TR05-021
September 30, 2005

Facetop on the Tablet PC: *Assistive technology in support of classroom note-taking for hearing impaired students*

David Stotts, Gary Bishop, James Culp, Dorian Miller, Karl Gyllstrom, Keith Lee

1. Background

Facetop is a collaboration tool conceived and implemented by University of North Carolina researchers led by Dr. David Stotts. It is a tool that combines video streams from two geographically separated users' computers into a translucent overlay on each user's desktop. One user's desktop is then shared, and the two users may then collaborate on work while being able to see each participant on their respective screens.

The video overlay showing both users has been highly successful in conveying a sense of presence not found in traditional video-conferencing or other collaboration techniques. It enables them to point, gesture, and interact with the computer as if they were sitting side-by-side. This has been used to great effect in pair programming, and shows some promise as an aid to deaf users.

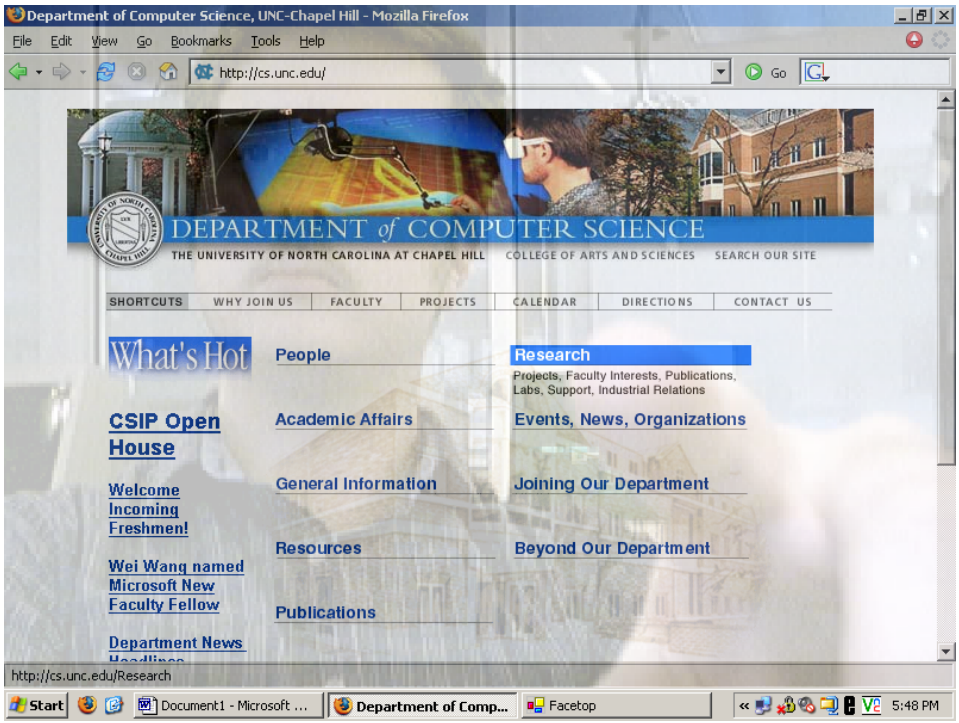

Figure 1: Windows implementation of Facetop

2. Current Full-function Implementation: Mac-based

Facetop was originally implemented on Mac OS X due to fundamental differences in OS X versus other common desktop operating systems. Thus, despite its effectiveness, it reaches a fairly small audience.

The system also has been primarily researched in the context of sharing text-based work, as the human brain has very little trouble distinguishing the moving, human-shaped images from static text. This does not necessarily address collaborative efforts at large since not all such tasks are textual.

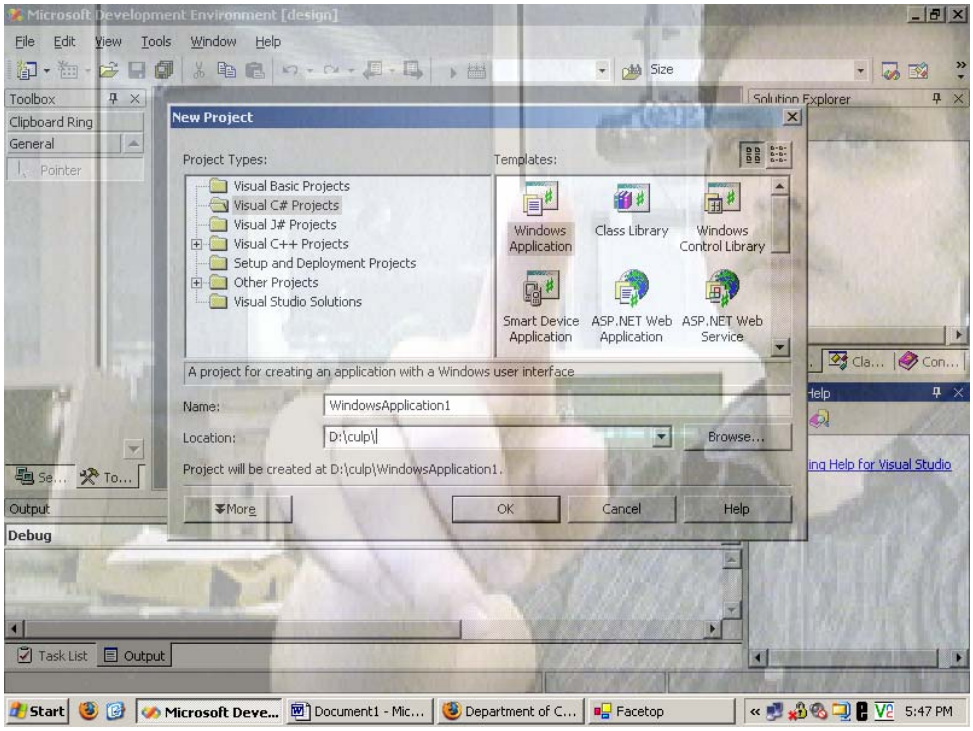

Figure 2: Windows version of Facetop, video slightly more opaque

3. Windows Facetop implementation

The immediate goal of this project is to create a Windows implementation of Facetop with an equivalent feature-set of the OS X version. This is motivated by two desires:

- To reach a larger audience, and
- To facilitate an identified need in medical imaging collaboration.

The motivation in the first reason is fairly obvious. Since more people use the Windows platform than the OS X platform, more people will be able to use Facetop if a Windows implementation exists. The impetus in the second reason is a medical group wishing to use Facetop technology to enable its doctors to collaborate on medical record examination -- records which are accessed solely on Windows terminals.

To date, Facetop has mostly been used to share textual data such as source code. The medical imaging project, however, entails sharing visual data showing human forms – including x-rays and other radiographs, sonograms, instrument charts, and other video streams. This is a very different usage environment from earlier Facetop experiments; with the non-textual data forms in the medical records, the translucent video overlay may interfere with the user's ability to distinguish between collaborating participants and patient data. Thus, researching this area is one of the key interests in developing Windows Facetop.

Current Status

Work on Windows Facetop has just started. Right now, single-user video overlay is implemented as well as desktop sharing. Video sharing and finger tracking are being developed. We are working in the .NET environment using C# to develop.

The Facetop video itself is rendered in a window by DirectShow. This window is the topmost on the screen and has variable translucency. All mouse events are being passed through this window using the WS_EX_TRANSPARENT extended window attribute. The video can be put in a more conventional window and moved and resized for uses other than full-desktop sharing.

Desktop sharing is being implemented using VNC. Video subtraction from the desktop is being done on the server side by ignoring alpha-blended windows.

Performance is good thus far. Latency is low and is appropriate for natural body language and gesturing. The main bottleneck is presently the Windows compositing layer putting the translucent video together with the desktop below.

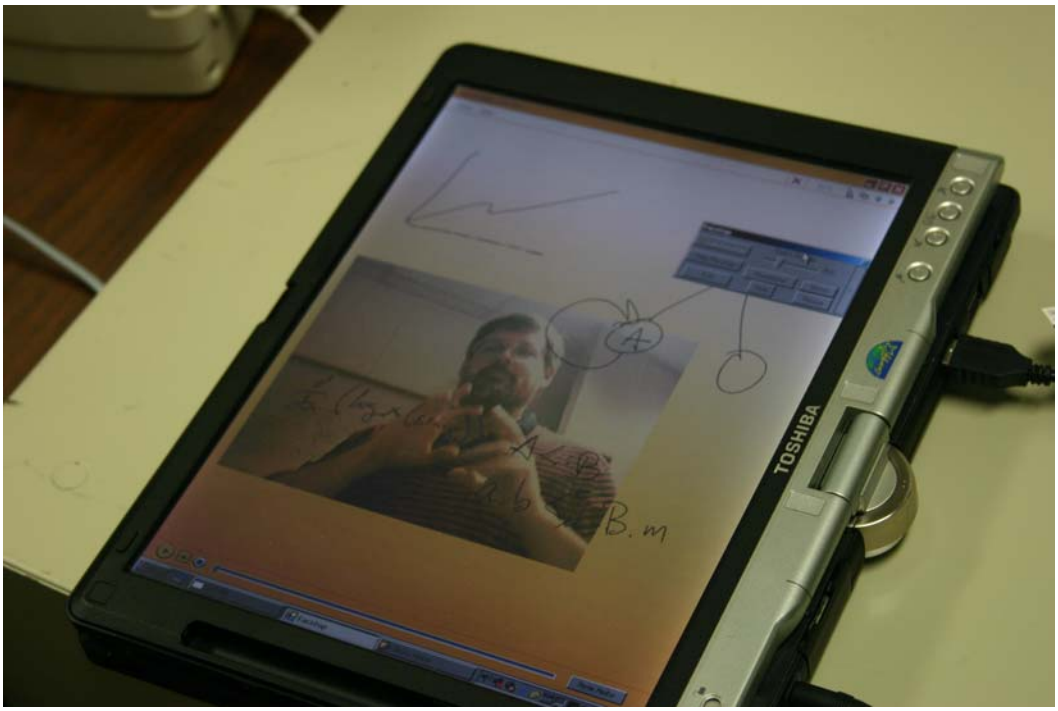

Figure 3: Tablet PC implementation of Facetop for hearing-impaired use

4. Tablet PC Facetop implementation

A secondary use of Facetop has emerged while developing the Windows implementation: its use as an assistive technology for the hearing impaired in the classroom or in other meeting environments where a signing interpreter is needed.

The widespread use of fast-paced visual aids (notably PowerPoint presentations) in many classrooms is a real problem for students trying to "hear" the lecture or discussion via signing interpreter. Those with full hearing really do not have much of a problem; they can follow the presentation by the professor visually and aurally, and still continue to listen if they need to divert their attention to paper for taking notes. Hearing impaired students, however, have their attention split thrice:

1. To the presentation,
2. To a sign-language interpreter to "hear" the professor, and
3. To paper or a computer for taking notes.

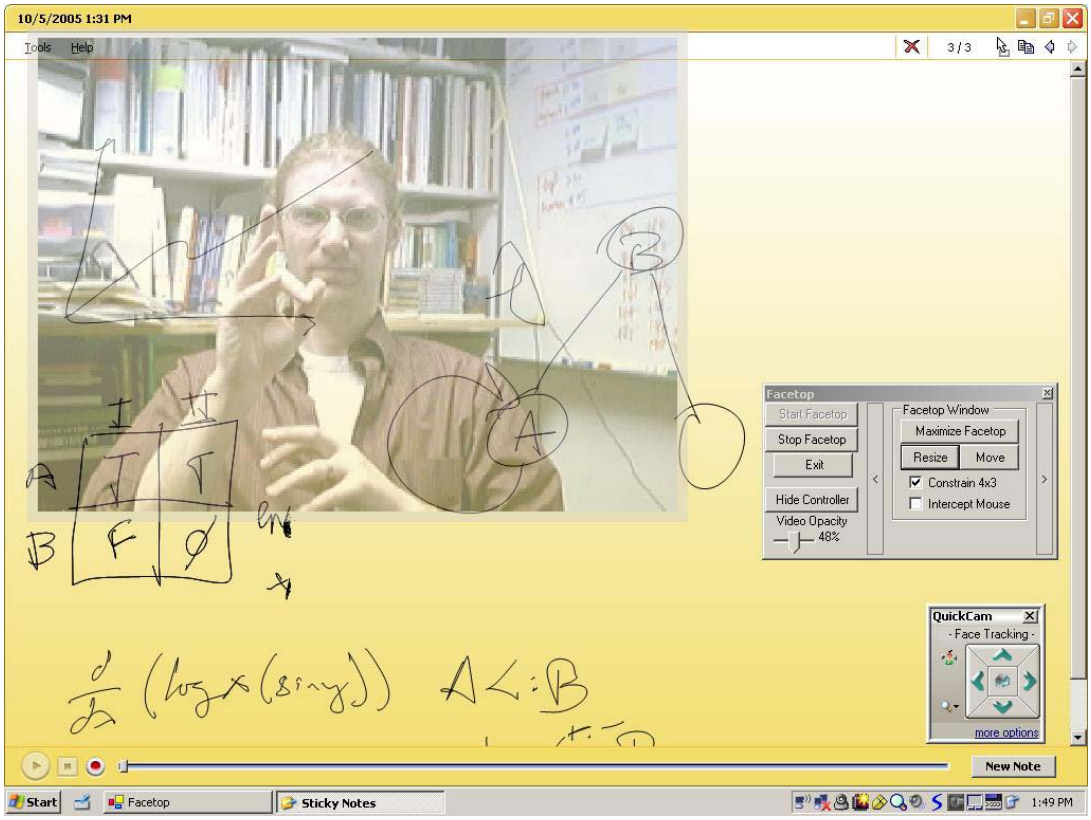

Figure 4: Screen shot showing components of the Tablet PC Facetop

Such a student must watch the screen for material, switch away and watch an interpreter to get details and explanation, and then switch away to write something down. The real problem is that last step – looking away from the interpreter -- which essentially cuts him or her off entirely from the class lecture and discussion.

We have an initial prototype implementation of a solution to this problem, show in figure 3. Our solution was suggested by a hearing impaired computer science major who participated in early usability trials of the collaborative Mac-based Facetop system. This student explained the basic problem to us, and suggested that Facetop provided the means for putting the torso of his signing interpreter right onto a note-taking platform.

We have adapted the single-user version of Windows Facetop and ported it to the Tablet PC. In the semi-transparent video window we show the interpreter (rather than the user self-image, as in single-user Facetop). Having the interpreter shown right on the tablet PC screen means a deaf student may simultaneously read an electronic copy of a presentation, watch the interpreter over top of it to get the class discussion information, and write notes directly on a digital ink application.

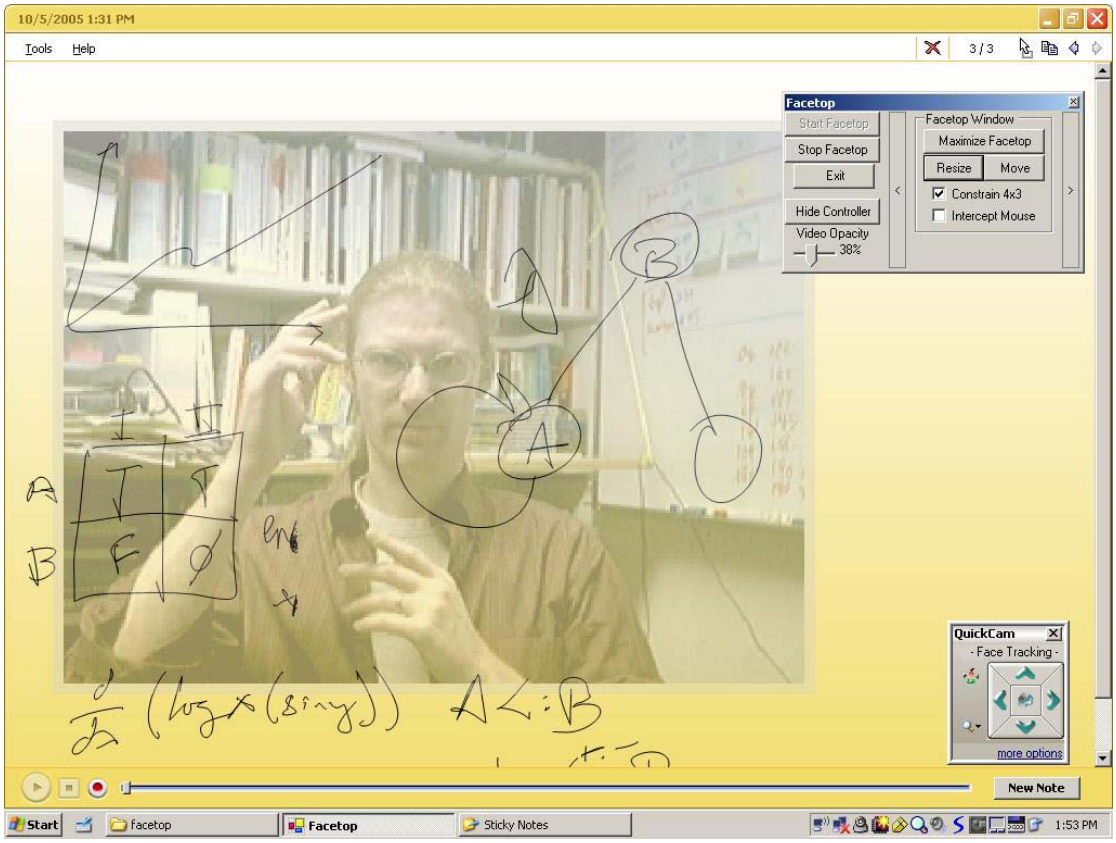

Figure 5: Tablet Facetop screen shot, showing larger video window

Implementation details

The details of this are similar to the normal Windows Facetop, but it is somewhat simpler due to no computer-to-computer video or desktop transport being needed. The setup is a tablet PC and a small camera facing an interpreter. In our current system, we have used a Toshiba Portege dating to about 2003, with a 30 fps USB video camera costing about $125.

The camera is placed near the signing interpreter, pointing at him/her. This image is translucently placed atop a portion of the tablet's screen. The current prototype has a control panel containing buttons for moving, resizing, and maximizing the video window, as well as varying the opacity of the video (see Figure 4). The user may use the stylus to move and resize this window and change

its opacity to meet his/her desires, and then use existing software to view the presentation and take notes.

It should also be possible to employ a second camera to capture any presentations or visual material which cannot be given to the student digitally ahead of class, such as chalkboard/whiteboard work or old transparencies.

Performance information at this stage is preliminary and is gathered from a 3-year old tablet platform. Note that is all the ensuing discussion, we expect some of the issue to be moot when executing on a current generation tablet.

Figure 6: Note taking on top of interpreter's semi-transparent video image

The video transparency is all done in software without the use of a graphics card or hardware accelerator. The expensive operation is compositing the camera image with some portion of the desktop. When the camera image window is the same size as the camera image (either 640x480, or 320x240 depending on driver settings), the compositing operation is fairly fast and we get a nice frame rate, approximately 20 fps. This is sufficient for understanding of signing motions without loss of fine rapid movements.

When the video window is scaled larger, performance degrades linearly to a point where the frame rate is too slow for good signing understanding. We seem to be able to get about 1.5 times the camera size before it becomes unsatisfactory. The performance degrades because the camera image is interpolated upwards to a larger pixel size, making more pixels to be handled in the compositing operation.

One solution to this slow performance (other than relying on newer, faster hardware) is to "burst apart" the camera image by spreading its pixels out regularly. Then instead of compositing the camera image with the desktop image, and doing alpha blending, we instead simply replace a desktop pixel with a camera pixel in all the appropriate places. Figure 7 shows this scheme. Our trials show that for a doubling of the camera image in both dimensions, we get no performance degradation. This is due to the fact that we are doing a simple pixel replacement rather than the more complex compositing operation, and doing that overwrite once for each pixel in the *original sized* camera image.

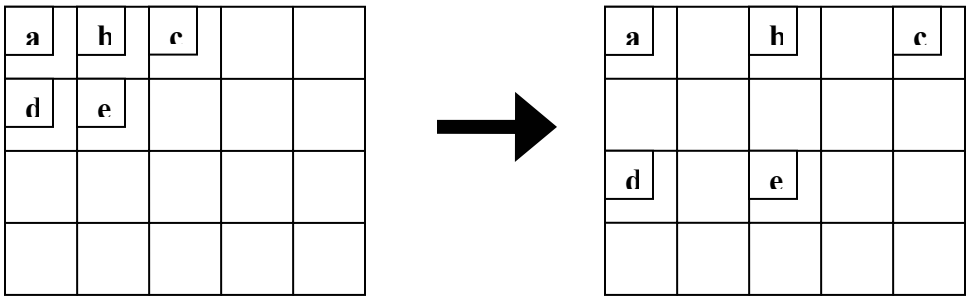

Figure 7: Screen-door transparency for larger image with no performance loss

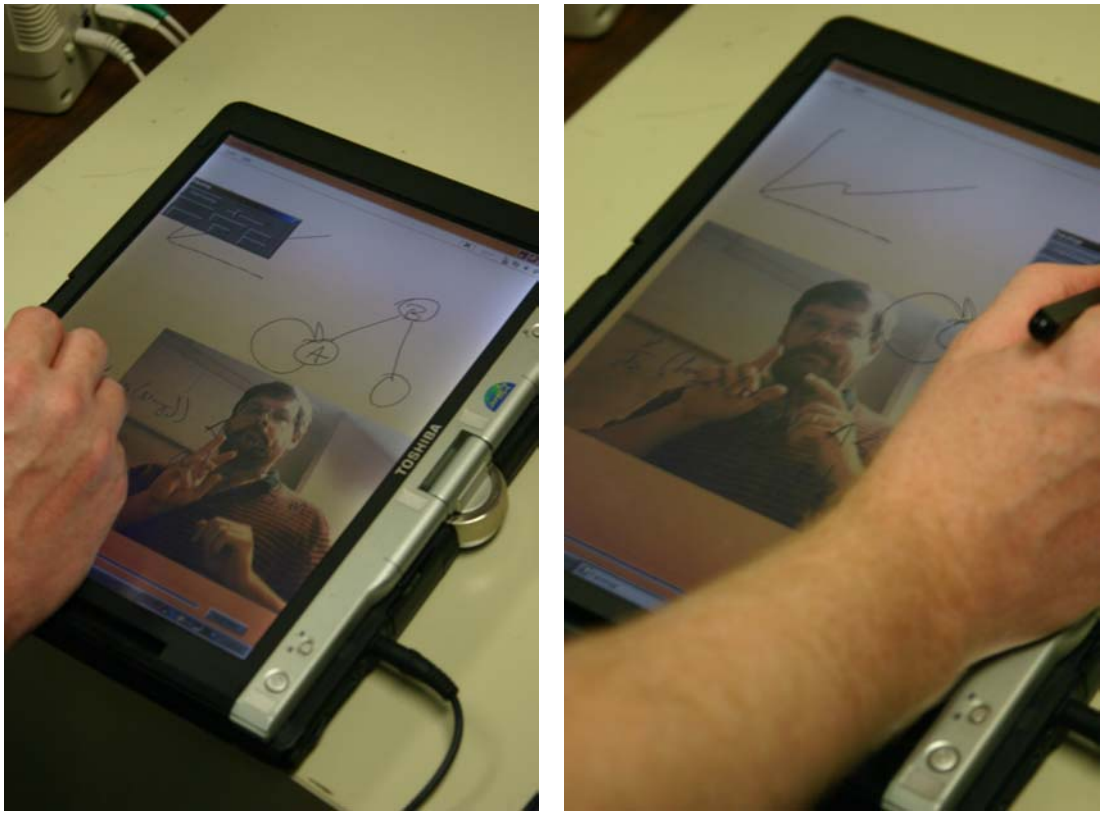

Figure 8: Moveable and re-sizeable video window

Discussion

Our solution should eliminate much of the attention switching that a hearing impaired user must do between sources of information. This, in turn, should enable him or her to get more information from each source and write any notes without missing anything.

We will be doing usability studies over the next few months to test this hypothesis.

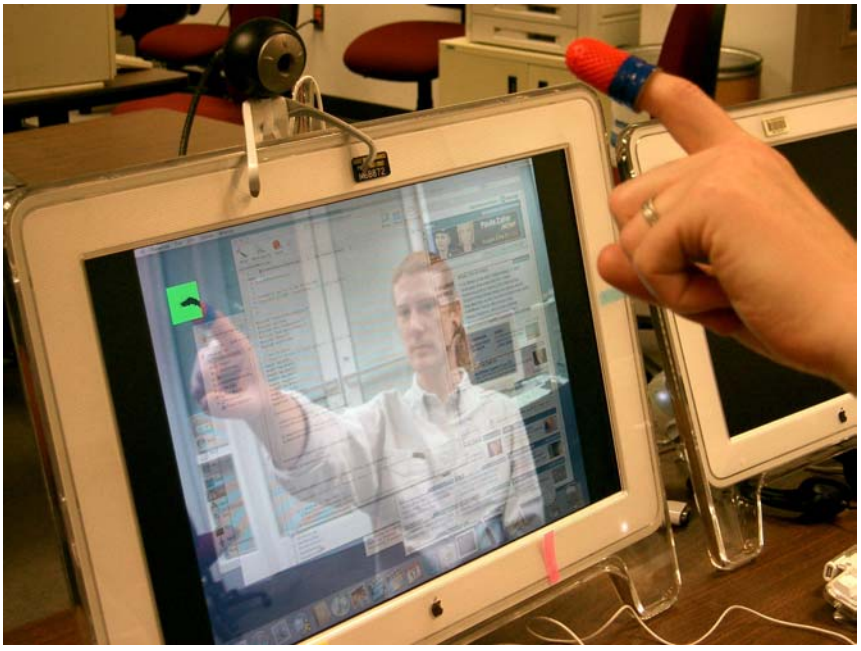

Figure 9: Original Mac-based Facetop, single-user mode

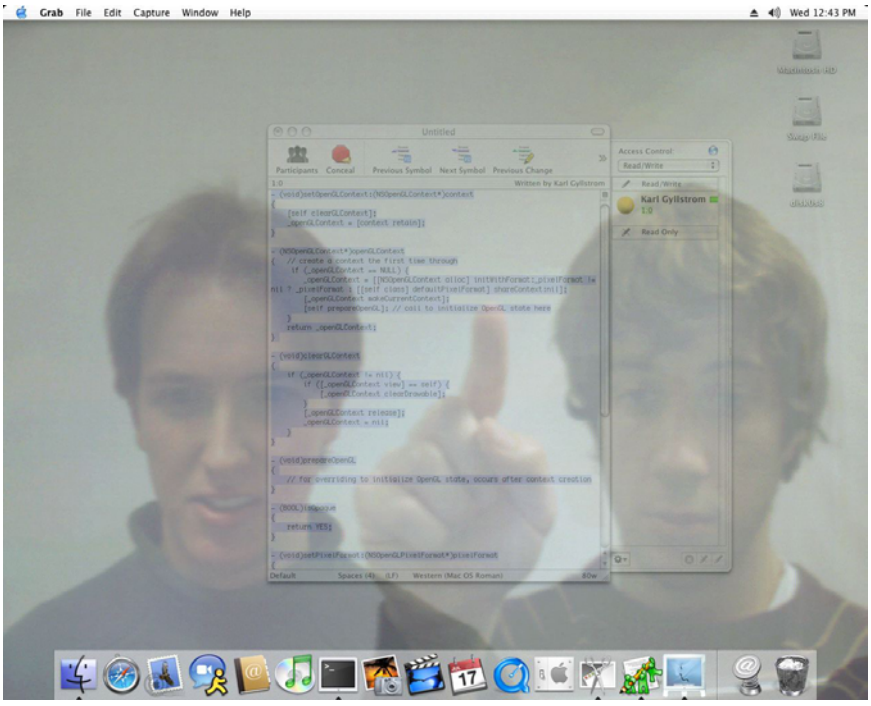

Figure 10: Original Mac-based Facetop, collaborative 2-user mode