CROWDED COLLABORATIVE VIRTUAL ENVIRONMENTS

Steve Benford, Chris Greenhalgh and David Lloyd
Department of Computer Science The University of Nottingham
Nottingham NG7 2RD, UK
+44 115 951 4203
{sdb, cmg, dxl}@cs.nott.ac.uk

ABSTRACT

We introduce a framework for supporting crowds of participants in collaborative virtual environments (CVEs). The framework is realised as an extension to our previous spatial model of interaction and aims to provide greater scalability and flexibility for communication between the inhabitants of virtual worlds. Our framework introduces an explicit crowd mechanism into CVEs in order to support the formation and activation of different kinds of crowd with different effects on mutual awareness and communication (achieved through the use of aggregation techniques combined with awareness adaptation). We present a demonstration application called the Arena - a shared space for staging on-line performances in front of a live audience.

Keywords

CSCW, Virtual Reality, Crowds

INTRODUCTION

Our aim is to introduce and demonstrate a framework for supporting crowds in collaborative virtual environments (CVEs). This framework is intended to support the development of more scaleable CVEs than has been possible before whilst maintaining a high degree of flexibility with respect to communication and participation. More specifically, we anticipate future CVEs that will support hundreds or even thousands of simultaneous participants engaged in real-time graphical, audio, textual and video communication with one another. Potential application areas of such technology include conferences, lectures, training and simulation. However, the key application area considered in this paper is that of entertainment and in particular, audience participation in on-line events of all kinds (music, drama, games, sports, exhibitions, parties etc.). Indeed, when coupled to development of Interactive TV platforms (i.e., fibre to the home connected to set-top boxes and advanced games consoles), our framework might eventually help in the creation of new forms of socially active mass entertainment - *Inhabited TV* if you like.

Our paper focuses on the specific issue of introducing a flexible and dynamic notion of crowds into CVEs. A crowd is an abstraction of a group of objects (typically people, but

perhaps also agents, artefacts and information) which allows them to be treated as a whole in some circumstances (e.g., for interaction at a distance) but as individuals in other circumstances (e.g. for interactions between members). Our framework provides CVE application developers with system level mechanisms to deal with the following issues:

- what effects can crowds have on awareness and communication in CVEs and how might they be represented in different media?
- how can we support individuals in joining and leaving crowds and in managing their interaction with them?
- what mechanisms can be provided to support the formation and disbanding of crowds?

Our approach is therefore concerned with the *mechanics* of managing awareness and communication among *human participants* in crowded CVEs. This contrasts with previous research into crowd behaviour modelling and simulation for applications such as safety analysis (e.g. [5]). Our work is concerned with participation in crowds as opposed to simulation of them and our approach anticipates that higher-level crowd behaviours will emerge if appropriate underlying mechanisms are provided. However, as we shall discuss later on, it may be beneficial to combine the participation and simulation approaches in the future.

At a technical level, our notion of crowds is realised as an extension to our previous spatial model of interaction [2]. This extension, called third party objects, combines spatial awareness mechanisms with information aggregation algorithms resulting in a flexible way of structuring shared virtual spaces which can naturally be mapped onto efficient underlying network protocols (e.g. the use of hardware multi-casting). Thus, our framework for realising crowds is not only concerned with representational issues (i.e. the user interface); it reaches right down into the network in order to manage the flow of information between different participants.

The remainder of this paper is structured as follows. The following section considers the motivations for introducing an explicit crowd mechanism into CVEs. We then introduce the underlying mechanism of third party objects as an extension to the spatial model. Following this, we discuss how third party objects can be used to create a variety of different kinds of crowds in CVEs. Finally, we present a demonstration application, implemented using the MASSIVE-2 system, called the Arena, which combines static and dynamic crowds with a structured space so as to create a venue for on-line performances.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

MOTIVATIONS FOR A CROWD MECHANISM

We begin by considering the motivations for introducing an explicit crowd mechanism into CVEs (i.e. for providing an additional level of technical support for crowd representation and management beyond just the ability for large numbers of individual participants to gather in one place).

Our first motivation is scale. Current CVEs support at most a few tens of simultaneous users. Although there are some exceptions, such as the NPSNET battle simulator which claims of the order of a hundred simultaneous participants [4], these have generally only been achieved by relying on there being highly predictable behaviours for objects (e.g. the movements of ships, tanks and missiles) and by reducing the potential for communication between participants (especially with regard to real-time audio). There are several dimensions to the problem of scale. First, can the network exchange rich information about many simultaneous participants sufficiently quickly and reliably so as to engender a sense of co-presence? This is a major limitation for systems based on unicast network protocols but will eventually become an issue even for systems which utilise more network efficient multicast protocols. Second, assuming that the network can deliver this information, can the computers involved process and render it? Third, even if the combination of network and computer can deliver and display the information, can individual participants make sense of it? (e.g., could one make sense of a thousand people speaking at once or view a thousand detailed embodiments at the same time?).

This leads us to our second major motivation, that of legibility and structure. Complex environments might be made manageable for users by introducing additional structures which group objects together, provide aggregate views of them and which might then be unfolded at a later time (e.g. on entering them). Some initial evidence for this is provided by recent work on enhancing the legibility of information visualisations through the introduction of districts and related features such as landmarks, edges, paths and nodes [3].

Our final motivation involves generating a sense of mass-presence for specific classes of application. Audiences play a key role in various real-world events such as a theatre, concerts, sports, exhibitions, fairgrounds, trade shows, rallies, demonstrations and even town centres and public spaces. A crowd mechanism might therefore enhance or even create a sense of audience presence and participation in CVEs and might open up opportunities for new forms of social interaction.

THIRD PARTY OBJECTS AND AWARENESS

Our framework for realising crowds in CVEs is based on our previous spatial model of interaction and in particular, on a recent extension to the model called *third party objects*. The spatial model defines mechanisms for the management of awareness and communication in shared virtual spaces [2]. To briefly summarise, the model considers a number of objects in a shared virtual space communicating through different media (e.g. audio, graphics, text and video). Instead of having each object transmit its information to all other

objects, basic connectivity is enabled through the concept of aura - a volume of space that delimits the presence of an object in a given medium. Thus, aura collisions lead to connections being established. The quality of any information which is subsequently transmitted (e.g. the volume of audio or the level of detail of graphics) depends upon the level of awareness that the observer has of the observed (awareness is a quantifiable concept in the model). This in turn is negotiated through focus and nimbus. Focus is a sub-space representing the attention of the observer and nimbus is a sub-space representing the projection of information by the observed. The observer's awareness of the observed is then some function of the observer's focus on the observed and the observed's nimbus on the observer. Aura, focus and nimbus may be medium specific, multi-valued, dynamically changeable and need not be strictly spatial in their definition (i.e. they need not be simple discrete volumes of space).

This basic model is limited in two main ways: its bi-lateral approach to interaction does not easily scale to large numbers of participants and, beyond a limited concept of adapter objects, it provides no support for introducing contextual factors into awareness negotiations (e.g. for representing the effects of the environment within which the observer and observed find themselves). Third party objects have been introduced in order to address these problems. A third party object is an independent object which affects the awareness between other objects. The basic scenario (in any given medium) is therefore now one of three objects, each with individual awareness relationships to the others (see figure. 1).

Figure 1: Introducing third party objects

Three general points should be noted about third party objects from the outset. First, all aspects of their operation as described below may be medium specific (e.g. they may operate differently in the audio medium than in the graphical, textual or video media). Second, as they are objects in their own right they may be embodied, mobile or fixed, dynamically or statically created and may apply their effects recursively to one another. Third, although they are most often described in spatial terms in this paper, they may operate according to non-spatial awareness relationships (i.e. one could define them in terms of arbitrary attributes of objects).

There are three key aspects to third party objects: their effects (i.e. what they do to awareness relationships between other objects), their activation (i.e. when and how these effects are brought into operation) and their creation and

destruction (i.e. how they are introduced to and removed from the environment). We now consider each of these.

The effects of third party objects

Third party objects can have two general kinds of effect on awareness (see figure 2) which may be applied in different combinations across different communication media.

Figure 2: the effects of third party objects

Adaptation involves the manipulation of existing awareness relationships between objects. In this sense, third party objects are a generalised notion of the adapters that were defined in the initial spatial model. These manipulations include attenuation (e.g. a barrier between objects) and amplification (e.g. increasing awareness between people who are accessing a common object).

Secondary sourcing involves the introduction of new indirect awareness relationships between objects in order to enable new transformed flows of information between them. Typically, secondary sourcing involves the consumption of information from an external group of objects, its transformation in some way and its subsequent re-transmission in order to provide a common view of the group. Various filters may also be applied at different stages of this process in order to reduce level of detail or to select key information. At the heart of secondary sourcing lies the problem of creating a single aggregate view or stream of information from a number of sources. We propose that there are three approaches to the aggregation problem:

- *selection* - switching between individual views or streams in some way (e.g. round robin, "loudest wins" etc.)
- *combination* - the direct composition of a new view from existing views (e.g. tiling multiple video windows).
- *abstraction* - generating an entirely new representation based on statistical information describing the sources (e.g. mapping the number of sources into the size of the representation, the level of their activity into colour etc.).

We will provide concrete examples of these classes of transformation when specifically discussing crowds later on.

The activation of third party objects

Next we consider the circumstances under which different combinations of these effects are applied. The activation of third party objects is based on the awareness relationships between the third party and the other objects involved. Thus, referring to figure 1, the activation of T depends on four possible awareness relationships: T's awareness of A and B respectively and their awareness of it. In figure 3 we identify three particularly interesting cases from among the various possibilities.

Figure 3: Activating third party objects

a) *membership* - cases where the third party is activated according to how aware it is of other objects. This is analogous to the idea of membership (i.e. the third party's awareness of an object expresses the degree of membership of that object). For example, one might become a member of a room by crossing its boundary.

b) *sharing* - cases where the third party is activated according to how aware other objects are of it. This is analogous to the idea of objects sharing the third party in some way and consequently, it having an effect on their mutual awareness

c) *hybrid* - cases where the effects of the third party depend on how much one object is aware of it and how much it is aware of another object. This turns out to be a useful case for crowds (see below).

Creation and destruction of third party objects

Given that they are independent objects in their own right, third party objects might be created and destroyed in any of the ways associated with normal objects. Thus, they might be static (i.e. a permanent part of a given environment) or dynamic (i.e. created or destroyed on the fly). It is also necessary to consider the issue of who creates and destroys them. Again, there are three cases to consider:

- *the system* - third parties might be automatically created or destroyed by the CVE operating system according to network and system load factors.
- *the application developer* - might act as an architect, defining the structure of an environment through third party objects in order to afford different modes of participation in a specific event. Thus, in this case, different third parties might be associated with different architectural units or features (e.g. making use of a library of different boundary types).
- *the end user* - end users might request the creation and destruction of third party objects themselves (e.g. dynamically creating a private sub-discussion).

Examples of third party objects

Before going on to consider how they can support the introduction of crowds into CVEs, we first briefly list a number of representative broader applications of the third party mechanism:

- *bounded rooms and buildings* - the creation of membership activated regions of space whose boundaries might have different effects across different media and which might be hierarchically combined to form arbitrarily complex virtual buildings.

- *floor control objects* - membership based third parties which attenuate direct awareness between members and replace it with a secondary source view based on an appropriate selection algorithm.
- *common foci* - objects which would amplify awareness between people who were using them (e.g., shared designs or information items in a shared visualisation).
- *group vehicles* - mobile third parties which would be capable of steering a group of people through a world and which would provide a shared environment for experiencing it.
- *load management groups* - a possibly invisible system level third party used to dampen down system load (i.e. reduce network traffic by replacing many individual awareness relationships with a single aggregate view).
- *data districts* - a cluster of data in an information visualisation that would be seen as an aggregate from a distance but could be unfolded if required. The district could also provide awareness of the presence of other individuals within that district.

This concludes our general introduction to third party objects in the spatial model. The following section now considers how they may be used to support crowds.

CROWDS AS THIRD PARTY OBJECTS

Crowds can be realised as a specific class of third party object which support potentially large groups of people (and possibly agents and other objects) in CVEs. We now consider the following aspects of crowds as third party objects: effects on awareness, representation, activation and membership, creation and destruction, mobility and generation and behaviour.

Effects on awareness?

We propose that, in general, crowds should have an asymmetric effect on awareness.

From the "outside", such as when perceived from a distance or from the perspective of a non-member, individuals within the crowd are hidden (adaptation of existing awareness) and, instead, are replaced with an aggregate view of the whole crowd (secondary sourcing).

On the "inside", individuals are able to interact with each other in the normal way (i.e. through their respective foci and nimbi). Typically, they will also be aware of those outside of the crowd on an individual basis. Indeed, in some cases the crowds may even amplify the awareness that those inside have of those outside such as in the case where people in an audience wish to be maximally aware of the performers at an event. Thus as a member of a crowd, I can communicate with nearby people who are also in the crowd, and can perceive those outside of the crowd in detail, although they may only perceive me in as much as I contribute to the aggregate view of the crowd.

Of course, given that they are third party objects, crowds may contain other crowds, thereby applying these effects recursively to one another.

Crowd aggregations

Developing appropriate aggregation techniques is clearly a critical issue for building useful and convincing crowds. Although the details of particular techniques will be application specific and are therefore beyond the scope of this paper, this section does propose an initial classification of approaches according to the two dimensions of general approach (i.e., selection, combination or abstraction as identified above) and the medium involved.

	select	combine	abstract
graphics	floor control/ scheduling policies	combine into hierarchical object	visualise stats (e.g. <i>Paradise</i>)
audio	as above	multiplex?	audio mixing
text	as above	digestifying	automatic abstracting
video	as above	tiling	morphing?
events	as above	multiplex	counting (e.g., <i>KMI Stadium</i>)

A summary of crowd aggregation techniques

Each entry in the above table refers to a possible technique for aggregating many sources in a given medium into a single output. This aggregate may then be filtered and translated to create a final representation in some (possibly other) medium. Note that we have defined an additional medium called events, which covers application defined events and protocols ranging from the general presence and location of objects through to specific events such as pressing an "applaud now" button on a user interface.

Several of the entries in this table suggest the application of well known techniques. For example, selection across any of the media could utilise a range of floor control and scheduling algorithms such as round robin, random selection, most active, currently active and so forth. It may also be possible to adapt existing text manipulation techniques for the text medium, including digestifying as used on newsgroups and automatic abstracting and indexing. Combination in the graphical medium might involve the creation of a new super-object whose parts are defined by the individual sources. Although not inherently scaleable in itself, this approach could be combined with automatic level of detail techniques. Video tiling provides a way of combining multiple video views into a single view, although this approach would appear less scaleable.

Two current systems can be associated with two of the entries. The Paradise project has been exploring the use of graphical abstraction techniques to produce aggregate views of groups of objects for use in distributed simulation [6]. Their approach generates statistics about graphical objects (e.g., the number present, their mean location and spread) and these aggregations are then used to generate graphical representations. Alternatively, the KMI stadium, developed by the UK's Open University, uses a threshold technique where a sufficient frequency of applause events coming

from different participants triggers a crowd applause event (which is eventually translated into the play back of an audio sample). This approach could also drive the playback of video samples or graphical animations.

Two further entries in table are worthy of special note as they would appear to pose great difficulties. These are the combination of audio signals and the abstraction of video signals. In the everyday world one never perceives a combination of audio signals that are not already superimposed. Other than having two ears (which allows some spatial separation), there is therefore no general way of distinguishing an individual audio signal from among a combination (especially a large combination). Conversely, although one might somehow use morphing techniques to blend video images together, in the real world one rarely perceives superimposed visual images (with the limited exception of partially transparent surfaces) and so, as humans, we have no apparatus for usefully dealing with an abstraction of visual signals. In other words, we suspect that the nature of our audio and visual perception will make the development of useful audio combination and video abstraction techniques especially difficult.

Activation and membership

We identify two general styles of activating crowds. First is a class of crowd whose effects are solely triggered by membership (i.e. on the crowd's awareness of other objects - figure 3 case a). Thus, the crowd can determine which objects are members and which are not and typically operates such that: members are normally aware of both members and non-members; non-members are normally aware of non-members; but that non-members are only aware of members through an aggregate view. Membership might be directly mapped onto spatial attributes such as proximity (i.e. one becomes a member of a crowd by crossing its boundary), but could potentially involve other non-spatial attributes. Our second class of crowd is based on the hybrid approach (figure 3 - c). This operates as for the membership based example with one key difference. Whether a non-member perceives individual members or not depends on how aware they are of the crowd. This allows people outside of the crowd to "unfold" it just by looking at it hard enough, even if they are not themselves members.

In essence, both classes of crowd use the idea of membership to determine whether or not an object contributes to the aggregate view. The difference is the basis on which that view is perceived by others - according to how aware they are of the crowd or vice versa.

It should also be noted that, just as awareness is potentially a multi-valued quantity in the spatial model, so then is the idea of membership. One can extend these examples to include multiple levels of membership which activate different combinations of effects.

Mobility

Crowds might be fixed or mobile objects in a virtual environment. Fixed crowds might be attached to various features of an environment such as a bank of seating in an auditorium. Mobile crowds introduce the further issue of the

relationship between crowd and individual navigation. There are several possibilities here:

- crowds might follow their members in some way (e.g. remaining at their mean position).
- crowds might navigate on behalf of their members (e.g. group vehicles which take many people on a shared ride).
- a combination of these where both the crowd and its members exert a pull on one another (i.e. where the crowd would follow its members but also slow them down if they moved too far away).

Creation and destruction

Crowds can be created and destroyed at all three of the levels identified previously, by the system, application developers and the participants themselves. However, the first two are of particular interest.

The system might automatically introduce crowd objects into an environment in order to manage system load by reducing the number of individual awareness relationships to be considered. This might be done to handle a sudden mass of new arrivals into an environment. Our experiences with the MASSIVE system suggest that people often move to new places together (e.g. a group of people might leave an environment together at the end of an event) and that such movements can cause intense bursts of network traffic (movements need to be conveyed to other objects, descriptions of new worlds need to be transferred across the network and so forth). The temporary introduction of a crowd object might help smooth this process and the crowd could then be removed once the major movements had settled down and a new phase of activity was underway.

Adopting a longer term view, the structure of a virtual environment might be used to predict where crowd objects could usefully be located. For example, in a persistent environment such as a virtual town, it may be useful to associate a crowd with key locations such as squares, major pathways and junctions. Indeed, recent research into the structure of virtual environment inspired by urban planning theory has pointed towards there being a direct causal relationship between the structure of a virtual environment, the navigation strategies employed by its inhabitants and the places where social encounters are likely to occur [3]. In short, given knowledge of the former two, it may be possible to predict the latter. Such knowledge would suggest in advance where crowd objects might most beneficially be introduced into the environment.

Application developers might introduce crowds in the form of different architectural features in the design of a virtual environment in order to pre-configure the communications that might occur within it (e.g. as a bank of seating in an auditorium, a hallway in a building etc.). This suggests the use of our framework to create a library of architectural components with associated crowd properties.

Agents, generation and behaviour

As a final note, it may be useful to be able to automatically generate or simulate crowds for virtual environments, even where there are relatively few human participants. One approach to this might be the use of autonomous agents

who join crowds and carry out simple actions in response to an event, nearby humans or even each other. Indeed, the use of such agents is going to be essential for initial testing of systems (see below). The automatic generation of crowds raises the question of whether it might be possible to introduce, control or re-inforce crowd behaviours among human participants. This suggests combining the kind of crowd mechanisms proposed in this paper with other crowd simulation techniques (e.g. [5]).

THE ARENA - A DEMONSTRATION OF CROWDS

In this section we present a prototype application of our framework. This prototype, called the Arena, realises a virtual space for on-line performance to a live audience. In this case, the performance is a simple interactive ball game between several participants. The Arena demonstrates the following features of our framework:

- the use of third party objects to support crowds.
- example graphics and audio aggregation techniques.
- fixed and mobile crowds.
- examples of crowds activated solely on the basis of membership and also "level of detail" crowds whose effects depend on an observer's awareness of them.
- hierarchical combination of several crowds with a bounded space so as to create a virtual space which offers its inhabitants different modes of participation depending upon where they are located within it.
- the use of agents which exhibit simple behaviours in order to generate crowds for testing purposes.

MASSIVE-2

The Arena has been implemented using the MASSIVE-2 system, a general purpose CVE which supports the extended spatial model of interaction and provides a platform for creating different kinds of third party object [1]. Like its predecessor, MASSIVE, the system allows users to interact using graphics, text and audio media.

MASSIVE-2 relies heavily on the use of multicast networking protocols in order to achieve efficient networking. Objects which are members of a third party (e.g. the members of a crowd) send information to one or more multicast groups associated with that third party. New objects which become members of the third party are automatically invited to join this multicast group. A further multicast group then allows new observing objects to request a state snapshot from all of the transmitting objects (i.e. there is a back-channel which allows arriving objects to request to catch up with the current state of play).

As third party objects move around an environment so they may recursively swallow each other up to form a mobile hierarchy of spatial regions, crowds and other kinds of group (e.g. crowds may enter rooms, vehicles may pass through regions etc.). Beneath the surface, this is mapped onto a dynamically evolving hierarchy of multicast groups. It is this highly dynamic use of multicast that allows MASSIVE-2 to achieve both scalability and flexibility.

Implementation of the Arena

We now describe the implementation of the Arena as an application of MASSIVE-2. The Arena combines several kinds of third party object: a bounded space and two kinds of crowd, static and mobile. The overall design of the Arena is shown in figure 4.

Figure 4: Overview of the Arena

The Arena is housed in a bounded space - a static, graphically embodied third party object whose effects are to completely attenuate awareness between members and non-members. Membership is achieved simply by crossing its boundary. Thus, those on the outside (participants F and G in figure 4) cannot hear or see what is happening on the inside and vice versa.

Within the Arena space there are two further third party objects, both of which are static crowds. These are used to locate the opposing supporters (the Reds and the Blues). They are membership based crowds which are created when the application is initialised and whose position is fixed. They support two kinds of aggregation algorithm:

- in the graphical medium, each crowd provides an aggregate embodiment whose size increases according to the (square root) of the number of its current members. Thus, as more people enter the crowd, so the size of its aggregate view increases. Our current crowd embodiment is similar in appearance to that of the individual participants so that the crowd takes on the appearance of a giant sized person.
- in the audio medium, each crowd sums the audio emissions of its members (i.e. mixes them together) and applies a low pass filter so as to distinguish the tone of the crowd from the tone of individual speakers.

The area between the static crowds represents the performance space. For the performance we have created a simple ball game (similar to the classic computer game Pong) where several participants bat a graphical ball backward and forwards over a net. Of course, they can talk to each other as well.

The presence of the two crowds inside of the Arena space gives rise to several different modes of awareness between its inhabitants. Referring to figure 4, C and D, who are in the same crowd, have normal mutual awareness of one another, have full awareness of the performers, B and A, but only perceive E in as much as they contribute to their crowd aggregation. As performers, A and B have full mutual awareness but only perceive the audience members, C, D and E through their respective crowd aggregations.

One or more dynamic crowds can be created outside of the Arena in order to handle the sudden outflow of participants at the end of the event. Although they use the same aggregation techniques, these differ from the static crowds inside the Arena in several respects:

- they are mobile such that, at any given time, they assume the mean position and orientation of their members.
- they can be dynamically introduced or just left waiting around outside of the Arena ready to "pick up" passers by.
- they work as level of detail crowds - i.e. whether the aggregate view of individual members are perceived by a non-member depends upon their level of awareness of the crowd. Thus, even at a distance, one can unfold these crowds by raising one's focus sufficiently.

In order to test and demonstrate this application we have also implemented some simple agent based crowd members who may occupy the Arena along side its human participants. These have been given the ability to follow various pre-defined paths through the environment (e.g. they can be sent off into or out of the Arena as a group) and also voice some simple chants (i.e. play back pre-recorded audio samples). It should be noted that, from a network point of view, these Agents operate just as human participants (i.e. they are separate entities in their own right who generate traffic according to their actions).

Figure 5 shows how the Arena appears from within the performance space and includes two red players, two blue players, the red and blue crowds (viewed as aggregates) and a scoreboard. Figure 6 shows how the Arena appears to a member of the red crowd. In this case, we can see several nearby individual members of the red crowd, the performers and a secondary source view of the more distant blue crowd.

SUMMARY

Our paper has been concerned with supporting crowds of participants in collaborative virtual environments. Specifically, we have introduced a framework for reasoning about and developing different kinds of crowds with different effects on spatial awareness and communication. This framework is based on an extension of our previous spatial model of interaction called third party objects. The key points of our framework are:

- crowds can play a significant role in configuring communication and awareness. In particular, they may introduce aggregate views of their members.
- the use of awareness to activate crowds supports crowds whose effects are based on membership and also on an observer's level of focus.

- crowds may be mobile or fixed and may be created and destroyed at the system level, the application development level or by participants themselves.

We have also presented a demonstration of our framework based around an application called the Arena - a virtual environment for staging different kinds of performance in front of an audience - which has been implemented using the MASSIVE-2 platform.

Having established this general framework, further research is now required. First, research is needed into alternative and more powerful aggregation techniques for different communication media. Second, it may be beneficial to create a "library" of higher level building blocks for creating different kinds of crowds. Such a library might represent a set of standard architectural building blocks for virtual worlds that could easily be accessed by application developers without the need for extensive programming. Third, trials are needed with significant numbers of participants (possibly combined with agents) in order to assess both the human and system implications of this approach (e.g. under what circumstances do people experience a sense of crowd presence and what is the impact on networking and computation). Finally, greater consideration needs to be given as to how this kind of framework might be realised using future public delivery platforms (e.g. a combination of cable to the home linked to so called set-top boxes or games consoles). Given that these various issues can be addressed, we anticipate that our framework may become a significant component in constructing mass participation social electronic environments and that, in the future, such technology could have widespread applications in many areas of life including arts, entertainment, leisure and culture.

REFERENCES

- 1 Greenhalgh, C. M. and Benford, S. D., *Introducing Regions into Collaborative Virtual Environments*, Internal report available from the authors of this paper (submitted to IEEE VRAIS'97).
- 2 Greenhalgh, C. M. and Benford, S. D., *MASSIVE: A Virtual Reality System for Tele-conferencing*, *ACM Transactions on Computer Human Interaction (TOCHI)*, 2 (3), pp. 239-261, ACM Press, 1995.
- 3 Ingram, R. J., Bowers, J. M. and Benford, S. D., *Building Virtual Cities: Applying Urban Planning Theory to the Design of Virtual Environments*, *Proc. VRST'96*, Hong Kong, July 1996, ACM Press.
- 4 Macedonia, M. R., Zyda, M. J., Pratt, D. R., Barham, P. T. and Zeswitz, S., *NPSNET: a network software architecture for large scale virtual environments*, *Presence*, 3(4), MIT Press, 1994.
- 5 Sime, J. D., *Crowd Psychology and Engineering: Designing for People or Ballbearings? Engineering For Crowd Safety* (Smith R.A and Dickie J.F., eds), pp119-131. Elsevier 1993.
- 6 Singhal, S. K. and Cheriton, D. R., *Using Projection Algorithms to Support Scalability in Distributed Simulation*, *Proc. 1996 International Conference on Distributed Computing Systems*, IEEE, 1996.