ShowMotion – Camera Motion based 3D Design Review

Nicolas Burtnyk, Azam Khan, George Fitzmaurice, Gordon Kurtenbach

Alias
210 King Street East, Toronto, Ontario, Canada
{ nburtnyk | akhan | gf | gkurtenbach }@alias.com
www.alias.com/research

Abstract

We describe a new interactive system for 3D design review, built to exploit the visual perception cue of motion parallax, in order to enhance shape perception and aesthetic evaluation. Traditional CAD applications typically use "bookmarked" static views for design evaluation. In our system, we replace static views with moving "shots" interspersed with cinematic visual transitions. Furthermore, users can access shots by picking object features on the 3D model, which invokes a spatial search over all shots, selects the ones deemed relevant to the feature, and plays back those shots. In this paper we describe the rationale behind this system, the main components, the implementation of the search technique, and the initial user evaluation of the system.

Categories and Subject Descriptors: I.3.6 [Computer Graphics]: Methodology and Techniques – Interaction Techniques; H.5.2 [Information Interfaces And Presentation (HCI)]: User Interfaces – Interaction styles, Input devices and strategies.

Additional Keywords and Phrases: interaction techniques, camera controls, 3D navigation, 3D viewers, 3D visualization, design review.

1   Introduction

Most traditional CAD applications do not provide sophisticated tools for interactively reviewing a 3D design. At the minimum, a designer can use the 3D scene navigation controls used in modeling for reviewing the completed model. "Bookmarks" are one improvement over manual navigation that allow users to create a collection of saved views that can be selected during a design review. The net result is a collection of still images with perhaps a computer generated camera animation between the bookmarks.

Thus the predominant visuals used in design reviews are static images – essentially a slideshow. Unfortunately, this emphasis on static images does not make use of the visual perception cue of motion parallax. As the literature suggests, motion cues are useful in 3D object assessment [Todd 2004, 1998]. In particular, motion cues are extremely important in the perception of subtle shapes and are of great value to aesthetic design reviews.

Our approach is to use motion instead of static images as the basis of composing and conducting a review (see Figure 1). The core idea is that design review content should be "alive" (i.e., in motion) most of the time. Obviously, using too much motion could be problematic but we have discovered that even small motions, such as a slow camera zoom-in[1], provides the necessary 3D sensation without producing a visually jarring and distracting experience. The use of subtle camera motions "to keep a scene alive" is a common technique in cinematography , from which we also borrow the notion of pleasing cinematic visual transitions, such as fade-to-black between shots, to further enhance the overall experience.

Value of Motion

We can begin to identify *when* one wants to use motion as well as the *value* of the motion for design review. Let us examine three existing tools and practices: (1) traditional navigation controls, (2) virtual turntables, and (3) bookmarks.

With traditional camera tools for scene navigation, users are fairly unconstrained in their motion to get to a new view. Unwanted views can often result during a design review particularly when the user has their focus of attention on the content of the discussion or if they are nervous. Inevitably, camera motion involves frequent clutching with a mouse and this produces a clunky, visually jarring, time consuming and distracting transition. Thus, the motion quality is quite low. Moreover, the visual experience is often very inconsistent. For example, moving back and forth between two views will produce different visual experiences each time since it is very difficult to exactly traverse the same camera path manually.
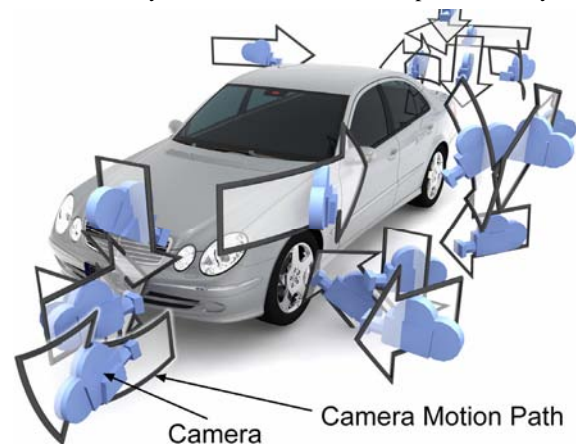

Figure 1. ShowMotion provides a collection of preauthored camera motion paths that can be selected and played by the user.

---

[1] The term "zoom" is commonly misused. Camera movement in and out from a subject is actually referred to as "dolly".

Virtual turntables offer a partial solution in that they can provide a consistent orbital motion path for object inspection. While this is useful for offering an overview of the 3D object, its value quickly degrades as the same information is being repeatedly presented. Most designers we have consulted do not make significant use of turntables. If they are used, it is only for a few full rotations and then it is turned off as it no longer provides any value, or because it is too annoying to have it playing continuously. This is an example of how good motion can quickly degrade to bad motion.

Bookmarks attempt to avoid some of these problems by offering consistent views and procedurally generated animation paths between the views. Unfortunately, in many respects, they produce low value motion when you don't want it (going between items of interest) and no motion when you need it (when you arrive at an item of interest). That is, bookmark transitions involve motion and then present a still image. Our approach is to eliminate the low value transition motion and replace these with cinematic cut/fade transitions. At an item of interest, instead of a still image, we introduce high value motion to sustain a user's interest and to assist understanding of the shape of the 3D object under review.

This paper focuses on how we can generate valuable motion and cinematic visuals at the right time in a 3D design review.

2 Background & Related Work

2.1 Basics of Design Review

We based our work on observations we have made in working closely with automotive designers in efforts to enhance their virtual design processes. Probably the most frequent activity, especially in the design of an aesthetic object, is looking at the object (e.g. looking from different viewpoints and examining details). Another major activity is the comparison of designs, for example, the automotive designer will compare the design of last year's model's taillight against the current design. Virtual design tools can be used in a manner to support these types of design review activities. However, despite the advances in computer display technology and interaction, designers report that looking at the physical object is still much more effective than viewing a virtual object on a computer screen. We feel that this is true for a number of reasons:

- Inspecting a physical object is easy — simply walk around or move your head. Even designers who are extremely proficient with virtual 3D navigation techniques report that a key advantage of physical models is that others who are not skilled in 3D navigation techniques, such as managers and potential users of the object, can inspect the model.

- Physical objects more accurately reflect the way the object is actually viewed in the physical world. For example, a physical car milled out of clay at full scale can be viewed in the same manner as walking up to or standing by the final car. A viewer does not have to take into account that the display of the object is a 2D projection. Moreover, in the physical realm, traditional 3D cues such as motion parallax and occlusion are naturally at play and help to evaluate the 3D shape.

- Physical objects allow for viewing from more angles and for more accidental viewpoints. Here a designer can see a model from a number of unexplored viewpoints when walking around the object or, for example, when bending down to pick up a dropped pencil.

- Designers report that physical objects allow for more "face time" on a model. When viewing a virtual model using traditional techniques, there is no pacing in the presentation that influences the amount of time that the images are viewed. This leads to shortened amounts of time that designers look at virtual 3D models. However, in a physical space, pacing of the presentation is slowed by the many physical constraints including the speed that humans move themselves through space and can reorient themselves, leading to more time spent looking at the model. More face time suggests a greater understanding of the 3D shape.

Above we outlined the problems with traditional virtual design review techniques, and the benefits of physical reviews. However, there are still shortcomings in both areas. For example, to properly evaluate variations of features against each other, the viewing experience must only expose the qualities being evaluated, and must otherwise not change from one variation to another. Otherwise there are too many changing factors to make unbiased decisions. The same is true for comparisons to other versions of the model or comparisons to competitive models. When comparing a design against dozens of models, a virtual design review may be the only practical way to attain the large number of identical viewpoints needed. As such, a primary feature of our system is to provide reproducible high quality motion.

In this paper we examine techniques which move beyond the accepted and common ways of supporting design review in the virtual realm (freeform navigation, turntables, bookmarks) and look at ways of more accurately reproducing the properties of review in the physical realm as well as augmenting it with techniques not available in the physical world. Specifically, we explore how movement of the virtual camera can be exploited to enhance the virtual design review process.

2.2 Related Systems

3D Navigation
A great body of research has studied the issue of efficiently navigating around complex virtual environments [Steed 1997; Mackinlay et al. 1990; Tan et al. 2001]. Some have attempted to achieve this through gestures [Zeleznik and Forsberg 1999], miniature overviews [Stoakley et al. 1995], constraints [Gleicher and Witkin 1992; Hanson and Wernet 1997; Bares and Kim 2001; Bares and Lester 1999] utilizing both hands [Balakrishnan and Kurtenbach 1999; Zeleznik et al. 1997] offering predictor based feedback [Chapman and Ware 1992] and sketching paths [Igarashi et al. 1998].

A variety of metaphors have been explored for navigating in a virtual environment such as "eyeball", "scene in hand", and "flying" [Ware and Osborne 1990]. The metaphor of a guided tour [Galyean 1995] has also been explored to pre-author camera paths and to signal interesting viewing vistas.

Camera techniques that use continuous depth sampling to modulate the flying speed in virtual environments have been studied [Ware and Fleet 1997]. Similarly, techniques such as HoverCam orbit objects at a constant distance and face the object's surface [Khan et al. 2005].

3D bookmarks have been used in many systems including the early 3D VRML environments [Rezzonico and Thalmann 1996]. The animated transitions between bookmark views, known as object constancy [Robertson et al. 1989], reduces the need to reassimilate the material once at the new view because the user saw "how we got there." We believe this is less important for very familiar shaped objects such as a car.

Rules can be defined for cameras to automatically frame a scene that follow cinematic principles such as keeping the virtual actors visible in the scene or following the lead actor [Christianson et al. 1996; He et al. 1996; Oliveros 2004]. Automatic path planning algorithms [Wan et al. 2001; Stoev and Straber 2002; Drucker and Zeltzer 1995] have also been used for virtual navigation as well as finding effective viewing angles for 3D direct manipulation [Philips et al. 1992].

The notion of a cinematic shot being applied to an interactive 3D environment has been illustrated in systems such as the Storyboard Frame Editor [McDermott et al. 2002]. In this work, an authoring system is designed to compose shots and transitions to compose 3D animation films. The editor automatically exports a set of constraints used by a solver to compute a satisfactory shot of the subjects in the 3D scene.

StyleCam [Burtnyk et al. 2002] has similar goals in providing a cinematic experience but has a more complex interaction model consisting of camera surfaces and animation clips. In StyleCam, the pre-authored motion occurs in the transition and when the user arrives at the point of interest (i.e., the camera surface) they are presented with a still camera view and are then able to manually move the camera to a limited degree. ShowMotion is similar to StyleCam in its use of cinematic transitions, but differs significantly in that it employs pre-authored camera motion at points of interest. The two systems have different goals too: StyleCam explorers the idea of an "interactive TV commercial" while ShowMotion explores the needs of design review.

In general, the related work on navigation techniques emphasize the framing of an object of interest or the act of getting to the object of interest. However, none of them focus on providing automated and reproducible motion at the object of interest. This is the main difference with our research.

Design Review
Some researchers have focused on providing a design review system. For example, Kremer's design review system [Kremer 1998] attempt to showcase 3D models, variations and engineering specifications (e.g., size, cost) in a unified environment. Specialized input devices such as the Flystick [Knopfle and Vob 2002] or the Boom Chameleon [Tsang et al. 2002] attempt to facilitate design review by easing the manual navigation.

Temporal thumbnails [Tsang et al. 2004] offer a way to quickly convey information about the amount of time spent viewing

specific areas of a virtual 3D model. Similarly, Singh and Balakrishnan [2004] show how data mining of previous camera moves might assist in future presentations. These types of visualizations could be useful for design managers to get an understanding of where current design activities are focused.

Depth Perception & Motion
Visual depth perception [Murray 1994] has been widely studied. While some applied researchers focus on creating taxonomies of viewpoint motion control techniques [Bowman et al. 1999; Bowman et al. 1997] others have studied the low-level neuroscience of 3D motion cues [Peuskens et al. 2004]. In particular, Cutting and Vishton [1995] studied the relative strengths of nine depth cues including occlusion, relative size, relative density, height in the visual field, aerial perspective, motion parallax (i.e., motion perspective), binocular disparities, convergence, and accommodation (see Figure 2). With the exception of occlusion, motion parallax ranks highly compared to other types of depth cues within distances relevant to design reviews.
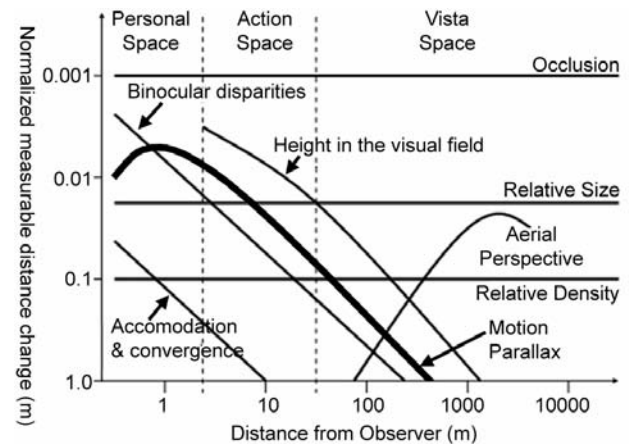

Figure 2. Relative depth cue strengths. Vertical axis indicates that when smaller changes are needed, the given cue is stronger.

Moreover, the aesthetic need for motion in interaction design within digital media has been studied [Bacigalupi 1998]. In this effort, Bacigalupi analyzes the aesthetics of movement in visual arts to explore how we map our 3D kinetic experience to make sense of 2D qualities in artifacts with the ultimate goal of providing a building block of concepts and vocabulary to link between designer, artifact and viewer experience.

Motion in general provides strong visual cues for the perception of shape and depth [Todd 2004]. For example, particles flowing over a surface, called Kinetic Visualization [Lum et al. 2002], can aid in highlighting the shape of an object that might not be readily visible with conventional rendering that uses lighting and view changes.

3 ShowMotion System

The ShowMotion system is built on top of an experimental 3D visualization application. This application allows extremely large, detailed 3D CAD models to be loaded and assigned materials. Fully detailed models can then be viewed at interactive frame-rates with high-quality real-time rendering. Interactive navigation is supported with standard

tumble/pan/zoom camera controls and bookmarks. ShowMotion augments this experimental application.

ShowMotion consists of an authoring mode, used to setup a design review environment, and a viewing mode in which users interact with the constructed environment to conduct a design review. The main entities that users author and interact with are: shots, transitions, shot sets, variations, and searches (Figure 3).
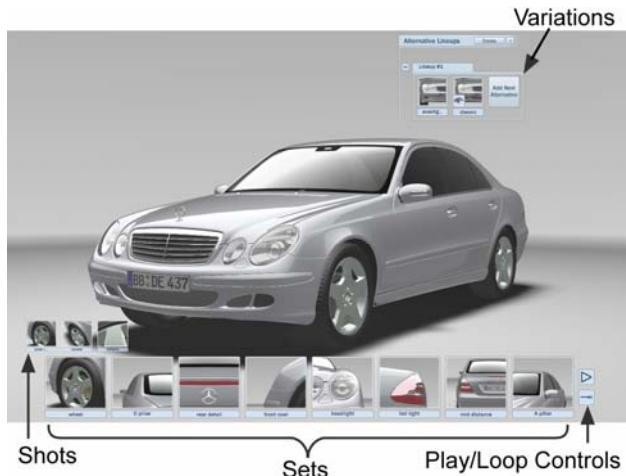

Figure 3. ShowMotion screen. The shot set selector is displayed in a row along the bottom of the screen.

Authoring: Shots, Shot Sets, and Transitions

The concept of a *shot* is borrowed from cinematography. It is simply a visual sequence of camera motion that runs for a short period of time, typically from 3 to 10 seconds. Predefined shot types are available which simplify authoring and provide high quality camera movements. In addition to these predefined camera movements, custom camera movement paths can be imported from other 3D animation software packages, such as Alias's Maya animation application. Still images are also possible, as are camera movements based on specifying starting and ending camera positions. Table 1 summarizes the properties required to set up a shot.

Shot Authoring Properties				
Shot Type	Camera Movement	Current Camera used as:	Duration	Speed
Predefined	Zoom in Zoom out Track left Track right Crane up + left Crane up + right Crane down + left Crane down + right	Starting Point Half-way Point Ending Point	Specify in seconds	Adjustable
Still	None	Viewpoint		N/A
Keyframe	Interpolated between Start and End Point	Starting Point / Ending Point		N/A
Imported motion	Any imported motion path	N/A	Adjustable	N/A

Table 1. Shot authoring properties for 4 shot types.

To create a shot using a predefined shot type, the author positions the 3D camera using standard tumble/pan/zoom camera controls, selects the type of shot to be created, its

duration, speed, and other properties. Depending on the type of shot being created, the system uses the current camera position as start, halfway, or end position of the camera in the shot. Once a shot has been created, a thumbnail of the first frame is presented in a *shot selector* at the bottom of the screen, as in Figure 3.

Shots can have two classes of look-at behavior: *fixed point* (a fixed target, while the camera moves along the side of the car, see Figure 4), or *fixed direction* (a moving target, see Figure 5).

Figure 4: Fixed look-at point.

Figure 5: Fixed direction uses a series of look-at points.

The shot selector allows the shots to be played back in left to right sequence by pressing a "play" button or accessed randomly by clicking on a shot thumbnail. A loop toggle button allows a shot set or the entire shot collection to be repeatedly played. At any time, users can click on the stop button to pause shot playback.

Our system has the ability to group shots into sets. This feature is used to group related shots together. For example, in an automotive design there may be many shots for each design feature, such as the headlights, grill, taillight, etc. The intention is that a shot set provides a rich set of viewings of a particular design feature that can be easily invoked. Clicking on the thumbnail associated with a set causes that whole shot set to be played back in sequence. Shots within a set can also be played individually by clicking on their associated thumbnail.
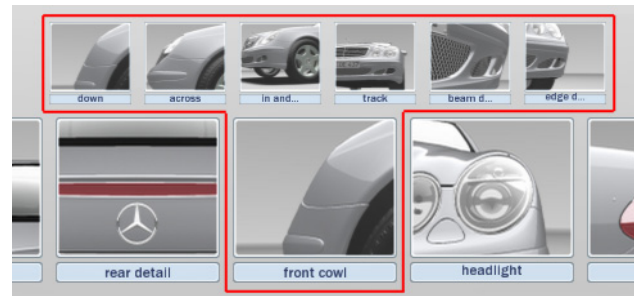

Figure 6. Shot sets contain individual shots.

The large thumbnails in the bottom row represent shot sets. The smaller thumbnails in the top row represent shots within the current set. An outline, shown in red, encircles the set thumbnail with its shots (Figure 6).

170

Transitions between shots are also supported. Transitions can be instant (a "cut") or a gradual fade to the next shot (a "fade-to"). Other transition types are possible but we currently limit our system to these two types. The duration of a transition can be specified by the author. The fade-to transition overlaps the end of the previous shot and the beginning of the following shot. For the first half of the fade, the view is incrementally mixed with black over time until the frame is completely black. Then, over the next shot, during the second half of the fade, the process is reversed until there is no black overlay in the frame.

Variations

In design reviews, comparing different designs is a desirable and common activity. Typically, several designs are viewed "side by side" so a design can be easily compared. Another comparison technique is serial presentation. While having the disadvantage of not being able to view the two designs simultaneously, it has the benefits of being able to view design variations from an identical view point. In our system we provide support for both of these activities within the context of camera based motion viewing.

Our system allows designers to construct a set of design alternatives and automatically switch between them. For example, in Figure 3, an auto design is shown with a set of alternate designs for the wheels, grills, and trim. These alternatives were constructed by a designer by creating the needed geometry and material properties, then grouping sets of these objects into particular design alternatives. When a button for a particular alternative is selected, the other alternatives in the set are hidden from view and the selected alternative is displayed. Using this, a designer can cycle through design alternatives evaluating their appearance.

Most importantly our system is configured so that alternatives can be changed while a shot or set is being played. For example, a designer can watch the "wheel" shot sequence, three times over, each time switching to an alternate design between sequences, or in the middle of a sequence, as desired. In this way a designer can have a rich set of viewings of each alternate design in a rapid, identical and reproducible way.

We also support "side by side" comparisons. For example, our systems screen can be split into two views with duplicate designs on either side. Then alternate designs can be selected for either model, allowing side-by-side comparison. The viewpoints of these two views are linked such that camera movements in one view are automatically applied to the other view. This allows identical manual review of objects. Additionally, a shot or set playback updates both views allowing for a rich set of viewings of the two different designs simultaneously in a "side by side" manner.

Click on thumbnails to navigate

Our system emphasizes a "click to navigate" design that is intended to remove the need for reviewers to understand and control typical high degree of freedom 3D camera controls such as tumble/pan/zoom controls by dragging or navigation via 6 DOF devices. For example, rather than navigating via these types of controls to a particular viewpoint, the user finds a shot corresponding to the viewpoint and goes to that viewpoint by a simple single click on its thumbnail.

Another benefit of the simple "click to navigate" approach is that this allows non-presenters to be successful in controlling the system and inspecting the design. We have observed that many formal design review presentation meetings breakup into situations where members of the audience (other designers and managers) have the opportunity to interact with the system while the presenter is busy talking to others. In this case, we have observed that many audience members are unable to successfully use the traditional navigation controls and quickly give up, most likely leaving the system at an awkward and low value viewpoint. In contrast, our intention with the "click to navigate" approach is that audience members will easily understand how to get to certain design viewpoints by watching the presenter and successfully operate the system themselves.

A final benefit of this simple navigation scheme is that since the system is never displaying awkward viewpoints or jarring visuals no matter who is controlling it, the presenter's design is always portrayed in a polished manner thus enhancing the impact of the presentation and reviewer perception of the design.

Shot Search - Click in scene to navigate

Burtnyk et al. [2002] reported with the StyleCam system that users have strong expectations to be able to click on objects in the 3D scene to have the system present information concerning the object or "to be taken to that object". For example, users expected when they clicked on a car's wheel that the system would either give them more information on the wheel or move the viewpoint to the wheel.

In ShowMotion we are attempting to fulfill this expectation by allowing an author to associate a shot or shot set with a particular location in the 3D scene or on the 3D object.

Since our current example involves a design review for a car (instead of say a shampoo bottle), it is our intention to have many shots (on the order of 50 to 100) defined for the interior and exterior of the car and the surrounding environment. Given this number of shots, it becomes difficult to manage and select any particular shot. Finding a specific shot during a design review can become disruptive and time consuming and this potentially breaks the conversational flow. Moreover, scrolling through a large collection of shots to find one that showcases a component that is already visible in the 3D view could be frustrating. Users tend to want to click on the component in the 3D view.

To support quick shot selection and direct navigation, we use a two-step heuristic. (1) Based on where the user clicked on the model, search neighboring space for shots in the vicinity. (2) Play the shot (or shot set) found, or, if no shots close to the selected spot are found, we currently take no action. Based on a spatial relevance metric (see next section for details) we playback the most relevant shots.

To facilitate "returning" from a search a user can click one of the thumbnails in the sequencer view. Typically, one particular thumbnail is an overview shot of the 3D object and this serves as the "home" button. This is especially useful if a user is zoomed-in to see surface details and wants to navigate to another region on the object.

3   Implementation

The ShowMotion system was implemented and evaluated on a high-performance PC running Windows XP with a dual-core AMD Athlon X2 processor at 2.2 GHz, 2 GB of DDR400 RAM and two NVidia GeForce 7800 GTX graphics cards in an SLI configuration. ShowMotion is built on top of an experimental 3D visualization application which is written in C++ and Python and makes use of OpenSceneGraph and OpenGL to achieve high levels of performance. We extended this application to build the ShowMotion functionality using Python (mostly) and C++ (sparingly).

3.1   Camera movement and transitions

Based on the camera movement type selected by the author, a Bézier path is created at a position relative to the camera's world space position when the shot is created. This path defines the position of the virtual camera in space over the duration of the shot. Camera movement is achieved by evaluating the Bézier path at the current time. Fade-to transitions are implemented by overlaying a black, screen-aligned quad and varying its opacity over the course of the transition.

3.2   Search Algorithm

The input to the search algorithm is a "target point" on the model surface. This point is derived by casting a ray into the scene from the screen coordinates under the mouse cursor and computing the intersection of this ray with the model. The goal is to find which shots are looking at or near the target point.
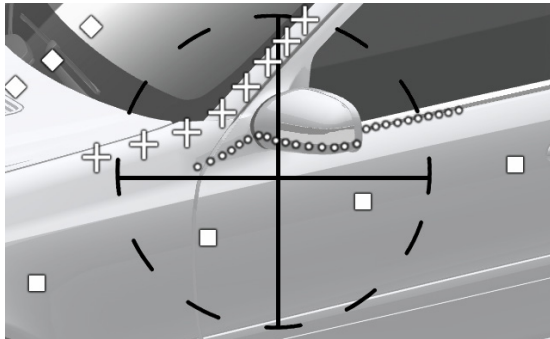

Figure 7. Algorithm visualization: squares, dots and crosses indicate individual shots and "look at" points; dashed target circle shows search clipping region and target center as initial user search point.

The search algorithm has four high-level steps:

1.   For each shot, compute the distance from each of its look-at points (explained below) to the target point.

2.   For each shot, compute its relevance, a function of the distances computed in step 1.

3.   Sort the shots in decreasing order of relevance, discarding those shots with zero relevance.

4.   Return the top N most relevant shots.

When a shot is initially created, the system computes a series of look-at points that encode what the shot is looking at (see Figure 7: squares, dots and crosses represent look-at points for different shots). The shot is sampled at a fixed frequency, and at each time step a ray is cast from the camera position along the viewing direction and its intersection with the model is computed and stored. In our implementation, we sampled each shot at a rate of 30 frames per second (fps). With an average shot length of about 8 seconds, we typically have approximately 250 sample look-at points per shot. Though the calculation of these intersections is relatively expensive, it need only be performed once, on shot creation and therefore does not interfere with the user experience during a search operation.

Given a target point and a shot's look-at points, we compute the relevance of the shot. First we calculate the distance d_i between each look-at point and the target point. Next, for each d_i we calculate a relevance value r_i from:

$$r_i = f(d_i) = \begin{cases} 0 & d_i > D_{max} \\ \\ \left(\dfrac{d_i}{D_{max}} - 1\right)^2 & otherwise \end{cases}$$

Note the use of the value D_{max} which is the distance beyond which we choose to effectively ignore shot look-at points by setting their relevance to 0 (see Figure 7 area outside the target circle is ignored). Finally, a shot relevance R_{shot} is calculated by summing the r_i's of a shot:

$$R_{shot} = \sum_i r_i$$

The following graph shows how the relevance of a shot look-at point varies as a function of distance from the target point.
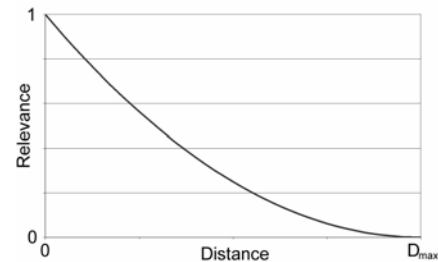

Figure 8. Relevance as a function of distance.

With a relevance value assigned to each shot, the N most relevant shots are returned. In our evaluation we chose to set N=3 which resulted in the user experiencing a small number of highly relevant shots when they clicked on the model. The system can also optionally return the most relevant shot set, which is defined when M of the top N most relevant shots belong to the same set (N=5, M=3 gave good results). Returning the most relevant shot set is most useful in scenes with a large number of authored shots where there may be a dozen shots relating to a given feature like the door handle. In such a scenario, returning the most relevant shot set gives the user a richer experience than simply returning the N most relevant shots.

It is, in general, possible for all shots to have a relevance of zero and our system takes no action in this case. Currently this is a limitation that requires future work to find an appropriate response.

4 Initial User Evaluation

To evaluate our system we asked three experienced industrial designers who work closely with major automotive design studios to design some sample review experiences based on the 2005 Mercedes-Benz E-Class model provided to us by Mercedes-Benz design studios.

We began by explaining to the designers the rationale behind our system and then worked with them to produce what they thought would be valuable shots and sequences. In this process they gave us feedback on our "motion based" approach in general and on the particulars of the implementation. Over the course of 4 hours a collection of 5 shots were authored, evaluated and discussed.

The designers in general agreed with our motivation and rationale, quickly understood the main concepts of our approach, and could think of many examples where moving the camera was valuable. The designers proposed three basic shot types: overview/establishing (viewing over the whole model), normal (viewing of a particular feature) and detail (close-up viewing of model features). We were able to use our predefined camera movement types (zoom-in, zoom-out, track left, etc) to satisfy all the desired shots, with the exception of an overview "turntable" like shot that was authored in Maya and imported.

One surprising result of this session was the length of the designer's shots. While we had expected shots to be relatively short (5 to 10 seconds), the designers created shots with very slow camera movement resulting in shot lengths of 30-60 seconds. In addition, we observed the designer intensely watching these slow moving shots. We believe that this was because the designers were in fact taking the time to evaluate the aesthetics of the auto design.

Ideas for many different shot types to enrich design reviews came up in the discussions: shots where the camera is stationary and the model moves to simulate large physical turntable for cars; shots customized based on an individual's height (for example: how a 6 foot tall person sees the car from the side); shots from particular viewpoints (for example, slowly moving down the side of a car), etc. The designers also pointed out that different designers have different viewing "tests", for example, a three-quarter view from 5 meters back, that could be programmed into the system. Finally, many real world usage context shots were suggested such as "view of the car as it drives by" or "view of the car as it passes another car".

Users found it intuitive to click in the 3D scene to select relevant shots and to navigate to points of interest. Additional testing is needed to determine if users wish the system to automatically infer what particular shot experience they want. For example, clicking on a particular point could imply many things: user wants a detailed view; user wants an overview from this point; user wants a single shot from this point or shot set, etc. Alternatively, rather than have the system try to infer what type of response the user wants, the user could input this information.

In summary the results of this evaluation indicate that there is value in our approach. Our belief is that as designers spend more time considering the use of shots and using our system, they will further develop methods of using camera motion and animation to enhance design evaluation.

5 Conclusion & Future Work

The ShowMotion system has demonstrated how camera motion can be used instead of static images as the basis of composing and conducting a design review. The main contribution of this system is the unique and compelling way of interacting with a 3D scene through a combination of cinematic shots, transitions, sets, search, and variations.

The well understood benefits of offering motion parallax as a foundation to the system was valued by the users. The "shot" metaphor is easy to understand and having a number of preset shot motions aids the designer in rapidly composing multiple and repeatable viewing experiences to evaluate their designs. Adding cinematic effects such as "fade to black" as transitions between shots improves the overall pacing and visual quality of the experience.

In addition, we have also found it effective to use shots as an alternative to manual camera navigation. This allows us to maintain the cinematic quality of the presentation instead of the often "jerky" and visually disruptive motion caused by manual camera movements.

With a large number of shots defined for a 3D object, our shot search algorithm and interaction technique assist in accessing shots and satisfying user expectations. We have found that the act of clicking in the scene to navigate by retrieving the most spatially relevant shots works surprisingly well. We wish to further refine the spatially relevant heuristics to automatically distinguish if a user is requesting an establishing, normal or detail view shot.

Initial feedback indicates that the shot sequencer and our single point-and-click interface work well in a group design review setting. In particular, when using a touchscreen display our "click to navigate" approach provides easy navigation and rapid turn taking while maintaining a cinematic, high quality visual presentation.

There are a number of future developments we wish to explore. First, we would like to create different collections or libraries of shots for classes of objects (cars, shampoo bottle, laptop, etc.). Second, we would like to provide ways of adjusting a shot collection based on the 3D object. For example, re-using a collection of shots from one car to another slightly bigger and different car model can involve scaling the shots in an intelligent non-proportional manner that factors in feature points. Thirdly, we would like to explore supporting in-scene animations such as moving lights, objects and dynamic environments. Fourth, we would like to create more pre-defined shot motions and conduct further user evaluation and testing.

References

BACIGALUPI, M. 1998. The Craft of Movement in Interaction Design. In *Proceedings of ACM AVI 1998*, pp. 174-184.

BALAKRISHNAN, R. AND KURTENBACH, G. 1999. Exploring bimanual camera control and object manipulation in 3D graphics interfaces. In *Proceedings of ACM CHI*. Pp. 56-63.

BARES, W. AND KIM, B. 2001. Generating Virtual Camera Compositions, In *Proceedings of ACM IUI'01*, pp. 9-12.

BARES, W. AND LESTER, J.C. 1999. Intelligent Multi-Shot Visualization Interfaces for Dynamic 3D Worlds. *Proceedings of ACM IUI '99*, pp. 119-126.

BOWMAN, D., JOHNSON, D. AND HODGES, L. 1999. Testbed environment of virtual environment interaction. In *Proceedings of ACM VRST*. pp. 26-33.

BOWMAN, D., KOLLER, D. AND HODGES, F.H. (1997). Travel in immersive virtual environments. *IEEE VRAIS'97 Virtual Reality Annual International Symposium*. pp. 45-52.

BURTNYK, N., KHAN, A., FITZMAURICE, G., BALAKRISHNAN, R. AND KURTENBACH, G. 2002. StyleCam: Interactive Stylized 3D Navigation using integrated Spatial and Temporal Controls, In *Proceedings of ACM UIST 2002*, pp. 101-110.

CHAPMAN, D. AND WARE, C. 1992. Manipulating the future: predictor based feedback for velocity control in virtual environment navigation. *ACM Symposium on Interactive 3D Graphics*. 63-66.

CHRISTIANSON, D.B., ANDERSON, S.E., HE, L.-W., WELD, D.S., COHEN, M.F., AND SALESIN, D.H. 1996. Declarative camera control for automatic cinematography. *Proceedings of AAAI '96* (Portland, OR), pp. 148-155.

CUTTING, J.E. AND VISHTON, P.M. 1995. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth, In W. Epstein & S. Rogers (eds.) *Handbook of perception and cognition*, Vol 5; Perception of space and motion, pp. 69-117, San Diego, CA; Academic Press.

DRUCKER, S. AND ZELTZER, D. 1995. CamDroid: A System for Implementing Intelligent Camera Control. In *Proceedings of ACM Symposium on Interactive 3D Graphics*, pp. 139-144.

GALYEAN, T. 1995. Guided navigation of virtual environments. *ACM Symposium on Interactive 3D Graphics*. pp.103-104.

GLIECHER, M. AND WITKIN, A. 1992. Through-the-lens camera control. *ACM SIGGRAPH 92*. pp. 331-340.

HANSON, A. AND WERNET, E. 1997. Constrained 3D navigation with 2D controllers. *IEEE Visulization*. pp. 175-182.

HE, L., COHEN, M. AND SALESIN, D. 1996. The virtual cinematographer: a paradigm for automatic real-time camera control and directing. *ACM SIGGRAPH 96*. pp. 217-224.

IGARASHI, T., KADOBAYASHI, R., MASE, K. AND TANAKA, H. 1998. Path drawing for 3D walkthrough. *ACM UIST*. pp. 173-174.

KHAN, A. KOMALO, B., STAM, J., FITZMAURICE, G. AND KURTENBACH, G. 2005. HoverCam: interactive 3D navigation for proximal object inspection, In *Proceedings of ACM Symposium on Interactive 3D graphics and games*, pp. 73-80.

KNOPFLE C. AND VOB, G. 2000. An Intuitive VR User Interface for Design Review. *Proceedings of ACM Advanced Visual Interfaces*, pp. 98-101.

KREMER, K. 1998. A Concept for Virtual Reality Tools for Design Reviews. *Proceedings of IEEE Visualization*, pp. 205-210.

LUM, E.B., STOMPEL, A. AND MA, K.-L. 2002. Kinetic Visualization: A Technique for Illustrating 3D Shape and Structure, In *Proceedings of IEEE Visualization 2002*, pp. 435-442.

MACKINLAY, J., CARD, S. AND ROBERTSON, G. 1990. Rapid controlled movement through a virtual 3D workspace. *ACM SIGGRAPH 90*. pp. 171-176.

MCDERMOTT, S. LI, J. AND BARES, W. 2002. Storyboard Frame Editing for Cinematic Composition. In *Proceedings of ACM IUI 2002*, pp. 206-207.

MURRAY, J. 1994. Some Perspectives on Visual Depth Perception. In *IEEE Computer Graphics*, Vol 28, No. 2, May 1994, pp. 155-157.

OLIVEROS, D.A.M. 2004. Intelligent Cinematic Camera for 3D Games, MSc. Thesis, University of Technology, Sydney Australia.

PEUSKENS, H., CLAEYS, K.G., TODD, J.T., NORMAN, J.F., HECKE, P.V. AND ORBAN, G.A. 2004. Attention to 3-D Shape, 3-D Motion, and Texture in 3-D Structure from Motion Displays. *MIT Journal of Cognitive Neuroscience*, vol 16, no 4, pp. 665-682.

PHILLIPS, C.B., BADLER, N.I., AND GRANIERI, J. 1992. Automatic Viewing Control for 3D Direct Manipulation. In *ACM Proceedings of ACM Symposium on Interactive 3D Graphics*, pp. 71-74.

REZZONICO, S. AND THALMANN, D. 1996. Browsing 3D Bookmarks in BED, WebNet 1996.

ROBERTSON, G., CARD, S.K., AND MACKINLAY, J.D. 1989. The cognitive coprocessor architecture for interactive user interfaces, In *Proceedings of ACM UIST 1989*, pp 10-18.

SINGH, K. AND BALAKRISHNAN, R. 2004. Visualizing 3D Scenes using Non-Linear Projections and Data Mining of Previous Camera Moves. In *Proceedings of ACM Computer Graphics, Virtual Reality, Visualization and Interaction in Africa*, pp. 41-48.

STEED, A. 1997. Efficient navigation around complex virtual environments. *ACM VRST*. pp. 173-180.

STOAKLEY, R., CONWAY, M. AND PAUSCH, R. 1995. Virtual reality on a WIM: Interactive worlds in miniature. *ACM CHI*. pp. 265-272.

STOEV, S.L. AND STRABER, W. 2002. A Case Study on Automatic Camera Placement and Motion for Visualizing Historical Data. *Proceedings of IEEE Visualization*, pp. 545-558.

TAN, D., ROBERTSON, G. AND CZERWINSKI, M. 2001. Exploring 3D navigation: combining speed-coupled flying with orbiting. *ACM CHI*. pp. 418-425.

TODD, J. T. (1998) Theoretical and biological limitations on the visual perception of three-dimensional structure from motion. In T. Watanabe (Ed.) *High-level motion processing computational, neurophysiological and psychophysical perspectives*. Cambridge MA: MIT Press, pp 359-380.

TODD, J.T. 2004. The Visual Perception of 3D Shape. In *Trends in Cognitive Science*, Vol 8, No. 3, March 2004, Elsevier, pp. 115-121.

TSANG, M., MORRIS, N. AND BALAKRISHNAN, R. 2004. Temporal Thumbnails: Rapid Visualization of Time-Based Viewing Data. In *Proceedings of ACM AVI 2004*, pp. 175-178.

TSANG, M., FITZMAURICE, G.W., KURTENBACH, G., KHAN, A. AND BUXTON, B. 2002. Boom Chameleon: Simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. *ACM CHI Letters*, 4(2). pp. 111-120. ACM UIST 2002 Symposium on User Interface Software & Technology.

WAN, M., DACHILLE, F. AND KAUFMAN, A. 2001. Distance-Field Based Skeletons for Virtual Navigation. *IEEE Visualization 2001*. pp. 239-245.

WARE, C. AND FLEET, D. 1997. Context sensitive flying interface. *ACM Symposium on Interactive 3D Graphics*. pp. 127-130.

WARE, C. AND OSBORNE, S. 1990. Exploration and virtual camera control in virtual three dimensional environments. *ACM Symposium on Interactive 3D Graphics*. pp. 175-183.

ZELEZNIK, R. AND FORSBERG, A. 1999. UniCam - 2D Gestural Camera Controls for 3D Environments. ACM Symposium on Interactive 3D Graphics. 169-173.

ZELEZNIK, R., FORSBERG, A. AND STRAUSS, P. 1997. Two pointer input for 3D interaction. *ACM Symposium on Interactive 3D Graphics*. pp. 115-120.