

A hand is visible in the bottom left corner, holding a laptop. The laptop screen shows a 3D virtual scene with a yellowish ground and a blue sky. In the top left corner of the screen, there is a small window with the text 'FPS 30'. The background of the entire page is dark and features a vertical line of light, possibly from a camera or a light source, creating a sense of depth and focus on the laptop screen.

GAME-LIKE NAVIGATION AND RESPONSIVENESS ~IN NON-GAME APPLICATIONS

REAL-TIME RESPONSE AND INTERACTIVE NARRATIVE
PROVIDE A GAME-LIKE EXPERIENCE IN TWO SYSTEMS:
BOOM CHAMELEON FOR EVALUATING VIRTUAL MODELS AND
STYLECAM FOR DEVELOPING ONLINE PRODUCT
MARKETING AND ADVERTISING.

[By Michael Tsang, George Fitzmaurice,
Gord Kurtenbach, and Azam Khan]

VIDEO AND COMPUTER GAMES AIM to give players a compelling interactive experience. Components that help shape this experience include navigation and action through the game world, real-time responsiveness, and a carefully crafted storyline. Many games give players explicit control as to the high-level outcome of the experience. Navigation and action allow players to move through the environment and interact with the virtual world.

Here we analyze two of our own research projects intended

for use in design evaluation and product marketing applications and reflect on the game-like experiences they provide. The first is Boom Chameleon, which incorporates intuitive, real-time navigation with rich interaction during the review of virtual models. The second is StyleCam, which allows user exploration of pre-authored interactive experiences for virtual marketing and product advertising. While both involve inspection of virtual 3D objects, each employs different approaches to giving users an engaging experience.

NAVIGATING IN
A VIRTUAL SCENE USING
THE BOOM CHAMELEON
(PHOTOGRAPH BY
AZAM KHAN).



Figure 1. Navigating with the Boom Chameleon. Moving the physical display changes the viewpoint into the virtual world. When the display is moved to the front of the space, the front of the virtual car is viewable; moving to the side shows the corresponding side of the car; and moving closer to the center shows a close-up the car.

Boom Chameleon

In traditional automotive design settings, design reviews are performed by teams of managers and designers gathered around clay models to inspect the model's physical characteristics. The designers take handwritten notes and even put marks directly on models to register comments or changes. In contrast, as more and more cars are designed using digital tools, the related design reviews increasingly involve examination of virtual models rather than physical ones. Yet automotive designers have told us that reviews and critiques of virtual models can be frustrating and cumbersome. For example, when performing a virtual design review, managers and designers typically gather before a large display. They critique the design from various viewpoints but must rely on the technician at the computer to operate the interface and move the 3D model to the viewpoints. As a result, some of them are frustrated by their personal lack of viewpoint control. This contrasts with the evaluation of physical models where everyone is free to move around and redirect the group's attention.

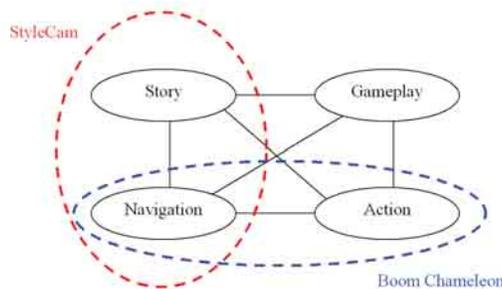
Other problems occur in both physical and virtual reviews [3]. First, information is often lost because

attention is divided between recording annotations and participating in the discussion. Secondly, both managers and designers feel misunderstandings would be reduced if notes and comments were created and shared by all parties simulta-

neously. To address these concerns, we developed the Boom Chameleon [7], a novel input/output device consisting of a flat-panel display mounted on a mechanical armature tracked in 3D space. The display functions as a physical window into 3D virtual environments that preserves a direct one-to-one mapping between real and virtual space; that is, moving the display in physical space changes the corresponding viewpoint in the virtual space.

Experiential components of games.

This configuration results in a spatially aware system [2, 6] where the 3D model appears situated at the base of the device, as if an empty picture frame could move around a fixed object. From the user's perspective, the view responds immediately to changes in the position



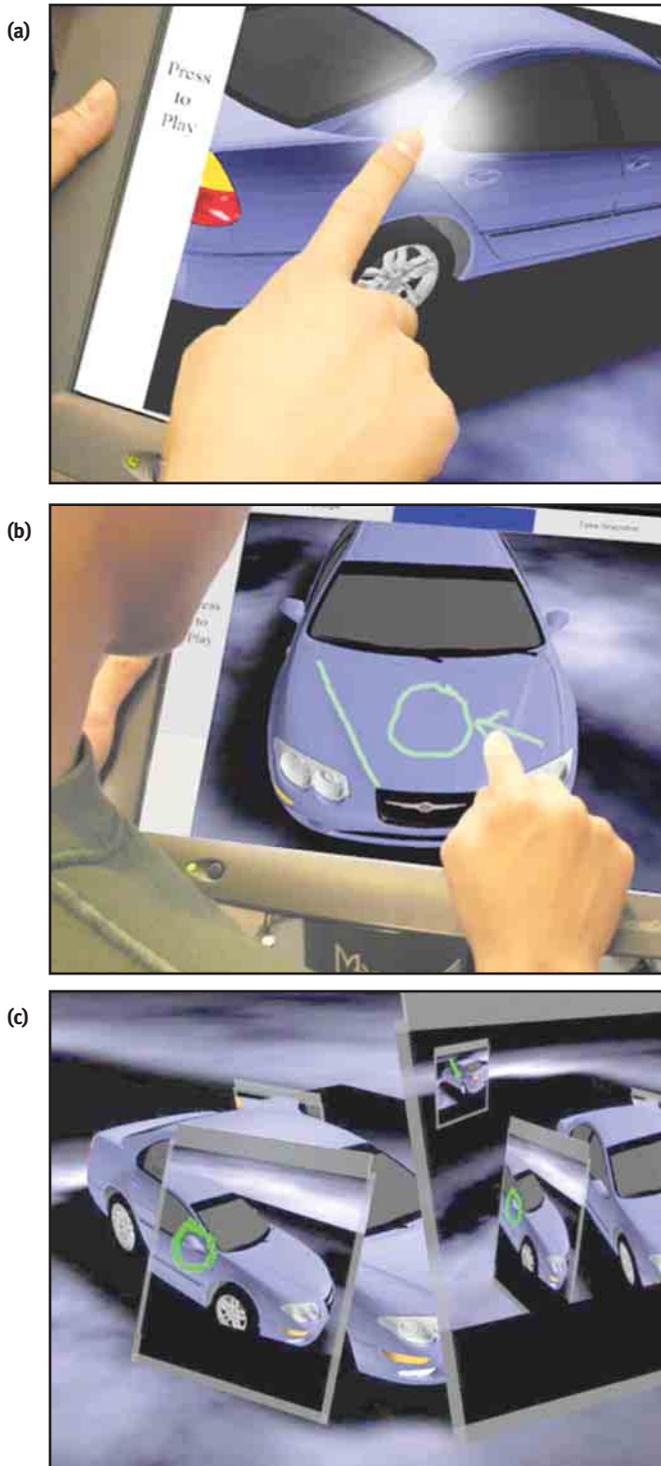


Figure 2. Boom Chameleon annotation tools: (a) the Flashlight for highlighting specific areas; (b) the Pen for drawing on objects; and (c) Snapshots for highlighting specific viewpoints and model features.

and conforms to user expectations. In order to capture the information expressed in design reviews, the Boom Chameleon is configured to record 3D viewpoint, voice, and pointing.

The display has a touchscreen overlay so users can point to objects of interest by touching the screen. A microphone is also attached to the display for recording users' comments as they move about.

When commenting on a design, users often need to indicate positions and draw shapes. The built-in Flashlight tool allows them to temporarily highlight areas of the model (see Figure 2a), and the Pen tool allows them to write objects into the scene (see Figure 2b). As the Pen tool cannot draw in empty space, a Snapshot tool is also included. To create a snapshot, users click on the "Take Snapshot" button, creating an image of the current viewpoint and placing it in the scene (see Figure 2c). Users then write directly on the snapshot image with the Pen. The system allows hiding and layering of snapshots to deal with the clutter resulting from having multiple snapshots in the same scene.

Our design philosophy is to allow users to express their thoughts and ideas with minimal overhead and planning. Toward this end, the Boom Chameleon constantly records all available streams of data, thereby making explicit capture of the information unnecessary. (A similar approach for 2D documents was first used by the Freestyle system [4] developed by Wang Computer.) As a result of this fluid approach, users need not switch modes to record different types of information. This arrangement allows creation of a richer and more fluid set of annotations compared to the traditional, more modal approach to capturing comments. Moreover, users may later view recorded sessions, including audio, markings, and changing viewpoint.

Hundreds of users have used the Boom Chameleon to inspect virtual models. Ranging from experienced 3D computer graphics artists to technology-shy executives, they have immediately understood the navigation mechanism and quickly learned to control the viewpoint. Using the annotation-enabled version for simulated design reviews, we have found that modalities [5] of navigation, pointing, and voice are all interspersed within the same short periods of time. In contrast, if annotation requires the user to explicitly engage different tools (such as to switch between navigating and pointing), the resulting behavior would be quite different. We also found that users like to speak during annotation sessions, usually in parallel with viewpoint movement or pointing gestures. Overall, voice annotations are the predominant style of commentary, along with "framing" the intended 3D view.

and orientation of the display (see Figure 1). Coupling virtual and physical space provides a simple method for examining 3D models that takes advantage of our innate skill manipulating physical objects.

Standard 3D viewing tools typically allow users to independently manipulate various parameters, including left/right, up/down, in/out, roll, pitch, and yaw. Our approach simplifies navigation by combining them into a single physical operation. The interaction feels natural

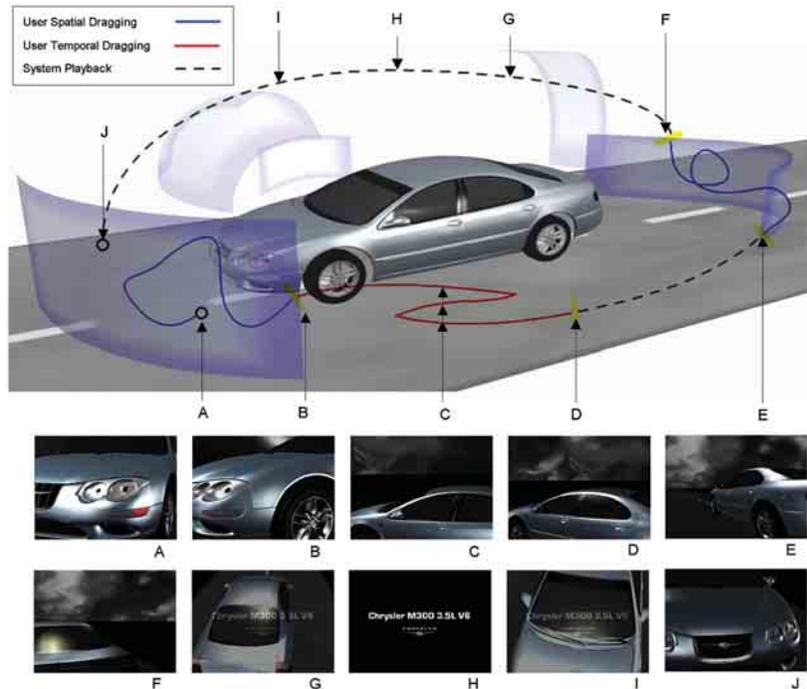


Figure 3. StyleCam experience: (top) system components and their response to user input; (bottom) what the user sees.

and gameplay elements of game design, interaction with the Boom Chameleon does involve game-like features. The system reacts to movement in real time, making it easy to get a “feel” for the navigation. Such responsiveness is similar to video and computer games’ real-time response to user input. Some games allow players to perform amazing feats that defy the laws of physics. The Boom Chameleon allows users to view virtual objects while keeping the interaction rooted in the physical world. The result is a mechanism that seems natural and meets user expectations. Consequently, interaction with the Boom Chameleon is reminiscent of games that adhere to physics-based rules and simulations of “real life” that use easy-to-understand metaphors for navigation. The Boom Chameleon also allows simultaneous navigation and action, like video and computer games in which players move through the game world while triggering actions. Another game-like feature is the ability to replay sessions.

StyleCam

Many manufacturers’ Web sites include both professionally produced 2D images and (increasingly) ways to view their products in 3D. Unfortunately, the visual and interactive experience provided by the 3D viewers often falls short of their 2D counterparts. For example, 2D imagery in automobile sales brochures typically provide a richer and more compelling presentation

Game-like features.

While a design review does not necessarily encompass the narrative

message they intend to convey to potential customers. In the worst case, target customers may end up disliking the product, a complete negation of the advertiser’s intention. StyleCam allows authors to ensure a positive customer experience while giving users the freedom to choose what they want to see [1].

Author vs. user control. Central to our research is the difference between authoring an interactive experience and allowing the user total viewing freedom. Authors of 3D Web content typically provide full camera controls (such as pan, tumble, and zoom). Yet by giving users complete control of the viewpoint, authors limit their own influence in the resulting experience. From an author’s perspective, ceding control represents a significant imbalance, as control of both viewpoint and pacing are lost, and the ability to persuade consumers is diminished. In contrast, movie directors control such major elements of their work as content, art direction, lighting, viewpoint, and pacing to define a movie’s visual style. StyleCam allows authors to specify a particular visual style (such as an experience in the style of a television commercial for the same product) while still granting some amount of user control. The system creates interactive viewing experiences where every frame is perfectly composed and rendered, even though the user chooses what to see. Balancing author and user control, StyleCam enables authors to develop the viewpoints and pacing for the user experience by incorporating an interaction technique that seamlessly integrates spatial camera control with temporal control during the playback of animations. To achieve this result we developed three main StyleCam ele-

than the interactive 3D experiences provided by the Web sites of some car manufacturers. Bridging this difference in quality is important if these 3D experiences are to replace or at least be on par with the 2D images.

When creating 2D images, photographers carefully control such elements as lighting and viewpoint to ensure users get the intended message. In contrast, commercial 3D viewers (such as QuickTime VR) typically allow users to view any part of the 3D model, possibly resulting in their getting lost in the scene, seeing the model from unflattering angles, missing important features, or simply experiencing frustration at the navigation mechanisms. Given the advertiser cannot control exactly which images users see in this scenario, they cannot ensure the 3D experience conveys the mes-

ments: camera surfaces; animation clips; and a unified user-interface technique.

Using StyleCam, authors define a set of viewpoints to give a stylized focus to what users see. These viewpoints, or camera surfaces, can be thought of as arbitrarily curved sheets defining the position and orientation of the camera (see Figure 3). Users may move their viewpoints anywhere on this surface to get a sense of the depth and detail of the 3D object being inspected. Moreover, the shape of the surface can provide for some dramatic camera movements (such as when it sweeps across a car's front grille). The idea is that authors should be able to conceptualize, visualize, and highlight particular viewpoints they deem important. Multiple camera surfaces can be used to convey multiple messages and focal points.

To support transitions and pacing between camera surfaces, StyleCam uses a "path," or animation clip, between the edges of camera surfaces triggered when users navigate to the edge of a camera surface. When the animation ends, users can resume navigating on the destination camera surface. One type of animation between camera surfaces is a simple interpolation of the camera's position when moving from one surface to another. Authors also have the stylistic freedom to express other types of visual sequences (such as 2D images momentarily placed in front of the viewing camera), as in Figure 3.

StyleCam allows two distinct types of behavior: control of the viewpoint and playback of animation clips. Viewpoint control is the navigation of space; animation control is the navigation of time. From the user's perspective, viewpoint control can be thought of as dragging the camera, while animation control is dragging the time slider. Users can transition back and forth between the spatial and temporal behaviors using a contiguous mouse drag. They pull the camera across a camera surface with a dragging motion; hitting the edge of the surface causes a smooth and immediate transition. Continuing to drag the mouse causes the position in the animation clip to change accordingly.

To get a sense of users' initial reactions to the StyleCam concept, we conducted an informal user study in our laboratory with seven participants with both technical and nontechnical backgrounds. We found the unified interaction technique worked for all of them, as they seamlessly switched between controlling the camera and controlling the pacing of the animations.

Game-like features. StyleCam allows authors to create an engaging game-like experience using interactive narrative. They might craft a simple storyline and related messages; users then make navigation choices involving nonlinear exploration. For example, an experience involving walking, flying, and glancing around an

object can be achieved through the appropriate selection of camera surfaces and animated transitions. Authors control the pacing by adjusting the length of the animated transitions and the size of the camera surfaces.

StyleCam can even be configured so when a user experiences a particular set of surfaces and animations, another set automatically becomes accessible. This layered progression allows an experience similar to how "levels" or "worlds" function in games; once a particular goal is achieved in a given world, access is granted to an entirely new world. Authors can also progressively reveal additional narratives to users while they navigate through a set of movie-like animated transitions. These transitions are like the prerendered "cut scenes" common in many video and computer games.

Conclusion

These research systems adapt and use elements of video and computer gaming to create a compelling user experience in the domains of virtual design and product marketing. Users interact with the designs and products they are evaluating while having an engaging experience. Our experience has shown us that the use of game concepts in non-game application domains can be effective in the construction of compelling interactivity. **C**

REFERENCES

1. Burtnyk, N., Khan, A., Fitzmaurice, G., Balakrishnan, R., and Kurtenbach. G. StyleCam: Interactive stylized 3D navigation using integrated spatial and temporal controls. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST'02)* (Oct. 27–30). ACM Press, New York, 2002.
2. Fitzmaurice, G. Situated information spaces and spatially aware palmtop computers. *Commun. ACM* 36, 7 (July 1993), 38–49.
3. Harrison, S., Minneman, S., and Marinacci, J. The DrawStream Station of the AVCs of video cocktail napkins. In *Proceedings of IEEE Multimedia Computing and Systems* (June 7–11, Florence, Italy). IEEE Computer Society Press, Los Alamitos, CA, 1999, 543–549.
4. Levine, S. and Ehrlich, S. The Freestyle System: A design perspective. In *Human-Machine Interactive Systems*, A. Klinger, Ed. Plenum Press, New York, 1991, 3–21.
5. Oviatt, S. Ten myths of multimodal interaction. *Commun. ACM* 42, 11 (Nov. 1999), 74–81.
6. Rekimoto, J. and Nagao, K. The world through the computer: Computer-augmented interaction with real-world environments. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST'95)* (Pittsburgh, PA, Nov. 14–17). ACM Press, New York, 1995, 29–36.
7. Tsang, M., Fitzmaurice, G., Kurtenbach, G., Khan, A., and Buxton, W. Boom Chameleon: Simultaneous capture of 3D viewpoint, voice, and gesture annotations on a spatially aware display. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST'02)* (Paris, France, Oct. 27–30). ACM Press, New York, 2002.

MICHAEL TSANG (mtsang@dgp.toronto.edu) is a graduate student in human-computer interaction in the Department of Computer Science at the University of Toronto.

GEORGE FITZMAURICE (gf@aw.sgi.com) is a senior researcher of the Interactive Graphics Research Group at Alias|Wavefront, Inc., in Toronto, Canada.

GORDON KURTENBACH (gordo@aw.sgi.com) is the director of the Interactive Graphics Research Group at Alias|Wavefront, Inc., in Toronto, Canada.

AZAM KHAN (akhan@aw.sgi.com) is a member of the research technical staff at Alias|Wavefront, Inc., in Toronto, Canada.