Social Spatial Genesis: Activity Centered Space Making

Joanna Crotch, Mackintosh School of Architecture, Glasgow School of Art
Robert Mantho, Mackintosh School of Architecture, Glasgow School of Art
Martyn Horner, Digital Design Studio, Glasgow School of Art

Abstract

Digital technologies and processes have been used to generate architectural form for over two decades. Recent advances in digital technologies have allowed virtual digital environments to be constructed from physical movement. But can a bridge that connects the physical and virtual realms be developed? Can this, currently arbitrary form making be grounded in human activity and subsequently be integrated in to real time, space, and place.

This research asks how space generated from the process of digital morphogenesis can be related to meaning beyond just the creation of form. Existing research asks how new form can be discovered, or what material and structural possibilities can be derived from form, through these morphological processes. The aim of this research project is to complete the loop, physical–virtual–physical, and to connect these digital processes to meaning through human activity. Its aim is to discover the consequences of generated spatial envelopes that are manipulated through digital morphogenesis and related to specific human activity, in the pursuit of possibilities for a digitally generated architecture that is socially engaged. This is not random form finding, wherein architecture tries to imitate biological processes or form, but form finding that is connected to a primary architectural concern, how is the architecture being used by humans.

1 Introduction

As we move into the twenty-first century, the possibilities that the new technological age can offer architects are far reaching. New and radical forms have been developed through digital processes. This research asks the question can function, or more appropriately activity, drive form. If so how can digital media be related to meaning beyond just the creation of form? With these questions in place, and recognising our position on the threshold of technological opportunities, what is the possibility of making spatial sense of current theoretical form finding and designing a new process that creates new architectural space grounded in real time and directly generated from specific human events? Having developed such a space, the research will continue with speculation into how form can be modified in order to accommodate the adjustment and fine-tuning of particular activities. We ask if a new methodology for the creation of adaptable space can be generated. Can this innovative form finding produce activity specific and adaptable space?

There are existing examples of form making utilising digital tracking and motion capture (Ramsgard-Thomsen 2005). Others are utilising the potential of genetic programming to generate 3-D shape grammars (Coates 2007), and self-organising maps (Ireland and Derix 2006). These studies have resulted in the development of new theoretical space, generated with either non spatial concerns or without the consideration of human interaction. The results are captivating but somewhat arbitrary. With questions regarding relevance in the making of Architectural Space how can we achieve a relationship with these methods of form finding and human event? Our curiosity anchored itself in finding relevance and following that we sought significance in the insertion of the physical interaction of humans in space. Our speculations have driven us to create a methodology that will utilise the tools and processes of digital morphological methodologies, but to use them in a loop, where physical activity and virtual artefacts inform each other. A close and thorough observation of specific physical activity creates data; this data transposed into a digital format generates a theoretical digital space. From patterns observed within the movement, mathematical formulae are written that allow manipulation of these spaces. This paper deals with the development of the process for collecting and digitizing movement information. The next stage of the research is to use the digital data in a feedback loop with a physical environment; this environment will be responsive and adaptable to the users input. New architectural space is generated from an activity, for an activity.

2 Data Selection and Collection Process

Our aim was to design a process that would test our speculations. The first step of the process is to select and record a chosen activity. This is then followed by the collection of appropriate and usable data. This essential data will form the basis for the establishment of the ‘movement zones’ and will be consolidated and utilised throughout the proceeding steps. This movement data is then translated into digital information from which 3-D spatial envelopes are formed. The following step involves analysis and extraction of movement information. Patterns observed are transposed into algorithmic formulae and a feedback loop is established. The reapplication of the algorithmically mutated information is reapplied, and manipulation of the spatial envelope will then occur. We propose repetition of this last stage, thus yielding a number of similar yet subtly different spatial envelopes.

2.1 The Activity

Currently, the project has reached a mid-point where the collection of data is complete, the transfer of the moving images into a digital format has been carried out, and the realisation of primary spatial envelopes is under way. This paper attempts to explain the process so far and reports on the possible subsequent phases and our expectations for the future of the research.

The research methodology begins with the digital spatial mapping of an activity. Several possibilities were considered, and preliminary recordings of these were made, viewed, and critiqued prior to our selection. The selection of this was based on many factors, but primarily we wished to use a very ‘everyday and ordinary’ human activity. We believed that repetition of certain actions was important, as recognition of pattern was going to be an element that would be examined in the following stages of the experiment. Without wishing to pre-empt our results a certain degree of speculation was necessary in this process. The team finally settled on the activity of preparing food and eating. It was also agreed that the introduction of more than one figure would enrich the data. Two figures performing the selected activity would provide social engagement and further ground the experiment in reality.

2.2 Set Design

The next step was the creation of the physical; a skeletal stage set had to be created that would frame the performance of the activity. Several factors were taken into account in respect of the 'set' in order to allow filming of all the movements to be recorded clearly and equally. The filming required a space big enough to accommodate the room set and the cameras, but more importantly one that also had controllable lighting conditions. A north lit disused gallery space was selected which fulfilled both requirements. The issues with lighting were initially recognised, but never fully resolved. It was important to try and eliminate shadow, as this would be picked up in the film analysis and could subsequently distort the results. While the chosen space did reduce some of the problems, lessons from subsequent phases revealed that the lighting conditions were not ideal.

Designs and drawing of the proposed 'set' were made. We were then able to discuss and agree on both the practical issues; for example camera positions and set lay, alongside those more creative elements, i.e., the choreography of the activity; and, more specifically, the movements that would be made by both figures. Minor adjustments to the design were made and construction of the frame took place.

2.3 Data Collection

The mode for collection of information raised further questions regarding the type of data we wished to produce as a result of the recording of the activity. The target was to produce a series of envelopes reflecting the actor's occupancy of space over time. It was anticipated that these envelopes would reveal the actors occupancy of any slice of 2-D, 3-D, or 4-D space. Having considered the alternatives alongside our particular expectations, the team conclusively agreed that 'motion capture' was not appropriate in terms of the spatial envelope that we wished to create. Motion capture produces a skeletal frame. The aim of this research was to capture a three dimensional space with a continuous skin, where a marker-less, non-invasive form of motion capture was sought. While the authors found examples of research into similar processes, they were technically sophisticated and mathematically intensive. (Tangkuampien and Chin 2005) (Li, Huang and Li 2007) A 'low tech' recording method was agreed upon, giving us the material most suited for our requirements. The plan was to install the 'set' or kitchen frame and accurately position cameras around the space, withdrawn sufficiently to allow a wide enough view of the 'kitchen,' and giving adequate overlap of views to ensure all movement could be detected from at least three cameras at any time. The design and making of the set was critical. All was contained within a three-meter cube, located with a three-meter square taped space on the floor. This taping would play a part in the future analysis of the material in order to correct the intrinsic camera distortion. We were aware that it had to be as skeletal as possible, as any solid object or surfaces would result in movement patterns being obscured (Figure 1).

Figure 1. The 'set.'

Minimal props were also included to allow the 'mime' to take place with as much accuracy as possible. The filming was reordered on four high quality DV cameras, three set at a height of three meters from ground level, one at fifty centimeters from the ground and located in the agreed locations (Figure 2 and 3).

Figure 2. Camera set up.

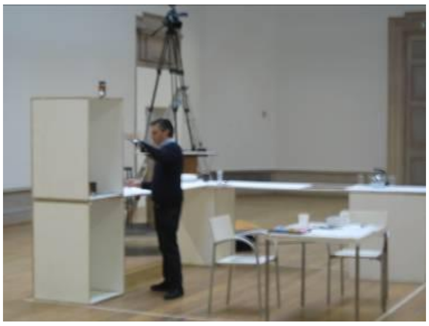

Figure 3. Rehearsal of the activity.

Once situated, their exact location was measured manually and recorded. Following the design, set up, and rehearsal of our set and activity, the filming process was a reasonably rapid one. Five live sequences were filmed during a one-day session (August 2008), with each sequence lasting eleven minutes. Standard filming procedures were adopted to provide a rigorous process.

2.4 Conclusions regarding Data Collection

Following the filming, we realised that despite care taken in the set design, decisions made have impacted the material collected, and have subsequently created more work to allow the spatial/movement information to be as clean and accurate as possible. One such problem arose with the white worktops; despite care with the lighting, shadows were cast on these surfaces and became part of the recordings—this had to be eliminated from the data. We would conclude that the set design is more critical than we had at first envisaged. Dark tones, more skeletal formwork, and fewer props would have provided a 'cleaner' set of digital results, with less work required to 'de-contaminate.' This knowledge will be useful if we are to repeat the experiment in a different format.

3 Data Processing

The film recorded the interaction between the two bodies mapped with video recording. The next stage transferred the physical into the digital—the goal being to produce a 'spatial envelope' of the activity over the full duration of the event.

From each of the four cameras, a 2-D image was derived for each actor at each moment in time. Using these 2-D images, a cone of space was generated for each camera; by intersecting these cones a 3-D shape for each actor was constructed for each frame. By placing the shapes for both actors together, a 3-D 'spatial envelope' of a single moment is generated. Once these individual moments are merged; a 'spatial envelope' of the activity is produced. This process requires a robust and translatable outcome at each step.

3.1 Video Transposing

The first step in the translation was to use an Avid console to translate the digital video format into four individual, fifty-six, four minute movies. From these four films, four sets of 1500 TIFFs, representing a typical one minute's filming, were extracted for study. Various processes were tried in ImageMagick to empty the background information to derive 2-D shapes of the human actors. The goal was to convert these TIFF images from the colour TIFF format into a white on black silhouette, two-colour image file. This process of isolating the human figure had two primary goals: each step needed to produce a digital format that would provide multiple options for the succeeding steps, while also allowing the process to be automated for processing the four complete films.

3.2 Figure Extraction

The first test was to use ImageMagick commands to segment the images of the actors from the empty background; this function having been written by a member of the team (Figure 4).

Figure 4. Empty Set – Background.

A sequence of carefully tuned and threshold RGB separation processes were tried (Figure 5).

Figure 5. Single Frame with Actor.

These simple techniques isolate the human actors from the background of the image, but shadows were still evident. Further use of a hue-discrimination filter removed most of the shadows cast on the white surfaces (Figure 6). These various filters were combined to produce a single white on black silhouette of the human actor with any objects moved by him and hence incorporated into his spatial occupancy.

Figure 6. White on Black Silhouette.

This filter process is easily varied and is being developed and adjusted as the research progresses. Different command combinations can be devised using this toolkit to reflect particular requirements or visual conditions. A minutes worth of this conversion process for one camera (1500 images) takes approximately ninety minutes on a normal PC.

3.3 Camera set up in Digital Space

The geometry of the camera acquisition process is required next. A manual model of the scene with its four cameras was constructed in Blender (a free modeling program) to determine numerically the positions of cameras, directions of view and scaling factors for the acquired images. Using the ability to simulate the operation of the cameras in the software while overlaying actual frames from the video in the view ports of the modeled cameras over a skeletal model of the scene, we could create a geometrical simulation of the setup and thereby derive quantities needed for the next stage. It was also determined that the intrinsic distortion of the cameras could be corrected or, within the accuracy of the process, effectively ignored (Figure 7).

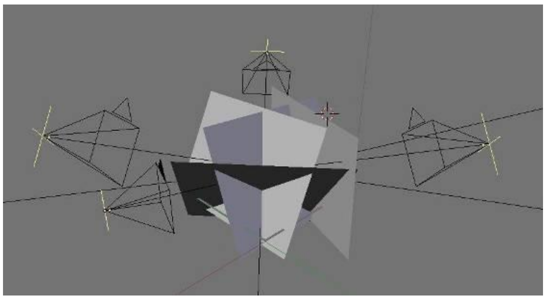

Figure 7. Camera set up in Blender.

3.4 Extrusion of Silhouettes into Forms

The mathematical process targeted by the next stage is illustrated by (Figure 8). The silhouettes of the occupants appropriately scaled and positioned according to the model, are projected from the camera locations into the space and intersected to form an approximation of the shapes of the participants.

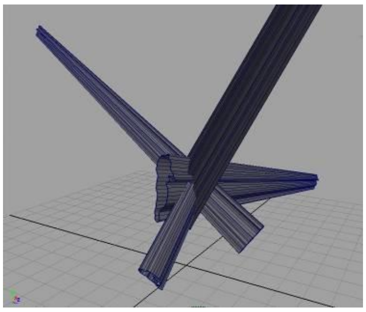

Figure 8. 'Cones of occupation' the concept.

A program was developed (in C for the best chance of speedy execution) which generated a 3-D 'bitmap' of each camera's 'cone of occupation' for a given frame of video. This bitmap is essentially a simple 3-D matrix of 1s and 0s which can be treated algorithmically and combined with other similar bitmaps or sectioned in various ways without using complex modeling software.

3.5 Intersection of Forms to Construct an Envelope

Various intersection operators were used on each set of camera cone bitmaps to cut down the volumes of the other cameras. The favored approach is to use a 'committee' system which takes the majority (three out of four) occupancy of the three cone bitmaps to identify the actual shape. This seems to minimize the ambiguity caused by furniture obscuring one camera's view while differentiating between multiple volumes. This produces a minimal set of solid shapes: when the shapes from all four cameras were combined, a unique ‘spatial envelope’ of the actors was constructed (Figure 9).

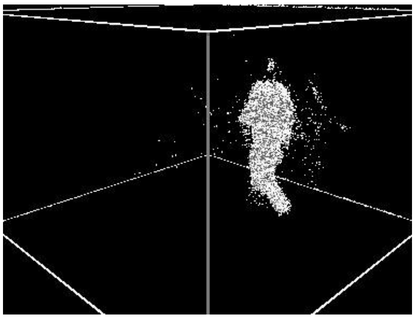

Figure 9. ‘The intersection of ‘cones of occupation during one moment.’

Straightforward filtering processes were then added to the program to remove isolated 'specks' and 'holes' to end with a 'solid' 3-D shape.

3.6 Envelope Manipulation

The 3-D bitmaps are easily combined and manipulated and at the resolutions used (300x300x200: a resolution of 1 cm in each axis) such processes are quite manageable, taking seconds rather than hours.

An example of such a combination is shown here: four frames from a video showing an integration of four seconds of movement of one participant (the accumulated personal envelopes for one hundred frames), viewed by ray tracing (within the same bitmap processing program), illuminated by a simple light model and rotated for inspection (Figure 10).

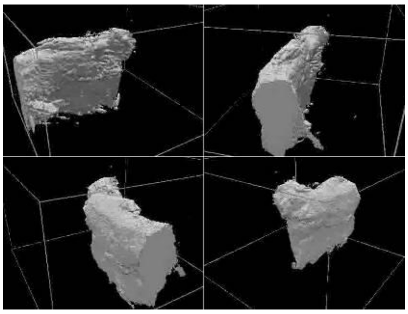

Figure 10. ‘The envelope of 4 seconds of movement’

4 Pattern Analysis

This envelope has been analysed to determine the patterns of spatial change. As the envelope fluctuates over the duration of the event, patterns of density and thinness have emerged. Areas of contained activity become thick with layers of repetitive motion, while the movements between locations are stretched. A planometric analysis was done to confirm these 3-D patterns.

5 Algorithmic Processing

The next phase of the project is to develop algorithmic expressions of these patterns. These algorithms will be utilised to manipulate the original spatial envelope. The number of iterations used in this algorithmic process is based on the structure and patterns of the initial activity. Each phase of activity lasts for approximately 3.5 minutes. Three distinct phases of the activity were observed, preparation, consumption and clean up. Each phase demonstrates different motion patterns and different rates of repetition of iterative movements. These phases each result in an algorithm, generating an algorithmic process for the transformation of the 'spatial envelope' of the activity. The patterns of movement in each phase determine the number of iterations for each algorithm. The 'preparation' phase algorithmic process will be run six times, which is based on the movement of each actor from the table to the various stations in the 'kitchen'. The 'consumption' phase algorithm will be used twenty times, to reflect the repetition of the actors arm movements. The 'clean up' algorithm will be run eight times, again derived from the movement between table and kitchen stations.

Following the extraction of these repeated movements and analysis, algorithmic formulae will be constructed. The algorithms will then be fed back into the spatial envelope data, and will be run to produce a transformation of the original spatial envelope.

This algorithmically transformed space will then be inhabited with the digital spatial map of the activity to test the interaction between the algorithmic influenced space and the spatial map of the activity. The algorithmically manipulated space will be assessed as architectural space, judged both for its appropriateness and architectural qualities. These complex digitally generated environments will then be transposed back to the physical, to produce a manifestation of space driven directly by the activity, which it is designed to contain. A new 'part physical, part virtual' architecture is born, an architecture which is based in physical activity and which explores spaces virtually with digital tools.

6 Further Development of the Research

The next stage of the research will explore ways in which the form-generated space could be developed to interact with users, and a feedback loop between human activity and digital space can be established. Once a methodology has been established to create digital envelopes directly from specific activities—with applied algorithms written in response to the space produced—the possibilities are endless, as are the activities that could be utilised to produce these new spaces. A variety of researchers have been working on creating systems in which human actors interact with the digital, seeking to find ways in which the human agency can utilize technology to shape experience (Fruhwirth, Hirschberg and Zedlachar 2007). This research pursues similar concepts, while also exploring how architectural space develops when the user's interaction influences successive generations of architectural enclosure. The research is speculative and proposes that this digitally generated space, tied to human activity can ground the new digital methodologies in the historical tradition of architecture.

The project would utilize a physical version of the architectural space developed using the event based digital morphogenesis process. Questions may arise for example 'as the architectural enclosure evolves, how does the user's behavior change in response to the emergent architecture?' and 'How can an architecture generated from human activity, which changes in relation to use influence future activity, be developed?' This research explores the possibility of an architecture, which is both dynamic and collaborative with its users, learning and teaching in a dialogue with events and their participants.

7 Conclusion

Current digital morphogenesis is concerned with generation of form; this research tests the possibility of connecting this generated form with specific activity. The criteria for judging the generated space is the relevance of this space to the specific activity that it contains and from which it was originally derived. The camera computation method of recording the movement associated with specific activity has been confirmed as the appropriate methodology, as it results in 3D spatial maps as opposed to linear tracking diagrams of motion. The spatial maps are then subjected to controlled manipulation resulting in unique spatial constructs directly relation to human activity. These new spaces open potential possibilities of an architecture whose relevance is in its direct relationship with meaningful social activity. Potentially the feedback loop between user and architecture could result in an architectural space that has shed cultural conventions and is a direct response to human spatial needs. The value of such architecture, where the relationship between the space and the activity which determines it and which it subsequently contains, seem obvious.

References

Coates, PS. (2007). Teaching Architecture through Algorithms – the First 15 years. In *Teaching and Experimenting with Architectural Design: Advances in Technology and Changes in Pedagogy*, eds. C. Spirdonidis and M. Voyatzaki, 105. EAAE Transactions on Architectural Education ed. 35.

Ramsgard Thomsen, M. (2005). Spatial Flux: interface as site. In *Space, Place & Experience in Human- Computer Interaction* Proceedings of Interact 2005, Rome, Italy.

Frühwirth, M., Hirschberg, U. and Zedlacher, S, (2007). Formotions*. In *Teaching and Experimenting with Architectural Design: Advances in Technology and*

Changes in Pedagogy, eds. C. Spirdonidis and M. Voyatzaki, 29-39, EAAE Transactions on Architectural Education ed. 35.

Tangkuampien, T. and Tat-Jun Chin. (Dec. 2005). Locally Linear Embedding for Markerless Human Motion Capture Using Multiple Cameras Digital Image Computing: Technqiues and Applications. Proceedings of DICTA 05: 498 – 505. Digital Object Identifier 10.1109/DICTA.2005.1578170.

Li, F., Huang, T. and Li, L, (2007). Method of Motion Data Processing Based on Maniflod Learning. In *Lectures Notes in Computer Science, Technologies for E-Learning and Digital Entertainment*, vol. 4469, Springer-Verlag, Berlin, Germany 248-259.

Ireland, T and Derix, C. (2006). The Poly-Dimensionality of Living: Communicating Space(s).Proceedings of eCAADe: 449-455. Volos, Greece.

Crotch, J., Mantho, R., and Horner, M. (2009). Space Making – Between the Virtual and the Physical. In the *International Journal of Architectural Computing*, vol. 7, issue 03, 403-414.

Crotch, J., Mantho, R., and Horner, M. (2009). Spatial Genesis - Event Based Digital Space Making. In *Joining Languages, Cultures and Visions*. Proceedings of CAADFutures: 188-199. Montreal, Canada.