DirectCam: A Gestural System for Animatic Creation

Noah Lockwood Patricio Simari
Patrick Coleman Karan Singh

University of Toronto

lockwood@dgp.toronto.edu, psimari@dgp.toronto.edu,
patrick@dgp.toronto.edu, karan@dgp.toronto.edu

In film production, pre-visualization begins with the creation of hand-drawn storyboards. These are edited together, along with temporary audio, to form a story reel to communicate shot placement, timing, and camera movement. Story reels are subsequently refined to create roughly-animated 3D animatics, which are better able to communicate shot composition with realistic spatial cues, but with coarse animation. The creation of the animatic can be laborious and time-consuming, however. This is in part due to the iterated feedback and refinement involving the director and animatic artists.

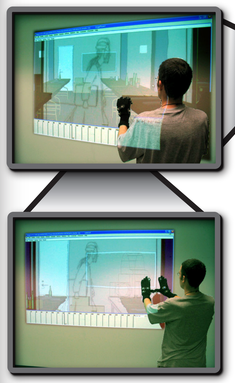
DirectCam simplifies this process by putting the task of animatic creation in the hands of non-technical users. A small, simple set of spatial hand gestures allows the user to directly create and manipulate 3D data without requiring technical knowledge of the underlying animation representation. Using a story reel as reference, characters, sets and other objects can be created and placed within a 3D scene. The camera can be directly manipulated using a through-the-lens perspective to achieve the desired shot composition, or from an outside point of view to help create crane shots. Motion and timing can then be specified to roughly animate the camera and scene elements over time.

The DirectCam System

In DirectCam, we track the user's hands in 3D space using a Vicon optical motion tracking system (above, center). Users wear a pair of gloves augmented with reflective markers (above, left). DirectCam has been designed primarily for use in front of a large, wall-size display, as hand motion can be scaled such that there is a close correspondence to perceived screen space motion. This also facilitates collaborative group presentations, discussion, and feedback as the animatic is created (above, right).

A small set of hand *postures* determines the system's current task. Associated with each posture are continuous *gestures*, subsequent motion of one or both hands. The motion of these gestures is mapped to different manipulations of 3D data, depending on the current task. The gesture and posture set is small and mutually distinct, and we have designed them as metaphors for the operations they invoke. The simple and expressive postures and gestures also allow collaborators to easily understand how the user is working with the system, as well as learn the system by observation.

Users perform three primary creative tasks: importing and manipulating models, manipulating the camera, and creating rough animation for both models and the camera. Each hand has an associated cursor to allow the user to select objects for manipulation. Users can trigger the appearance of visual *shelves* that allow for easy access to additional assets and operations.



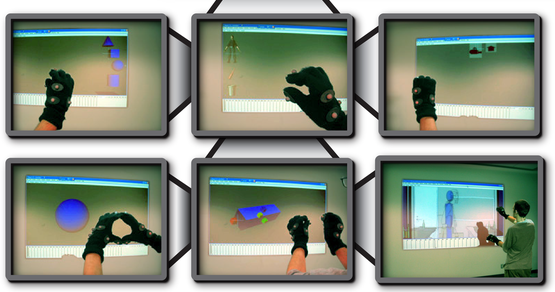
At left, we see a user load a set and story reel (above), and then manipulate the camera with a framing gesture to align the view to the drawn image (below).

Object and Camera Manipulation

Complex models can be imported from a model shelf, while free-form object creation can be accomplished using simpler 3D shapes from the primitive shelf. Models are selected with a pinching posture of the dominant hand and moved directly into the scene--hand position and orientation are directly mapped to the object. Object scaling is accomplished by selecting and moving a scaling handle with the non-dominant hand. Hierarchical relationships can also be specified to allow for creation of moderately complex objects.

Inspired by the workflow of live-action directors, the through-the-lens perspective is directly manipulated using a two-handed framing posture. Horizontal and vertical gestures are directly mapped to camera pan and tilt, respectively. A handle on the overhead camera shelf can be grasped and moved in 3D to manipulate the camera's position. A *pull-back* gesture also transitions to an external view where the main camera can be directly manipulated from afar. The through-the-lens view can be restored with a corresponding *push-in* gesture.

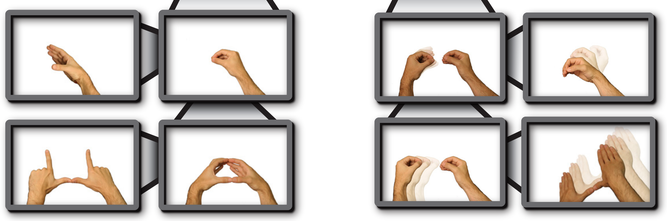
Below, we see the user interacting with DirectCam to complete various modeling tasks. Clockwise, from top-left: The user reaches to the right to cause the primitive shelf to appear. The user reaches to the left to cause the model shelf to appear. The user reaches up to cause the scene shelf to appear, for scene file access or camera selection for direct manipulation. A user assembles a hierarchy to represent a character for primitives created via the primitive shelf. The user demonstrates a scaling manipulation. The user invokes a special-purpose gesture to create the commonly-used primitive sphere. Postures and gestures for these tasks are described in more detail to the right.



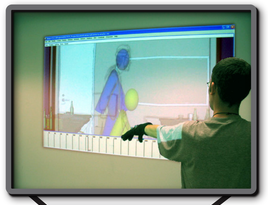
Postures and Gestures for DirectCam

Below-left we see the manipulation and modeling postures for DirectCam. Clockwise, from top-left: An open posture releases all selections. A pinch posture selects the object behind the hand's cursor. Two closed hands in a ball shape create a primitive sphere. A framing posture invokes through-the-lens camera control.

Modeling and manipulation gestures are shown below and to the right. From top-left, clockwise: Pinching a second object while the dominant hand holds a selection creates a parent-child relationship. Rotation and translation of the dominant hand map to object manipulation. Manipulation of the framing posture causes the camera to pan and tilt. Pulling away the non-dominant hand scales an object along the corresponding axis.



Animation



Users create animated transitions of object transforms by selecting a current frame, editing the scene, and then setting a keyframe value for the current frame. To reduce mode errors caused by mistaken gestures, animation and temporal manipulations are performed by the non-dominant hand, while direct manipulation is invoked with the dominant hand. The animation timeline can be grasped and directly scrubbed to reach a particular frame or to play back the current animation (top, at right). After the user manipulates an object or camera to a desired state, keyframes can be set with a grasp-and-toss gesture towards the animation timeline (center, at right). In addition, keyframe sequences can be re-timed by directly grasping and manipulating them on the timeline. Nonlinear timing changes can be made using a hand-tapping technique to set relative timing between pairs of adjacent keyframes (bottom, at right). In the image above, the user sets a keyframe on the character model to match the drawn frame in the reel.

Evaluation and Conclusions

To evaluate DirectCam, five test users who were unfamiliar with conventional 3D animation software were instructed to perform a variety of tasks. The users were successful with the interface with a minimum of instruction and found it simple, easy, and fast to create scenes of their own design. In addition, a more technical evaluation was performed by an experienced user, where an animatic was created using the story reel and 3D assets from an existing film. Within a few minutes, a rough animatic was created that successfully communicates the motion and composition of the final shot in the actual film. Images from the story reel and corresponding frames from our animatic are shown below.

DirectCam allows users to more expressively create and communicate by focusing on simple hand postures and gestures with clear metaphors that are easily learnt and understood. Users felt that our techniques for object manipulation, camera control, and animation were expressive enough to create the intended scenes, while remaining intuitive enough to support fast learning. Furthermore, our techniques are appropriate for a group setting, as users are able to observe, discuss, and learn while operating the system.

Acknowledgements

We would like to thank Autodesk for the donation of the Maya animation software, within which we have built our system. We would also like to thank Chris Landreth, Copper Heart Entertainment, and the National Film Board for providing the story reel and models from the film *Ryan*.