2009

# Modelling Groups of Plausible Virtual Pedestrians

Christopher Peters
*Coventry University*

Cathy Ennis
*Technological University Dublin*, cathy.ennis@tudublin.ie

Carol O'Sullivan
*Trinity College Dublin, Ireland*, carol.osullivan@cs.tcd.ie

## Recommended Citation

Peters, C., Ennis, C. and O'Sullivan, C. (2009) Modelling Groups of Plausible Virtual Pedestrians. *IEEE computer graphics and applications*, Year: 2009, Volume: 29, Issue 4. doi:10.1109/MCG.2009.69

# Modelling Groups of Plausible Virtual Pedestrians

Christopher Peters [1,2†], Cathy Ennis [2] and Carol O'Sullivan [2]

[1] Department of the Digital Environment, Coventry University, Coventry, United Kingdom
[2] Graphics, Vision and Visualisation Group, Trinity College, Dublin, Ireland

## Abstract

*In the domain of real-time urban vistas, a modern vision of a vast populace inhabiting a sprawling metropolis is replacing that of the ghost-towns of past. Working towards this ideal, we describe the methodology we are using to model dynamic crowd scenarios and an exemplar based on it. Video corpus is consulted for informing the modelling process and perception based evaluations are conducted on the resultant scenarios with the goal of improving the visual plausibility of the crowd, rather than correctness of simulation. Using this methodology, the exemplar focuses on modelling small groups of pedestrians. While it is generally assumed that the addition of groups can improve the plausibility of crowd scenarios, little is known about how humans perceive the results. We shed light on these issues, demonstrating the practical application of the methodology in a real-time crowd animation system.*

## 1. Introduction

Crowd simulation is enjoying considerable success in a number of applied domains, most notably in evacuation scenarios where simulated crowd behaviours can help to improve the safety of interior building designs. However, not all applications involving virtual populace have the over-arching goal of realistic simulation; in many cases, it is necessary that the crowd merely be *perceived by the viewer* to be realistic. In many of the latest movies or video games involving large numbers of virtual actors, liberties can be taken in the display of those far away or otherwise obscured from the eye, if it does not diminish the viewing experience noticeably. For example, simulation level-of-detail [OCV*02] may be reduced or collision avoidance calculations forgone in order to allow a larger crowd to be simulated, or enhanced behaviours for individuals deemed most likely to occupy viewers' attention.

In doing this, a guiding methodology is desirable to ensure the system is developed with the proper end goals in mind. After all, when defining pedestrian behaviour, it is tempting for us as humans, to presume we know exactly how our species behaves. Thus, a designer might attempt to define behaviours based solely on intuition without recourse to real-world examples, or may consider the issue purely as an engineering challenge, where success is determined only by the fulfillment of technical constraints, such as proper collision testing. These situations are likely to produce results akin to a *clockwork* crowd simulation, where everything behaves the way the designer feels it *ought* to, but which may bear little resemblance to the real-world situation or meet viewers' expectations (although we note with interest that these last two aspects need not always correspond either).

In order to meet these challenges, we have adopted, as part of our efforts in constructing a crowd simulator called *Metropolis* (see Figure 1), a guiding methodology for a corpus-based, perceptual approach to crowd modelling. This article describes the methodology and in particular, an exemplar for group modelling, as follows: In the next Section, we describe important related work and provide context for our own. We provide details of the overall approach in Section 3 and proceed to describe a concrete example of the analysis (Section 4), synthesis (Section 5), and evaluation (Section 6) of group behaviours. We conclude by discussing implications of the work in Section 7.

---

[†] Research conducted while at Trinity College Dublin

Manuscript submitted (10/2008).

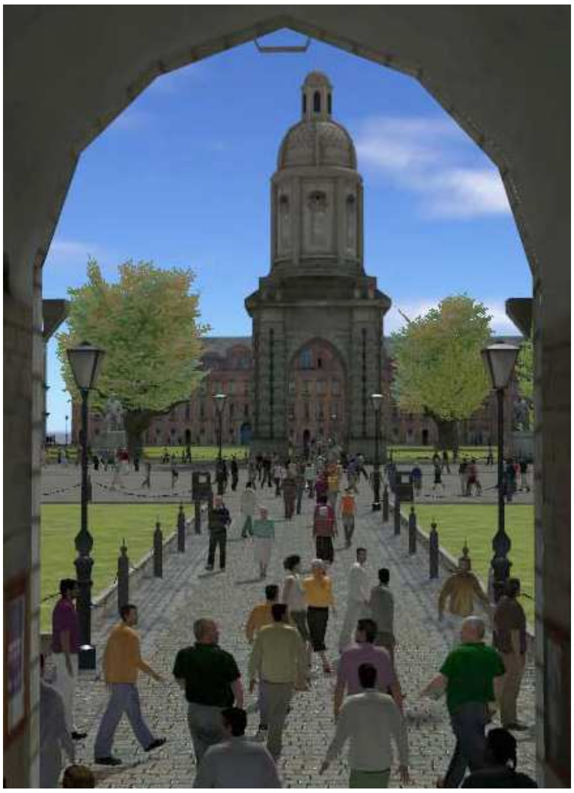


**Figure 1:** *A screenshot from the Metropolis project, where the aim is to recreate a populated virtual city.*

## 2. Related Work in Crowd Animation

A key challenge in human crowd modelling has been to impart human-like individuality to members of the crowd, in order to move away from the appearance of a particle system [HBJW05] or behaviours ideally suited to flocks or herds of animals [Rey87]. There have been many impressive attempts to do this by endowing individuals with simplified human-like models of perception, decision-making and other capabilities (see for example [ST05] [PAB07]) or producing variation in the crowd based on personality factors [DAPB08].

### 2.1. Group Modelling

Groups can be defined as semi-permanent collections of individuals sharing navigation goals who attempt to maintain spatial cohesion with each other, reflecting relationships defined in the underlying simulation. The group concept allows for easier management of a crowd hierarchy by creating an intermediate level of representation between the individual and the crowd. Most work in this area has been concerned with how groups are modelled in the underlying simulation, or related issues such as how groups can form based on individuals' traits [MT01]. To be differentiated from this, is the aim of providing the *impression* of pedestrian group structures to the viewer, for example, small groups of friends or acquaintances seen walking together in formation down the virtual street. In [Hos02], pairs of individuals maintain formation by making navigation decisions based on a voting process: each vote is cast according to the fulfillment of a specific constraint, such as avoiding an obstacle, keeping a separation distance from neighbours or maintaining formation.

### 2.2. Annotation for Informed Environments

In addition to group modelling, [Hos02] also defines skeletal splines aligned with walkways, called *ribbons*, which provide explicit information for groups to use, such as the two major directions of travel on the walkway. Such navigation aids, placed inside the environment description, fall into a general category of annotations [DHR98] added by the designer during the construction process. The resulting environment description contains not only rendering data, but also geometric, semantic and spatial partitioning information for informing pedestrian behaviour [FBT99],[TD00], thus transferring a degree of the behavioural intelligence into the environment.

### 2.3. Corpus-based Reconstruction

More recently, case-based approaches have used a real corpus. Lerner et al. [LCD07] constructed a database of example trajectories from a video of pedestrian motions and, at runtime, queried it for the closest matching example, given the current situation, to be used as the new trajectory. Lee et al. [LCHL07] recorded crowd videos from an aerial view in a controlled environment and manually annotated video frames with environment information. Multiple individuals were tracked in a semi-automatic manner to determine their trajectories and inform a pedestrian movement model, which created a crowd similar to that observed in the video. Both cases involve manual or semi-automated annotation of a crowd corpus, using the results to directly generate behaviours, although perceptual studies assessing the impact of results on viewers were not reported.

## 3. Overview

As is evident from a review of the literature, crowd modelling can be regarded as covering three general tasks: (1) *a priori* analysis of a crowd corpus to inform the modelling process, (2) synthesis of crowd behaviour using a simulation model and (3) *a posteriori* evaluation of the resultant crowd behaviours. Our methodology highlights the necessity for all three of these tasks to be adopted when creating models.

First of all, this ensures a real-world grounding exists for the model. A glance at real crowd corpus often reveals strange behvaiours, easily neglected if a corpus is not consulted. Only more recently [LCHL07] [LCD07]

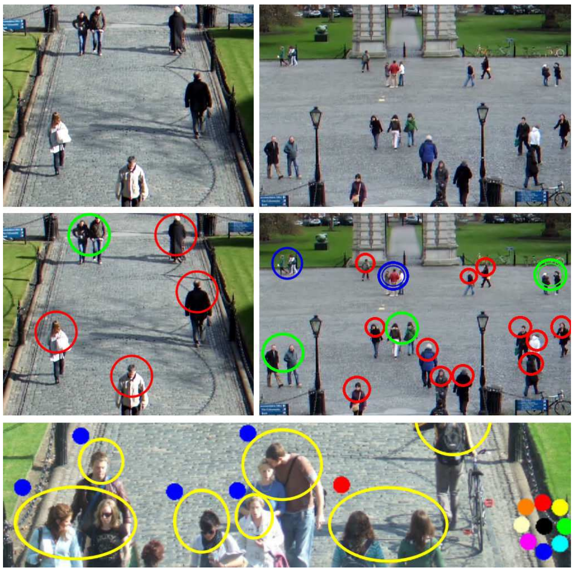


**Figure 2:** *Frames from our video corpus (top row) of two prototypical walking areas and corresponding results of annotation (middle row), in this case, for the purpose of assessing the number and size of groups. Here, groups are circled in red, green or blue according to the size of the group; static groups receive double circles. More complex forms of annotation (bottom row) are also possible, here for group orientation.*

[EPO08] has such data been reported to have been formally consulted as part of the modelling process. We discuss corpus analysis for the specific case of group modelling in Section 4.

Secondly, perception based evaluation, that is, establishing the impact of synthesised crowd behaviours on viewers' judgements of realism, ensures that updates to the model reflect improvements over previous versions. Overall, this suggests a process of iterative corpus analysis, model enhancement and evaluation, each stage focused on specific factors of interest at a global (crowd and scene) or local (individual and group) level. We provide an example of a perceptual evaluation for groups at the global level in Section 6.

Finally, the synthesis task is also crucial. An important element here is consideration of the environment in which the behaviours take place. Unlike simplified grid-like cities based on a block structure, our environment is based on a European-styled city, containing many twisting paths of variable width and arbitrary obstacles. This environment must be tagged with behavioural information in order to inform the simulation techniques used for group control techniques, as described in Section 5.

## 4. Corpus Analysis

While the overall appearance and distribution of a crowd in a real scene may seem perfectly natural, a focus on individual behaviour can highlight seeming peculiarities. It is not unusual in real-life, for example, to witness individuals or groups inexplicably conducting a sudden u-turn, walking back in the direction from which they came. While this could be regarded as irregular according to one's own view of what ought to constitute pedestrian behaviour, experience shows it to be as much a part of real pedestrian behaviour as those which we regard as being more normal. Corpus analysis helps us to better understand the way things are, to become aware of what actually constitutes real-world behaviour.

A corpus may consist of photographs of crowd scenes, or more desirably, video. We have collected a video corpus of two prototypical areas in Trinity College - an open area offering freedom of movement, and a more restricted corridor (see top row of Figure 2). Even when an analysis of static scenes is required, a video corpus is advantageous as it allows the viewer to better gauge the composition and dynamic of groups, which may be difficult to discern from a single image. There are different ways in which the same corpus can be analysed and interpreted: usually, only those factors of relevance to the current study are considered rather than attempting to consider all possible aspects of the corpus at once. This is particularly relevant when the process is conducted manually, so as to simplify and speed it up.

### 4.1. Annotation of Global Factors

One option for corpus analysis is to *annotate* whole scenes, a process in which one or more annotators manually add semantically relevant descriptions (see middle and bottom row of Figure 2). When used in this situation, the term annotation is to be differentiated from the same term when applied for modelling purposes (Section 2.2): here, annotation of a real-world corpus seeks to analyse real behaviour, rather than to directly inform synthesised behaviour. Generally, the annotation process should be as simple and efficient as possible and consider only those factors of relevance to the study. This is particularly important for a video corpus, where many different frames may need to be analysed. We have adopted an efficient annotation process, using simple graphical mark-ups, that enables a quick glance to detect important aspects over a whole scene (see Figure 2). For example. the middle row of Figure 2 illustrates how groups are highlighted by drawing a colour-coded circle around them, where the colour-code indicates the group size.

We use the term *global factors* to refer to the composition of the crowd as whole. This is in contrast to fo-

**Figure 3:** *Examples of groups of two, three and four walking, in each case, (left) perfectly abreast and (right) in staggered formation. In our corpus, the staggered formation was more commonly adopted, especially in constrained areas or when the crowd density was higher. Bottom row illustrates the dynamics of this formation.*

cussing on the behaviour local to a specific individual or group in a particular situation (see next Section). From the global perspective, we assess factors relating to individuals and groups in relation to the context of the overall crowd scene. For example, some reports have estimated that individuals or groups of size two tend to occur most frequently in pedestrian situations, whereas groups of three and so on, tend to be much less frequent [Hos02]. We have adopted a global factor analysis to establish group size norms in our scenes by annotating a number of scenes as described above, and have used this to inform our hypotheses for conducting the perceptual study described in Section 6.

### 4.2. Local Factors

In contrast to the global approach, we may instead choose to focus on more specific factor. We could consider, for example, the dynamics of individual and group behaviour in specific situations, such as formations adopted to avoid obstacles and individual behaviour within such formations. While annotation can also be used in these situations, empirical observation of specific sequences also highlights important modelling issues.

As a specific example, Figure 3 depicts a standard formation that groups of two, three and four people tend to adopt when walking. While in this *abreast* formation, as we shall refer to it here, it can be seen that individuals do not always walk exactly shoulder to shoulder, but rather adopt a slightly staggered formation. This allows the formation to reduce its frontal aspect while still allowing group members to stay within each others interpersonal space; the resulting *cold shoulder* of the leader appears to be socially acceptable under these circumstances. Futhermore, the distance between each group member in this formation often does not remain constant, even in the absence of obstacles and other walking constraints. This case illustrates a subtle behaviour that may have important consequences for simulation, allowing groups to maintain a more flexible version of a formation, even while walking in densely crowded areas. While groups may change formation to maintain their distance from their neighbours, another possibility, as seen in the corpus, is for the group to split in order to navigate an obstacle and merge afterwards (Figure 4). Both of these observations have been used to inform the creation of the group model in Section 5.2.

**Figure 4:** *Example of splitting and (top and bottom rows) merging behaviour for groups of two, three and four. In the middle and bottom rows, the sub-structure within groups also becomes evident: in the bottom row for example, the group of four splits into two pairs.*

## 5. Group Synthesis

Group modelling consists of two important components: firstly, embedding information into the environment that supports the addition of crowd behaviours (Section 5.1) and secondly, programming the actual behaviours that allow for group control (Section 5.2). In order to facilitate the definition of the environment and crowds within it, we have developed a specialised tool called *MetroPed* (see Figure 5). This tool creates an abstraction for those wishing to define new environments or create pedestrian scenarios without the necessity to have knowledge of the underlying source code. MetroPed scenarios and behaviours can

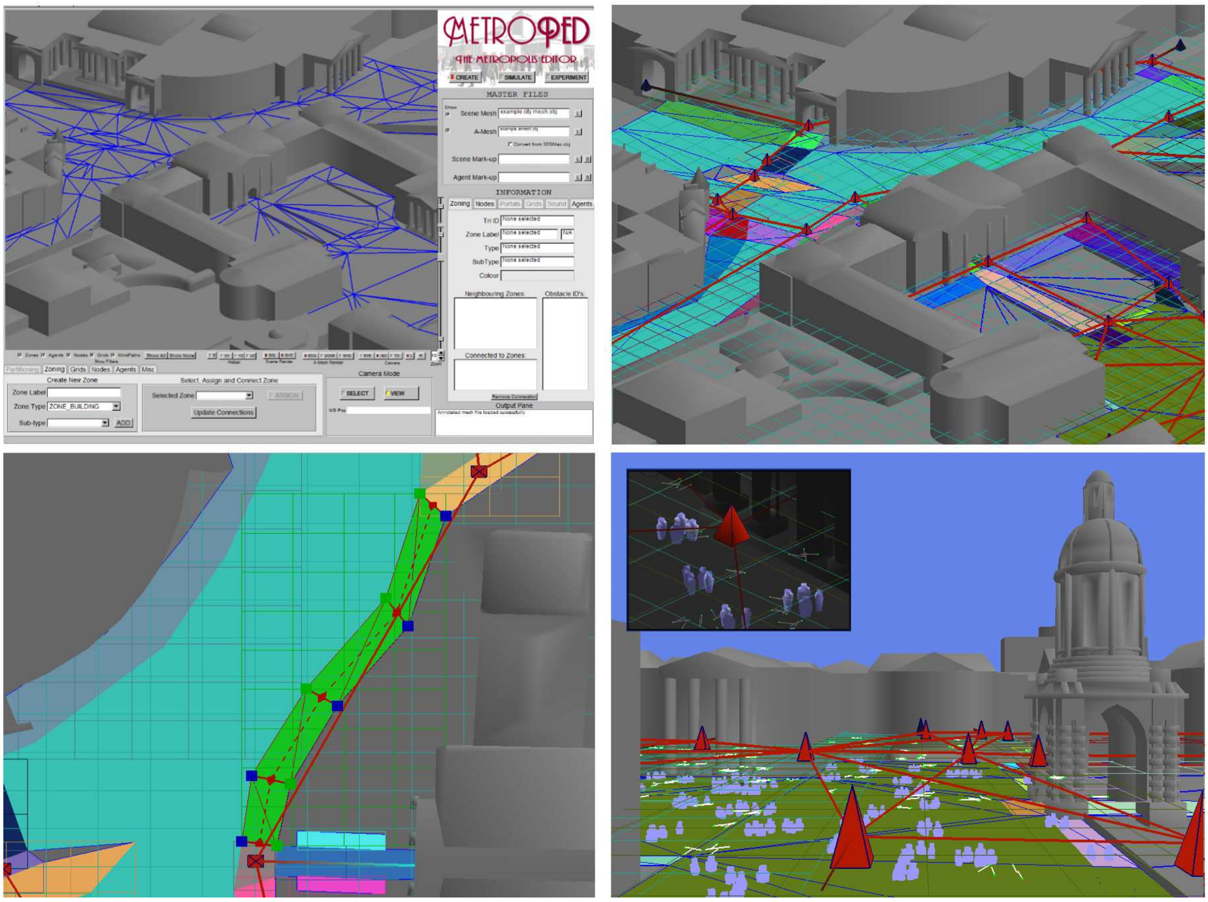


**Figure 5:** *The MetroPed interface. Top-Left: An initial mesh, shown in blue, is loaded into the editor with a simplified representation of the city environment as a backdrop.The key purpose of MetroPed is to allow the designer to tag the mesh with behavioural details. Top-Right: Walking zones and paths are defined on the mesh. Bottom-Left: Low-level paths are automatically generated to support intra-path navigation. Bottom-Right: Individual pedestrians or groups may be placed within the environment and simulated.*

be transferred seamlessly into the Metropolis visualisation engine. In addition, MetroPed provides simulation capabilities and supports the creation of scenarios for use in experiments (such as the study described in Section 6).

### 5.1. Environment Structure and Annotation

We refer to the process of adding behavioural details to the environment specification as *environment annotation*. For the city environment, we have defined three general categories of walk areas: pedestrian corridors, intersections connecting corridors, and open areas. The environment annotation process consists of a number of stages:

1. The primary purpose of MetroPed is to help the designer tag a *ground mesh* with relevant behavioural information for use by the simulation system. The ground mesh is created in a 3D modelling package and imported into MetroPed. The mesh geometry specifies roads and walkways and excludes areas that are of no behavioural significance. The mesh must be well-formed, abiding by a number of rules in order to allow MetroPed to automate some tasks, such as the generation of intermediate path points.
2. Given the ground mesh, the first task is to define *zones*. Each zone is defined as either a pedestrian *intersection*, pedestrian *corridor*, pedestrian *open area* or a *road*. As the designer specifies zones and zone types on the ground mesh, MetroPed automatically generates connectivity information at both the geometric and zone level. Intermediate path points are automatically calculated for pedestrian corridors (see Figure 5, bottom-left for example). These points upport within-zone pathfinding,

allowing pedestrians to navigate from an entrance edge of a corridor zone to an exit edge.

3. Each zone is associated with a *grid* containing pedestrian position, enabling efficient nearest neighbour searching. Grid cell size varies depending on the density of pedestrians.
4. Global pathfinding nodes support global navigation and related operations, such as determining locations where new pedestrians may be generated or removed, or container nodes representing building entrances.
5. Obstacles can also be placed in positions corresponding to graphical objects in the city environment, for example, telephone boxes, lampposts and bins, in order to ensure they are avoided by pedestrians during the simulation.

Once the environment has been defined, individuals or groups of pedestrians are placed in the scene using the interface and provided with initialisation details such as navigation goals. Individuals are also assigned behaviour and appearance templates, specifying their average walk-speed, personality factors based on the popular *OCEAN* or *Five-Factor* model [Wig96], and a link to their graphical representation in order to ensure consistency of appearance with behaviour. For example, a slower walk-speed could be assigned to an older individual than to a younger one.

### 5.2. Group Control

Group control is concerned with the simulation of pedestrians in order to ensure that they form cohesive groupings, while also adhering to more basic pedestrian behaviours, such as path following, avoidance of other pedestrians and containment within the path area. These behaviours rely on the information embedded in the environment, as described in Section 5.1. For example, in order for a pedestrian to stay on a path, it needs to know the boundaries of that path, information made available through the zone definition procedure. To avoid other pedestrians, an efficient look-up of neighbours is needed; again, this is based on a zone's grid structure defined during the environment annotation phase. During simulation, pedestrians keep track of their current zone, navigation goals and group membership. While this basic repertoire allows individuals to navigate their environment and reach goals, it is not enough to produce natural looking humanoid groups in most cases. Here, we describe, in a general manner, additional concepts of importance for creating the appearance and representation of group structures. These reflect observations made from our local factor analysis (Section 4.2). We suggest these concepts as additions to existing methods for individuals to enable small group structures.

### 5.3. Cohesion

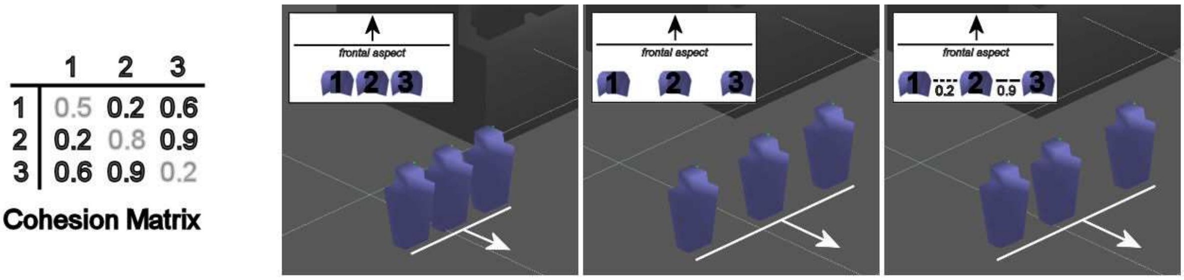
Group cohesion relates to the propensity of group members to maintain close spatial relationships with each other; the concept has been noted as an integral part of group simulation since the earliest models [Rey87]. Whereas most models have assumed that cohesion factors between all group members are equal, we model cohesion individually between different members and use it as a control for formations adopted by the group (Section 5.4) and splitting and merging behviours (Section 5.5).

Defining different cohesion values between members also implicitly creates sub-groupings within groups (see Figure 6). Sub-groups are an important concept: In groups of three, two individuals may be especially engaged in conversation, leaving the third to walk slightly ahead or behind. In groups of four, a substructure of two pairs may become evident (Figure 4, bottom row), or alternatively, there may be a cohesive central pair while the others walk ahead and/or behind (Figure 3, middle row, right). We define a square *cohesion matrix* for each group, with dimensionality equal to that of the group size. Entries in the matrix represent the *level of cohesion* between the two indexed group members. Level of cohesion specifies the tendency of two members to maintain a spatial relationship relative to each other (i.e. to match their speed and heading), although it is not a direct specification of the distance to be maintained between them.

### 5.4. Formation Representation

During the definition stage, a number of formations are specified by the designer for each group size. Each of these *formation templates* represents a discrete formation that the group may adopt, giving a position to each pedestrian: for example, standing perfectly abreast or in a line. Each formation template is associated with a *minimum frontal aspect*, which is a measurement of the width of the formation in its direction of travel when in its most compressed state. Frontal aspect is used to determine whether a formation can fit through a constricted space. In a pedestrian corridor for example, frontal aspect allows for a trivial comparison to take place with the width of each segment to determine if a group can navigate the corridor in their current formation.

Each formation template is associated with a state in a Fuzzy State Machine and each formation state is connected to one more allowable formations into which they can change. Although each template has a discreet representation, intermediate formations can be specified as the interpolation between two specified templates when a fuzzy state has been adopted. For example, if the current formation template specifies an abreast formation, but a narrower formation



|   | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.5 | 0.2 | 0.6 |
| 2 | 0.2 | 0.8 | 0.9 |
| 3 | 0.6 | 0.9 | 0.2 |

**Figure 6:** *Illustration of group cohesion. A cohension matrix (left)is specified to establish a spatial relationship between group members, each entry having a value between 0.0 and 1.0 (diagonals represent extroversion). Minimum frontal aspect of a formation template is calculated as the width of the formation in its direction of travel when in its most compressed state (second image). The frontal aspect can be compared to path width to establish if compression or a change of formation is required (third image). Cohesion values allow sub-groups to emerge within groups (final image) and determine how likely groups members are to split of temporarily.*

is required, the group may adopt a formation interpolated between the abreast and single file formations, producing a staggered formation of pedestrians with a reduced frontal aspect.

Each group can be considered to be a root node in a binary tree structure. As long as the group remains in formation, it continues to be represented by a single root node. However, it is possible for the group to temporarily split, in order to avoid an obstacle for example. When this occurs, two child nodes are created in the binary tree to represent the two new subgroups. Each sub-group has a formation appropriate to its size. For example, if a group of three splits into a pair and an individual, then one child node represents the pair and the other represents the individual. In this case, the pair represented in the child node may only adopt formations appropriate to a group of size 2. When sub-groups merge, child nodes merge back into parents [[ref Figure of merge and split and corresponding tree]]. Sub-group structure within groups of pedestrians can often be witnessed in the corpus (see for example, Figure 4, middle and bottom rows).

### 5.5. Formation Changes

At a basic level, formation changes are controlled by comparing frontal aspect in the direction of movement with impending obstacles and constrictions. Groups can compress as necessary in their current formations in order to navigate constrictions unless the minimum frontal aspect is less than the constriction width. When a formation change is required, a new one is selected by searching the formation templates' state machine for the nearest connected formation with a suitable minimum frontal aspect. In the worse case, a single file formation may be adopted in order to allow each pedestrian to pass individually.

Another option for the group is to split and, later, merge. This may happen when members of a group are not highly cohesive with others. A split causes a new sub-node to be created in the group binary tree representing the new member. Split members try to regain cohesion with their group with an urgency proportional to their cohesion level.

## 6. Perceptual Study

In the previous Section, we described how a number of specific behaviours observed from a video corpus (Section 4.2) can be applied to pedestrian modelling. Recalling from Section 3, an important part of our methodology is to obtain feedback about how new modelling techniques may affect viewers' perceptions of scene plausibility. In this case, one viable approach would involve evaluating the specifics of the techniques applied. For example, one could test whether viewers tended to find, over a number of different scenarios, a group split and merge behaviour more plausible than a formation change maintaining cohesion. However, as part of our methodology, we also highlight the importance of evaluating more general aspects (Section 4.1) of the scene in order to obtain higher-level indicators before embarking on in-depth modelling. In relation to groups, a more general and perhaps initially pertinent question, is whether the addition of group structures tends to improve the overall plausibility of resulting crowd scenes, and under what situations. Certainly, the addition of appropriate groups will most likely make the crowd simulation more *realistic*, by providing a more natural intermediate structure between the individual and the crowd, but little research has focused on how such additions will affect the visual *plausibility* of the crowd. This is the general question we address in this Section. We present participants with a variety of short animations, each containing pedestrians organised into different ratios of individuals, pairs and

**Figure 7:** *Example scenes from a sample of animations that were rated by participants in the perceptual study. These images corresponding to group ratios 1:1:1 (left), 2:1:1 (center) and 1:1:2 (right).*

groups of 3, and find out how realistic they are perceived to be.

### 6.1. Method

We generated a variety of short, 2-3 second animations in the Metropolis crowd system, using a crowd generation feature of MetroPed (see Figure 7 for screenshots from the animations shown to participants). In each animation, the virtual pedestrians simply walked forward at constant speed: their starting positions were set to ensure that no collision avoidance or turning maneuvers were necessary. This ensured that local factors, such as turn velocity, or avoidance distance, could not effect viewer ratings. Scenes were created according to two general categories: *groups* and *no groups*. For the *groups* category, each animation consisted of a varying number of groups and individuals, specified by a group ratio *S:P:T*, with *S* as the proportion of single individuals in the scene, *P* the proportion of pairs, i.e. groups of size 2, and *T* the proportion of triples, i.e. groups of size 3. We enumerated seven group ratio combinations (see Table 1). Through trial and error, we found 6 pedestrians to be an appropriate choice as a base value for generating pedestrian numbers from the ratios, as, given the scene area, this figure produced manageable populations of pedestrians for both high and low ranges. Thus, for a ratio of 1:1:1 (see Case 1. in Table 1), we obtain 6 singles, 6 pairs and 6 triples (6S, 6P, 6T), giving a total of 36 pedestrians.

The *no groups* category of animations contained solely individuals. The total number of individuals in these animations were matched with the maximum, average and minimum number of pedestrians over all the scenes in the *groups* categories, in order to create coherency in the overall number of pedestrians in each animation and ensure participant responses would not be biased by the total number of individuals in each category. An equal number of *groups* and *no groups* scenes were created.

| | Ratio | Groupings | Total Characters | Hyp. |
|---|---|---|---|---|
| 1. | 1:1:1 | 6S, 6P, 6T | 6+12+18= 36 | Pl |
| 2. | 1:1:2 | 6S, 6P, 12T | 6+12+36= 54 | Im |
| 3. | 1:2:1 | 6S, 12P, 6T | 6+24+18= 48 | Pl |
| 4. | 1:2:2 | 6S, 12P, 12T | 6+24+36= 66 | Im |
| 5. | 2:1:1 | 12S, 6P, 6T | 12+12+18= 42 | Pl |
| 6. | 2:1:2 | 12S, 6P, 12T | 12+12+36= 60 | Im |
| 7. | 2:2:1 | 12S, 12P, 6T | 12+24+18= 54 | Pl |

**Table 1:** *Group ratios used for the* groups *category of animations. The third column shows the numbers of singles (S), pairs (P) and triples (T) when 6 is chosen as the base value, giving the resulting total number of characters column 4. In each case, we hypothesised that the ratios would be (Pl)ausible or (Im)plausible*

Thirty three participants (7F, 26M), aged 17 to 30, were seated in front of a computer screen and provided with an instruction sheet. A virtual representation of a familiar scene was shown and they were instructed that some of the short animations they were about to view would reflect aspects of real pedestrian scenarios, while others were synthetically generated. A total of 42 animations were shown to each participant: between each trial, a blank-screen was displayed for 5 seconds, participants were given time to note their judgement for the animation they had just witnessed, after which a preparation message signalling the next trial alerted participants to the next display.

Based on the video corpus described in Section 4.1, we hypothesised that some of the group ratios in Table 1 would be more plausible for viewers than others. Specifically, we hypothesised that cases biased towards higher ratios of individuals and pairs would be more plausible than those biased towards more pairs and triples. Therefore, in the final analysis, we conducted comparisons between three conditions: *plausible groups*, *implausible groups* and *no groups*.

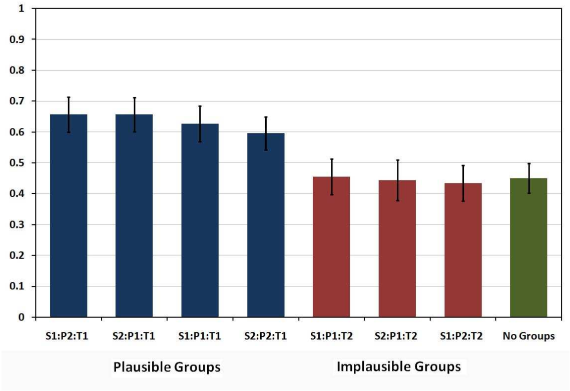


**Figure 8:** *Graph showing 'real' ratings for each group ratio, with 1 representing a rating of 'Real' and 0 representing a rating of 'Synthetic'.*

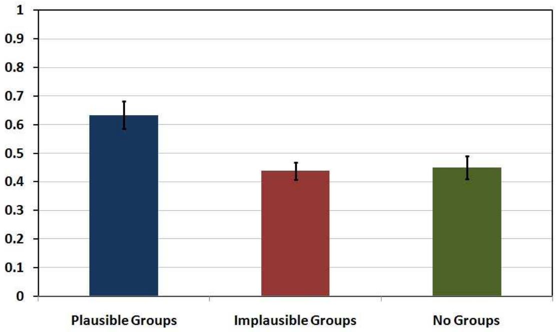


**Figure 9:** *Graph showing 'real' ratings for the three main conditions, with 1 representing a rating of 'Real' and 0 representing a rating of 'Synthetic'.*

### 6.2. Results and Discussion

We found no significant differencesbased on the ordering of the images within the experiment blocks in the responses of the participants, implying that there were no ordering effects within the individual experiments.

A two factor *ANOVA* with repeated measures for each of the three pedestrian population levels was conducted for the *no groups* condition. There was no significant difference in the responses of the participants to changes in pedestrian population alone, suggesting that any variance in participant responses to the group stimuli would likely not be influenced by the change in number of pedestrians in each scene (see column 4 of Table 1). This result allowed us to consider the three *no groups* cases as a single condition.

There was also no significant difference when performing a two factor analysis across the *groups* conditions and *no groups* conditions. This was expected, since the *groups* condition was composed of both hypothesised plausible and hypothesised implausible group ratios.

From Figure 8, it can be seen that the ratios 1:2:1 and 2:1:1 were rated as real more often than any other group ratio, followed by 1:1:1 and 2:2:1. This was followed by the *no groups* condition, and the *implausible groups* ratios, 1:1:2, 2:1:2 and 1:2:2. Post-hoc analysis was then performed using a standard Newman-Keuls test for pairwise comparisons among means. This test concluded that participants perceived the group ratio 2:1:1 as real more often than group ratio 1:1:2, but that all other combinations were perceived as real an equal amount of times.

We also conducted a 2 factor ANOVA analysis, with repeated measures design, for all of the cases (i.e. all seven group ratios and the additional *no groups* condition). This highlighted a main effect of the grouping method adopted i.e. the 8 cases of no groups, 1:1:1, 1:2:2, 2:2:1, and so on ($F(7, 224) = 3.6349, p < 0.001$).

The overall results of our comparison between the three primary conditions, *plausible groups*, *implausible groups* and *no groups*, can be seen in Figure 9. The results indicate that those animations in the *plausible groups* category were rated as more real than those in the *implausible groups* and *no groups* categories ($F(2, 64) = 8.6173, p < 0.05$), suggesting that the addition of groups to crowds can increase the plausibility of the scene, provided that care is taken in terms of the ratios of individuals and groups added.

In the context of our methodology, this result represents an important practical insight for synthesis: it may suggest, for example, that emphasis should be placed on modelling individual and pair behaviours, and ensuring that scenes tend to contain greater numbers of individuals and pairs and fewer groups of 3.

## 7. Outlook

Crowd modelling is a challenging task: conglomerations of individuals express structured and chaotic characteristics. It is hoped the methodology presented here will aid in this task by highlighting to the designer the importance of establishing practical links between corpus analysis, synthesis and perceptual evaluation for creating plausible crowds. We have proposed an exemplar for the methodology for the case of group modelling, highlighting how it can be applied at both global and local levels. The global approach provides indicators as to the plausibility of general concepts at the crowd or scene level, after which local elements, such as individuals or specific groups, can be modelled in more detail, either through annotation or case-based study.

Naturally, there are many other important aspects of group behaviour that have not been covered here. These include conversational behaviours as group members walk and talk, the formations adopted by groups when they are standing still, and the dynamics of how a group may form from individuals in the first place. However, we are confident that our methodology is flexible and can also be applied to investigating many other situations. It has already proven to be a useful guide for our investigations of contextual factors in static crowd scenes [EPO08].

A further intriguing question that we note is whether strange, but realistic, behaviours witnessed in the corpus may in some cases have the effect of reducing the perception of realism in the scene. An intriguing possibility is that more realistic situations may in some cases appear less plausible to viewers than synthetic ones that have been tweaked to match viewer expectations. While further investigation is necessary to answer such questions in detail, we suggest our methodology as the first step towards approaching such a problem.

Perhaps the most important element of future work on the horizon for this domain involves semi- or fully-automated processing of crowd corpus. This is currently a painstaking manual effort: the use of computer vision technologies for automating aspects of the process would not only reduce the burden on annotators, but would also enable a larger corpus and more impressive range of behaviour to be analysed.

## Acknowledgements

We wish to thank the rest of the Metropolis team, in particular Dr. Simon Dobbyns, for their efforts and assistance with the Metropolis visualisation engine.

## References

[DAPB08] Durupinar F., Allbeck J., Pelechano N., Badler N.: Creating crowd variation with the ocean personality model. In *AAMAS '08: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems* (Richland, SC, 2008), International Foundation for Autonomous Agents and Multiagent Systems, pp. 1217–1220.

[DHR98] Doyle P., Hayes-Roth B.: Agents in annotated worlds. In *AGENTS '98: Proceedings of the second international conference on Autonomous agents* (New York, NY, USA, 1998), ACM, pp. 173–180.

[EPO08] Ennis C., Peters C., O'Sullivan C.: Perceptual evaluation of position and orientation context rules for pedestrian formations. In *APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization* (New York, NY, USA, 2008), ACM, pp. 75–82.

[FBT99] Farenc N., Boulic R., Thalmann D.: An informed environment dedicated to the simulation of virtual humans in urban context. In *Proceedings of EUROGRAPHICSŠ99* (1999), pp. 309–318.

[HBJW05] Helbing D., Buzna L., Johansson A., Werner T.: Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation Science 39*, 1 (2005), 1–24.

[Hos02] Hostetler T.: *Controlling steering behavior for small groups of pedestrians in virtual urban environments.* PhD thesis, 2002. Supervisor: J. Kearney.

[LCD07] Lerner A., Chrysanthou Y., Dani L.: Crowds by example. *Computer Graphics Forum 26*, 3 (2007), 655–664.

[LCHL07] Lee K. H., Choi M. G., Hong Q., Lee J.: Group behavior from video: a data-driven approach to crowd simulation. In *SCA '07: ACM SIGGRAPH/Eurographics symposium on Computer animation* (2007), pp. 109–118.

[MT01] Musse S., Thalmann D.: Hierarchical model for real time simulation of virtual human crowds. *IEEE Transactions on Visualization and Computer Graphics 7* (2001), 152–164.

[OCV*02] O'Sullivan C., Cassell J., Vilhjálmsson H., Dingliana J., Dobbyn S., McNamee B., Peters C., Giang T.: Levels of detail for crowds and groups. *Computer Graphics Forum 21*, 4 (2002), 733–741.

[PAB07] Pelechano N., Allbeck J. M., Badler N. I.: Controlling individual agents in high-density crowd simulation. In *SCA '07: ACM SIGGRAPH/Eurographics symposium on Computer animation* (2007), pp. 99–108.

[Rey87] Reynolds C.: Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics 21*, 4 (1987), 25–34.

[ST05] Shao W., Terzopoulos D.: Autonomous pedestrians. In *SCA '05: ACM SIGGRAPH/Eurographics symposium on Computer animation* (2005), pp. 19–28.

[TD00] Thomas G., Donikian S.: Virtual humans animation in informed urban environments. In *CA '00: Proceedings of the Computer Animation* (Washington, DC, USA, 2000), IEEE Computer Society, p. 112.

[Wig96] Wiggins J.: *The Five-Factor Model of Personality: Theoretical Perspectives.* The Guilford Press, New York, 1996.