Group Interaction in a Surround Screen Environment

Michael Hoch, Detlev Schwabe
ZKM | Institute for Visual Media
Lorenzstr. 19, 76135 Karlsruhe, Germany
[mpunkt | detlev]@zkm.de

Abstract

This paper describes a setup using a surround screen environment (Extended Virtual Environment EVE dome) that we used to explore group interaction in real and virtual space. We have created a prototype framework to explore different modes of group interaction, the position and motion of users in real space are tracked using a vision-based interface that allows the activities of real crowds to be monitored. In the virtual space, we use a simple behavioural animation system that serves as a testbed to generate virtual group and crowd behaviour. Exploring different kinds of dynamic relationships between real and virtual groups gives insight to possible directions of group interaction.

1. Introduction

Our aim is the exploration of new boundaries between real and virtual space and between real people and virtual representations. The intention is to create new techniques for allowing individuals and groups to navigate through and interact with mixed reality shared environments with other participants. In computer graphics the behavioural animation of crowds and individuals with flocking and scattering actions, has been studied for several years [14][15][19][20]. With more powerful computers being available at home and permanent access to computer networks, new virtual environments will emerge that will be populated with avatars, i.e. virtual representations of the user. When creating such worlds, the different modes of interaction with such figures have to be explored and evaluated, especially the control of crowds or groups of virtual agents [13].

In the case of collaborative virtual environments, a framework for supporting crowds to provide greater scaleability and flexibility for communication between the inhabitants of the virtual worlds is described in [3]. It introduces an explicit crowd mechanism with respect to mutual awareness and communication between the users within shared virtual environments. Another field of research is the interaction between autonomous crowds, virtual agents, and real participants, i.e. building an appropriate interface between real space and the virtual world to explore different kinds of interrelations. In [2] an application called the "Internet-Foyer" is described as an example of such a mixed-reality environment that supports the awareness of the occupants of physical and synthetic space. Here, the physical space as well as the virtual space in represented by using a *window-into-a-world* paradigm.

As an extension to the approaches described, we want to support and explore mutual awareness of participants in real space within a mixed-reality virtual environment. In this paper we describe a prototype using a surround screen setup (EVE dome) that allows the exploration of group interaction in real and virtual space. We create a prototype framework to explore different group modalities of interaction. The EVE dome provides a context for research into a vision system for tracking individual and group movements, interactively linked to the visualisation of virtual, computer-generated individuals and crowds (humanoid or other). The user’s position and motion are tracked in real space using a vision-based interface that also allows the activities of real crowds to be monitored. For the virtual space a behavioural animation system is used, that serves as a testbed to generate virtual group and crowd behaviour.

In the following paragraphs we first comment on the importance of supporting body and space within an interface. Next, we describe the EVE surround screen environment. Thereafter, the vision-based interface is described, followed by a description of the virtual environment and the real-time package used for rendering and behavioural animation. We then explore the dynamic relationships between real and virtual groups and individuals. Finally, we present some results with users experiencing the system.

2. Supporting Body and Space

Spatial relationships between individuals within a defined area is important in terms of personal space and group interaction. In the research of public interaction, territories, personal spaces, and proximity are issues [7]. For example, if two single individuals are present, the interest will concentrate on the straight line distance. Proximity will signal different situations, and within common interaction, entering and intrusion can occur simultaneously and together with corrective reactions can direct the behavior of the individuals (see [7]). In perception theory it is known that the perception of artifacts and distances between them is not related to some objective measurement but to the relative reach of a person's arm within the environment [11], i.e. to the body. Objects are only perceived by being situated within the environment. This *being in situ* is critical to create an awareness of the body within a space. Also, the periphery (the environment that surrounds a person) is important in the perception of objects, i.e. one cannot perceive a vase without being aware of the table, the ground etc.

In systems intended for multiple users many of these issues are lost. Traditionally interaction is limited to a single user paradigm by means of mouse and keyboard or other single user input devices. This paradigm works well for zapping TV channels and numerous computer games and applications (one person has the control). But in interactive experiences, the audience that is just watching the interaction being performed by the controlling person is excluded from the experience. It is a difference being in a CAVE [6] or in front of a responsive workbench [10] with or without a tracker. In networked communication in virtual worlds, the representation of the above mentioned issues mostly focus on the virtual space within these worlds and not so much on the user's body and space in the real world. Supporting group interaction has been a subject of research in CSCW for several years, but, only recently the interest in spatial approaches to CSCW that support awareness and a notion of periphery has grown (see [2] for an overview). Again however, these approaches deal mostly with the support of interaction paradigms within collaborative virtual environments.

In this paper our focus is more on the real world: In the proposed environment we explore different modes of group interaction and explicitly incorporate the notion of body movements and real space. The goal would be a dynamic fluent group interaction as opposed to the common single-finger interaction (via mouse). The EVE dome projection environment described below, is used as the platform for building the testbed. Within this environment it is possible to have either only one or more than 20 people interacting with a virtual world. A camera setup allows the tracking of the spatial relationships between the users in real space and between individuals or groups in real and virtual space. An explorative study of the application of such interaction paradigms can give fruitful feedback for future group interaction systems.

3. The Extended Virtual Environment EVE

The Extended Virtual Environment [4] EVE is a unique implementation of a *window-into-a-world* paradigm, originally developed in 1992 and presented to the public at IMAGINA 1993. An enterable 3/4 of a sphere projection dome with a diameter of approximately 12 meters contains a turnable stereo-projection device in the center. The stereo projection is accomplished by using two projectors with polarized lenses. Passive polarized glasses are used for viewing the stereoscopic imagery. In addition to the video projectors, the device contains a loudspeaker system and an infrared camera (see figure 1). In the standard application, an observer stands inside the dome and wears a small head-mounted infrared light pointer.

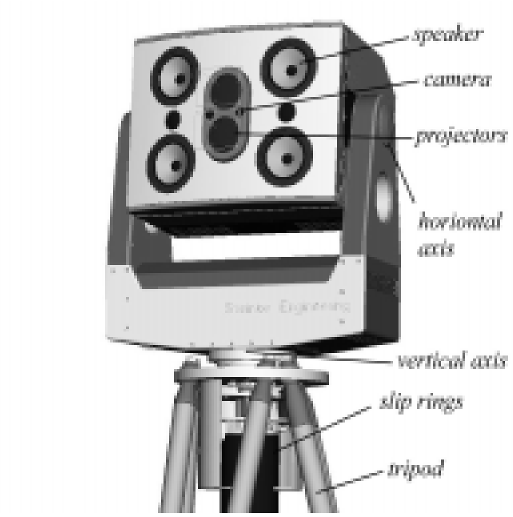

Figure 1: The EVE stereo-projection device features two rotational degrees of freedom, two video projectors (with polarized lenses), a stereo audio system and a built-in infrared camera for tracking purposes

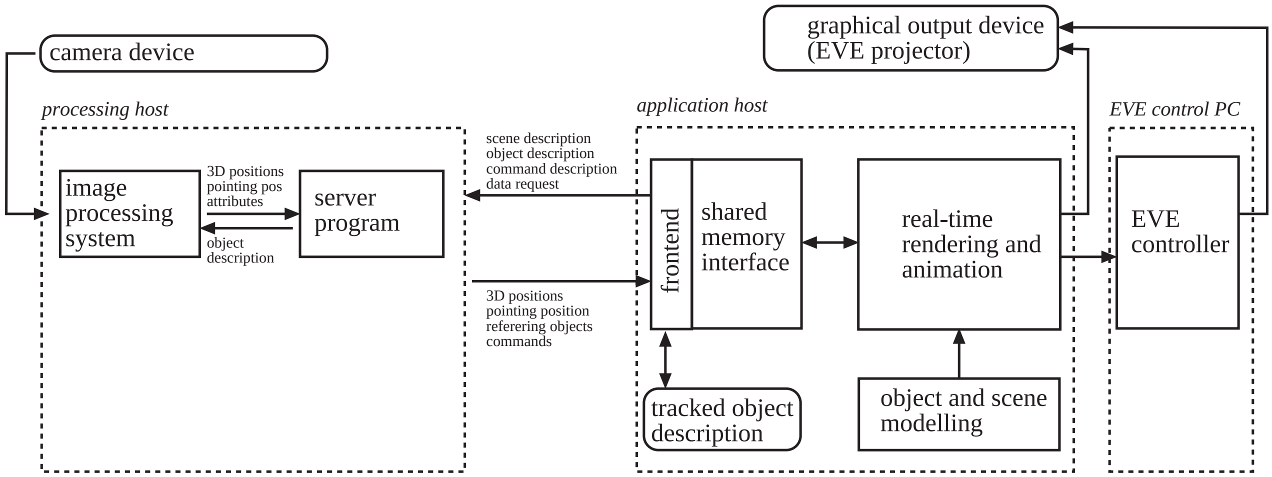

Figure 2: Architecture of the mTrack tracking system and application program for real-time rendering and animation within the EVE environment

While the user is turning his head in order to look into different directions, a dedicated PC grabs the images from the infrared camera, tracks the motion of the infrared lightpoint on the projection surface and calculates and transmits the acceleration and deceleration values to the motors which control the orientation of the projection device to keep the projected images centered on the lightpoint. Two absolute-angle sensors measure the current orientation of the projectors and deliver these values to the actual application which can then render an updated view. The application programmers interface is implemented as a shared memory interface between a separate process on the application machine, which continously reads the current state of the angle sensors and the actual application. For non-standard applications, an external control interface is available to control the turning of the projection device directly from the application program. For the work described here, the projection is completely controlled by the application.

4. Vision Based Interface

For tracking the movements of individuals and groups in the EVE projection environment in an unencumbered way, we use a vision-based interface that is based on luminance-level segmentation and blob analysis. The used system *mTrack* is divided up into the recognition part consisting of the image processing system, a server program, and the library frontend with the application program (see figure 2 and [8], [9]).

The image processing system tracks the user via an infrared camera setup that is mounted on the ceiling of the dome. It consists of a standard black and white camera and 12 standard halogen lamps. The lamps are placed in a ring around the camera and are covered by an infrared filter. During initialization the system receives a description of the objects to be tracked. Thereafter, it continuously sends position data and other calculated information of the segmented objects to the server. The server connects the application with the image processing system and updates the current states by an event-driven loop. Upon request it continosly sends data to the application The information that needs to be extracted is: position, velocity, direction of movements, distance to other participants, number of participants, and the possible detection of interaction between participants. The extraction of the position data is achieved by tracking regions based on luminance and blob segmentation. The other parameters are then calculated with respect to this data.

Tracking regions. For getting update rates higher than 10 frames per second, which is essential for establishing a direct feedback between a user and graphics [17], we had to find a compromise between sophisticated tracking algorithms and simplicity to achieve the desired frame rate on standard PCs. A simple and robust approach is the tracking of colored or greylevel regions. If the objects to be tracked are of high contrast against the background, a simple segmentation algorithm together with a blob analysis is sufficient to reliably track these objects at a high frame rate.

Detecting individual behavior. As basic parmeters that are detected for the individuals, we determine the position, the direction of movements, and the speed for each individual by using the tracking algorithms outlined above. Furthermore, it might be of interest to detect the intersection of the direction for two individuals (see figure 3a). As a simple model of the notion of territories and personal zones described in [7], or like the notion of *aura* and *focus* described in [3], a dynamic zone of attention can be defined, that is dependent on the direction and speed of movement of the individual. This zone may later

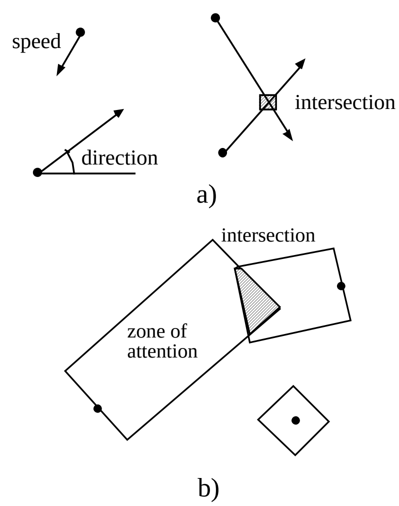

Figure 3: a) Detecting position, direction, speed, and intersection of movement. b) Detecting simplified zone of attention.

be used to detect interaction between real participants as well as between real and virtual participants (see figure 3 b).

Detecting group behavior. In our system the detection of groups in real space is directly mapped onto the spatial attribute of proximitiy, i.e. it is limited to a simple aggregation of individuals that comply to a certain distance threshhold. The threshhold is variable and depends on the number of people being present in the interaction area.

5. The Virtual Environment

We found the EVE dome to be a perfect analogy to an under-water observatory which resides on the sea bottom. The projection serves as the observation window into the surrounding virtual sea world (see figure 4). Schools of virtual fish move through the virtual world, driven by behavioural rules and interacting with the people inside the dome. For realizing this virtual environment an proprietary software programming package for real-time rendering and animation has been developed. In the current state the package mainly supports importing scenes, models and keyframed animation from a commercial 3D modeling and animation software. It further supports the creation of flocks, i.e. large groups of similar individuals

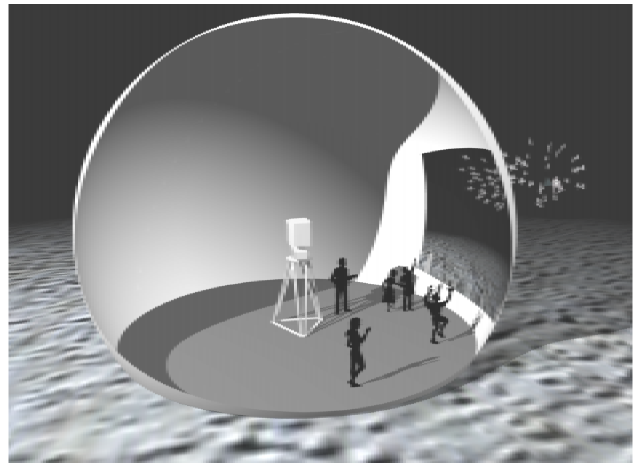

Figure 4: The EVE Dome - Extended Virtual Environment. The figure illustrates the dome situated inside a virtual world. The projection acts as a window into the virtual environment.

and behaviours, which control the basic motion of each of the individuals. The system also provides a special class of objects which interface with the external tracking system and represent the visitors position and other parameters inside the virtual world.

5.1 Behavioural Animation

The implementation of the behavioural system closely follows the work of Reynolds [15], who set up the principals of behavioural modeling for the purpose of animating groups of individuals. In his work he suggested a method for controlling the motion and behaviour of large groups of entities by means of defining a set of simple acceleration rules which applies to every member of a group. A complex, yet controllable behaviour of the whole group emerges from the application of this local rule set. Many adoptions and applications of the original paradigm of artificial flocking have been seen over the past ten years. Especially in the motion pictures it has become a standard technique to animate fish, birds, wild stampedes, swarms of bats or even armies of martial horsemen galloping into a battle [1] [5] [12] [16] [18]. In computer games it is also a widely used technique for secondary animations to create more interesting and realistic looking virtual worlds.

A behaviour of a flock in our system consists of a set of well-combined acceleration rules. Each rule adds a certain amount to the overall accelleration vector which then drives the individuals into a certain direction. Currently, the package provides the basic behavioural rules for flocking i.e. *aggregation* (staying close to immediate neighbours), *alignment* (maintaining the same speed and heading as the neighbours) and *avoidance* (not bumping into neighbours). Other rules exist for floor and ceiling avoidance as well as for maintaing a certain speed (which is usually used to prevent the individuals from coming to a complete stop). New rules can be added by deriving them from a base rule and implementing a new acceleration formula (it is planned to have a scripting language like TCL for creating new rules in the future. Currently all extensions must be hard-coded in C++).

A behaviour consists of a prioritized list of weighted rules. The most important rules (usually all rules which deal with avoidance or collision) are at the top of the list. When calculating the total acceleration of a behaviour, the list of rules is processed top to bottom. The acceleration of each rule $\vec{a}_i(t)$ is calculated and the resulting acceleration vectors and their magnitudes are summed up separately. This is continued until either all rules have been processed or the sum of the acceleration magnitudes exceeds a predefined maximum acceleration value $a_{\max}$. This process can be formalized as follows: Let a_B be the total magnitude of acceleration and $\vec{a}_B$ the total vectorial acceleration of the behaviour B

$a_B := 0$; $\vec{a}_B := \vec{0}$;
foreach rule i do
  $a_B := a_B + w_i |\vec{a}_i(t)|$;
  $\vec{a}_B := \vec{a}_B + w_i \vec{a}_i(t)$;
  if $a_B \geq a_{\max}$ break;
endfor

By additonally weighting each rule with a factor w_i the overall behaviour can be fine-tuned and new behaviours can be created or adopted to new types of flocks without the need to adjust parameters of individual rules.

5.2 Virtual Fish

For our experiments we created a simple virtual scene consisting of a representation of the EVE dome, a flock of 40 fish, five objects serving as stand-ins for the recognized visitors, and a *target fish* continously moving on a predefined motion path (see figure 5). For the behaviour of the fish, the following basic elements have been integrated:

attraction, i.e. the fish should move towards a specific recognized visitor,

avoidance, the fish should maintain a certain distance to the visitors,

fleeing, in the event of fast movements of the vistors or if the visitors are standing close together

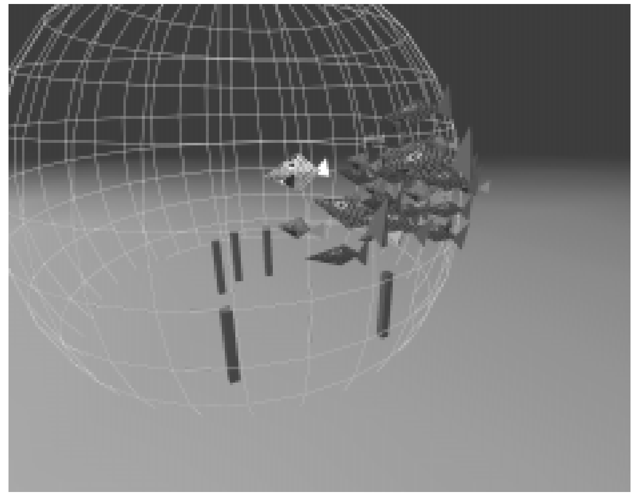

Figure 5: External view of the dome, the five stand-in objects, the keyframed fish and the flock of other fish currently following him

following the target fish, in the case that there is no visitor to be attracted to.

6. Results

To explore the proposed interaction paradigms of the prototype described in this paper we analysed differing numbers of people in real space and different modes of interaction between the virtual and real world. For the first tests described in this paper, we did not move the projected imagery, i.e. it remained a static window into the virtual world as shown in figure 4. This restriction allowed us to explore single ways of interaction separately without dealing with the complexity of a fully animated virtual world together with a (physically) moving window in the EVE environment. This has the disadvantage that the participants might not get fully immersed in the environment. However, the special setup of the EVE dome still gives a strong effect of immersion. By entering the darkened space, the user as well as the physical space get transported. We tested the environment with one, two, three, four and five people separately. Thereafter, we also made tests with situations of 5 to 10 people being present in the interaction space. The results presented here have been collected in an ad-hoc and intuitive manner.

First, we explored if it is possible to reinforce mutual awareness between human participants by having two participants in real space approaching each other. When the participants came close enough to form a single group in the environment, the fish were fleeing, which had an intensifying impact on the awareness of the people. On the other hand, we found it difficult to reproduce such results when more then three people where present in real space.

This might indicate that the interaction paradigm (fleeing fish, attraction) is too simple for a more complex situation. It might as well stem from the environment itself which is merely a dark and empty space with a strong focus on the projected visuals. Adding sound or adding other environmental artifacts, or, having a fully projected surround environment, e.g. by using four or more projectors, might overcome this limitation. Another problem we encountered, is the virtual scenary: The underwater world with fish influenced the expectations of the participants and limited the variety of interactions that people were willing to try. Either a more abstract, or a more human-to-human-like communication mode might deal with that. An advantage of the familiar underwater seaworld is, that the transportation, i.e. "the extent to which the participants perceive that they have left behind their local space and have entered into some new remote space" [2], did work very well, further enhanced by the influence of the EVE scenario itself as described above.

One limitation of the scenario is that the center part of the interaction area is covered by the projector, i.e. an area of 1.5 meters diameter is unaccessable and limits the movement to a circular space around that area. We did not encounter this as being a major limitation. However, we do believe that it influences the motion and, with it, the way the user is represented in the virtual world as well as the user's awareness of the environment. But, since we wanted to explore the applicability of different modes of interaction between real and virtual space, we did not consider this as to be a major problem.

As a general observation, we can say that the interactive nature of the environment made the visitors more active, i.e. it was unlikely that people would come in and stand still. Also, we found that feedback is an important issue in our environment: Because the tracking system currently only tracks up to five people, the participants had difficulties in linking their movements to events presented in the virtual scenary in situations where more than five people where present. To overcome this problem we used a simple graphical representation of the participants in virtual space, a thin pole, that became visible when being close to the projection screen. Hence, a participant could reasure himself of being tracked by walking towards the screen.

7. Conclusion

In this paper we have presented a prototype for exploring different modes of group interaction in real and virtual space. An exploration of situations with differing numbers of people present in the interaction space has been undertaken using an underwater virtual world. A more thorough investigation and interrogation of participants would give a more profound insight in the applicability of the proposed scenario. However, these first results clearly show some strengths and weaknesses of the approach and give hints for further investigation. For future work we would like to test the scenario with more human-to-human like communication models, which would require a more appealing virtual world as well as more complex behavior patterns of the virtual creatures.

Acknowledgements

The project presented in this paper has been supported by the ESPRIT long-term research project 25379 *eRENA* in the field of *Intelligent Information Interfaces*: i[3] (http://www.i3net.org).

References

[1] Batman Returns (1992), Motion Picture, Warner Bros., 1992

[2] Benford, SD, Brown CC, Reynard GT, et. al. (1996), "Shared Spaces: Transportation, Artificiality and Spatiality", Proc. ACM Conference on Computer Supported Co-operative Work (CSCW'96), Boston, pp. 77-86, ACM Press, 1996.

[3] Benford SD, Greenhalgh C.M, and Lloyd D (1997), "Crowded Collaborative Virtual environments". Proc. ACM Conference on Human Factors in Computing Systems (CHI'97), Atlanta, Georgia, US, March 22-27.

[4] Blase C, Morse M, Blunk A, et. al. (1997), "Hardware, Software, Artware : Confluence of Art and Technology. Art Practice at the ZKM Institute for Visual Media 1992 - 1997", Cantz Verlag, 200 pages + CD-ROM.

[5] Cliffhanger (1993), Motion Picture, TriStar Pictures, 1993

[6] Cruz-Neira C, Sandin D J, DeFanti T A, "Surround-screen projection-based virtual reality: The Design and Implementation of the CAVE", Computer Graphics Proceedings, Annual Conference Series, volume 27, August 1993, pp 135-142

[7] Goffman E (1971), "Relations in Public: Microstudies of the Pulic Order", Basic Books.

[8] Hoch M (1997), "Object Oriented Design of the Intuitive Interface", 3D Image Analysis and Synthesis, Proceedings, Erlangen November 17-19, Infix 1997, pp 161-167.

[9] Hoch M (1998), "A Prototype System for Intuitive Film Planning", Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), April 14-16, Nara, Japan 1998, pp 504-509.

[10] Krüger W, Fröhlich B (1994), "The responsive workbench", IEEE Computer Graphics and Applications, May 1994, pp 12-15.

[11] Merleau-Ponty M (1992), "Phenomenology of Perception", Routledge.

[12] Mulan (1998), Animated Motion Picture, Walt Disney Productions, 1998

[13] Musse S, Babski C, Capin T, Thalmann D (1998), "Crowd Modelling in Collaborative Virtual Environments", ACM VRST'98, Taiwan.

[14] Renault O, Magnenant-Thalmann N, Thalmann D, “A vision-based approach to behavioural animation”, Visualization and Computer Animation 1, 1990 pp 18-21
[15] Reynolds C W (1987), Flocks, Herds, and Schools: A Distributed Behavioural Model, Computer Graphics, Volume 21, Number 4, July 1987
[16] Stanley and Stella (1987), Computer Animation, Electronic Theater, SIGGRAPH 1987
[17] Stary C (1996), “Interaktive Systeme. Software-Entwicklung und Software-Ergonomie”, Vieweg Informatik/Wirtschaftsinformatik 1996.
[18] The Lion King (1994), Animated Motion Picture, Walt Disney Productions, 1994
[19] Tu X, Terzopoulos D, “Perceptual Modeling for Behavioural animation of fishes”, In Proc. 2^{nd} Pacific Conf. on Computer Graphics, Bejing, China, 1994
[20] Tu X, Terzopoulos D, “Artificial Fishes: Physics, locomotion, perception, behaviour, Computer Graphics Proceedings, Annual Conference Series, volume 28, August 1994, pp 43-50