The Scanning Force "MAcroscope":
a didactical version of the Scanning Force Microscope

G. Torzo[‡], B. Pecori[§], P. Scatturin[†], G. Delfitto[‡]

[‡] *CNR, INFM and Dipartimento di Fisica, Università di Padova*
[§] *Dipartimento di Fisica, Università di Bologna*
[†] *Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università di Padova*

Abstract

We describe an home-made device simulating, at macroscopic level, the behavior of a scanning force microscope, that we suggest to use for didactic purposes aimed at clarify genesis and meaning of the electronic images, frequently found in textbooks.

Introduction

It is a common experience the game of trying to find out the value of a coin inside our pocket by hand inspection only. Few of us, however, have probably spent time considering how complex and sophisticated are the mechanisms involved in this "simple" action of our fingers collecting data and our brain transforming data into a "coin image".

Our world is full of images. Some of them, like the image of a coin, are part of a general background knowledge which has been kept growing since our birth in a way that is substantially the same as that of our ancestors and, appearing "natural", do not rise any curiosity about the mechanisms generating the image .

On the other hand the modern technologies, investigating microscopic structure of matter (e.g. in physics or biology), produce today images of different kind, associated to "objects" quite far from our perceptive capability (viruses, atoms…), whose production mechanism is totally mysterious for the common reader.

The lack of information on the production processes of such images leads often the non-experts to "wrong" interpretations, by assigning to the imaged objects characteristics that are meaningful only at a macroscopic level[1].

---

[1] This incorrect extrapolation of macroscopic properties to microscopic objects is often induced into the readers by the frequent use of artificial colors and shadows.

These considerations brought us to design a device that could be used, even with students with poor physics background, to help understanding the basic processes through which images are normally built from electric signals acquired by sensors[2].

The leading idea is to use a device working at macroscopic level with techniques similar to those used investigating the microscopic world.

It might appear useless to produce an image of objects that can be seen by eye, but we believe it useful to offer the students the possibility of comparing the "natural" image with the "reconstructed" one, allowing them to follow step-by-step the image-generating process, thus reaching a better understanding of the different nature of the two images.

1. How are topographic images produced?

Electronic images may be of different type, but quite often they are "bit-mapped", i.e. they are a *matrix of dots* of different color (or gray-scale), that may be printed or displayed on a computer screen. Each dot (or "pixel" in electronic language) is a numeric value in the matrix stored in the computer memory.

Topographic images are special images, that give a quantitative representation of the surface of an object, where the color of each point "measures" the distance z of that point from a geometric plane parallel to the average surface of the object ($z=z_0$).

For example in the gray-scale LUT ("look-up-table" = correspondence between colors and numeric values in the matrix) the brightness may be proportional to the distance from the reference plane, so that the darker tonality corresponds to deeper surface sites and lighter tonality to higher surface sites.

From this point of view SEM or TEM images are *not topographic* images, even if they give some information on the surface geometry. A more familiar example of non-topographic image is that obtained by rubbing a pencil onto a thin paper sheet pressed against a corrugated sample (e.g. a coin): the result gives an information on the sample structure which, however, is not topographically reliable[3].

---

[2] The system sensor + ADC+ computer is the typical way of producing microscopic images in the modern technology.

[3] The 3D representation is reconstructed in our mind by the image, according to our previous knowledge of the sample, in a process similar to that induced by the famous Escher's picture "*High and Low*" which leads the observer to see a paradoxical perspective.

Typically the topographic matrix I=z(x,y) maps a square region of the surface, each pixel representing the average value of the surface height within a square whose side depends on the image resolution. The surface is sampled in a square lattice of points equally spaced in the x and y directions along the reference plane.

If we pick up only a row of the matrix we obtain a vertical cross-section of the topographic image. By plotting this set of numbers versus their position along x, with segments joining the nearby dots, we get an x-line profile. Putting in the same plot all the x-line profiles corresponding to the various y positions (each one shifted upward of the same quantity z) we obtain a 3D representation of the surface that is named "wire-frame"[4].

The azimutal "angle of view" of this 3D image is determined by the shift value z, while a polar rotation of the angle of view may be obtained by applying a shift x to each x-line profile.

Wire frame images may finally be converted into more "realistic" surface plots by coloring the facets enclosed within contiguous segments, and by "hiding" (i.e. not tracing on the screen) those parts of the surface that appear on the back of overlaps. These kind of transformations are commonly performed by modern software dedicated to 3D tracing and rendering.

In figure 1 we report three images obtained with commercial SFM operating at constant force (square matrix of 256x256 pixel where lighter dots indicate larger z values).
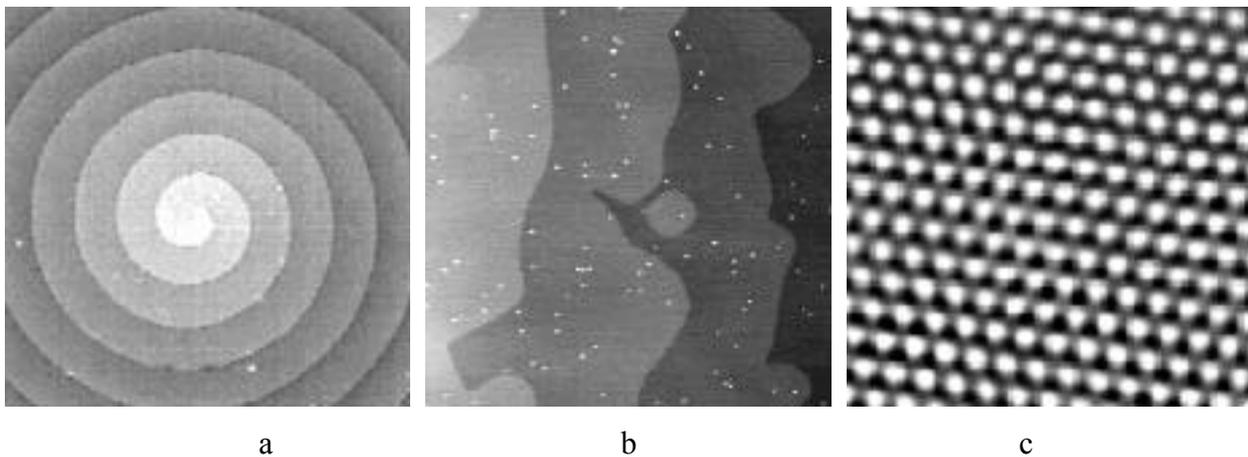

| a | b | c |

Figure 1: SFM images a: SiC (15x15μm), b: GaAs (3x3μm), c: Mica (60x60 Å)

The first image shows a lattice defect (screw dislocation) in a Silicon Carbide single crystal: a spirally terrace few microns wide. In the second image are evident single atomic layers steps in a Gallium Arsenide single crystal cut along 001 plane, with Indium Arsenide nanostructured

---

[4] Normally the wire-frame images collect both x-line and y-line profiles.

overlayer (white dots). In the third image, postprocessed by filtering the noise on the original image, are evident the atomic sites and the hexagonal symmetry of mica.

2. Working principle of the "macroscope"

Our device [5]is based on the technology of "stylus profilometers". The commercial profilometers work in one dimension, generating a line-profile of the investigated sample, while our device must scan a surface, by moving the probe along a series of parallel lines (the raster scanning normally used by SFM when storing bitmapped images in a matrix).

In our didactic laboratory we were routinely using a two-channel interface with easy software for data acquisition[6]: therefore we chose this simple system to first build a profilometer, using as vertical displacement probe (z-axis) a tip soldered at the end of a cantilever whose deflection was measured by strain gauges, and as horizontal displacement sensor (x-axis) a trolley whose driving screw rotation angle was measured by a potentiometer (Figure 2).
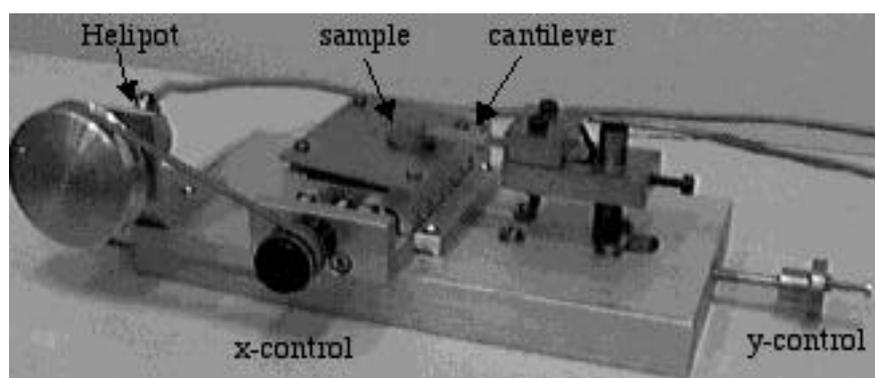

Figure 2: The macroscope

Our device was able to record two series of values (vectors) corresponding to the signal $V_Z=V(z)$ of the vertical displacement of the tip following the sample topography, and to the signal $V_X=V(x)$ of the potentiometer measuring the displacement along the x axis. On the computer screen was

---

[5] Named "macroscope" following the suggestion of prof. Cesare Ascoli, an Italian pioneer in the field of scanning probe microscopy, during a private talk with one of the authors.

[6] We used the interface ULI and software "RotaryMotion" from Vernier (Portland , Oregon, USA), but similar performance may be obtained with the cheaper CBL interface for Texas Instruments graphic calculators (see Appendix C).

therefore possible to trace the plot $V_z=f(V_x)$ reproducing the sample profile along a line parallel to x.

A series of scans (each one made for a different position along y-axis) contains information on the whole sample surface, but if we placed them all together in the plot $V_z=f(V_x,y)$ they would result unreadable, due to the overlap.

We therefore used a simple "trick" in order to make visible the sample topography, and to produce a "pseudo-3D" image. The trick consists in mounting the sample onto a tilted plane so that every displacement along the y-axis produces a vertical shift $V_Z=V(y \tan\alpha)$

The tip-cantilever system (z-axis displacement sensor) is mounted on a metal holder that may be moved along the y-axis by rotating a driving screw, while the sample-holder performs the x-axis scan.

The x-axis movement is obtained using a cheap trolley[7]: whose knob is ganged to the axis of a potentiometer by means of a pulley.

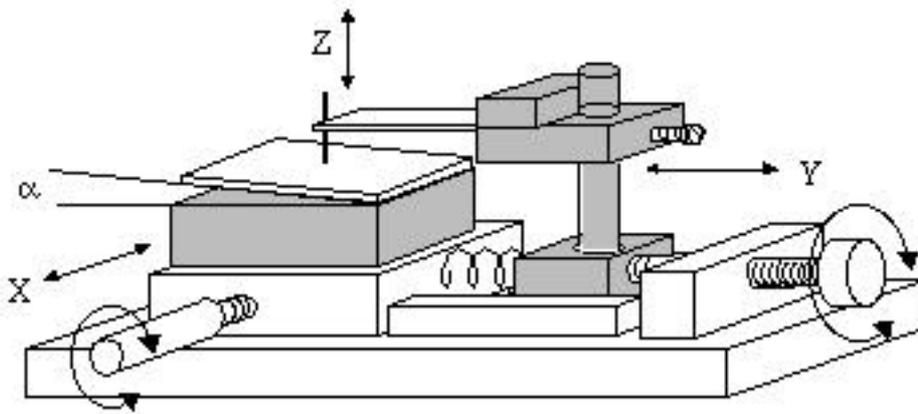

Figure 3: Schematics of the device

The maximum scan width is limited by the range of x and y screws (in our case about 20 mm).

A sketch of the device is shown in figure 3.

---

[7] We used a Domino 50, produced by Schluderbacker

3. Some images

To give an idea of the capabilities of this device we present in figure 4 some images of coins obtained using 40 lines of 100 points per line; with a sampling frequency of about 4 points per second a total acquisition time of about 10 minutes is required.
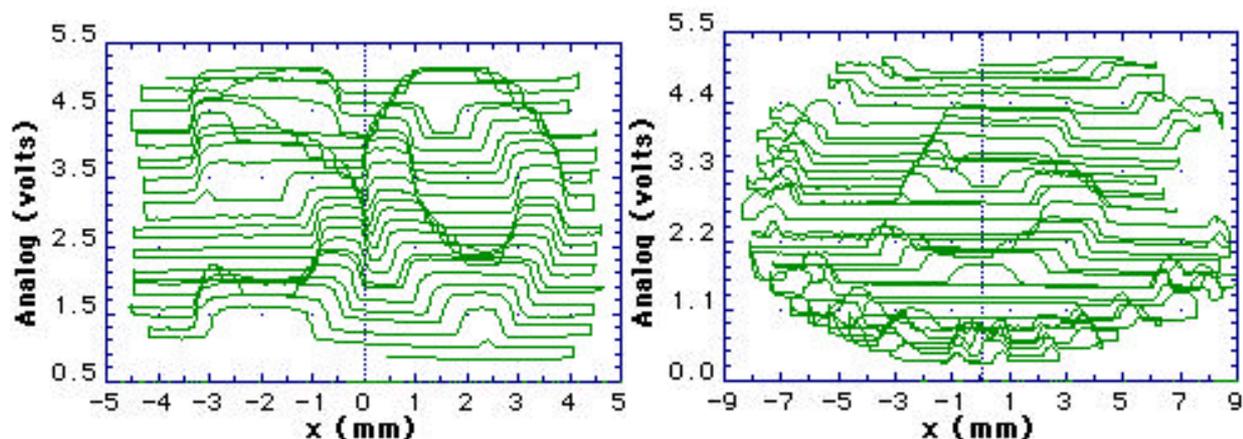

Figure 4: Profiles of 50 and 5 Pfennig coins

A photo of the sample coins (50, 5 and 1 Pfennig, chosen for their small diameter and simple design) is shown in figure 5.

Figure 5: The sample coins and the cantilever

The images shown in figure 4 were taken using a wool-needle, with a not too sharp tip (whose radius of curvature R=0.28 mm is shown in figure 6a), thus reducing the risk of chokes and lateral deflections of the cantilever when crossing sharp steps on the sample. Our device is affected by this problem because it does not have the feedback mechanism that allows SFM to work at *constant*

force: in that case when the tip meets a sample relief, the sample holder is suitably displaced along z to keep constant the cantilever deflection.
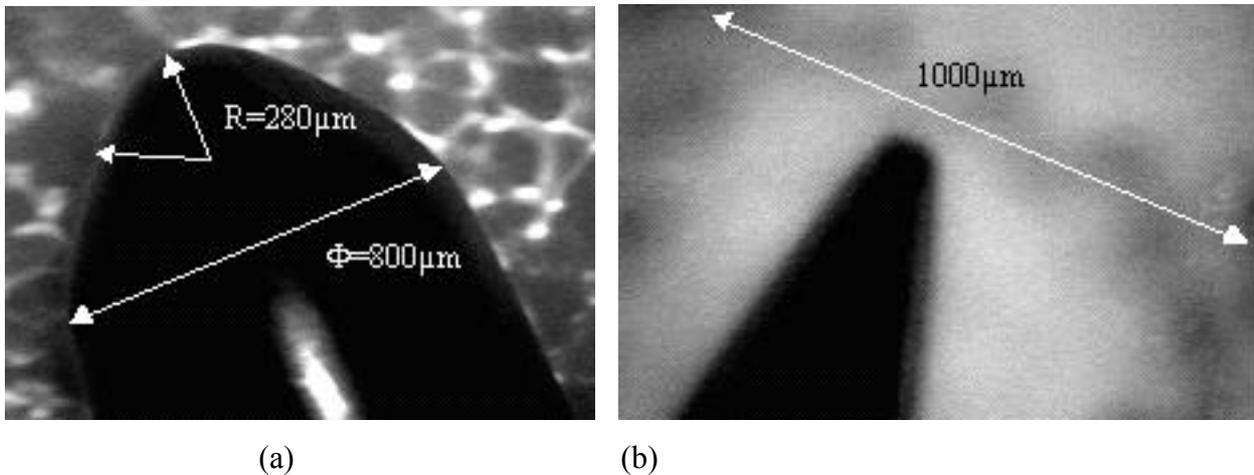

(a)   (b)

Figure 6: Photomicrographs of a wool-needle tip (a) and of a sewing-needle tip (b)

Obviously the lateral resolution is as better as smaller is the tip radius of curvature. To show this effect, that in the language of probe microscopists is named "tip convolution"[8], we recorder two images of the same sample first using a sharper sewing-needle (R=0.08mm) and then a ball-bearing sphere (R=1.5mm) as probe tips.
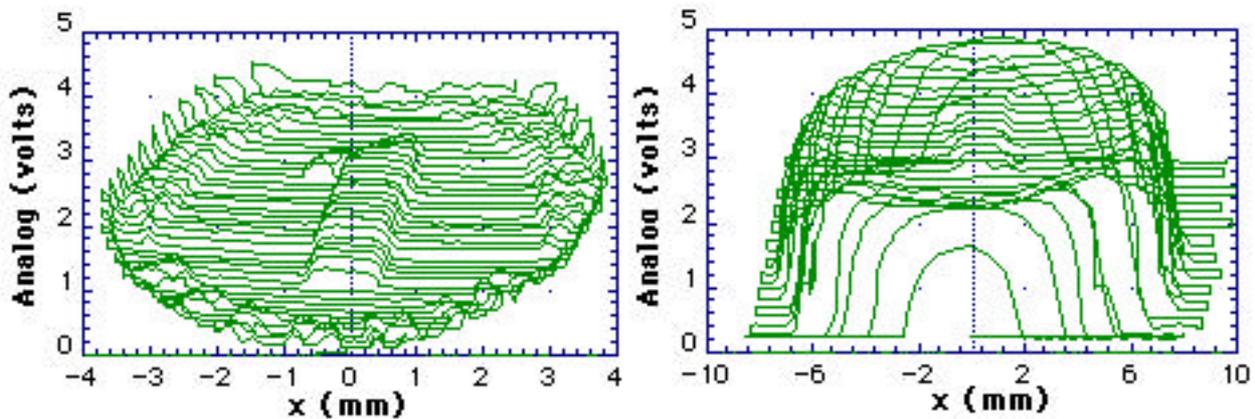

Figure 7: Images of the same sample (1 Pfennig coin) taken with different tips (see text))

Using a curvature radius larger than the coin thickness we can detect the sample borders (the tip can cross steps slightly smaller than R), but we spoil our lateral resolution. Using a smaller curvature

---

[8] See G.Torzo and D. Cerolini: *SFM image reconstruction reducing tip artifacts*, Microscopy and Microanalysis, May 2000.

radius we get better resolution, but the probe-sample interaction becomes important (note that in our devices the applied force increases with z), so that at the scan end the sample results slightly damaged (see Appendix B).

4. The Z-detector working principle

The strain gauge is made of a thin film resistance incorporated into a plastic strip to be glued to the object whose strain has to be measured.

The geometry of the resistive film track gives to the sensor vectorial properties, with maximum sensitivity along one direction. Applying to the film a tensile (compressive) strain increases (decreases) its resistance. The change in resistance dR is proportional to the relative length change ϵ =dL/L, with a gauge factor G=(R/R)/ 2.

Figure 8: Strain gauge

By gluing two identical strain gauges to an elastic cantilever, and by connecting them in a Wheatstone bridge with two fixed resistors R, we obtain a force sensor (see figure 9 and Appendix A).
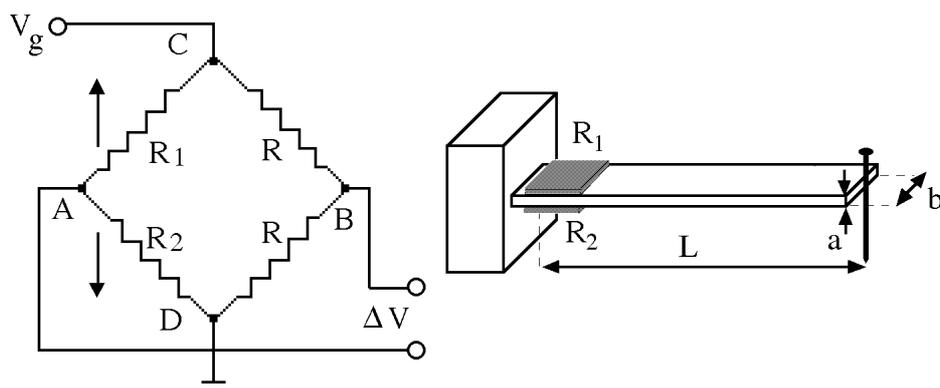

Figure 9: The Wheatstone bridge configuration of the strain gauges resistances R_1 and R_2, and sketch of the cantilever geometry. The actual probe is a needle soldered at one end of the cantilever.

When the tip is pushed upward, with a vertical shift z, the cantilever deflection produces the relative change + in lower strain gauge and – in the upper one. The value of the shift may be

written (see Appendix B) $\Delta z = (2L^2/3a)\epsilon$, where a is the cantilever thickness and L its length, so that the relative resistance change is $\Delta R/R = G\epsilon \approx \Delta z(3a/L^2)$. For example for $L \approx 30$ mm, $a \approx 0.3$ mm we get $(\Delta R/R)/\Delta z \approx 10^{-3}/$mm, a sensitivity comparable with the thermal drift of metallic resistances $\beta = (\Delta R/R)/\Delta T \approx 10^{-4}/°C$.

This explains why we use *two* strain gauges instead of one: besides the gain of a factor 2 in sensitivity, with matched gauges we make the system unaffected by thermal drifts.

The voltage V_A at the strain gauges midpoint depends on the bridge bias voltage V_g and on the values of the two resistances $(R-\Delta R)$ and $(R+\Delta R)$: $V_A = V_g/2 + V_g \Delta R/2R$. The fixed resistors midpoint voltage is $V_B = V_g/2$, so that the bridge output signal is $\Delta V = V_A - V_B = V_g \Delta R/2R = V_g\epsilon$.

With $V_g \approx 1$V and $\Delta z \approx 0.1$mm we get a signal $\Delta V \approx 1$ mV, requiring a gain of about 1000 in order to match the input voltage range of the used interface (0-5V).

The signal measuring horizontal displacements is provided by a potentiometer, biased at 5V, driven by a pulley linked to the micrometer axis through a rubber band .

Typical resolution of Helipot is of the order of 10^{-4}, introducing an uncertainty smaller than that due to the ADC (10 bits): e.g. with a trolley moving 2 mm per each turn of the driving screw, the 5mV sensitivity gives a theoretical resolution of 20μm, better than the limit imposed by the curvature radius of the thinnest usable needles.

The images shown in figures 4 and 7 report the x-axis calibrated in mm. How do we calibrated it? We simply measured the change ΔV_x corresponding to a displacement Δx known to calculate the conversion factor $k_x = \Delta x/\Delta V_x$ (mm/V).

To calibrate the y-axis we measure the change in the Vz signal produced by a known shift Δy along a flat region of the sample surface: the conversion factor is $K_y = \Delta y/\Delta V_z$ (mm/V).

For the z-axis we have $K_z = K_y \tan\alpha$, where α is the tilt angle between the sample surface plane and the x-y scan plane.

5   Comparing the macroscope to commercial SFM microscopes

The first obvious difference between macroscope and SFM microscopes is in the order of magnitude both of the resolution and of the maximum sampled area.

Lateral resolution in SFM is much better (about 100nm) not only due to thinner probe, but also because the x,y displacements are much more accurate thanks to the piezo scanner replacing the trolley. Also vertical resolution is much better: with SFM the accuracy may be better than 0.1nm.

The degraded resolution in the macroscope (we loose about 3 orders of magnitude) is balanced by a saving of more than 2 orders of magnitude in the cost (300$ against   60.000$).

The second difference is in the way of image recording: in SFM we get matrix $z_i(x_i,y_i)$, while the macroscope produces a sequence of vectors $z_i(x_i)$, each one containing approximately the same x_i values and the correspondent z_i values are increased each time of the amount   z =tan   y due to shift along y.

The third difference consists in the detected signal: we record the cantilever *deflection* instead of the *vertical displacement that keeps constant* the cantilever deflection. This technique (named *constant height mode* as opposed to the more common *constant force mode*) is rarely used in SFM, where a servo-system varies the cantilever-sample distance (piezo-scanner computer driven) in a feedback loop. The lack of servo-system in the macroscope sets an upper limit to the maximum height H of steps, on the sample surface, that may be detected with a tip of given curvature radius R: typically H<R. As a consequence the recorded image is distorted proportionally to the "aspect-ratio" H/B , where B is the base of the sampled relief. The imaged relief has *rounded borders* with curvature radius R and *oversized base* (B'= B+2R). Therefore best images are obtained using a tip with the minimum R larger than maximum H.

6. Conclusion: The macroscope in physics teaching

The analysis of the images recorded with the macroscope, of the steps required to obtain them, of the technology involved in the process and of the underlying physical knowledge may give several suggestions for teaching.

The macroscope is an example of the "translation process" that is essential in any representation of real world.

It offers the students a concrete context where they may decode the modeling process implicit in the representation .

The object profile (the goal of our representation) is modeled in terms of tip position in a co-ordinate system x, y, z ; these co-ordinates are related to the cantilever deflection due to the force acting between tip and sample; this deflection is related to an electric signal "read" by a device and translated into a bi-dimensional image by a software application.

Schematically we may resume the principal steps in our translation:

Object Profile (1) -> Tip Position (2) -> Force Applied to Cantilever (3) ->

-> Cantilever Deflection (4) -> Sensor Strain (5) -> Resistance Change (6) ->

->Electric Signal (7) -> Computer Handling of the Signal (8) -> Visible Image of the Profile (9)

To understand the logic of this process one does not need to know all the technical details: it is still possible to detect the various steps and to evaluate their effect on the final result.

The macroscope structure, being directly visible, allows to follow the various steps of the measurement up to (4), and comparing the image on the screen with the sample, that is "naked eye" visible, allows to control that the "invisible" process of translating the electric signal is effective and reliable.

Analyzing the image generation process, enlighten the role of the interaction between observer and observed, that is characteristic of any measurement.

Comparing different images obtained with different tips (i.e. different coupling between profile and cantilever deflection), one may really "see" the accuracy limits on the information that can be obtained from the interaction with the investigated sample.

This analysis highlights the differences between what we can know about the coin with the macroscope (its profile within the limits of the device sensitivity) and what we can know looking at the coin, i.e. exploiting the light signal translated into a colored 3D image through a complex sequence terminating in our brain.

The technical difficulties that must be overcome in order to obtain a readable image with the macroscope may be exploited by the teacher to illustrate an interesting set of "tricks" available to physics researchers. Some of these (as using the tilt to get a pseudo-3D image) are easy understandable; others are less immediate (as the use of bridge configuration to increase sensitivity and stability, or the relation between cantilever deflection and measured strain) but they may be however understood in their mainlines.

These technicalities however give the students the opportunity of thinking about the links between physics and technology, not generically, but by giving concrete evidence, within an experiment, to the application of theoretical knowledge to technical problems.

Last but not least, we remember the didactic potentialities of a device that is a "scaled-up" version of one of the most powerful modern instruments for investigating the surface physics. A comparison with commercial SFM and with the SFM-produced images may give the students the

skill necessary to recognize and fully appreciate problems and results of the advanced research in this field.
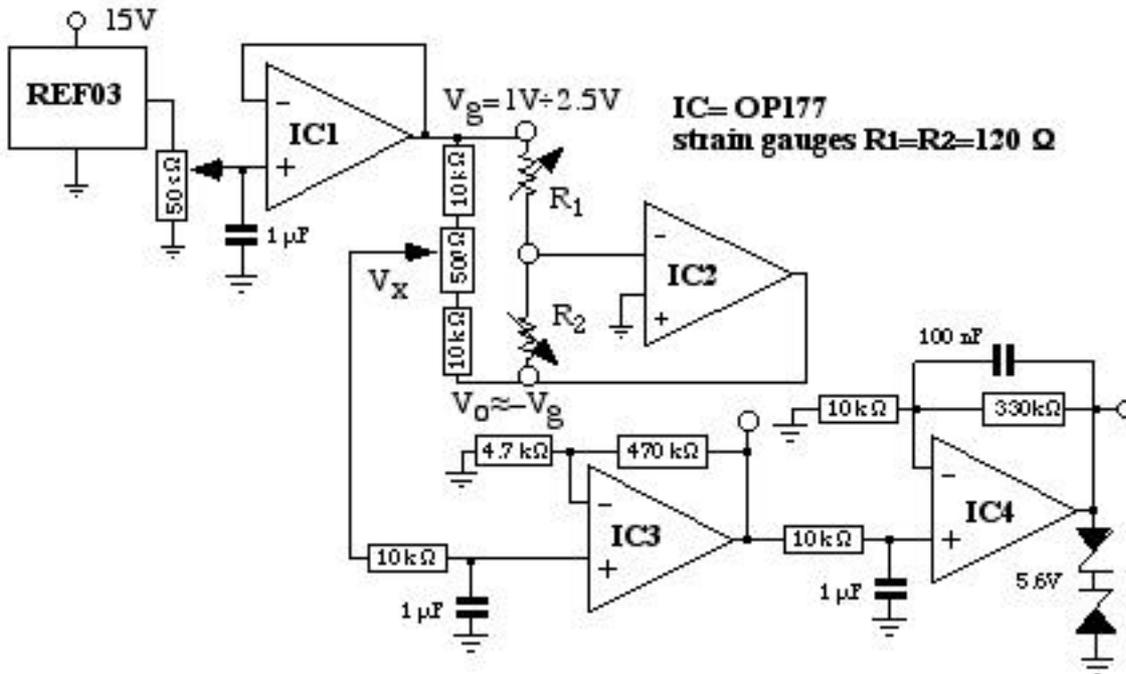
Appendix A: Electronics of the z signal detector

Figure 10: Schematics of the Vz signal amplifier

The bridge is biased at constant voltage V_f (2–5 V) by IC1 and IC2. One arm is made of the two strain gauges (120 Ω) and the other is made of two fixed resistors (10k Ω) and a balancing potentiometer (500 Ω). The strain gauges midpoint is kept at virtual ground by the operational amplifier IC2. The bridge is therefore biased by two signals: V_f and $V_O=-(R_2/R_1) V_f$, where $R_2=R+G$ and $R_1=R-G$ are the gauges resistances, and $G = \Delta R/R$ their relative resistance changes.

The output signal $V_x=(1/2)(V_f+V_O) \approx 2 \Delta V_f$, taken at the sliding contact of the potentiometer, is amplified by IC3 and IC4.

Appendix B: Calculus of the strain ε produced by the cantilever deflection Δz.

The strain in elastic bodies is described by the relation $\varepsilon = 1/E (F/S)$ relating $\varepsilon = \Delta L/L$ to the applied force F, the section S and the Young modulus E (for steel E $\approx 22.10^6$ N/cm^2).

We are looking for a relation between the displacement Δz of the free end of a cantilever and ε.

Thinking the cantilever (horizontal in equilibrium) as a beam of fibers, a deflection produces an increase of the fibers length on the convex side and a decrease on the concave side, with respect to the median fiber.

Now, the length dx of an arc P-Q of the median fiber may be written as product of the curvature radius and angle : dx = .

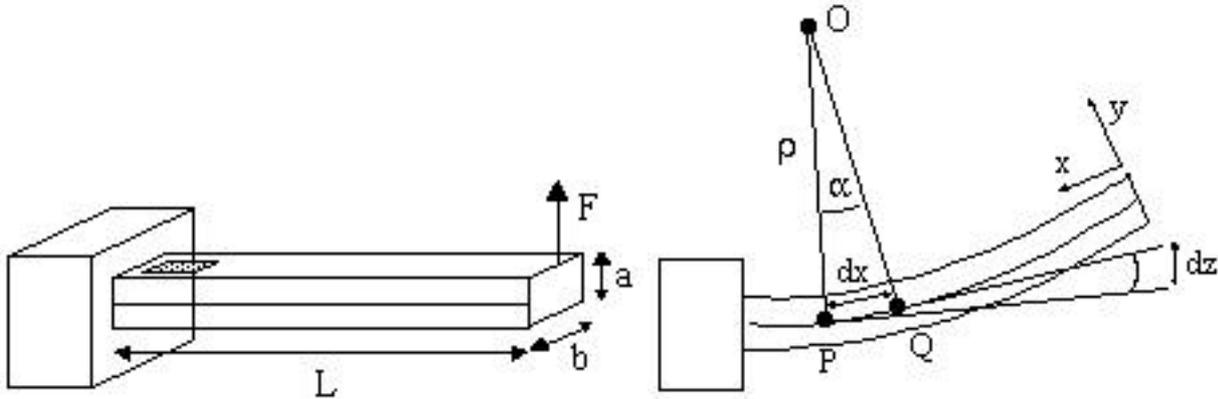
Figure 11: Schematics of the cantilever deflection

The relative length change, at the generic distance y from the median fiber is therefore $=\pm y/$, corresponding to an applied tangent force dF=EdS(y/), where dS=bdy is the cross section of the fiber.

A cantilever deflection corresponds to an applied torque given by the integral

$$= y\,dF = y^2(E/\ )dS = (E/\ )j,$$

where $j = y^2 dS = a^3 b/12$.

Because the angle between tangents to the fiber in P and Q is ,the differential contribution to vertical displacement is dz = xdx/ , where x is the distance of P from the cantilever end (tip position). By substituting we get dz = (/Ej)x dx

Neglecting the cantilever weight, the equilibrium between applied torque Fx and the restoring elastic torque gives dz=(F /Ej) x^2 dx

The total displacement is obtained by integrating along the cantilever length:

z= (F /Ej) $L^3/3$=(F/E)($4L^3/ba^3$).

The strain changes along x growing with curvature 1/ , reaching a maximum close to x L, where the gauges are glued. Here $=\pm(a/2)/ =\pm(a/2)(\ /Ej)$, and the torque is =FL.

From z=(F/Ej)$L^3/3$ we get (FL/Ej)= z/(3/L^2) and finally $=\pm(a/2)(FL/Ej)= z(3a/2L^2)$.

Higher sensitivity (= / z) can be attained by increasing thickness a and reducing length L, but

this also increases the tip to sample force F= $Ea^2b/6L = E$ z $ba^3/4$ L^3.

With a=0.3mm, b=10 mm, L=30mm, the applied force with a deflection of 1 mm is 0.25 N 25 grams, already sufficient to slightly scratch a metal with a sharp needle.

Appendix C: Imaging with CBL

To show the resolution achievable with the cheaper CBL interface we report in figure 12 the plot appearing onto the LCD screen of a TI89 graphic calculator after an acquisition using the sample coin of figure 4b, and the simple software reported in the Appendix.
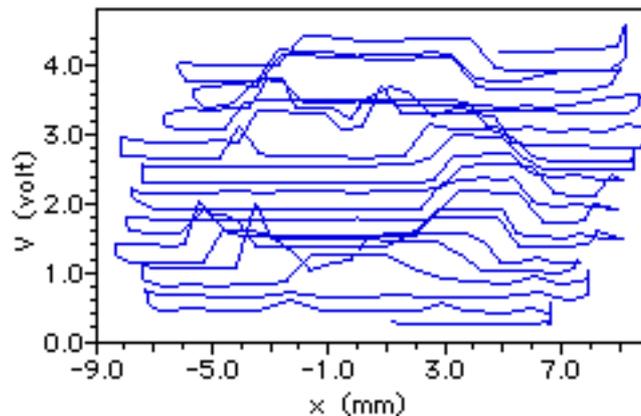

Figure 12: Image obtained with CBL and TI 89

The source code used for obtaining the image of figure 12 contains the following 26 lines

```
Prgm
ClrIO
Request "time step(s)",dt
Expr(dt)    dt
Request "sample number",nump
Expr(nump)    nump
nump*dt    tsamp
PlotsOff
ClrDraw
Disp "Ready ▶ Press ENTER"
Pause
Send{0}
Send{1,1,14,0,0,0}
Send{1,2,14,0,0,0}
Send{2,1,0}
Send{3,dt,nump,0}
seq(n,n,dt,tsamp,dt)    lt
Send{5,1,0,1,nump}
Get lx
Send{5,2,0,1,nump}
Get ly
PlotsOff
ClrDraw
NewPlot 2,2,lx,ly,,,,5
ZoomData
EndPrgm
```