

TELE IMMERSION

A SEMINAR REPORT

Submitted By

VARUN V

In partial fulfillment for the award of the Degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

SCHOOL OF ENGINEERING

COCHIN UNIVERSITY OF SCIENCE & TECHNOLOGY,

KOCHI-682022

AUGUST-2008

DIVISION OF COMPUTER ENGINEERING
SCHOOL OF ENGINEERING
COCHIN UNIVERSITY OF SCIENCE & TECHNOLOGY
KOCHI-682022

Certificate

Certified that this bonafide record of seminar entitled

“TELE IMMERSION”

Done by

“VARUN V ”

of the VIIth semester, Computer Science and Engineering in the year 2008 in partial fulfillment of the requirements to the award of Degree of Bachelor of Technology in Computer Science and Engineering Of Cochin University Of Science and Technology.

Mr. SACHITH RAJAGOPAL

Seminar Guide

Dr. DAVID PETER .S

Head Of Division

Date:

ACKNOWLEDGEMENT

Although a single sentence hardly suffices, I would like to thank almighty god for blessing me with his grace and taking my endeavor to a successful culmination.

I also express my gratitude to **Dr. David Peter, HOD** and my seminar guide for providing me with adequate facilities, ways and means by which I was able to complete this seminar. I express my sincere gratitude to him for his constant support and valuable suggestions without which the successful completion of this seminar would not have been possible.

I thank **Mr. Sachith**, my seminar guide for his boundless cooperation and helps extended for this seminar. I would like to extend my gratitude to all staff of Department Of Computer Science for the help and support rendered to me. I have benefited a lot from the feedback, suggestions and blessings given by them.

I would like to extend my gratitude to all of my friends for their help and support rendered to me in the various phases of this seminar.

Abstract

TELE IMMERSION

Tele-immersion is aimed to enable users in geographically distributed sites to collaborate in real time in a shared simulated environment as if they were in the same physical room. Tele-immersion is aimed to be used in different areas, such as 3D CAD design, entertainment (e.g. games), remote learning and training, 3D motion capturing. We define tele-immersion as that sense of shared presence with distant individuals and their environments that feels substantially as if they were in one's own local space. One of the first visitors to our tele-immersion system remarked "It's as if someone took a chain saw and cut a hole in the wall [and I see the next room]." This kind of tele-immersion differs significantly from conventional video teleconferencing in that the user's view of the remote environment changes dynamically as he moves his head.

Tele-immersion is a technology to be implemented with Internet2 that will enable users in different geographic locations to come together in a simulated environment to interact. Users will feel like they are actually looking, talking, and meeting with each other face-to-face in the same room.

This is achieved using computers that recognize the presence and movements of individuals and objects, tracking those individuals and images, and reconstructing them onto one stereo-immersive surface

Such approaches are geared toward exploration of abstract data; our vision, instead, is of a realistic distributed extension to our own physical space, this presents challenges in environment sampling, transmission, reconstruction, presentation, and user interaction. Other approaches that concentrate on realistic rendering of participants in a shared tele-conference do not employ the extensive local environment acquisition necessary to sustain a seamless blending of the real and synthetic locales.

Tele-immersion presents the greatest technological challenge for Internet2etc.

TABLE OF CONTENTS

Chap no.	Title	Page No.
	LIST OF FIGURES	ii
1.	INTRODUCTION	1
1.1	Early Developments	2
2.	REQUIREMENTS FOR IMMERSIVE TELE CONFERENCE SYSTEMS	3
3.	HOW TELE IMMERSION WORKS	4
4.	SHARED TABLE ENVIRONMENT	5
5.	NOVEL VIEW SYNTHESIS	8
6.	TELE CUBICLES	14
6.1	How tele cubicle works	
7.	COLLABORATION WITH I2 AND IPPM	17
8.	PRESENT RESEARCH	18
8.1	Media technologies	18
8.2	Soft	19
8.3	Sandbox	19
9.	CONCLUSION	23

LIST OF FIGURES

No:	Name	Page no:
1.	Tele Immersion Implementation	4
2.	Left and right camera view of stereo test sequence	7
3.	Test maps extracted from stereo test sequence	9
4.	View adaptive synthesis in virtual 3D scene	11
5.	Images of convergent stereo rig	13
6.	Tele cubicles	14
7.	Working of tele cubicles	15
8.	Vision based 3D reconstruction of image	18
9.	Architecture of SANDBOX	20
10.	Selection of Range	21

1.INTRODUCTION

According to *Jason Leigh*

“The term Tele-immersion was first used ... as the title of a workshop ... to bring together researchers in distributed computing, collaboration, virtual reality and networking”

According to *Watsen & Zyda*

“It enable the interaction between geographically remote participants within a shared, three-dimensional space.”

In the past people only dream about communicating geographically but the advancement in telecommunication along with advancement in media techniques make it possible. But still there was struggle to make them collaborate in a real time world, like efforts to have users share

the same physical space, during their meetings, conferences, etc. National Tele-Immersion Initiative – NTII team leads the way to make all these things possible. They are working on projects to have users share the same physical space in a real time world, as if they are sitting in front of each other in the same room. In this regard Advanced Network & Services played a vital role, to bring together the experts in this field close together. This team is led by Jaron Lanier, who was one of the pioneers in development of Virtual Reality (which according to him is “the brain anticipates a virtual world instead of the physical one”) in 1980’s. National Tele-Immersion team started their work in middle of 1997 and the collaborating schools were Brown University, Providence Naval Postgraduate School, Monterey, University of North Carolina, Chapel Hill and University of Pennsylvania, Philadelphia.

1.1 EARLY DEVELOPMENTS

In start the main aim of the team was to take into account the ultimate synthesis of media technologies for the scanning and tracking of three dimensional environment. Based on vision based three dimensional reconstruction with the help of new advancement in fields like media technologies, networking, robotics. In May 2000 whole the hectic efforts of the team cope up with some success, with first demonstration of three years long work. National Tele-Immersion Initiative team led by virtual reality pioneer Jaron Lanier, conducted which at one stage was just imagination. This effort led to the thinking which could change the way we communicate over long distances, people could feel each other submerge together in the same physical space.

The experiment was conducted in Chapel Hill led by UNC computer scientists Henry Fuchs and Greg Welch. It linked UNC Chapel Hill, the University of Pennsylvania in Philadelphia and Advanced Network & Services at New York. Researchers at each place could feel themselves in the office of their colleagues hundreds of miles far apart. The apparatus of the test consisted of two large walls, projection cameras and head tracking gear. One screen was at left side of Welch and other was on right. Through left wall Welch can see his colleagues at

Philadelphia and through other of New York. He can peep in and out and images change accordingly, like when he leaned forward images grew larger and become smaller when he moved back. At each target site there were digital cameras to capture the image and laser rangefinders to gather information regarding the position of the object. Computer then converted them into a three dimensional information which was then transmitted to Chapel Hill via Internet², where computers were mounted to reconstruct the image and display that on the screen.

To some point it seems that Tele-Immersion is another kind of Virtual Reality but Jaron Lanier is of other view. According to him "virtual reality allows people to move around in a pre-programmed representation of a 3D environment, whereas tele-immersion is more like photography. It's measuring the real world and conveying the results to the sensory system," he says.

2. REQUIREMENTS FOR IMMERSIVE TELECONFERENCE SYSTEMS

To meet the requirements of immersion, it is absolutely necessary to use a large display that covers almost the whole viewing angle of the visual system. In addition, the large display has to be integrated into the usual workspace of an office or a meeting room. Thus, the most practicable solution is a desktop-like arrangement with large flat screens like plasma displays with a diagonal of 50 inch and more. Starting from such a desktop-like system and taking into account results from intensive human factors research , further requirements on the presentation of the scene can be formulated as follows :

- conferees are seamlessly integrated in the scene and displayed with at least head, shoulders, torso and arms in natural life-size
- all visual parameters of the scene and the different sources have to be harmonised

- the perspective of the scene is permanently adapted to the current viewpoint of the conferee in front of the display (head motion parallax; look-behind effect)
- eye-contact between two partners talking to each other has to be provided
- gaze from one conferee to another has to be reproduced in a sufficient manner such that everybody can recognise who is looking at whom (e.g.: who is searching for eyecontact)
- voice of a conferee must come from the same direction where he is positioned on the screen.

3. HOW TELE-IMMERSION WORKS

Fig 1. Tele Immersion Implementation



Above figure is a nice description of the Tele-Immersion implementation. Two partners separated by 1000 miles collaborate with each other. There is a sea of cameras which provide view of users and their surroundings. Mounted Virtual Mirrors provide each user a view how his surrounding seems to other. At each instant camera generated an image which is sorted into subsets of overlapping trio. The depth map generated from each trio then combined into a single view point at a given moment.

4. SHARED TABLE ENVIRONMENT

A very attractive way to meet the above requirements is to follow the principle of a shared table environment. It is based on the idea to position the participants consistently in a virtual environment around a shared table.

At the transmitting side the conferee in front of the display is captured by multiple cameras and a 3D image of the conferee is derived from this multiview set-up. The 3D images of all participating conferees are then placed virtually around a shared table. Ideally, this is done in an isotropic manner in order to obtain social symmetry. Hence, in the case of three-party conference shown in **Fig. 2** the participants form an equilateral triangle. In the case of four parties it would be a square, an equilateral pentagon for a five-party system, and so on.

At the receiving end this entirely composed 3D scene is rendered onto the 2D display of the terminal by using a virtual camera. The position of the virtual camera coincides with the current position of the conferee's head. For this purpose the head position is permanently registered by a head tracker and the virtual camera is moved with the head. Thus, supposing that the geometrical parameters of the multi-view capture device, the virtual scene and the virtual camera are well fitted to each other, it is ensured that all conferees see the scene under the right perspective view, even while changing their own viewing position. As a consequence, they can also change the view knowingly in order to watch the scene from another perspective, to look behind objects or to look at a previously occluded object. Moreover, all deviations of the conferees' position from a default position, are picked up by the multi-view capture devices. Thus, again supposing well fitted geometrical relations, the 3D image will be moved equivalently in the virtual world and, as a consequence, the other conferees can follow the resulting perspective changes at their displays. These circumstances ensure a natural reproduction of eye-contacts and body language in the case of direct face-to-face communication between two partners as well as a natural perspective of this bilateral communication from the position of the third conferee.

Last, but not least - the isotropic scene composition and the resulting symmetry enable that the displays can be also placed symmetrically between the partners (i.e. at the middle of the direct viewing axis). Thus, the display works similar to a mirror. Hence, all portrayals appear well balanced in natural life-size at the displays and a psychological dominance of particular participants is avoided.



Fig 2. Left and right camera view of stereo test sequence

5. NOVEL VIEW SYNTHESIS

From the image processing point of view, the main difficulty of the shared table is to obtain an 3D image of the conferees for placing it into the virtual environment. It is well known that the 3D shape of an object can be reconstructed if multiple camera views are available and the respective cameras are calibrated. Often, the depth structure is then represented by 3D wire-frames. However, to achieve a natural impression, a wire-frame technique requires a lot of triangles and vertices. Such a detailed wire-frame can only be obtained by complicated image analysis algorithms which suffer from a high computational load. Moreover, real-time rendering of detailed wire-frames is only possible with high-power graphic stations. This all leads to a system complexity which is not desired for the application under study. To overcome these problems, several authors consider model-based methods tailored for special situations; e.g. face and human body models for video-conferencing. However, these methods assume a priori knowledge about the object to be modeled. Although such an approach works well for some low end applications, these techniques obviously reduce the visual realism as generic representations are applied to natural objects.

Thus, they are not yet usable for tele-immersion. A much more attractive approach is a novel view synthesis on the basis of implicit intermediate viewpoint interpolation. Here, the 3D object shape is not reconstructed explicitly, but virtual views are directly calculated from the real camera images by exploiting disparity correspondences. In this context, a very efficient method is the so-called incomplete 3D representation of video objects (IC3D) proposed in [1]. In this case, a common texture surface as the one shown in the left image from **Fig. 3** is extracted from the available camera views - e.g. the two views from **Fig. 2** -, and the depth information is coded in an associated disparity map as depicted on the right side in **Fig. 3**. This representation can be encoded like an arbitrarily shaped MPEG-4 video object, where the disparity map is transmitted as an assigned grey scale alpha plane. For synthesis purposes the decoded disparities are scaled according to the user's 3D viewpoint in the virtual scene, and a disparity-controlled projection is carried out. Basically, the original left and right camera views, and also any views from positions on the axis between the two cameras, can be reconstructed. **Fig. 4** shows some examples for this

synthesis process. Note that the 3D perspective of the person changes with the movement of the virtual camera. One benefit of this technique is its low.



Fig 3. Texture and disparity maps extracted from stereo test sequence

complexity and high stability compared to algorithms using complete 3D wire-frames. In particular, the rendering of the viewpoint-adapted video object is quite simple and requires a

very low and constant CPU time. Due to these properties, it becomes realistic to implement the 3D representation of natural object (e.g. conferees in an immersive tele-conference) as well as virtual view synthesis in real-time.

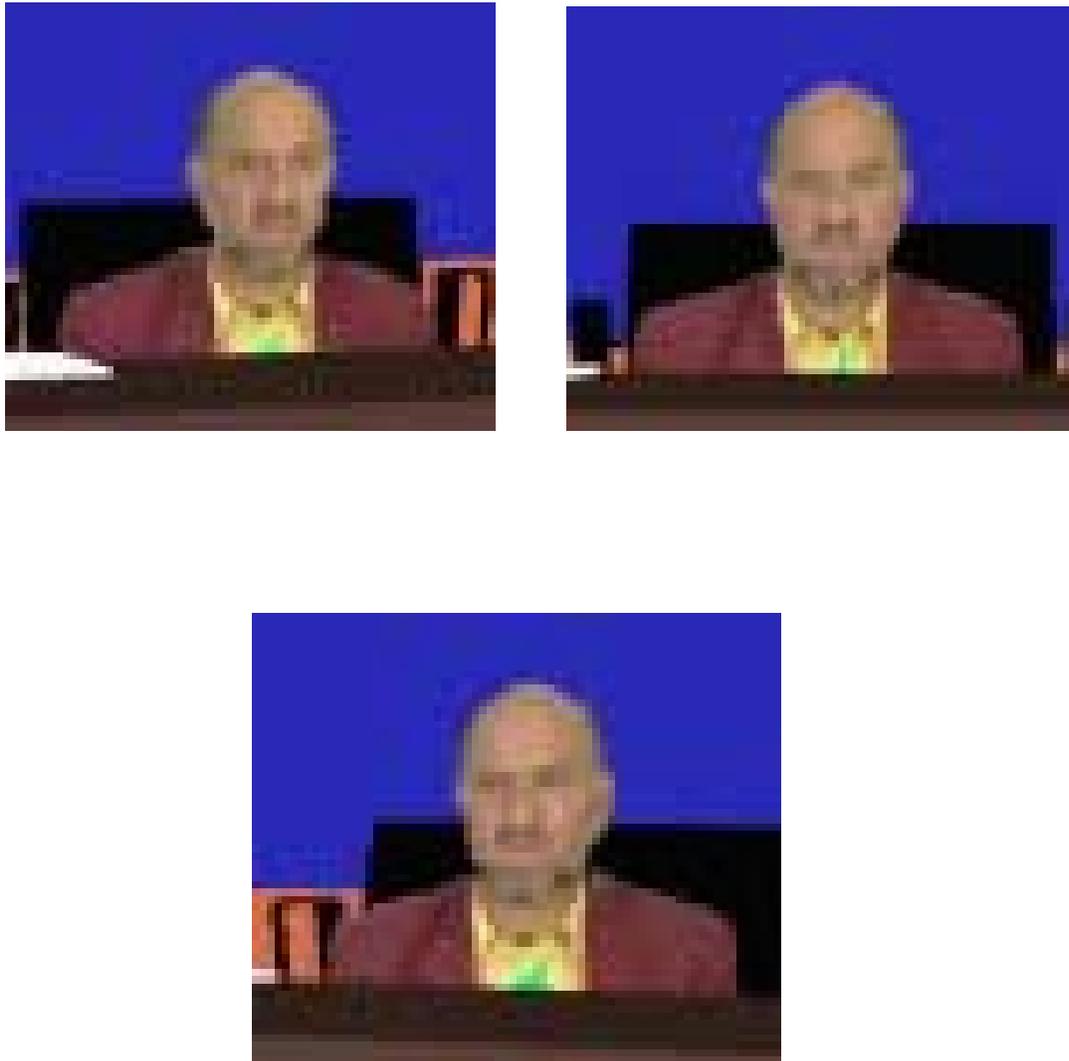


Fig 4 :View-adaptive synthesis in virtual 3D scene (based on representation data in Fig 3)

The conventional IC3D technique has originally been developed for parallel camera set-ups. In this simplified case the vertical component of the disparity vectors is always zero and only horizontal displacements have to be processed. That is the reason why the IC3D approach

works with one and not with two disparity maps (see **Fig. 6**). On the one hand, this limitation to one disparity map is essential as long as MPEG-4 is used for coding because the current MPEG-4 profiles do not support the transmission of more than one disparity map.

On the other hand, the restriction to parallel camera set-ups is no longer possible in immersive tele-conferencing scenarios. An immersing system requires large displays and short viewing distances. Therefore the distance between the cameras becomes quite large and the cameras have to be mounted in a strongly convergent set-up in order to capture the same location in front of the display. Nevertheless, the techniques of IC3D can be extended to this generalised situation, although disparity correspondences are 2-dimensional in strongly convergent camera set-ups. To explain this IC3D extension in detail, **Fig. 5** shows images of a convergent stereo pair. The mounting of the cameras is similar to the one sketched. The left image refers to a camera at the top of the display, whereas the other has been captured from its left border. It is well known from epipolar geometry that the disparity correspondences follow so-called epipolar lines which can be derived from the fixed geometry of the convergent camera set-up (see black lines in **Fig.5**). Due to this epipolar constraint, 2-dimensional disparity correspondences can always be projected onto 1-dimensional displacements along epipolar lines. In addition, it is possible to warp the images in such a way that the epipolar lines become horizontal. Basically, this means that the cameras of the convergent set-up are virtually rotated until they would form a parallel set-up.



Fig 5 : Two images of a strongly convergent stereo rig referring to a real conference setup.

6. TELE-CUBICLES

“A tele-cubicle is an office that can appear to become one quadrant in a larger shared virtual office space.” Initial sites were UIC, UNC, and USC, as well as one in the New York Area. The main idea behind this work came directly from the Tele-Immersion meeting on July 21, 1997 at the Advanced Network Office. At the meeting each participant university (UIC, NPS, UNC, Columbia, and USC) brought its individual designs of cubicles and together immersed the user and the desk. One of the striking results of the meeting was the discovery of how future immersive interfaces look like, and what were the needs and requirements at that time to make this impossible looking task in the past, into reality.

5.1 HOW TELE-CUBICLE WORKS

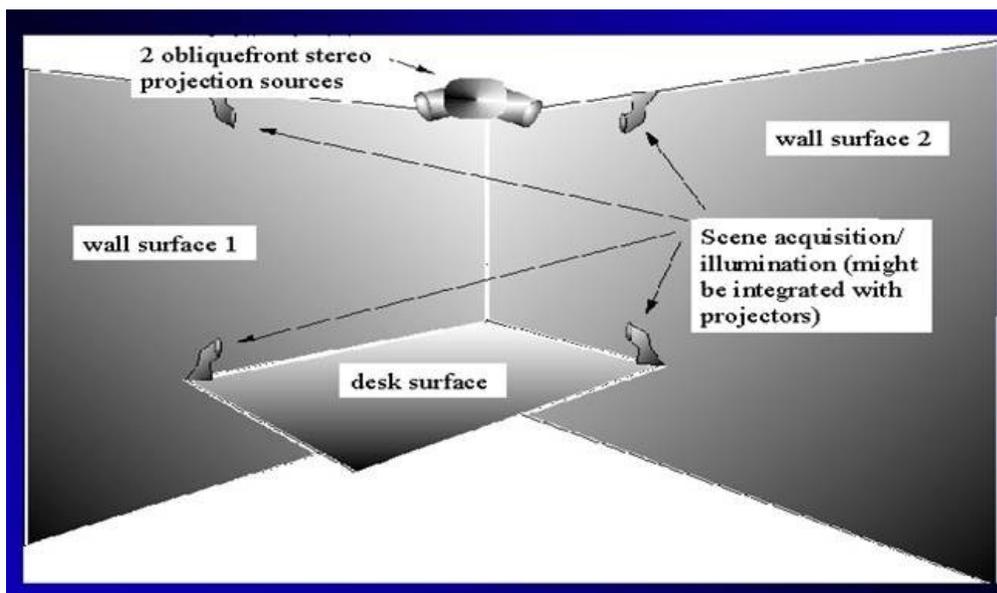


Fig 6. Tele Cubicles

The apparatus consists of:

- desk surface (stereo immersive desk)

- two wall surfaces
- two oblique front stereo projection sources (might be integrated with projectors)

As illustrated (in fig 6) the three display surfaces meet each other in the corner to make a desk. At the moment four tele-cubicles can be joined to form a large virtually shared space. The walls appear to be transparent passage for other cubicles during this linkage, and the desk surfaces join to form a large table in the middle. Objects at each place can be shared for viewing across the common desk and through walls can be seen the colleagues at other end and their environment.

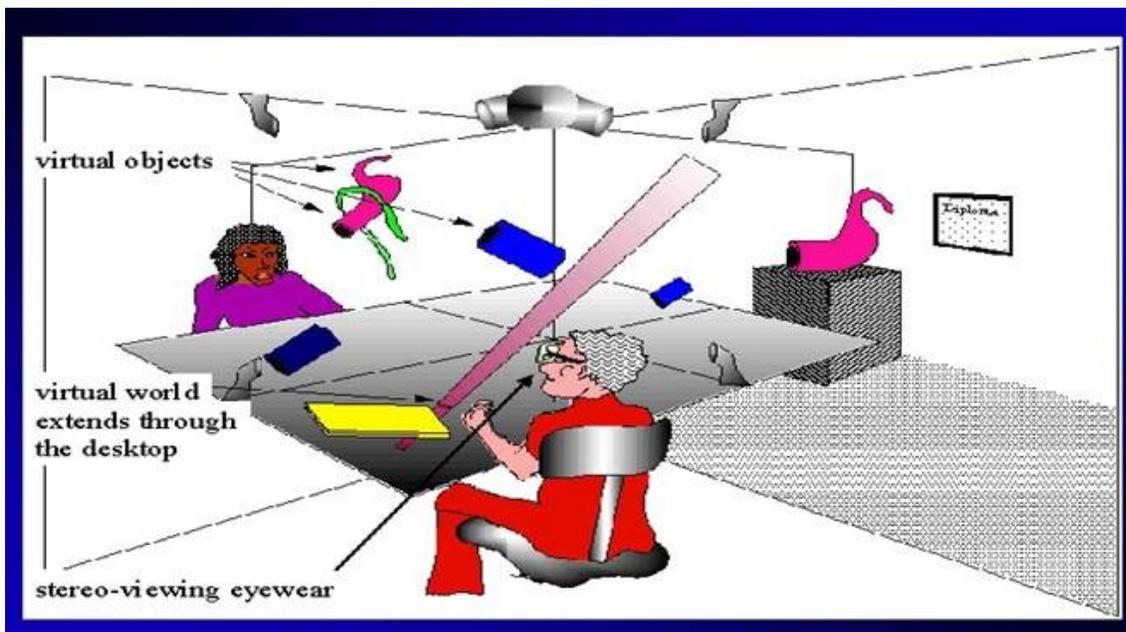


Fig 7. Working of tele cubicle

Fig 7 describes how the participants so far away share the same physical space, through common immersed stereo desk and can see each other environment, virtual objects place in the others environment, across the walls which looks like transparent glasses when cubicles connected together. So the virtual world extends through the desktop. The short term solution at

that time was to have remote environment pre-scanned which lead towards the goal which was obviously to have environment automatically scanned.

In the early years there were some limitations in the task as each partner university did not have the same techniques to present itself to others. Various modules like Sketch, Body Electric, and Alice were the results of the first year development, but they were not of much success as there were not much technologies available at that time to integrate them.

The hectic efforts in this regard initiated a project called Office of the Future. In this project the ideas which were discussed in the July, 1997 meeting coined together. The approach was to use the advanced techniques in computer vision field to capture the visible objects in the office like furniture, people, etc. The capture images were then reconstruct and transmitted over the network to the other remote site for display.

7. COLLABORATION WITH I2 & IPPM

To cop up with the problems like communicating speed and better transmission of data over the network, Tele-Immersion team collaborated with Internet2 and Internet Protocol Performance Metrics. Main problem as obvious was that today's internet is not fast enough to transmit data, specially when you need to transmit a huge bulk of data across the internet about people and their environment. The experiment conducted at Chapel Hill used 60 megabits per

second and good quality tele-immersion requires 1.2 gigabits per second. To make it possible 160 USA universities and other institutes started a research project which should provide high reliability, and the propagation delay and queuing should be as less as possible. This could lead to revolutionary Internet applications and to ensure quick transfer of services to everyday growing network. All the members of the research collaborate on:

- Partnerships
- Initiatives
- Applications
- Engineering
- Middleware

Abilene proves to be the backbone behind this research, with initiative of providing a separate network capability. The aim was to update the cross-country backbone of 2.5 gigabits per second to 10 gigabits per second, taking into consideration of achieving a goal of 100 megabits per second.

8. PRESENT RESEARCH

8.1 Media Technologies

Media Technologies use vision based three dimensional reconstruction from set of images using multi-baseline stereo algorithms for extraction of information. Trinocular Stereo Reconstruction Algorithm is used for this purpose. 3D ray intersection yields pixel depth and then media filter the disparity map to reduce outliers and produce depth maps of 320x240 (1/z, R,G,B).



Fig 8. Vision based 3D reconstruction of image

8.2 SOFT

Soft was the first technology construction of images without the need of recompilation and customization of objects. It was standard tele-cubicle implementation for blend real and synthetic worlds. Researchers worked to enable existing and future 3D application. They develop baseline distributed virtual reality platform to make collaboration and application sharing easy.

DIS - multicasts small update packets which contain positional and event information.

CAVERNsoft - provides a persistent, shared memory distributed over a user constructed software topology with a relaxed consistency model.

Bamboo - uses *lots* of multicast groups to perform area-of-interest management.

8.3 SANDBOX

Overview

Stands for Scientists Accessing Necessary Data Based On eXperimentation.

It was developed as a subset of NASA's FIFE scientific database using the CAVE(tm) virtual reality theatre. It allows the researchers to retrieve data from scientific database with the help of virtual reality tools. User can investigate data by the usage of tools, by setting and placing tools of their choice in Sandbox to retrieve data. Like they can have data of specific day by choosing the specific temperature.

Architecture

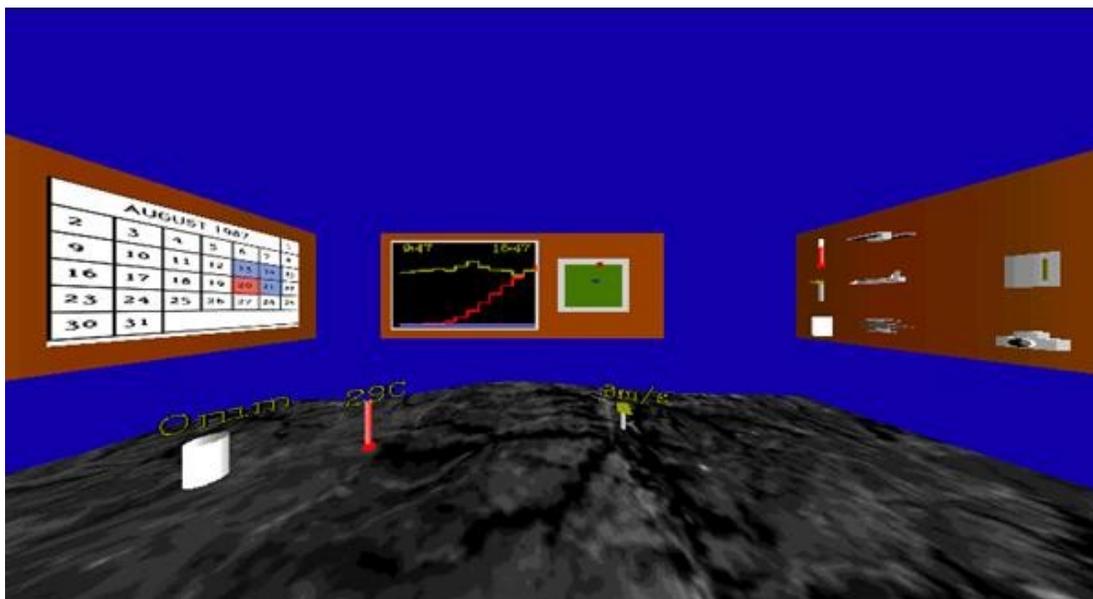


Fig 9. Architecture of sandbox

- Left Hand Wall: thermometer, wind-sock, and water beaker. They are linked to the columns in the relational database.
- Centre: LANDSAT satellite, airplane, helicopter. They all are linked to the graphics files
- Right Hand Wall: notepad, camera. They are linked to meta-data.

- On Pallet: instruments seen on pallet are 3D and they are animated to improve their recognisability.

There is a 20 km square patch of Kansas which serves as a environment for the user.

Working

The instruments placed in the virtual environment helps the investigator to interact with the system.

- a wind-sock measuring wind speed and direction
- a beaker measuring rainfall
- a thermometer measuring temperature
- a camera displaying a photograph taken of a site
- user can change the settings on virtual instruments using the menu

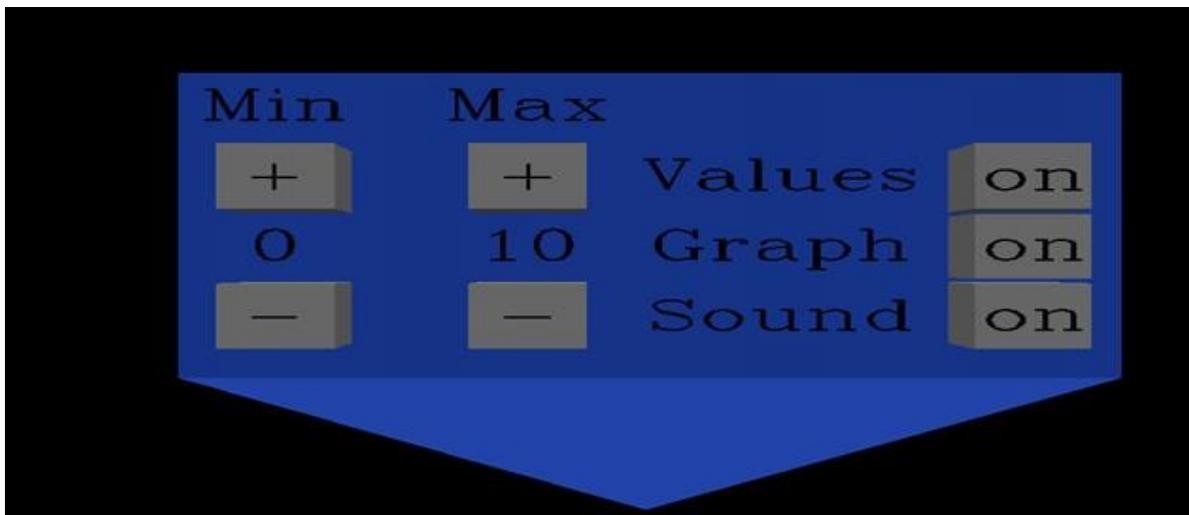


Fig.10 Selection of Range

For example, if a user wants to have data on 28th of August, at 28C, and have wind speed 9m/s. He can just select the settings from the menu, and have experimental data of the specific date.

He can have full choice of selection between maximum and minimum values as described in the Fig 10.

9.CONCLUSION

All this relies on the advancement in emerging technologies, most heavily on the ability of Internet to ship data across different networks without delay. In this regard Internet2 is the key. Both projects are going hand to hand. According to one of the researchers of Tele-Immersion Team, Defanti, "such technology would enable researchers to collaborate in fields such as architecture, medicine and astrophysics and aeroplane design. The beauty of it is that it allows widely separated people to share a complex virtual experience. You might be testing a vehicle," says Defanti. "You want to smash it into the wall at 40 miles per hour and put your head by the cylinder block. Say there's a guy from Sweden and you have to prove to him that it doesn't move by 3 centimetres or more. That kind of stuff works."

In the years to come it will be one of the major developments. You could visit each other environment, but one thing which is far behind to achieve is the physical contact of individuals at each end.

So it can be summarized as:

- Collaboration at geographically distributed sites in real-time
- Synthesis of networking and media technologies
- Full integration of Virtual Reality into the workflow

REFERENCES

1. www.advanced.org/teleimmersion.html
2. <http://www.cs.unc.edu/%7Eraskar/Office/>
3. <http://www.cs.brown.edu/research/graphics/research/telei/>
4. <http://www.cis.upenn.edu/~sequence/teleim1.html>