Elsevier Editorial System(tm) for CIRP Annals - Manufacturing Technology
Manuscript Draft

Manuscript Number: 2010-A-16R3

Title: Integrated visual nanometric three dimensional positioning and inspection in the automated assembly of AFM probe arrays

Article Type: 2010, Volume 1

Keywords: Assembly; Visual Inspection; Localization.

Corresponding Author: Prof. Michele Lanzetta (2), PhD, MEng

Corresponding Author's Institution: University of Pisa

First Author: Michele Lanzetta (2), PhD, MEng

Order of Authors: Michele Lanzetta (2), PhD, MEng; Martin L Culpepper, Associate Professor

Manuscript Region of Origin: USA

UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA MECCANICA, NUCLEARE E DELLA PRODUZIONE

SEZIONE PRODUZIONE

Cirp Secretariat
Address: 9 rue Mayran, 75009 Paris - France
Telephone: +33 1 45 26 21 80
Fax: +33 1 45 26 92 15
E-Mail: cirp@cirp.net

Pisa, February 26th 2010

Subject: GA 2010. Revised article submission for STC A

Dear Sirs,

I am pleased to submit the enclosed article, title:

Integrated visual nanometric three dimensional positioning and inspection in the automated assembly of AFM probe arrays

written with prof. Martin L. Culpepper from the Laboratory of Manufacturing and Productivity of the Massachusetts Institute of Technology.

Sincerely yours,

Michele Lanzetta

Via Diotisalvi 2 – 56122 Pisa – Italia
tel. (+39) 050 2218.000 fax (+39) 050 2218.065 – amministrazione tel. (+39) 050 2218.062 fax (+39) 050 2218.069
Cod. Fisc. 8000367 050 4 – Partita IVA 0028682 050 1
Sezione Produzione: Via Bonanno Pisano 25B – 56126 Pisa – tel. (+39) 050 2218.141 fax (+39) 050 2218.140

Answers to Reviewers' comments:

> Some parts are unclear and have to be improved from this point of view. For example,
a description of an AFM probe should be included.
There is a great confusion between figures 1 and 2 cited by the text.

The overall text has been reviewed for clarity, particularly regarding the references to figures 1 and 2.
A SEM micrograph of an AFM probe has been inserted in Figure 1.

> What the reason of finding the position of the gripper, if the real interesting
measurement is the position of the probe ?

True, we are not interested to know the position of the gripper, because there would always be a grasping error. Instead we need to calibrate the relative camera-gripper position in order to be able to correct THAT grasping error.
This concept has been stressed in the revised version of section 4.

> What happens in figure 3 if the number of defective tips is between 2 and 7 ? Does
it loop ?

As indicated in the main text:
"this inspection scheme (...) requires taking another sample of 5 rows if at least 2 tips are found missing or defective and so on according to [8]"

Reference 8 is
ISO 2859-1:1999, Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.

So, yes, it does loop.

For better clarity, in the modified version of the above sentence, "and so on according to 8" has been replaced by "and loop according to 8", thank you.

Multiple sampling schemes, by definition, involve the three cases summarized in Figure 3 (accept - reject - take another sample). The whole scheme is very difficult to be included in a single figure, for instance the number of samples changes at each iteration. Reference 8 contains all the details for the interested reader.
We have specified in the modifed version of Figure 3 the meaning of the feedback arrow, by adding: "take another sample".

> The aim of the paper has to be given at the end of the introductive paragraph
instead of in the middle of section 2.

The following sentence has been added at the end of the introductive paragraph, after "it seems that the integration [of localization and inspection] has not been proposed, particularly in the micro scale":
"This is the main task of current paper focusing on the automated assembly of silicon chips, as described in the next paragraph."

> Figure 1 is not very clear and explicit. The authors are invited to try to clarify it. The same remark can be made for Figure 2 and Figure 3, mainly about the signification of all the arrows.

The clarity of Figure 1 has been improved by additional captions and by adding a SEM micrograph of an AFM chip.
The meaning of looping arrows in Figure 2 has been clarified in the main text. The colour of some arrows and the caption of Figure 2 has also been changed.
Figure 3 has been dealt with as already explained.

Note about figures: figures are duplicated because they have been provided to the editor both as images and in raw format.

Integrated visual nanometric three dimensional positioning and inspection in the automated assembly of AFM probe arrays

Michele Lanzetta[1] (2), Martin L. Culpepper[2]

[1]Department of Mechanical, Nuclear and Production Engineering, University of Pisa, Italy

[2]Laboratory of Manufacturing and Productivity, Massachusetts Institute of Technology, USA

Abstract

This paper presents the design of a monocular three dimensional artificial vision system attached to a 20× microscope lens for precision and microsystems applications. Possible uses in assembly include: positioner calibration, sensor-based part handling, positioning, and inspection in the nanometric range. The developed image acquisition method – along one direction (in steps of 100 nm), the depth-from-focus algorithm and subpixel interpolation (of 5 acquisitions for concurrent localization and inspection), allow to overcome the physical optics limitation achieving a resolution under 200 nm. The vision strategy and algorithms, described in the paper, have been validated by handling an AFM probe array by a micropositioner.

Keywords: Assembly, Visual Inspection, Localization

1. Introduction

After three decades from the introduction of artificial vision in the automotive and electronics industry, the next challenge is three dimensional vision. Sensors and particularly vision in production have been reviewed in [1], [2] and [3]. Despite the growing interest of microsystems and their assembly, not many vision applications are available in the literature, particularly regarding 3D vision [2].

In any given application, the main tasks of a vision system (both in two and three dimensions) can be summarized as recognition [4] (including OCR), inspection (also measurement/gauging [3, 5]) and localization (tracking if in motion [6]). Assembly can be considered a combination of all three: for instance an artificial vision system inspects mating [4, 5] by identifying and by checking the relative position (and inclination) of the parts involved. However it seems that the integration of all the three tasks, in the sense of using the same hardware and images, has not been proposed, particularly in the micro scale. This is the main task of current paper focusing on the automated assembly of silicon chips, as described in the next paragraph.

2. Assembly automation

Although Atomic Force Microscope (AFM) probes are of widespread and increasing use in laboratory equipment, for direct writing, templating of proteins, viruses, and carbon nanotubes and more applications, their assembly is still carried out manually.

This actual industrial problem concerns automating the assembly of AFM probe arrays to an interface (Figure 1) for fast exchange and secure attachment to the scan head [7].

Table 1: AFM probe array and optical system specifications.

Figure 1: Grasping and handling of the AFM chip (SEM inset) under a camera before gluing.

The examined AFM probe arrays are essentially chips derived from a diced silicon wafer with cantilevers and pyramid tips with the features listed in Table 1. In the present process, each chip is glued to a wedge with magnets for mounting. However the leveling and alignment to the scan head after mounting each chip with its wedge is a tedious 30 minutes manual operation.

In this paper a new assembly strategy is proposed replacing the wedge with a specifically designed interface to refer each chip with the required accuracy without manual leveling and alignment.

Figure 1 shows the spatial configuration of the developed system: 1 row of the chip with 2 cantilevers and pyramid tips out of 110 are displayed (55.000 in a chip); 2 of the 16 tips of a row that can be actually captured by the camera are projected on the sensor. The automated assembly of the chip to the new interface is carried out by a gripper attached to a positioner and is assisted by the vision system in the following tasks:

- measure the chip grasping errors, expressed by the position of the O_G reference system with respect to the O_P reference system,
- provide the necessary correction acting on the O_P reference system and
- check positioning (in closed loop) before assembly.

Figure 2 shows the looping interaction between the vision system and the positioner in the steps outlined above. The integration between localization and inspection is also shown by the clear arrows path.

According to the proposed assembly strategy, photocurable glue is pre-deposited on the interface and radiation is activated after the chip reaches the correct alignment and leveling.

Figure 2: Assembly phases for the AFM probe array, showing material in/out (red arrows) and the vision and positioner (cyclic) interactions.

In the preliminary system design, the positioning error (Table 1) has been equally shared between the mechanical tolerances (gluing, glue aging, interface design, mating etc.) and the optical performance (3D positioning error). The first aspects would benefit from the maximization of the second, which is dealt with in this paper. However, it is well known that the resolution of an optical system is upper limited by the wavelength of light. With industrial matrix cameras, which operate with visible light, despite the sensor resolution, the minimum optical resolution is in the order of 200 nm. For this reason, no industrial application of vision systems are available at this resolution and under, although they could have tremendous interest in the growing micro- and nanosystems area.

Because of the involvement of artificial vision, it has also been requested to provide a certification of the number of defective tips on the chip. The second challenge faced in this work has been the integration of 3D visual positioning and inspection in order to minimize the assembly time. This has been carried out by minimizing the number of image acquisitions and the chip handling as described.

3. Vision system design

The main steps followed to determine the hardware configuration are summarized in this paragraph.

Vision has been selected for its flexibility compared to other position sensors. It is able to provide the 3D position and shape of the part for localization and inspection at the same time. In this specific configuration, vision is able to assess (Figure 2): the grasping errors, the product defect and the assembly errors. Probably no other single sensor of comparable cost is able to provide the same amount of information in the same amount of time.

The design of a new vision system starts from the definition of the environment and of the target. Because of the limited accessibility to the small assembly stage and the large lens required to achieve sufficient magnification (Table 1), current configuration involves a monocular system.

Another crucial aspect in microscopy is direction and intensity of lighting. Confocal lighting has been selected, so that camera and light have the same direction, and no accurate manual alignment is necessary.

As for the intensity, the basic idea is to take advantage of the flat reflecting surfaces of the square pyramid tip faces. Preliminary tests have shown that within a wide range of inclination of camera and lighting (between 55° and 70°) with respect to the chip in order to be perpendicular to the tip face (Figure 1), tips have a high contrast ($90 < \text{graylevel} < 240$) with respect to the background (image mode: 40).

A monochrome camera has been used because the color information are not relevant for the image analysis of the target, monochrome sensors of the same size have higher accuracy and 1/3 of data is to be processed.

4. Localization strategy

3D vision in industry is still an open problem and strictly depends on the system configuration. In this case, camera and positioner are calibrated. Alignment of positioner and camera to produce positioning errors under 1 pixel would require an angular accuracy of $3.5° \times 10^{-4}$, so it is reasonable to consider their relative position unknown.

Because the relative position between camera (reference system O_C in Figure 1) and positioner (O_P) is unknown, it also needs to be calibrated in order to be able to provide the necessary correction of the grasping error. According to the proposed model, both tasks, positioning and calibration, can be carried out concurrently.

The chip can be moved by the positioner on a user defined trajectory. To minimize alignment errors the pivot of the positioner for the 3 rotational axes coincide with the chip center O_G.

To determine the 3D position of the chip after grasping, 3 coordinates and 3 angles, we need to know the inverse transformation $\mathbf{T}^{-1}$ from the camera coordinates $\vec{X}_C$ to the positioner coordinates $\vec{X}_P$. This transformation is represented by 12 coefficients and requires a minimum of 6 non aligned points [6]. So although we have 16 control points in a row (= tip Centers of Gravity), we need at least 2 rows (or more, and solve the system in the least squares sense). The transformation

$$\vec{X}_C = \mathbf{T} \cdot \vec{X}_P$$

is invertible also with current single-camera configuration, because the depth information is provided by the focus.

There are two cases.

a) If the relative position between camera and positioner does not change, which means that we are confident that the camera support is sufficiently stiff compared to the system resolution, we determine the transformation $\mathbf{T}_*$ at the setup phase, by acquiring the reference position $\vec{X}_P^*$ using at least 2 rows of a reference chip (aligned manually). The acquisition strategy is moving the chip at assigned steps along only 1 axis, e.g. Y_P.

   For each new chip, we determine the correction

$$\Delta\vec{X}_P = \vec{X}_P - \vec{X}_P^* = \mathbf{T}_*^{-1} \cdot \left(\vec{X}_C - \vec{X}_C^*\right)$$

   also using 2 rows to determine the actual position, compute the distance from the reference position directly in camera coordinates, and apply the inverse transformation.

b) If there is a risk that the relative position between camera and positioner changes, like in the current case, and in general to repeat the setup for ordinary maintenance, than we need to determine an additional transformation $\mathbf{T}_G$ to separate the component of the grasping error from

$$\mathbf{T}_* = \mathbf{T}_i \cdot \mathbf{T}_G$$

   for i=0, and the correction will be given by

$$\Delta\vec{X}_P = \left(\mathbf{T}_i \cdot \mathbf{T}_G\right)^{-1} \cdot \left[\vec{X}_C - \left(\mathbf{T}_i \cdot \mathbf{T}_G\right) \cdot \vec{X}_P^*\right]$$

   for i≥1. In this case the acquisition strategy will require moving at assigned steps along 2 positioner axes, e.g. Y_P and Z_P, (because X_P is approximately parallel the sensor plane) both during setup (i=0) and for each new chip to determine current $\mathbf{T}_i$.

The camera calibration parameters (image pixel sizes) are determined in the system setup phase by multiple acquisitions of the chip and averaging its (known) size in both directions. With a fixed focus lens, they will not change during use.

5. Chip inspection

Product defect types are summarized in Figure 3: missing, damaged or misplaced tips for a maximum of 5% (2750 on each chip), in addition to tips planarity with a 3.2 μm tolerance. Basically, the majority of tips is used for calibrating the system and locating the chip. For the inherent parallel manufacturing process, systematic errors are expected. The minority of defective tips, if present, are detected by a statistical approach.
With current configuration, the two tasks, the positioning and the inspection of the chip, can be dealt with concurrently, but to reduce the assembly time the number of acquired images needs to be minimized.

Figure 3: Overview of chip defects with image examples (insets), inspection rules and multiple sampling scheme according to [8].

The inspection of 100% of tips would require the acquisition of more than 3,400 images. By reducing the processing time by a reasonable $1/10^{th}$ by the system engineering, this would take 5 min./part. For the product inspection by sampling according to standards [8], for an Acceptable Quality Level of 4%, in a single sampling scheme 315 tips (= 20 rows, images) need to be inspected and a maximum of 21 tips must be defective. Considering that with an image processing time of 1 s the total inspection time of 20 s per sample is not acceptable, a multiple sampling scheme is selected. In cluster sampling, this inspection scheme only requires 5 random rows to be analyzed and requires taking another sample of 5 rows if at least 2 tips are found missing or defective and so on according to [8]. The inspection algorithm with actual data is summarized in Figure 3. It should be noticed that for misplacement no absolute position is necessary from the localization module, but 2D row information are used and the Euclidean distance is computed, so inspection can be a stand alone module. For the cantilever planarity assessment, instead, 3D information on the 5 rows is necessary; as explained, 2 rows in 2 directions were already required by the localization algorithm.

The inspection algorithms described have shown a reliability of 100% in hours of testing 1 image/s. This is a direct consequence of the algorithms with higher resolution than required for inspection, which have been developed for localization purposes.

6. Image analysis algorithm

To allow the chip localization and inspection, several modules have been developed. The most relevant are briefly described.

1. Image binarization at fixed threshold. Considering the constant direction of light (confocal) and the constant reflectivity of the surface (same material, parallel manufacturing process) the appearance of the tips is consistent. In the setup phase, an automatic training programs determines the optimal threshold. Experiments have shown that the appearance of binarized tips does not change significantly with lighting: the tip area changes less than 5% within a range of 20 graylevels over 102 (actual selected threshold).
2. Blob analysis. To achieve subpixel accuracy, the position of the tip is associated to its center of gravity, because it is well known that region based properties are more accurate than edge or point measures. At current magnification, a row of 16 tips is visible.
3. Depth-from-focus. In addition to the magnification limitations, the small field depth is another drawback for vision system in microenvironments, but it is one of the benefits from the point of view of the depth-from-focus criterion [9] and it is exploited here.

   In simpler words, the basic principle of the algorithm developed is an autofocus where by focusing we determine the target distance. The degree of focus is determined from the image gradients and is calculated as the mean graylevel of a square surrounding box of size 120 pixels around the center of gravity of each tip, after applying a 2D discrete convolution using a Laplacian of Gaussian filter with a fixed threshold of 0.002. Experiments have shown that the degree of focus is over 0.015 (Figure 5), with unfocused images under 0.001. The original image has a mean graylevel of 30-35 with a maximum > 230.
4. Approach phase. Before reaching the working distance of the objective, to accelerate image acquisition and processing the image maximum (> 150) and low resolution (0.8 megapixels) are sufficient, so this phase is essentially limited by the positioner speed.

7. Experimental

The actual system components depicted in Figure 4 are listed here and have the features in Table 1.

Figure 4: Experimental setup and interaction among the system modules.

The camera is a 2.5" CMOS USB from Pixelink, model B777U with resolution of 2592×1944 at 8 bit/pixel and 7 Hz.

The lens is Navitar model Precise Eye Ultra C-mount, with Numerical Aperture 0.42, field depth 3 μm, 1× adapter and lighting.

The optical resolution determines the resolution of the camera sensor, which has been selected 400% higher.

The optical parameters in these extreme magnification conditions are fundamental to determine the positioning strategy. The minimum perceivable focus (reached by the approach module) determines the start and end position of the (exhaustive) localization strategy (along Y_P): ± 10 μm across the position with the maximum degree of focus.

Image analysis by non-compiled (interpreter) software and motion actuation takes 1 s with a dual 3 GHz processor. The processing at given steps for 5 rows is required for localization and inspection.

The micropositioner is model F-206.S from PI, a calibrated 6 axis hexapod with nominal resolution of 33 nm.

8. System resolution

The maximum (= the mean) of the distribution of the focus along Y_P or Z_P, determines the tip position. The small red circles in Figure 5 are calculated and represent the maximum likelihood estimation assuming a Gaussian distribution, as predicted by the theory [9]. Increasing the acquisition number slows down the localization process, but despite the noise increase, qualitatively shown from left to right, accuracy increases. To quantitatively estimate the system resolution, the distances between successive tips in a row are considered, conservatively assuming a perfect chip and that errors are only due to the localization system. In particular from 44 tests summarized in Table 2 (3 examples are displayed in Figure 5), the margin of error in the estimate with a confidence level of 95% is well below the pixel. This also shows that the vision algorithms and the localization strategy implemented and tested have preserved the maximum theoretically achievable optical resolution.

Figure 5: (a) 3D Map with the degree of focus for each of the 16 tips in 30 images spaced 500 nm along Y_P (Figure 1). 2D detail (b-c) with the focus profile at different step sizes (and number of images) for the 16 tips (color codes).

Table 2: Estimates of the distance between successive tips in a row as a function of the number of images (acquisition steps along Y_P).

9. Conclusions

We have described the artificial vision issues of an automated assembly system in the nanometric range.

Not only the automated system is able to reduce the positioning time from half hour to minutes (after engineering), but the redesigned assembly process is also integrated with inspection, by only adding the acquisition of 1 chip row to the 4 required by localization. The optical lower boundary of resolution in microscopy has been overcome by multiple acquisitions required by the two tasks according to the described localization and sampling strategies.

For their size and repetitive geometry, AFM probe arrays can be used as patterns in vision or positioner calibration tasks in microsystems.

The examined hardware configuration and the developed system can be exploited for different assembly tasks, including: handling (without inspection), inspection (without handling), camera calibration (for a new inspection system), and positioner calibration.

Acknowledgements

The authors would like to thank Mr. M. Nelson (Nanoink Inc, Skokie, IL), Mr. L. Coletta (MVI Inc., Avon, MA), Mrs. A. Watral and prof. J.-H. Chun (MIT), and prof. M. Santochi (Univ. Pisa) for their support.

References

[1] Santochi, M, Dini, G., 1998, Sensor Technology in Assembly Systems, CIRP Annals, 47/2:503-524.

[2] Van Brussel, H., Peirs, J., Reynaerts, D., Delchambre, A., Reinhart, G., Roth, N., Weck, M., Zussman, E., 2000, Assembly of Microsystems, CIRP Annals, 49/2:451-472.

[3] Weckenmann, A., Jiang, X., Sommer, K.-D., Neuschaefer-Rube, U., Seewig, J., Shaw, L., Estler, T., 2009, Multisensor data fusion in dimensional metrology, CIRP Annals, 58/2:701-721.

[4] Hesselbach, J., Schöttler, K., Tutsch, R., Berndt, M., 2006, Assembly of Hybrid Microsystems Using an Assembly System with 3D Optical Sensor, CIRP Annals, 55/1:11-14.

[5] Watanabe, A., Sakakibara, S., Ban, K., Yamada, M., Shen, G., Arai. T., 2005, Autonomous Visual Measurement for Accurate Setting of Workpieces in Robotic Cells, CIRP Annals, 54/1:13-18.

[6] Lanzetta, M., Tantussi, G., 1996, Monitoring Critical Points in Robot Operations with an Artificial Vision System, Proc. 4th Int. AMST Conf., Udine (Italy), 02-03/09/1996:593-600.

[7] NanoInk, Inc., 2009, 2D nano PrintArray™ Data Sheet, http://www.nanoink.net/d/datasheet_2DnanoPrintArray.pdf, last acc. Mar. 2010.

[8] ISO 2859-1:1999, Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.

[9] Pentland, A.P., 1987, A New Sense for Depth of Field, IEEE Trans. on Pattern Analysis and Machine Intelligence, 9/4:523-531.

Table 1

Table 1: AFM probe array and optical system specifications.

chip size:	1×1 cm^2
positioning error:	0.5 μm/0.03°
pyramid tip #:	55.000
tip height:	7.5 μm
tip base size:	4×4 μm^2
rows/spacing:	110/90 μm
columns/spacing:	500/20 μm
lens external ∅:	25 mm
Working Distance:	20 mm
viewfield:	320×240 μm^2
image pixel size:	122.79 nm

Figure 1

Figure 1: Grasping and handling of the AFM chip (SEM inset) under a camera before gluing.

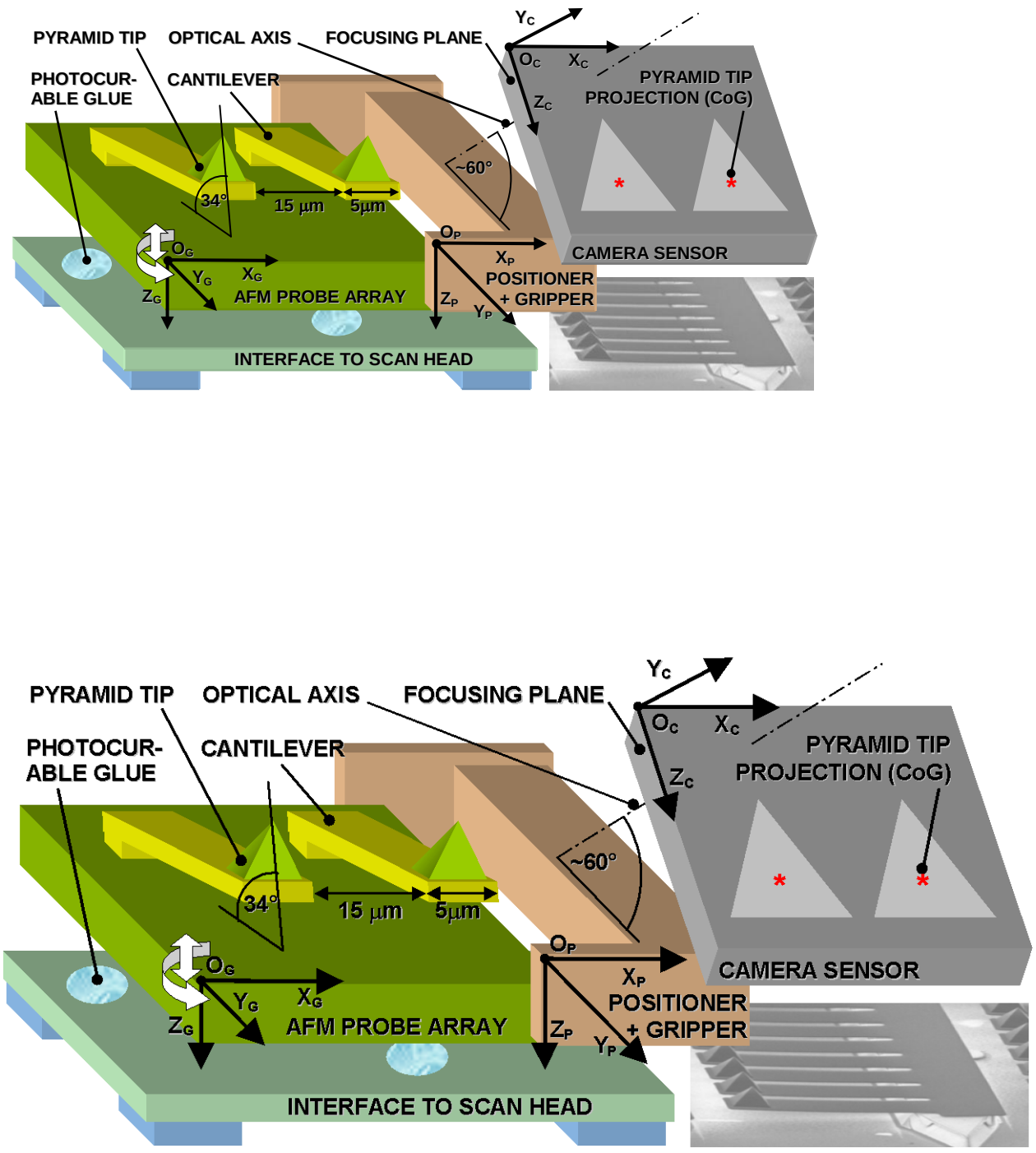
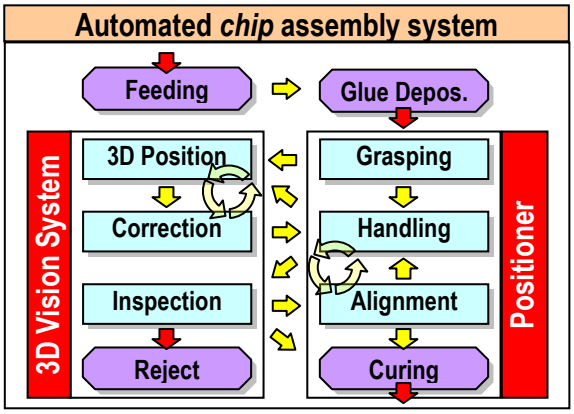


Figure 2

Figure 2: Assembly phases for the AFM probe array, showing material in/out (red arrows) and the vision and positioner (cyclic) interactions.



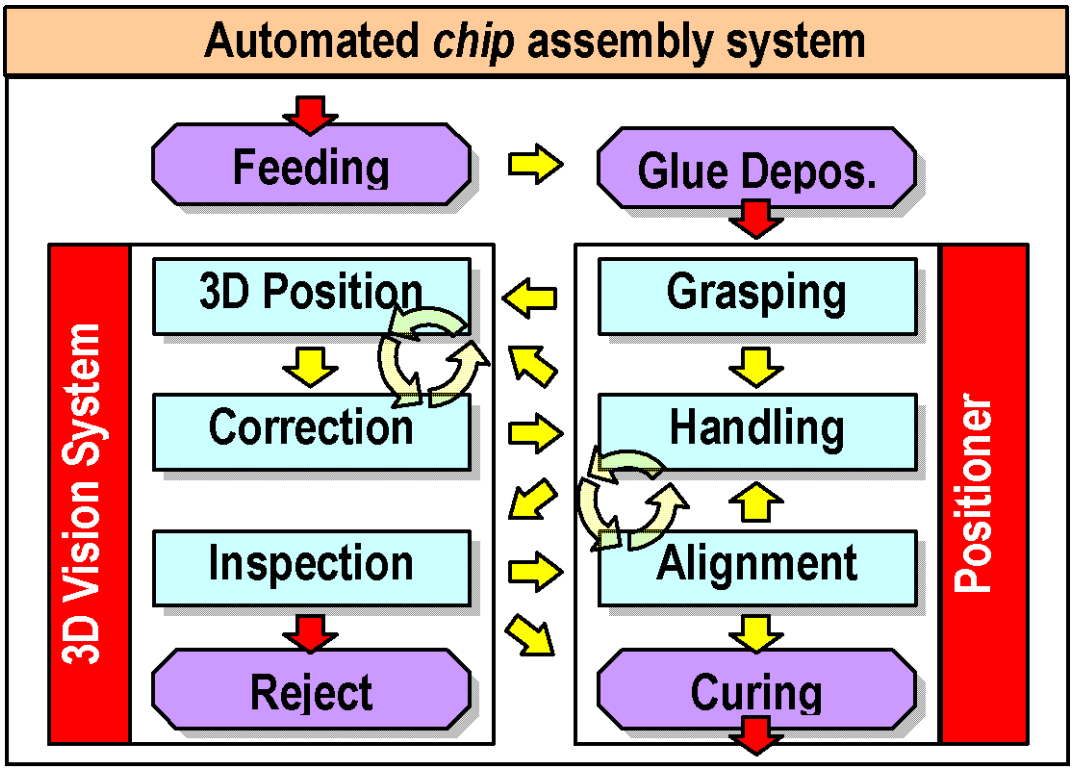
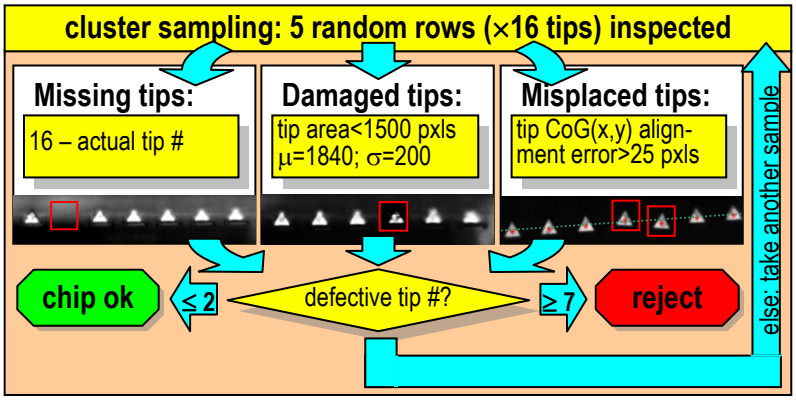


Figure 3

Figure 3: Overview of chip defects with image examples (insets), inspection rules and multiple sampling scheme according to [8].



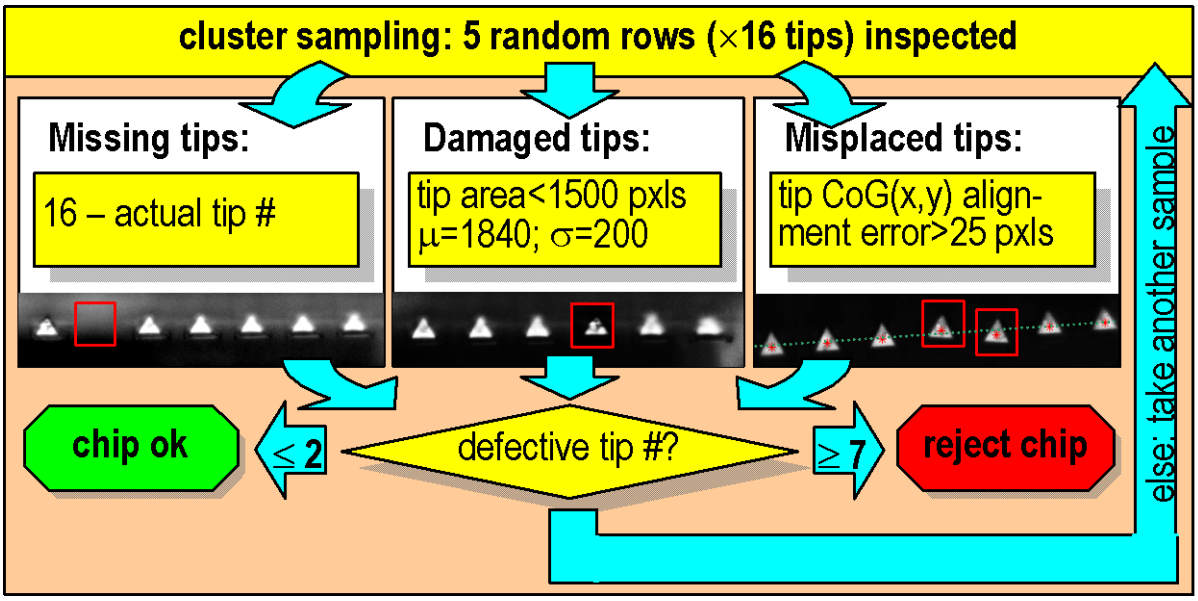
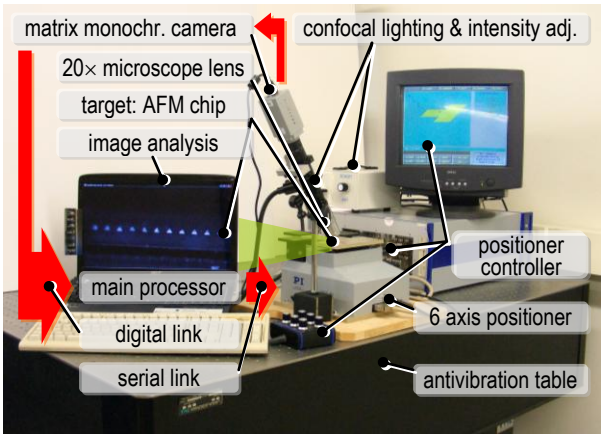


Figure 4

Figure 4: Experimental setup and interaction among the system modules.



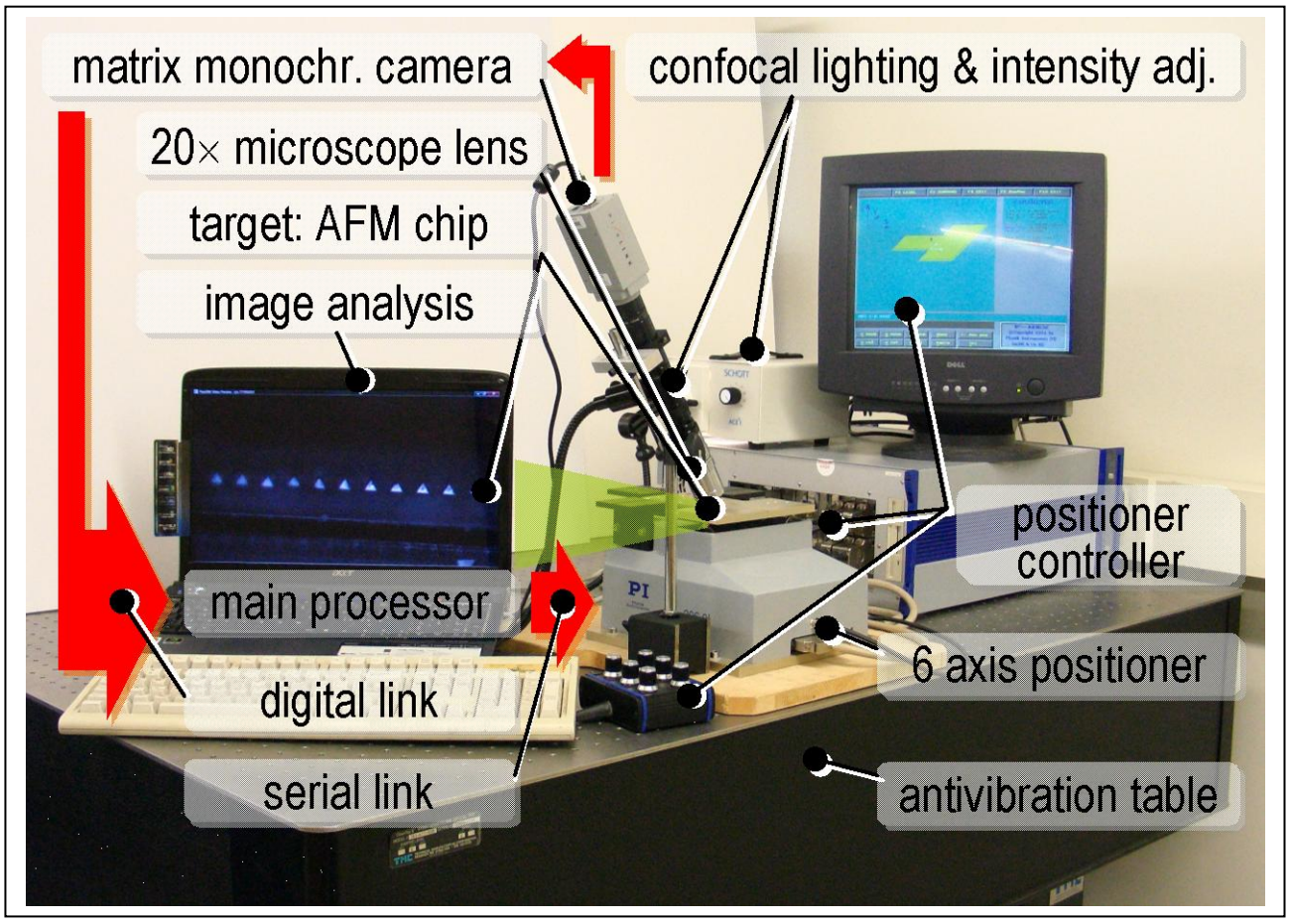
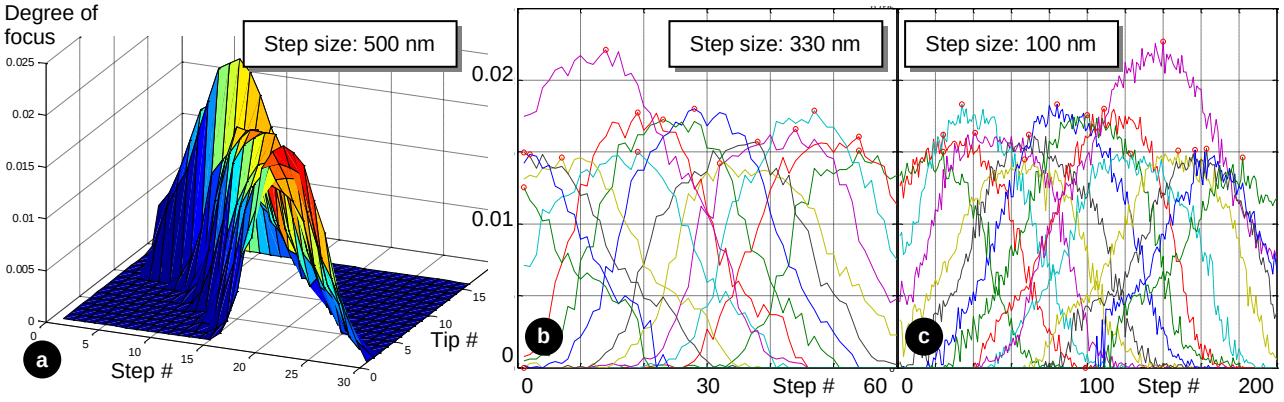


Figure 5

Figure 5: (a) 3D Map with the degree of focus for each of the 16 tips in 30 images spaced 500 nm along Y_P (Figure 1). 2D detail (b-c) with the focus profile at different step sizes (and number of images) for the 16 tips (color codes).



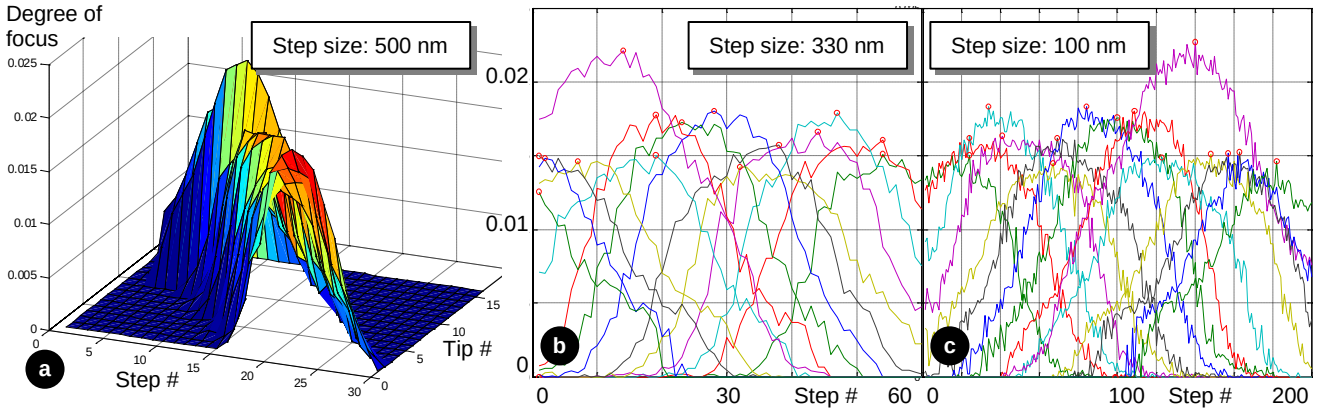


Table 2

Table 2: Estimates of the distance between successive tips in a row as a function of the number of images (acquisition steps along Y_P).

Steps #	30	60	200
CI @ 95%	±124 nm	±72 nm	±50 nm
Std. dev. [pxls]	2	1.1	0.8
CI with 3 repetitions	±23 nm	±13	±8 nm
Same as above [pxls]	0.16	0.09	0.07