding CNTs to STM-tips has been chosen as an example. The process flowchart shows the single tasks used for accomplishing this process in Fig. 1. The core tasks deal with the assembly process itself, however, the need for a "warehouse" process for automated localization and registration of the parts needed for assembly has been derived. Although the warehouse tasks consist of many subtasks and can therefore be a source for process failures, it has been shown that this task is essential for successful process automation on the micro- and nano-scale.
However, these warehouse tasks also offer chances for improving nano-assembly, due to differentiation between wasted and usable parts and exact determination of the tool center point.

In the second section, the control system architecture for realising assembly tasks on the micro- and nano-scale has been presented. The main components are the high-level control unit (HiLeC), a sensor server for distributing sensor data to the high level control. low-level control units (LoLeC) and the image recognition and object tracking software transmit measured and computed data to the sensor server. The HiLeC is used for the actual implementation of the automation routines and for storing the parts entities.
Emphasis has been laid on the "Localization step" for finding the STM-tips using image recognition with generalized models. It has been shown how generalized models can be created by blurring sharp models. The results of the localization step using sharp and blurred models respectively object images have been presented for a setup consisting of six STM-tips with very different qualities. The measurements revealed that localization using blurred models and blurred object images is very promising, as the relevant recognition score for parts with a high degree of individuality can be improved significantly. Moreover, the influence of the SEM's specimen current and scan time on the object tracking score has been described. Based on these measurements, the consequences on process automation have been described. As a conclusion it can be said, that the warehouse task is of significant importance for successful process automation on the micro- and nano-scale. We would like to thank the German Ministry for Education and Research (BMBF) for funding this project under the number 16 SV 2276.

REFERENCES

Böhringer, K. F., Fearing, R. S., and Goldberg, K. Y. (1999). Microassembly. In *Handbook of industrial robotics*, pages 1045–1066. Shimon Y. Nof (ed.), 2nd edition.

Clevy, C., Hubert, A., and Chaillet N. (2006). Micromanipulation and Micro-Assembly Systems. *The IARP - IEEE/RAS - EURON Joint Workshop on MICRO & NANO ROBOTICS (av. online at: http://iarp06.robot.jussieu.fr/)*.

Fearing, R. (1995). Survey of sticking effects for micro parts handling. *IROS*, 02:2212.

Jähnisch, M., Hülsen, H., Sievers, T., and Fatikow, S. (2005). Control system of a nanohandling cell within a scanning electron microscope. In *Proc. Int. Symposium on Intelligent Control (ISIC'05) / Mediterranean Conference on Control and Automation (MED'05)*, pages 964–969, Limassol, Cyprus.

Reimer, L. (1998). *Scanning Electron Microscopy – Physics of Image Formation and Microanalysis*, volume 45 of *Springer Series in Optical Sciences*. 2nd edition.

Sievers, T. and Fatikow, S. (2006). Real-time object tracking for the robot-based nanohandling in a scanning electron microscope. *Journal of Micromechatronics - Special Issue on Micro/Nanohandling*, 3(3-4):267–284(18).

Sievers, T., Jähnisch, M., Schrader, C., and Fatikow, S. (2006). Vision feedback in an automatic nanohandling station inside an sem. In *Proc. Int. Optomechatronics Conference on Visual/Optical Based Assembly and Packaging (SPIE Optics East)*, volume 6376, page 63760B, Boston, MA, U.S.A.

Stolle, C. (2007). Distributed control architecture for automated nanohandling. *Proceedings of the 4th Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO)*, pages 127–132.

Wich, T. and Fatikow, S. (2007). Assembly in the sem. Robotics Science and Systems Conference - Published online http://www.me.cmu.edu/faculty1/sitti/RSS06/RSSWorkshop.htm.

Wich, T. and Hülsen, H. (2008). *Automated Nanohandling by Microrobots*, Chapter in *Robot-based Automated Nanohandling*, Fatikow, S. (ed.), Springer Verlag.

Wich, T., Sievers, T., and Fatikow, S. (2006). Assembly inside a scanning electron microscope using electron beam induced deposition. In *Proc. Int. Conf. on Intelligent Robots and Systems (IROS'06)*, pages 294–299, Beijing, China.