Version 7.0 (Approved version)

Nanomanipulator Proposal

Russell M. Taylor II

Table Of Contents

The task of making controlled surface modifications with a Scanning Tunnelling Microscope (STM) is sufficiently complicated that it requires an effective and powerful user interface to the microscope. The processes of making changes are not well understood and will require extensive experimentation to understand.

The goal of this system is to approach the ideal interface for the scientist - presence on the surface itself. The system will make use of a Head-Mounted Display (HMD) and Force-Feedback Arm (ARM) to present an immersive virtual environment in which the user is projected onto the surface and given the ability to act at the atomic scale. It will give the user the power to modify the real world from within a virtual world. This will require that the virtual world be carefully registered with the real world at all times and that the model presented to the user be consistent with the external environment. The purpose of this system is to transport the user into a nanoworld and to provide a means for the user to make changes in this environment. It seeks to create teleoperation at the atomic scale.

1 Thesis statement

The success of an immersive virtual reality interface in allowing planned changes to a surface with an STM will show that a virtual world can be used as a powerful and effective interface to a real world device.

2 Background

2.1 What is an STM?

The Scanning Tunnelling Microscope (STM) was invented around 1978 by G. Binning and H. Rohrer at the IBM Zurich Research Laboratory. It was originally designed to aid in understanding the growth, structures, and electrical properties of very thin oxide layers.

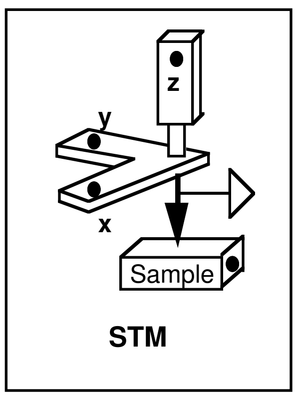

Figure 1: a Scanning Tunnelling Microscope

An STM consists of a piezoelectric positioning element, a conducting (usually metal) tip and a conducting sample (the surface under study). In our instrument, the piezo elements are arranged as three orthogonal bars, each of which controls one axis. As voltages are applied across the crystals, they change their lengths. Since the tip is rigidly attached to the crystals, they can be used to position the tip relative to the sample. A bias voltage is applied to the sample with respect to the tip. At very close range (on the order of a few Angstroms), a tunnelling current to flows between the tip and the surface. This current depends exponentially on the inverse of the distance between the tip and the sample. In our configuration, the X and Y piezoelectric crystals are used to raster the tip across the surface and the Z crystal is controlled by a feedback circuit which attempts to maintain a fixed tunnelling current. A scan of the surface proceeds by repeatedly moving the tip in X and Y and then reading the voltage on the Z axis to determine surface height.

While the best resolution of the Scanning Electron Microscope (SEM) is around 100 Angstroms, the STM can be positioned to within 0.1 Angstrom and can observe features on the order of 1 Angstrom. This means that the STM can view surface features that are two orders of magnitude smaller than those viewed by a Scanning Electron Microscope.

In addition to its ability to see atoms, the tip of the STM can be used as a local probe to modify the surface. This feature makes the STM potentially useful for nanolithography. There are at least two ways this can be accomplished. The first is to physically contact the surface with the tip, which causes large and unpredictable modifications to both the tip and the surface. The second, more controlled, method is to apply a voltage pulse between the tip and the surface. Since the distance between the two is so small, even moderate voltages produce a large electric field. This field can cause material to be moved from the surface to the tip or vice-versa. Other, more exotic methods include irradiating the surface with radiation whose frequency depends on characteristics of the sample surface or of the tunneling gap while applying voltage to the tunnelling gap.

In layman's terms, the STM can be thought of as a very tiny junk yard crane that has a camera mounted on the bottom of the magnet. The crane operator can scan the crane around the junk yard to look for something interesting to pick up. He can also look around to get an idea of where he would like to put things down. However, as soon as he picks something up he blocks the view of the camera so he has to move around blind until he can put the junk down again somewhere. While he is carrying the junk around, he needs some sort of map to tell him how to move around in the junk yard.

2.2 What is the state of the art?

The state of the art for STM data visualization during data collection is examining greyscale images where dot brightness corresponds to surface height at each point in the image. For later offline viewing or publication, most visualization is done using various graph routines on personal computers. These packages draw a connected line for each scanline the tip has made, possibly doing hidden line removal on areas that would be occluded. Color is used effectively to show either surface height or other surface properties. For presentation, shaded surface images are sometimes computed. See the Edstrom paper, the Komuro paper, the Driscoll paper, the Stoll paper, the Snyder paper, and the Magonov paper for more information.

Recent work by Dr. Joe Lyding at the University of Illinois in collaboration with the Beckman Institute has resulted in an AVS interface module that gives data from the STM scans to other AVS modules for high-quality interactive rendering of surfaces using the Gouraud shading model. This results in shaded images from a given point of view at rates on the order of one image per minute. This information was obtained from a personal interview.

E. Besenbacher et. al. at the University of Aarhus, Denmark, have taken another approach. They still draw the images on a personal computer, but they make images from successive scans and put them onto videotape for later viewing. This allows them to view surface dynamics that happened during the experiment. The dynamics cannot be viewed while the experiment is in progress, but nonetheless they have found the motion to be very useful. See the Bedenbacher paper for more information.

3 System Design

Unlike the optical and electron microscopes, which form 2D images of the sample, the STM scan produces 3D data (a 2D array of heights) and therefore is well matched to the HMD. By using head tracking and computation, a viewpoint can be computed for the user from any position, at any scale relative to the surface. The overall system design is shown on the next page.

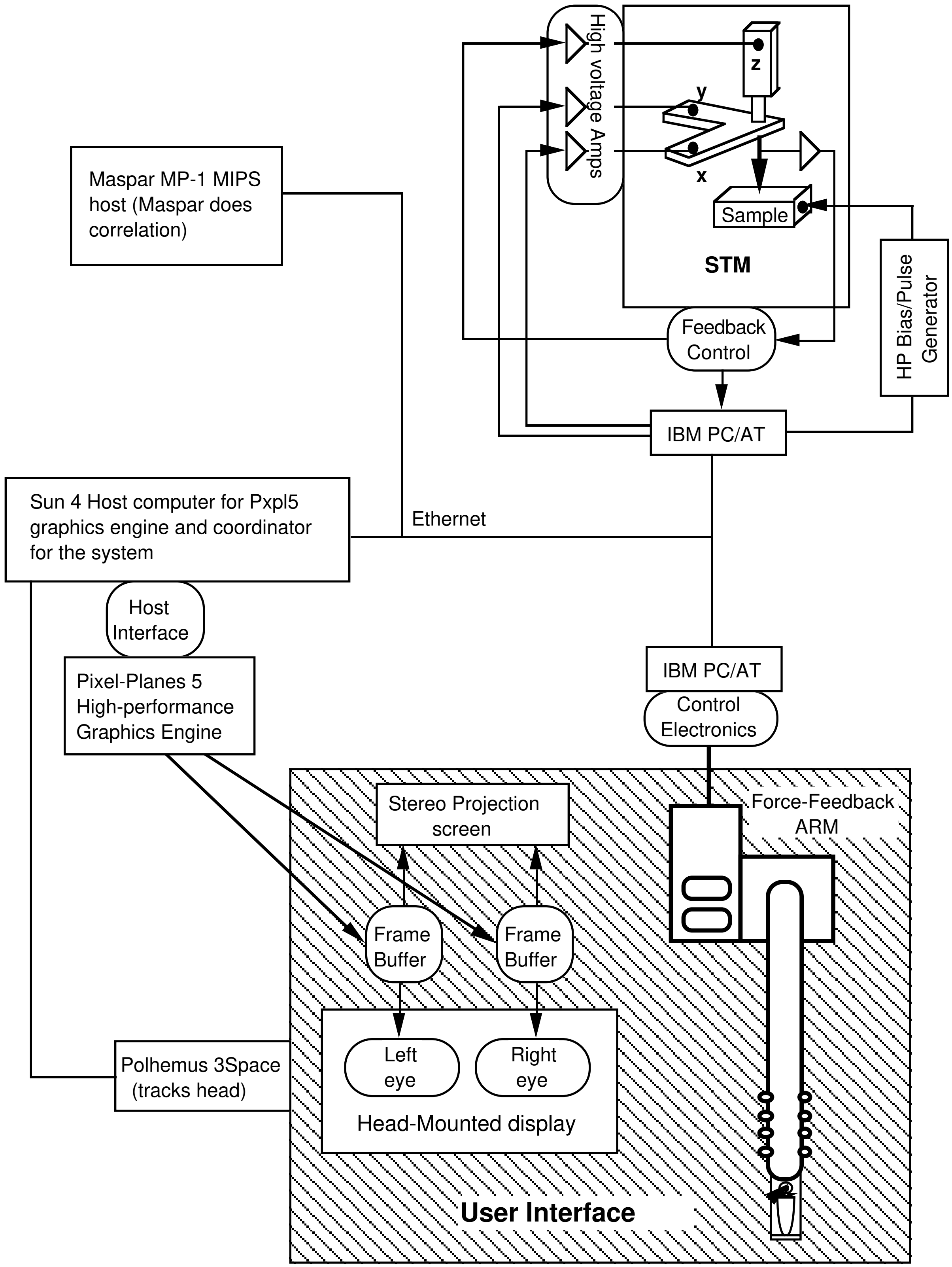

Figure 2. Hardware system diagram of the proposed Nanomanipulator system.

The Nanomanipulator system will comprise several parts, including a massively parallel image generation subsystem, a real-time control subsystem, a massively parallel correlation engine, and a user interface subsystem. Due to the varying requirements of the different system parts, this will require using a distributed set of heterogeneous processors. These subsystems will be communicating sequential processes that talk to each other over a Local Area Network (LAN). This means that communication and synchronization between processes are major parts of the overall system design.

For the purposes of this dissertation, the STM is modelled as a black box with X, Y, bias, and pulse inputs; Z and tunnelling current outputs. There may also be additional voltage-controlled inputs and output added to the system. If so, they will also be handled by the system and controllable by the user.

The STM must be further characterized in terms of its effect on the surface and its sources of error. While these parameters do not have to be optimized by the system, they certainly have to be known in order to demonstrate that the instrument is behaving as advertised. In particular, the behavior of the tip-sample interaction must be well understood to understand the effect that the user is having on the surface.

The user's "presence" on the surface is modulated by the transfer functions of the STM, sampling, smoothing, and HMD system. These have effects both going from the surface to the user and from the user back down to the surface. Going from the tip to the user the feedback circuitry affects the closeness of the tip tracking; the coarseness of sampling and reconstruction causes both smoothing and aliasing. Going from the user to the tip the DAC granularity and piezo hysteresis affect the actual position of the tip with respect to the position indicated by the user.

3.1 User Interface

The user interface section comprises several existing hardware and software systems. It makes extensive use of work from the Head Mounted Display group, the GRIP/ARM group, and the Pixel-Planes group.

3.1.1 HMD

The current HMD that we use is made by Virtual Research. This is a stereo color display helmet that fits snugly on the head and is tracked by a Polhemus tracker. The HMD group has written a software library to handle the interface between the Polhemus and user code.

The HMD group has also built several software libraries for creating virtual environments. These libraries use the tracker and PPHIGS libraries to handle the interface to the tracking and display devices. They also encapsulate the geometry of the display helmet. On top of this are libraries to provide control panels and menu systems. There is also a sound library that generates sounds for the user via a Macintosh computer. The remote connection to the Macintosh is via a network-transparent device library that I wrote while I was on the HMD team.

The images that are generated for the HMD are produced by Pixel-Planes 5 (Pxpl5.) This is a massively parallel graphics engine that was developed at UNC-CH under the direction of Henry Fuchs and John Poulton. In addition to the hardware itself, the Pixel Planes team has created the PPHIGS library as a programmer's interface to the machine. This is a graphics library modelled after the PHIGS standard.

It will also be possible to view the data on a high resolution rear-screen projector located right in front of the ARM. It is currently possible to see the view from either eye. It may later be possible to view in stereo on that screen as well (by duplicating the image on the high-resolution frame buffer and using a CrystalEyes device to do frame-sequential stereo).

3.1.2 ARM

The ARM is an Argonne III Remote Manipulator that is presently interfaced to an IRIS 4D GTX workstation via A/D and D/A cards. It will be migrated to a PC controller during the course of this investigation.

The software interface to the ARM is via a software library that I wrote while I was on the ARM project. I have integrated this library into the HMD tracker library so that application code can treat the ARM as a normal tracker that has the additional capability of generating forces.

3.2 STM Control

The interface to the STM will be via A/D and D/A convertors on a 486 - based personal computer. Controls will be provided for X and Y positioning and data will be available from the Z position and tunnelling current. In addition, a Hewlett-Packard pulse generator will be connected via its HPIB bus to allow control of the tip bias and pulse characteristics. The software to drive the A/D and D/A converters will have to be written, but an existing system that I wrote for the IRIS can hopefully be ported.

If improved electronics are made for the STM, the system will provide control for the desired current set point and any other voltage-controlled parameters as well.

3.3 Drift correction and correlation

There are two phenomena that cause gross changes in the position of the surface from the point of view of the STM. The first is thermal expansion of the materials that hold the sample and tip in position. The resulting drift is fairly slow and regular, measured against the time period of a surface scan. The second phenomenon is tip changes as a result of applied voltage pulses. These changes are of unpredictable size and direction and will happen whenever we are trying to perform surface modifications. In order to offset these changes, correlation between successive frames must be performed.

The drift correction/correlation system receives sample and time data from STM module and computes a model of the drift in the system and of the sudden changes. It then informs the user interface and control modules of drift parameters for compensation.

We have in the department a MasPar MP-1 massively parallel computer that can be used to compute correlation for the large changes due to tip modification. Kalman filtering may be used for tracking the thermal drift given the drift as a vector. It may be possible to use local differences rather than full correlation to compute the drift.

I will need further study in correlation techniques and the magnitude of the drift before I can determine exactly what will be required.

3.4 Communication

The several host computers to be used in the system are all connected to our building Ethernet LAN, which provides about a 1 Mbit/second sustained path between hosts on the network. This bandwidth is adequate to handle the data coming from the STM, which is limited in scan rate by the electronics and piezoelectric crystal resonances. It is also adequate to pass the correlation parameters as they are computed. The current segmentation of the problem seems to fit well in the networked configuration.

In order to establish the connections between the hosts, I will use the fourth implementation of network software that I initially wrote for a distributed systems class. The system has been optimized for fast connection times and low latency as part of several projects in the past. Measurements of data communications on the resulting connections have shown about a 5 ms delay for data traversing the network rather than being read from a local data connection.

Conversion of binary data from one architecture to another will be done by a common-format conversion library I wrote as part of the ARM project. This library has been successfully used by several projects on a variety of architectures (including Decstation, VAX, SPARC, IRIS) and has proven to be both reliable and far faster than ascii data transfer. The floating point format is the hard part, as the "network standard" integer format conversion routines already exist.

4 Why is the proposed system better?

This system is better than current systems because it puts the scientist **on** the surface **in** control **while** things are happening. Each of these three parts is important to performing Surface Science experiments.

4.1 HMD (on the surface)

The Head-Mounted Display (HMD) that is part of the system provides several improvements over conventional viewing techniques. The first is that it provides a view of the surface as a surface, rather than as a sequence of curves. It provides a Phong-shaded color surface that is also colored smoothly according to height or other information.

In addition, the fact that there are two displays (one per eye) allows the scene to be rendered in stereo, so that the surface looks like it has actual depth. It does this in much the same way "Viewmaster" devices do, by presenting a picture from slightly different points of view to each eye. The resolution of current HMDs is much lower than that of conventional CRTs, but we hope to have improved displays by the end of this year. Note that it is also possible to view the scene on the hi resolution projection screen when it is important to see details. This view is presented directly in front of the user.

Because the user's head is tracked , the viewpoint can be changed as the user moves around. This is a very powerful tool in allowing the user to explore the surface since it is possible to change the viewpoint in a natural fashion, simply by moving around.

When R. Stanley Williams tried out the Nanomanipulator prototype for the first time, he exclaimed "I put it on and immediately it's there - a 3D object!" This expresses the power of the HMD as a viewing system - rather than interacting with a device to view an object, the HMD allows a user to seemingly interact with the object directly.

Referring to his videotaped images of the scanned surface, F. Besenbacher writes "...one can record *STM movies* and thereby visualize in real time and space dynamical processes on metal and semiconductor surfaces. Such information, which cannot be obtained by any other means, is very decisive for a full understanding of both the growth mode of reconstructed phases and the resulting static structure." This indicates the usefulness of a display that can show "real time" evolution of the surface. It would also be possible to capture frames of data and later show accelerated- or retarded- time evolution of the surface.

4.2 ARM (in control)

The force-feedback ARM acts as the user's hand in the virtual world. The ARM is fitted with switches that can be used to activate menus and specify locations and with dials that let the user specify analog parameters to the system. These are used to provide a user interface that lets the user control scanning locations, pulse width/height, and other system parameters.

In addition, the ARM is capable of generating forces for the user to feel. This allows the user to actually feel where the surface is, as if he had a tiny screwdriver that he could poke on the real surface. This provides a haptic sense of where the surface is in space and enhances the feeling that there is an actual surface present. Ming Ouh-young showed that the addition of force feedback to a system can decrease task completion time by up to a factor of two.

By providing a software feedback loop that includes the user, the Nanomanipulator will allow direct tip placement by the scientist. The user can position the tip at an arbitrary location in x and y and feel a force that represents the height of the surface at the given location, thus using the tip as a probe to feel around on the surface under study.

4.3 Real time (while the experiment is happening)

Since all of the above interaction is possible while the surface is still under the microscope and being scanned, the user can direct the study of the surface based on data just obtained. If part of the surface looks more interesting and merits further study, the user can specify that region as the active region and get faster scanning there. If there is some question whether an area contains a certain feature or if it contains only noise, it can be rapidly scanned multiple times to see if the feature is robust. As Dr. Joe Lyding said of his interactive system, "I used to look at the computer data later and say 'I wish I had done that.'"

While attempting to make changes to a surface, the user can specify a small area around the place they are attempting to modify and have the system scan it repeatedly while modifications are made. This means that the user can react to mistakes in the modifications as they are occuring and can immediately see when they have been completed. This should dramatically decrease the task completion time.

In other systems, where the user cannot see a clear model of the surface until after all the data is collected, an algorithm must be run to obtain all of the desired data and perform the modifications. Allowing the user into the loop means that they can apply their pattern recognition, intelligence, and knowledge of the system under study when it is needed - while the modifications are occuring.

One reason that time is so critical to task completion is that the surface will drift away if it takes too long to figure out what to do, due to thermal drift in the STM. Since the areas under study are so small (2000 Angstroms on a side, maximum) even small movements due to temperature changes can move the area under study out of the range of the microscope in a fairly short time. This means that it is crucial to give the scientist maximum power to act during this time.

5 Why is this an interesting Computer Science problem?

This system is interesting because it pushes the frontiers of several areas of computer science and attempts to link them all together with each other and the real world.

Generating an image of the surface in real time for the user requires an image generation system that is capable of generating 128x128x2 triangles to tessellate the surface. These triangles must be shown to 2 eyes, and the scene must be updated about 20 times per second. To allow surface reflections and coloring information, these triangles must be color-interpolated between vertices and Phong shaded. In addition, it must be possible to move up to 500 of the vertices per frame to allow for new data coming from the STM. On top of that, it is important that the system have low lag. This pushes the limit of existing graphics engines and results in a required display rate of 1.3 million triangles per second for the surface alone, not counting the models that are used for interaction with the system.

Since tip changes due to voltage pulses cause the surface to move about unpredictably with respect to the tip, the surface itself must be tracked. The only thing we can use to track the surface at this scale is the tip itself, so information must be derived from successive scans of the surface. This will be done through the use of correlation of successive scans. Optimum system performance demands that these studies be performed on a grid of 128x128 samples in under ten seconds. This will push the limits of correlation algorithms and may require the use of massively parallel computing, because the correlation of two NxN surfaces is order N^4. For 128x128 surface, on the order of 300 million operations will be required. Using the 4096 processors in the MP-1 reduces this to 65K per processor.

In order to maintain proper control of the STM as it is scanning and to maintain a record of when each scan point was taken to allow drift studies, the process that controls the STM must be a soft real-time control process ("soft" because actions must be taken within specified time limits but disastrous results do not occur if a deadline is missed). It must ramp the voltages on the X and Y control lines at the proper rates to raster the image and then scan the height at that location at a known, specified time with respect to the movement. The results of the scans must also be sent to the user-interface process in a timely manner. In addition, in "tip control" mode, the system must become synchronously locked with the user interface process and also provide a fast update rate for the user, while at the same time insuring the proper maximum scan rate on the STM itself.

The tasks of presenting the user with all the possible information needed in a low-resolution head-mounted display and of getting all the command information required without the use of a keyboard are formidable user interface problems. The user wants to see the surface itself, pulse profiles, length measurements, and other information about the system. The user also needs to be able to specify operating modes, positions and areas of interest, and system parameters - all from within the virtual world. This will require a well designed and powerful user interface.

The work described above is too much for a single computing engine. Fortunately, there are several places where the problem can be split, thus allowing the problem to be handled by several asynchronous or synchronous processes communicating over low bandwidth (Ethernet) connections. The different processes will run on heterogeneous architectures, which raises the issues of data transfer from one host to another. The system will need to operate on distributed, heterogeneous processors all working together via message passing. This raises problems of partitioning the problem into pieces and running the various parts on suitable machines, as well as determining the minimum data flow between distributed parts of the application and dealing with synchronization issues. Fault tolerance and gentle failure modes will also be required from the system to prevent damage to the STM tip.

Possibly the most compelling reason this system is so interesting is that it gives the user the power to modify the real world from within a virtual world. This requires that the virtual world be carefully registered with the real world at all times and that the model presented to the user be consistent with the external environment. The purpose of this system is to transport the user into a nanoworld and to provide a means for the user to make changes in this environment. It seeks to create teleoperation at the atomic scale.

6 What I'll Show and how I'll know

The current models of the surfaces under study are the line drawing and the static view of a tessellated surface. I believe that the dynamic, real time viewpoint offered by this system will provide additional insight to the scientist that was not available from earlier visualization methods. This will imply that a virtual world is a more effective model of the surface than earlier methods.

If the virtual world provides insight to the surface scientist that was not found by other visualization methods, then this method is more effective than previous methods. This means that the scientist must discover surface features or behaviors that are scientifically useful and verifiable and not simply artifacts of the visualization method (ie: they must be both real and useful). This is still a subjective criterion that is subject to influence by how the scientist views the system.

I believe this system will also show that a virtual environment is an effective method of closing the system feedback loop around the scientist. This embedding of the scientist in the experiment will allow him to perform planned modifications to the surface in real time. This implies that a virtual world can be an effective part of a feedback system that allows a user to control real objects, in particular the surface under study. Since I am using a model of the STM as a black box with voltage inputs and outputs, this method should generalize quite easily to other instrumentation, especially instrumentation that generates grids of height information.

According to Stan Williams, “As yet, there is still not even a universally accepted theory of how the STM perturbs a surface as it collects an image, although everyone recognizes that it does. The underlying mechanism for how and why an atom can be foreced to jump from a surface to the tip or vice versa is totally controversial at the present time. These are issues that will take many years to straighten out. It is not reasonable to expect you to make a large advance in the physics of STM. ... there are a lot of detailed experiments that still need to be done that are not presently being done. This is because they are so tedious with the present types of human interface that we have.” “The state of the art now is such that if you can make almost any type of modification to any type of surface under direct control, i.e. manually writing on a surface with an STM, you have achieved something no one else has done and also made a significant advance to the state of the art. In that regard, modifications to graphite or molybdenum disulfide are quite acceptable. However, to really push the envelope I would like to try a metal such as platinum silicide or nickel disilicide... The way to demonstrate that you have succeeded is to have before and after STM images of the surface proving you have made a change to the surface... Also, making several such features to demonstrate how reproducible the technique is is always good.”

If user can apply the planned changes to a surface, this means I have shown that the system can indeed allow planned changes to a real world object. If the changes cannot be made reliably, then I will examine the reasons for the failure and explore further methods of overcoming the problems and the costs associated with the solutions. The test for system success will be planned modifications of graphite or molybdenum disulfide. Of course, I will be more pleased if we can perform modifications on the metallic surfaces as well.

7 Problems and Solutions

7.1 Control of the STM

In order to measure 0.01 Angstrom resolution over a scan area of 2000 Angstroms, we would need to be able to measure features to 18 bits of precision and within 1/2 LSB accuracy. I will embed a PC controller with 16 bit A/D converters to allow us to sense the required resolution on a 500 Angstrom scale. This is a significant improvement over the 12 bit converter we are presently using.

In order to be able to properly correlate the thermal drift, it is necessary to know precisely when each sample was taken. It is also necessary to move the tip at carefully controlled rates to avoid overtaxing the feedback control. These both require a fine-grain clock, which is not normally available on a PC/AT. I intend to make use of the secondary timer chip on the PC to drive an interrupt routine to maintain the finer timing information. If this is not sufficient, then it may be possible to purchase or build a higher-resolution clock card to put in the PC that will give more precise time information.

When the user is controlling the tip directly, it is necessary for the STM control to be inside the user-interface feedback loop. This requires very fast response to position requests in order to maintain a 20 Hz update rate for the user. The current solution is to allow 1 frame of lag between sampling the surface and responding to the user. The embedded PC and faster feedback electronics may allow us to remove this lag. In addition, it may be possible to micro-raster the tip around the last sampled point and be able to provide the information without having to scan the tip again.

In order to effect controlled material transport at the atomic scale, we will need to be able to generate pulses on the order of 10 nanoseconds in duration. The pulse duration, amplitude, and shape must be repeatable so that we can have control over the electric field imposed by the pulse and perform experiments on the effects of these parameters on material transport. Once the pulse has been created, it must also be transferred over some sort of transmission line to the tip/sample junction without distortion. I am working on an interface between the system and an HP pulse generator that will create the required pulses. Vern Chi and I have installed a new low-capacitance sample holder and a coaxial transmission line to reduce the pulse degradation. There is still an issue of how to model the impedance from the tip to the pre-amp as part of the pulse transfer.

The tunnelling current we are measuring is on the order of 1 nA. This makes our instrument very susceptible to electrical noise, either coupled capacitively to the signal lines or electromagnetically. We also have the problem of driving the analog signals over medium distances back to the A/D converters on the backplane of the IRIS workstation, which is not the cleanest electrical environment. The proper solution to this problem is to place the A/D conversion as close as possible to the signal source. Towards this end, I will install a PC as an embedded controller near the electronics and then ship data to and from the PC digitally over an Ethernet connection.

7.2 Thermal Drift

Small changes in the temperature of the STM as it is operating cause small thermal expansion in its various parts. Since we are examining surface features as small as atoms, these expansions can have drastic effects on our data. The expansion makes it appear that the surface is drifting around slowly with time. This problem shows up two ways. The first is as a shear in the data as we are taking it and the second is that the surface tends to drift off while the tip is being used for modifications. This is a problem because the tip cannot be used both for scanning and for manipulation at the same time. Since we present the user with a model of the surface that is intended to match the real surface, this means that we must somehow compensate for the drift.

The proper solution to the problem of drift is to purchase a more modern STM that has thermal drift compensation. STMs on the market today can provide drifts of as little as 1 Angstrom per hour. If this solution is not possible, then there are other approaches to a solution, listed below.

One method that is obvious will be to insulate the STM environment thermally. This will hopefully decrease temperature gradients in the air around the instrument, although it will not affect heat flow in the instrument itself.

The primary method for controlling the drift will be to attempt to characterize it and then offset it in software. In order to characterize the drift, several scans of the surface will be taken and the time of sampling carefully determined. Then, a correlation study will be run on the MasPar to determine how much the surface drifted between scans. I hope that the drift rate is linear enough to be modelled by interpolation, but any polynomial evaluation could be used. Once the drift rate has been characterized, data points can be modified as they are taken to match those in the canonical frame of reference.

Once the Nanomanipulator is capable of making controlled modifications to the surface, it will be possible to make fiducial marks, which should make the correlation study go much faster, as they will be known landmarks to navigate by.

7.3 User Interface

7.3.1 Information Flow

While the image of the surface itself is a high-bandwidth channel to the user, there are other types of information that the user needs to know. The user wants to be able to control the state and parameters of the environment and send commands to the STM controller while immersed in the virtual environment. The severely limited resolution of the HMD and the restricted working volume of the ARM will act to restrict the choices here.

Direct manipulation has proven to be a powerful metaphor for dealing with objects in a virtual world. Using the ARM to manipulate the surface, tip, and other objects in the environment will allow the specification of locations, distances, and actions. Work here with particle systems code and at Brown university (one of our collaborators in the Gang of Five) has indicated that placing control "handles" on objects can allow a natural and powerful method of control over various parameters.

There are some things that do not map easily into direct manipulation, and taken to the extreme it can result in a plethora of confusing handle objects littering the space. In order to deal with the complexity of handles, a menu system to control mode changes, etc will be placed in the environment. Control panels will be available from the menu to allow varying parameters that do not have a natural handle, such as the tip bias voltage and pulse duration. Some of these quantities can also be mapped to dials that are present on the ARM itself.

Since there is extremely limited display resolution in the HMD, it will not be possible for all of the information required by the user to be present at all times on the screen (such as the waveforms of pulses if we can measure them and cross sections of the surface itself). There are two attacks to make on this problem. The first is to place the different information at different locations in the virtual world and have the user select among them by looking in different directions. Another tactic is to only have some of the information up at any given time and to allow choosing the currently displayed information as different display modes (exploration mode, measurement mode, manipulation mode.) The data from the current mode can be displayed as a heads-up display in the user's eye space. Another option would be to have a support person to help with the interface, in a "mission control" capacity. They can answer questions and modify settings, acting as a "man behind the curtain." A keyboard interface will thus be provided to allow this to happen.

Yet another problem is to determine what kinds of information the scientist will want to record for later perusal and how best to capture that information during the session. The scientist will probably want snapshots of the surface, cross-sections of the surface, measurement information about features, and information about pulses and their locations on the surface. He will probably also want video segments with voiceover notes. These capabilities must all be taken into account in the system user interface design, but they are not in the domain of this dissertation.

7.3.2 Update Rate

Past experience with HMDs has shown that HMD tasks become more difficult at update rates below about 20 Hz. Although this is closely tied to lag, for me, the system ceases to be an environment in which I am immersed and instead becomes a display that the I am wearing below about 20 Hz. This update rate must be maintained to both eyes of the HMD, requiring that the scene be drawn twice each 1/20th of a second.

I plan to achieve the required update rate through the use of a three-rack Pxpl5 system. This should allow display of a grid that has 128 points on a side at the required rates. I estimate that this rate is achievable using the Pphigs display-list library, which is optimized for drawing triangles. A two-rack system has demonstrated the required rates when drawing a grid of 100x100 samples.

In addition to display of the scene, new height data is arriving from the STM at a rate of about 500 samples per second, which is 25 data points per frame. Since each data point affects the normals at the 8 surrounding points, and each normal affects the 8 surrounding triangles, this means that the system must be able to perform on the order of 1000 triangle updates per frame while still maintaining the 20 Hz update rate. Since the Pphigs library is display-list oriented, it is not fast enough to make changes to many triangles per frame. This is made especially problematic since all changes to the display list take place on the sun4 host and are then sent via the VME bus to the display engine. This makes Pphigs completely inadequate to handle the required number of triangle changes that must occur each frame.

In order to handle the high volume of changes to the data set, I will write code that executes on the Graphics Processors (GPs.) This will allow the work to be distributed among the processors, with each doing a small part of the work. In addition, only the point update information has to flow from the host to Pxpl5 over the VME bus. The rest of the work is done locally on each GP. In order to maintain the display rate, it will be necessary to build an extra display list on each GP and to hand this list to Pphigs to be displayed. This requires propagation of point changes into the display list, with each point going to the several triangles that share it as a vertex.

7.3.3 Lag

Lag in any feedback system can cause overshoot and oscillations, and the Nanomanipulator system is no exception. There is lag in several parts of the system, but it is largest in the tracking and image generation subsystems. This has been and continues to be a large problem in HMD applications.

The good news is that since it is such a problem, both the Pxpl team (Just In Time Pixels, Mega) and the HMD team (FastTrack Polhemus) have initiated projects aimed at studying and reducing the lag in various system components. The ARM group is also studying the possibility of putting the ARM device into a mode that allows it to report positions without being asked, thus saving one network transmission. In addition to this work, it is possible that some sort of predictive tracking can be used in the system to attempt to minimize the effects of lag.

8 Plan of Attack

8.1 Building the system

I have already begun work on a prototype system under the name Microscape. This has given me insight into the problems I will encounter with the Nanomanipulator and has pointed towards some possible solutions, which are presented above. I intend to use the prototype system as a testbed for additions and improvements to the system, and then to do a total redesign for the Nanomanipulator. There are several subsystems that I have not yet begun work on, and these are described below.

Several system requirements point to the use of an embedded PC as the STM controller. I have travelled to the University of Illinois at Urbana-Champaign to study the controller that they use, and I plan to have an embedded controller for the ARM in place here by October 30. Using the knowledge gained from this installation, I intend to have a PC-based STM controller in place by November 21.

In order to design the correlation subsystem, I will have to study filtering and image processing, and then correlation proper. I have begun independent study of image processing concepts using Dr. Coggins' book. Once this is completed, I will need to do research into correlation algorithms proper and finally implement the system. I plan to have completed reading the book by October 30 and will then estimate completion for this phase of the project.

I plan to begin designing the Nanomanipulator system in early December and hope to have a working base system up by early January, with modules added as needed thereafter. In light of upcoming hardware moves and upgrades, as well as my course load, I anticipate that this date may slip well into the Spring semester.

8.2 Oral Examination

Having only recently decided to work with both an STM and a feedback control system for my dissertation, I find myself ill prepared for an examination on related topics at this time. To remedy this, I have undertaken independent study of control theory and correlation algorithms and am auditing a course on Quantum Mechanics. I will then try to take a course on control theory in the Spring semester, either at UNC or a nearby university. I will need to wait until I have completed much of this study before undertaking my oral examination, although Fred Brooks and Gary Bishop have suggested that the committee recommend further study in control systems rather than testing me in that area. Therefore I propose taking the examination at the beginning of the Spring semester, 1993. I will be in touch with the committee members individually during the Fall 1992 semester to discuss appropriate areas of questioning.

8.3 Experiments

Once the Nanomanipulator system has been completed to the point that it can control surface modification, the modification tasks will be attempted. This should occur sometime during the Spring 1993 semester. The exact modification tasks will be determined while Stan is here in the early Spring after he has been exposed to the system and we have talked about some possible work. One possibility is the formation of electron interference transistors.

8.4 Dissertation

Writing of the dissertation will proceed in parallel with the above work. The research and system sections will be started when the work that they describe has been done. The background section and overall formatting can be started immediately.

References

Besenbacher, F., F. Jensen, E. Lægsgaard, K. Mortensen, and I. Stensgaard, "Visualization of the dynamics in surface reconstructions," Journal of Vacuum Science Technology. B 9 (2), Mar/Apr 1991, pp. 874-877.

Binnig, G. and H. Rohrer, "Scanning Tunnelling Microscopy", Helvetica Pysica Acta. Vol 55 (1982), pp 726-735.

Coggins, James, Image Pattern Recognition: A Unified Approach with Multidisciplinary Applications, Pre-print copy.

Driscoll, Robert J., Michael G. Youngquist, and John D. Baldeschwieler, "Atomic-scale imaging of DNA using scanning tunnelling microscopy," Nature, Vol 346, 19 July 1990, pp. 294-296.

Edstrom, Ronald D. and Maria A. Miller "Scanning tunnelling microscopy and atomic force microscopy visualization of the components of the skeletal muscle glycogenolytic complex," Journal of Vacuum Science Technology, B 9 (2), Mar/Apr 1991, pp. 1248-1252.

Holloway, Richard, Henry Fuchs, and Warren Robinett, "Virtual-Worlds Research at the University of North Carolina at Chapel Hill," Proc. Computer Graphics '91, London.

Kobayashi, A., F. Grey, R. S. Williams, and M. Aono, "Mechanism of Nanometer-scale Silicon Trench Formation by a Scanning Tunnelling Microscope", (Private copy)

Komuro, M., S. Okayama, O. Kitamura, W. Mizutani, H. Tokumoto, and K. Kajimura, "Nanometer Structure Fabricated by FIB and its Observation by STM," Microelectronix Engineering 6 (1978), pp. 343-348.

Magonov, S. N., G. Bar, E. Keller, E. B. Yagubskii, and H. J. Cantow, "Atomic scale surface studies of conductive organic compounds," Synthetic Metals, 40 (1991), pp. 247-256.

Mamin, H. J., S. Chiang, H. Birj, P. H. Guethner, and D. Rugar, "Gold deposition from a scanning tunnelling microscope tip", J. Vac. Sci. Technol. B 9 (2), Mar/Apr 1991, pp 1398-1402.

Ouh-young, Ming, "Force Display In Molecular Docking," University of North Carolina at Chapel Hill, TR90-004, February, 1990.

Snyder, Eric J., Mark S. Anderson, William M. Tong, R. Stanley Williams, Samir J. Anz, Marcos M. Alvarez, Yves Rubin, François N. Diederich, and Robert L. Whetten, "Atomic Force Microscope Studies of Fullerene Films: Highly Stable C60 fcc (311) Free Surfaces," Science, Volume 253, 12 July 1991, pp. 171-173.

Stoll, E. P., "Picture processing and three-dimensional visualization of data from scanning tunnelling and atomic force microscopy," IBM Journal of Research and Development, Vol. 35, Nu. 1/2. January/March 1991, pp. 67-77.