The purpose of this project was to investigate the feasibility and cost of developing an inexpensive videoconference center in the CSDL conference room. Tools such as NetMeeting provide low cost solutions for small conferences but are rather limited in the functionality they provide. Higher end solutions are available in which users in a large meeting hall participate in conferences between many locations with multiple cameras at each location, but for the purposes of the CSDL, the cost of such systems is prohibitive. In this paper I outline the results of my investigation, identify three possible hardware solutions and describe some preliminary work in designing an interface for the system.

Categories and Subject Descriptors
H.5.2 [User Interfaces]:

General Terms
Human Factors

Keywords
Videoconferencing, Class Project, CSDL Conference Room

1. INTRODUCTION
This paper presents my findings in investigating the options for developing a video conferencing capability for the conference room in the Center for the Study of Digital Libraries (CSDL). Videoconferencing offers the ability for a much richer interaction between collaborators located in separate facilities, however, commonly available videoconference facilities tend to be either too expensive and to far removed from normal work practices to be practical or else to low tech and simplistic to provide an effective alternative to "being in the room."

At one end of the spectrum, videoconference centers offer specially designed desks with built in microphones and a "push to talk" button that can cause a high end camera to pan, tilt and some to focus on the person in that particular seat. The costs of such a system prohibit its installation at modestly sized research labs. The cost of these systems can typically be justified only by much larger institutions who can then rent time in the center to smaller groups. But the use of this sort of a conference room requires careful advance planning, and takes those who will participate in the conference away from their usual working environment. This disruption in work practices makes a teleconference an inconvenience at the very least.

Microsoft's NetMeeting stands at the other end of the cost (and quality) spectrum of video conferencing solutions. The software is free, and the hardware (depending on what one is willing to settle for) can be obtained for anywhere from $30-$180. The conference can be held between any two people with a broadband Internet connection. This solution is limited, though, by the fact that it typically can be used to connect only two locations and only two cameras.

In the remainder of this paper I outline a web based video conferencing system tailored to the needs of the CSDL though one which I feel also has the potential to serve as a general model for a relatively low cost video conferencing solutions for small and medium sized research and development groups. I begin by briefly describing the CSDL conference room to get a feel for how a video conference system might be deployed. I then move to survey some of the hardware options and then proceed to describe some interface design considerations and present a mock up of the potential. I conclude by describing the next steps that will be needed in actually implementing this system.

2. The Conference Room
In developing a solution for transforming the CSDL conference room into the CSDL videoconference room, there needs to be a relatively low cost solution that still provides adequate coverage of the entire conference room. Specifically, there are three primary areas of activity: a SMART board which is used for PowerPoint presentations and other computer based demonstrations, a white board running along the length of the left side of the room (looking toward the front) and the people sitting around the table. One possible way to adequately cover this space is to provide one camera looking forward, one camera looking backward toward the table and two on the right wall looking toward the left and right half of the long marker board. An additional fifth camera can be added to photograph paper documents. Generalizing from this observation we find the need for a solution that allows for multiple cameras to be available at a single location. While this falls far short of the ability of high end videoconference centers to direct the camera directly at the
individual speaking, it would make a notable advance over the single camera model of NetMeeting, allowing the focus of the video conference to shift as the activity of the discussion moves around the various areas of the room. This raises design questions regarding who will control which camera is displayed and how.

3. Survey of Hardware Options

One of the most critical (and costly) elements in the videoconference center will be the video cameras used. Two major features of cameras are relevant to their suitability in creating a low cost videoconference center. First, is the ability to access the camera directly over a network connection. This offers a considerable advantage over non-web accessible cameras since it removes the need to install multiple drivers on a computer and select between the camera input devices. It also means that separate processors (the cameras rather than on a single computer) can be responsible for capturing and transmitting the video feed from each camera. Most of the cameras mentioned below are web accessible with built in web-servers and internet based control and configuration systems. Second is support for panning, tilting and zooming the camera. This feature might allow individuals in a video conference to focus their attention (and their colleagues’ attention) on a particular individual or area of the room. This would allow for greater precision in selecting an area of focus than would four statically placed cameras as mentioned in the description above. One concern with using this approach in a small conference room where individuals are often seated around a table is that, unlike large auditorium style rooms, it is unlikely that there will be a single vantage point from which a camera with pan, tilt and zoom functionality could effectively view all the members of the room. Other factors, such as the noise of the motor and the ability of multiple participants to control the location of a single camera also factor into considerations about the suitability of such a camera. A number of other features also merit consideration, though they have a significantly smaller impact.

I have identified the following five cameras as options for use in the CSDL conference room:

<table>
<thead>
<tr>
<th>Camera Model</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panasonic KX-HCM280</td>
<td>$1200</td>
</tr>
<tr>
<td>Sony EVI-D70</td>
<td>$900</td>
</tr>
<tr>
<td>Panasonic KX-HCM10</td>
<td>$250</td>
</tr>
<tr>
<td>Axis 205</td>
<td>$180</td>
</tr>
<tr>
<td>D-Link DCS-900</td>
<td>$100</td>
</tr>
</tbody>
</table>

The Panasonic KX-HCM280 and the Sony EVI-D70 are both clearly high end cameras. They are included in this list largely to provide an upper bound consideration on what equipment might be useful. Both are pan/tilt/zoom camera with an optical zoom of more than 18x. Both these cameras have nearly 360 degree pan angles and greater than 90 degree tilt angles, greater than 18x zoom and can be ceiling mounted. One possible use for these cameras would be to mount them in the center of the conference table, allowing a single camera a clear view of all the individuals in the conference. This approach, has the notable disadvantages that the people are very likely to be looking away from the camera (toward the screen for instance). This would destroy any face-to-face feeling that might be found in a videoconference and seriously undermine the usefulness of being able to point the camera directly at an individual. In light of this, the increased costs of these two cameras seem to offer very little advantage over the other cameras listed in terms of functionality (though they may have better image quality). One possible advantage is that the Sony camera supports very low light conditions – the most common case for presentations.

The Panasonic KX-HCM10 also offers pan and tilt control (though not zooming) at a much more affordable price. It can pan through 120 degrees and tilts from 0 to negative 45 degrees. This camera could be incorporated into the conference room layout described above by mounting it on the wall opposite the whiteboards. From this vantage point it could be panned to cover both halves of the whiteboard and a large portion of the conference room table, providing a view of the table in addition to the one provided by the camera mounted in the front of the room. If this camera can offer comparable image quality to fixed cameras and if the issue of having only one user able to control the camera at once is considered unimportant, then using this camera offers a solution that is competitive in terms of cost to using two of the D-Link cameras to view the whiteboard and that would be cheaper than using two of the Panasonic KX-HCM10 cameras.

The Axis 205 and D-Link DCS 900 are both fixed field of view cameras. Both come with built-in web servers so the video feed can be accessed directly via the Internet. The feedback for these cameras provided at Amazon indicated that the Axis 205 had surprisingly good image quality while the D-Link camera performed more poorly than expected. In addition to simple performance, the Axis web site provides developer documentation that describes the HTTP API, Windows programming using ActiveX components and how to tailor the embedded scripting using C or PHP. I have not found such information to be readily available for the D-Link camera. The added quality and availability of technical documentation seems to offer a distinct advantage in the case at hand.

With this available hardware, the conference room setup with four cameras as previously described could be implemented for a hardware cost of as little as $400 using the D-Link DCS-900 cameras and assuming that each camera could be easily connected to a 100 Mbps LAN. If the cameras are upgraded to four Axis 205 Cameras, the price jumps to $720. This increased price would seem to be justified if the reported difference in image quality is correct. The ability to read the whiteboard and to see presentations at the front of the room will be greatly impacted by the video quality and the ability to access the Axis 205 via well documented APIs will add a degree of flexibility and customization that is not possible with the D-Link camera. These factors seem to me to justify the increased cost of the Axis 205. Using the three camera alternative by replacing two of the Axis cameras with the Panasonic KX-HCM10, the cost of the cameras needed would come to $610.

4. Initial System Design

Having sketched the hardware aspects of setting up a videoconference room for the CSDL, the next step is to outline some of the issues in designing the user interface for this setup. In this section I present an interface design for a standalone application to support this type of videoconference setup. An alternative approach would be a to implement the interface as a
purely web based application. While I do not discuss this here, many of the issues presented below will remain the same for a web based interface.

The major goal in designing the user interface for the videoconference tool was to keep the video as the central element and allow other interface elements to recede to the background or to be moved out of the way for the majority of the conference. Additionally I wanted the video itself to be easily positioned by those participating in the conference. It is expected to be the case that conference participants will often be more interested in presentation slides or an Internet resource they are exploring than they are in seeing what another conference member is doing. Or perhaps a participant would rather observe the conference in the background while focusing his attention on more pressing issues (e.g. writing a term paper). To facilitate these objectives the interface is composed of a main “Control Center” window that can be used to handle the major conference configuration and monitoring tasks and separate “Participant” windows that show the video feed from a given participant in the conference. Multiple participant windows can be shown and arranged on the display in a way that most closely meets the needs of moment.

The control center, shown in figure 1 is divided into two main areas. The top half of the display provides information about and control over the local cameras and audio. The video feed from the local cameras can be monitored in a tabbed pane. In this pane, each tab corresponds to one available video feed from the local conference room. The text on each tab can be highlighted red to indicate that one of the participants in the conference is currently viewing the video from that camera. Below the monitor pane are two check boxes. The first allows a conference participant to enable or disable the audio feed from this particular conference room. Currently, only a single audio transmission from each participant in the conference is envisioned, though this may be unduly restrictive, especially for large conference rooms in which a single microphone cannot reliably record all the members present at one particular location. One possible workaround (within the current design) would be to have each location with a microphone register as a participant in the conference and use a single video feed for multiple participants. This seems less than ideal, however, and further thought, guided by feedback from actually using the proposed system, is needed. The second check box allows a participant to decide if he wishes to allow the other conference participants to select which video feed they see (with multiple participants being able to see different video feeds simultaneously) or if he wishes to allow participants to view only the video feed that he selects.

Figure 1: Proposed Control Center Window

The bottom half of the control center display presents a list of the conference participants. Next to each participant are two buttons, one that causes the video monitor for that participant to be displayed, and another that enables or disables the audio stream from that participant. The menu bar provides facilities for initiating a conference call, inviting participants and adding and configuring the cameras at the local conference site.
Figure 2: Proposed Participant View Window

The participant window simply displays the video feed from a single participant. If the participant has chosen to allow others to select the video feed they would like to see, users can click on the tabs at the bottom of the display. In the hypothetical window shown in figure 2, the participant being observed here is named Maryland, there is a camera that is focused on a table (presumably in a conference room) and two other personal cameras one focused on Ben S. and the other focused on Ben B. This participant monitor window can be moved to a convenient location on the display in order not to interfere with other tasks that might be more pressing that watching what is happening in Maryland.

One might move Ben and Ben off to the corner and continue working on something else, periodically glancing down to see if anything interesting is happening.

5. Future Work
So far the work that has been done is largely conceptual in nature. The obvious next step is to acquire one or two cameras in order to begin working with some of the technical issues more concretely. Once these issues have been addressed, a final decision can be made on the hardware that would be needed (or desired) to adequately meet the needs of developing a local, low cost teleconferencing center. In addition to further development on the technical aspects of the system, some preliminary work can be done to evaluate the usability of the proposed interface. Both usability studies (including paper-based Wizard of Oz style, and lightweight interface prototype with recorded video) and further analysis of the needs of potential users of the system could be done at this time. As progress is made in developing the system, further evaluation could take place. Additional considerations might be given to the idea of adding cameras with support for remote control of pan, tilt and zoom functionality.

6. Conclusions
In summary, it seems possible to convert the CSDL conference room into a relatively simple but highly functional videoconference center for a cost between $400 and $720 dollars plus the expense of a document camera (not discussed here) and any necessary upgrades to the networking infrastructure. Purchasing the necessary hardware incrementally will mitigate the risks of unexpectedly poor video quality, difficulties associated with transmitting the video over the available network bandwidth (on both the sending and receiving ends) and other currently unresolved technical challenges to implementing a videoconferencing system.