Man with the Movie Camera

An Approach to Synthetic Cinematography for Built Environment
Takehiko Nagakura, Panagiotis Chatzitsakyris
Massachusetts Institute of Technology
takehiko@mit.edu, panosarc@mit.edu

This paper discusses shortcomings of current animation software, and introduces a new type of digital tool that helps an architect quickly construct a cinematic spatial representation. The novel approach is that this tool looks at architectural space as a condition populated with mobile human inhabitants, takes advantage of cinematic conventions a trained film-maker would deploy and automates the process of camera work and montage. A prototype implementation demonstrates its application to a small house designed by Le Corbusier.

Keywords: Animation; Digital Cinematography; Architectural Visualization; Citrohan House.

Introduction

Casa Malaparte (1943) is a villa built by Curzio Malaparte and Adalberto Libera in Punta Massullo on the Italian island of Capri. As this is considered an architectural masterpiece of the twentieth century, numerous conventional architectural representations of this house are published, and its aerial photo is probably the most notable and widely circulated (Figure 1).

In 1963, a French Nouvelle Vague director, Jean-Luc Godard, used this house as an important shooting location for a part of his movie Le Mépris. A comparison between his cinematic representation populated with Bridget Bardot and the other deserted architectural photography of an impeccable composition makes it clear that the director’s method intuitively conveys certain comfortable empirical qualities of the place, while the other imposes a reading of the artificial geometry against the wild landscape from a single dominant vantage point (Figure 2).

As a complicated love story unveils, the mind of the movie audience unconsciously builds a uniquely rich spatial understanding through different casual camera angles, the up and down movement of the main characters on the roof and its stepped approach, as well as the carefully composed juxtaposition of scale of the figures, building and vast landscape.
Computer graphics animation tools recently have become available to architects at relatively low cost and without extraordinary technical requirements. However, most contemporary productions of architectural animation do not reach the level of sophistication where architects’ essential spatial ideas and sense of place are truly conveyed through moving images. This presents a sharp contrast to films that professional film-makers produce. As seen in Le Mepris and recent Hollywood special effect movies, they often seem superior to architectural walkthrough productions, more easily connect to minds of audience, and instantaneously communicate spatiality and atmosphere of the locations.

One of the reasons for this shortcoming is architects’ lack of training and understanding of cinematic language, which is necessary to translate an architectural form into a cinematic form. This language, largely built for narrative films with characters in space, controls attributes such as camera movement, sound effect, mise en scene as well as film editing, and has been developed by various film-makers throughout the past century. Early compilation of such conventions is already found in Film as Art (1933, 1957) by Arnheim. His meticulous analysis includes methods of framing and montage. Framing has to do with the selective delimitation of the viewable space through the pictorial boundary, and montage refers to method in which various individual shots are combined into a longer sequence. Using different framing and montage strategies potentially provides a director with different representation of the same event. Elaborations of specific stylistic development includes montage theory by the constructivist film-maker Eisenstein. Numerous pedagogic manuals are available today for beginning film-makers (Figure 3).

It is true that architects have long exploited sense of moving in architecture as a major driver of spatial design. But, their primal means of spatial representation had been limited to sketches, plans, sections, perspective, scale models and others. Using these static media is not very suitable for representing such ideas as development of scenes through building circulations and space as event generator of live inhabitants.

Recent digital technology solved the mechanical problem of production of synthetic motion graphics,
and an architect today can pick up an animation software package, move a synthetic camera in a geometric model, and edit the sequences in non-linear editing software. Moving characters can also be inserted into space. However, these software tools do not provide any disciplinary basis for how an architect should move the camera and how he or she should edit the sequences in order to communicate with the audience. For architects, therefore, a way of harnessing with these missing methods could be an important solution to their problem. Trained film-makers, in contrast, would know camera work conventions such as pan, dolly, and subjective camera. They also know ways of assembling a sequence through such typical methods as rhythmic montage, shot-reverse-shot or an uninterrupted long sequence (Figure 3).

The movies we see at theaters and television dramas usually are visual interpretations of underlying narrative stories. A story is broken down to a series of events, and major components of an event are acting characters and its location. Cinematic representation sometimes puts emphasis only on interacting characters with little concern to the background. A typical shot-reverse-shot sequence is a method to alternate close-ups of two characters in conversation. Little is informed about space here. On the contrary, often an introductory portion of an event describes its location before the characters come into play. A typical sequence is somewhat in the middle of these two extremes, a synthesis of interacting characters and a spatial context with different degrees of emphasis. Automation of a cinematic composition primarily for describing interacting characters has been discussed and explored by previous groups such as Li-wei He, et al (1996). However, analysis and automation of a cinematic composition for describing spatiality has been largely left unexplored even though many cinematic conventions clearly help descriptions of spatial contexts and locations as seen in Le Mepris.

A digital tool for synthetic spatial cinematography

This paper introduces a new type of digital tool that augment current animation software by algorithmically implementing those basic cinematic methods specifically useful for spatial representations. Its aim is to help an architect quickly construct a credible animated spatial representation without professional understanding of cinematic conventions. It focuses on algorithmic automation of camera placement and movement as well as montage, and takes advantage of moving inhabitants as reference to describe spatiality. To facilitate its process, the software works in two main phases: individual clip generation and montage. It first lets a user populate a given spatial environment with mobile figures, and uses the points of view of these inhabitants as well as their observers to generate a set of component video clips. In the second phase, the image content of these clips is analyzed, evaluated in reference to the visibility of objects of interest, and edited into a stylized sequence.

Population with mobile figures

The approach here is to develop an animated spatial
Depiction with reference to the movement of figures inside. The software treats a given geometric model of an architectural space as a stage set where main characters interact with the environment. The user first declares these characters, and specifies a circulation path of each character inhabiting the space (Figure 4). The software then inserts a synthetic figure in the geometric model that walks on the specified path at a natural pace.

Camera placement and movement
In a typical cinematic production, there are conventional camera placements including a camera placed on the head of a moving figure, another dollying on the side of the figure oriented towards the figure, and a fixed aerial one placed at a corner of the space observing the figure. The software automatically places these cameras as listed below (figure 5). Some provides points of view of these mobile figures and the subjective sense of moving in and around the space at a human eye level in a natural pace, while others observe the interaction between the figures and the spatial context and reveal the sense of event, juxtaposed scale, and function. This process creates multiple clips of animations, each made from a particular camera.

- Point Of View Camera: This is a camera that is located on the head of the moving figure. It is originally targeted straight ahead of the moving direction and may consequently move and focus on the specified objects of interests in the space.
- Close Character Cameras: Each of these cameras is placed about half meter away from the moving figure and frames approximately the upper part of the figure’s body. Each frames the front, side, back or one of 45 degree views of the figure.
- Far Character Cameras: These are similar to Close Character Cameras but are located further away from the character (about 2m) and frame the whole figure.
- Panoramic Cameras: These cameras are placed in reference to the space in which the characters move. Each camera is static and located at a high corner position of the space, and its target follows the moving figure.

Objects of Interest
Every designed space has objects of interest, that is, certain spatial forms that its architect considers important for the experience of the space. An object of interest may be a well-positioned picture window that frames a spectacular view, a carefully designed spiral staircase or a wall with an impressive pattern of marble stone. The software tool lets the user specify these objects of interest in the given geometric model and assign a numbered weight to indicate a degree of interest. For instance, an object with degree 2 could be less interesting than another with degree 3. It then uses these weight numbers in two different processes. First, visual focus of a character would be attracted by the objects of interest such that the camera put on their head will turn towards a weighed object in proximity. Second, at any point in the time line of animation, clips made from different cameras are evaluated and prioritized in reference to

Figure 5
Software automatically places mobile cameras (left) and static panoramic cameras (right)
the visibility of these weighed objects. At the end of
the generation of these clips, the software analyzes
each one of them and checks if the view of a camera
is interrupted such that a camera accidentally goes
outside of the wall and no longer views the interior
space concerned. It then checks if any part of the clip
includes a view of any objects of interest and what
weight number the objects have. Based on these
analyses, at any point in the timeline, the clips are
prioritized (Figure 6).

Stylistic Montage
The automated final phase is montage and to edit
these clips and assemble them together into a final
sequence. The user can choose from a list of pre-set
montage methods, which provide the user with vari-
ous versions of the representation of the same space.
In the current implementation, the tool will use three
distinct montage styles: Rhythmic, Analytical and
Monoshot.

Rhythmic montage is based on a theory of mon-
tage made by Eisenstein (1949). According to this,
the visual power of the sequence stems from the
rhythmical juxtaposition of radically different shots.
Every next shot attempts to contract the previous
shot through a switch of the framing: a close shot is
followed by a long or panoramic shot and vice versa.
Figure 7 is an example from Eisenstein's Battleship
Potemkin (1925) that influenced many followers
such as Orson Welles.

Analytical montage makes a composition of the
sequence only from methodically selected detailed
shots. The whole panoramic view of the event is ei-
ther never shown or it is shown just once at the end
of the sequence. Here the timing is not rhythmic but
dependent on the complexity of each shot. By only
revealing parts of the whole, this technique encour-
gages the audience to fill in the gaps by its own ex-
pectations. Russian director Timoshenko was a pio-
near of this technique in the 1920s. Among the many

Figure 6
Clips made from different
cameras. Mesh shows the ob-
jects of interests identified in
a frame.

Figure 7
A Sequence from Eisenstein's
Battleship Potemkin. Long
and close up shots alternates.
Directors that incorporated analytical sequences in their films, Brian de Palma offers a clear example in his film, Dressed to Kill (Figure 8).

Monoshot style involves the complete absence of cut and paste editing. The whole sequence consists of only one uninterrupted shot. As the camera moves through the space, it continues to follow the character both in space and in time. The framing of the character changes because the camera’s relative position to the character changes from front to side, side to rear, and so forth. To achieve this style, multiple cameras around a figure are selected in an order, and their positions are interpolated to provide a smooth, uninterrupted transition from one clip to another. Directors like Alfred Hitchcock, Orson Welles and Quentin Tarantino have often used this technique in their films. Figure 9 shows an example from Tarantino’s Kill Bill.

Each of these styles is provided as a timeline template. And the template of final animation is filled by sliced parts of the clips derived from cameras that are selected in consideration to the requirement of each style as well as the priority derived from the objects of interest (Figure 10).

**Application and example: Citrohan House**

An architect may use this automated animation tool in different stages of a design project. A designer can
use this tool for a generation of presentation visualization. Quick production in a cinematic style and conventions familiar to general audience is helpful for the final stage of a design project. Changing parameters of editing styles and paths for mobile figures can also provide various alternative solutions quickly.

Since the tool works reasonably fast and with operational simplicity, another scenario is for an architect to use it to improve the initial animation by changing the spatial design, especially the portion identified as the objects of interest, and to create another animation sequence for comparison. Making this feedback loop realizes an architectural design that truly takes into account a sense of moving in the space and other motion-based spatial characteristics such as development of scenery. This tool then becomes a potentially very interesting, new means of design exploration. Ideally, an architect can keep developing a spatial form in parallel with revising its cinematic representation, until they reach a meaningful equilibrium.

To illustrate the use and benefit of this proposed software, a prototype 3D Studio Max plug-in, EventD², has been made by Panagiotis Chatzitsakyris (2005), and tested on a small unmaterialised residential design project by Le Corbusier (Citrohan House, 1929). A user first inserted a character moving from an end of the living room near the entrance to the gallery in rear and then back to the living room (figure 4, left). Then the objects of interests were specified to include every window that frames the view of the sea as well as Mondrian’s painting placed on a living room wall.

The Rhythmic style preset quickly generated the sequence illustrated in the left column of Figure 11. Every shot there is sliced to three seconds segments, and the selection of the montage segments favors the ones including the windows of interests. In comparison, specifying instead the staircase as an object of interest generated the right column of Figure 11. As many of the cameras were rotated and targeting at the staircase this time, the montage sequence reveals the staircase from cameras placed in different locations that alternate throughout the sequence. Choosing the MonoShot style preset creates yet another different sequence that has no interruption in camerawork. Toggling among different editing parameters of a style preset enables exploration of various film-making conventions while producing credible alternatives.

After previewing these alternatives, the user may develop the formal design of the space. For instance, if the feedback suggests a more sense of enclosure with a figurative object in the center, some windows can be replaced with opaque surfaces and the form of the staircase can be further elaborated. The new design can then quickly be tested through generation of a revised walk-through sequence.

**Conclusion**

The test on Le Corbusier’s Citrohan House demonstrates that it is possible to build a sophisticated digital animation tool that allows a quick, automated generation of a credible cinematic representation of architectural space. This is achieved by algorithmically implementing montage styles and camera placement conventions. The process takes advantage of moving inhabitants as reference to cinematic spatial visualization, and generates a sequence in a format familiar to ordinary audience who consumes the language of commercial films and television dramas. Also the demonstration shows that a selection of different objects of interest and montage style produces a dramatically different sense of space and conveys different spatial characteristics.

This digital tool is useful for quickly producing a credible cinematic presentation of a designed environment. Because of its speed and automation, an architect may well use it in the context of a design in progress. An architect can get a quick visual feedback of movement in the space, and the spatial form can be revised accordingly to generate the next cinematic representation.

The future development immediately targets
implementation of other montage styles and other types of cinematic camera placement. Another area of automation opportunity lies in the generation of a mobile figure's walking path. This is manually specified by the user now, but authors believe that there are correlations between a type of spatial form and a walk through path that appropriately highlights experience of the space type. For example, an elongated space and a concentric space typically suggest different walk through paths for cinematic representation. Analysis of conventional character movement in various film sets and locations as well as sophisticated form recognition algorithm will be required to make a meaningful development for a tool to help such a path generation.

References

Arnheim R.: 1957 (Originally 1933), Film as Art, University of California Press, London.