11
Interactive Scene Walkthrough Using a Physically-Based Virtual Camera

Russell Turner, Jean-Francis Balaguer, Enrico Gobbetti, Daniel Thalmann

Laboratoire d'Infographie
EPF Lausanne
CH 1015 Lausanne, Switzerland

One of the most powerful results of recent advances in graphics hardware is the ability of a computer user to interactively explore a virtual building or landscape. The newest three-dimensional input devices, together with high speed 3D graphics workstations, make it possible to view and move through a 3D scene by interactively controlling the motion of a virtual camera. In this paper, we describe how natural and intuitive control of building walkthrough can be achieved by using a physically-based model of the virtual camera's behavior. Using the laws of classical mechanics to create an abstract physical model of the camera, we then simulate the virtual camera motion in real time in response to force data from the various 3D input devices (e.g. the Spaceball and Polhemus 3Space Digitizer). The resulting interactive behavior of the model is determined by several physical parameters such as mass, moment of inertia, and various friction coefficients which can all be varied interactively, and by constraints on the camera's degrees of freedom. This allows us to explore a continuous range of physically-based metaphors for controlling the camera motion. We present the results of experiments using several of these metaphors for virtual camera motion and describe the effects of the various physical parameters.

Introduction

With the existence of graphics workstations able to display complex scenes containing several thousands of polygons at interactive speed, and with the advent of such new multiple degree-of-freedom interactive devices as the Spaceball, Polhemus 3Space, and DataGlove, it is now possible to create graphics applications based on a full 3D interaction metaphor in which the specification of the camera motion is given in real-time.

Interactive virtual camera motion is an important technique in a number of different computer graphics areas: animation, scientific visualization, CAD and virtual environment applications all make considerable use of interactive virtual camera control in a three-dimensional environment to give the user the possibility of moving inside computer-generated virtual scenes (Baum et al. 1990, Magnenat-Thalmann and Thalmann 1986, Shinagawa et al. 1990). In architectural CAD applications, this gives both the architect and the client the ability to walk through virtual buildings and inspect them from any angle, allowing the architect to explore complex spaces and the client to provide feedback early in the design process (Brooks et al. 1986, Greenberg 1974).

In such systems, camera control using various input devices must be as natural and intuitive as possible so that the user is no longer conscious of it and can concentrate on the application task. However, the relationship between device input and virtual camera motion is not straightforward. Usually, a mathematical function or *filter* must be placed between the raw input device data and the virtual camera viewing parameters. This filter is usually associated with some sort of real-world metaphor such as *flying vehicle* metaphor, or *eyeball-in-hand* metaphor. Several recent papers have proposed and compared different metaphors for virtual camera motion control in virtual environments using input devices with six degrees of freedom (Ware and Osborne 1990, Mackinlay et al 1990). These are usually based on a kinematic model of control, where the virtual camera position, orientation, or velocity is set as a direct function of an input device coordinate.

One way to obtain more intuitive interactive camera control is to shift from a kinematic to a dynamic model, and to use some sort of physical laws to interpret the device input data. One such physical model that we propose is an interactive camera control metaphor based on physical modeling of the virtual camera, using forward dynamics for motion specification. This model is motivated by the following hypotheses:

- the interaction is natural because humans are used to physical "Newtonian" behavior
- it is a general parametric model with continuously variable behavior
- the parameters have physical meaning and are easy to understand

In this paper we present a physical description and mathematical derivation of the physical camera model. We then give our subjective impressions of the interactive "look and feel" of the model with different parameter settings using the Spaceball as an input device. We then present some examples of how this model could be used to aid interactive camera motion in a specific domain. Finally, a description of the implementation on a Silicon Graphics Iris workstation is given.

The Physical Camera Model

Like all physically-based modelling in computer graphics, the physical camera model is motivated by the assumption that human beings are best-equipped to deal with environments that resemble the natural world. Since the virtual camera is not normally seen, we are concerned with modeling not its physical appearance but rather its interactive response to the user. A natural virtual camera model is therefore most appropriately some kind of rigid body with mass that can be pushed or driven in various directions such as a moving vehicle or a real movie camera.

There exist a variety of real-world vehicles and camera mounts capable of moving in three dimensions from which to obtain an intuitive interaction metaphor. Although it would obviously be possible to physically model all of these, we have chosen to model an idealized real-world camera which is manipulated by a human being who exerts forces and torques on it. In this case, the behavior of the camera--the *virtual camera metaphor*--is determined by the mechanical properties of the camera. These properties can then be considered the parameters of our parametric virtual camera model.

Although real cameras and moving vehicles are complicated mechanical devices, their motions are determined for the most part by a few simple gross physical properties. Therefore, we have constructed an idealized physical model consisting of a single rigid body attached to a massless camera mount. The camera mount consists of three gimbals and three rails, one in each of the x, y, and z local coordinates, resulting in three Cartesian and three rotational degrees of freedom. Each of the gimbals and rails exerts friction and elastic forces on the camera.

The important mechanical properties of this model which affect its motion are its mass, its moments of inertia, and the coefficients of friction and elastic forces imposed by the camera mount.

The mass parameter specifies the amount by which the various forces will change the camera's linear velocity over time. The moment of inertia parameters affect, in an analogous way, the camera's response to the various torques. Usually a real camera's degrees of freedom are constrained in some geometrical way. For example, it may be placed on a dolly or a railway or tripod, constraining its linear motion to two, one and zero dimensions respectively. Likewise, the angular degrees of freedom may be restricted in one, two, or three axes by locking the gimbal bearings on the camera mount. Friction forces tend to reduce the linear and angular velocity of the camera over time and to oppose the applied forces and torques. There are several types of phenomenological friction forces used by physicists to model the dissipation of kinetic energy. We have found two, viscous friction and static friction, to be useful in our model. Viscous friction is proportional and opposite to the direction of motion, bringing the camera eventually to rest. Static friction is a constant force opposing the applied force, active only when the camera is at rest or below a threshold velocity. We have also found it useful to add an elastic parameter, in the form of a Hookian spring, to each degree of freedom. This elasticity is only important when the camera is at rest (i.e. static friction is active) and results in a more naturalistic transition from the static state to the dynamic state and back.

Motion Under External Driving Forces and Torques

Fortunately, the simplified camera model is quite easy to analyze physically and, from the point of view of classical mechanics, is a well understood problem (Feynman et al 1963). The general motion of a rigid body such as a camera can be decomposed into a linear motion of its center of mass under the control of an external net force and a rotational motion about the center of mass under the control of an external net torque.

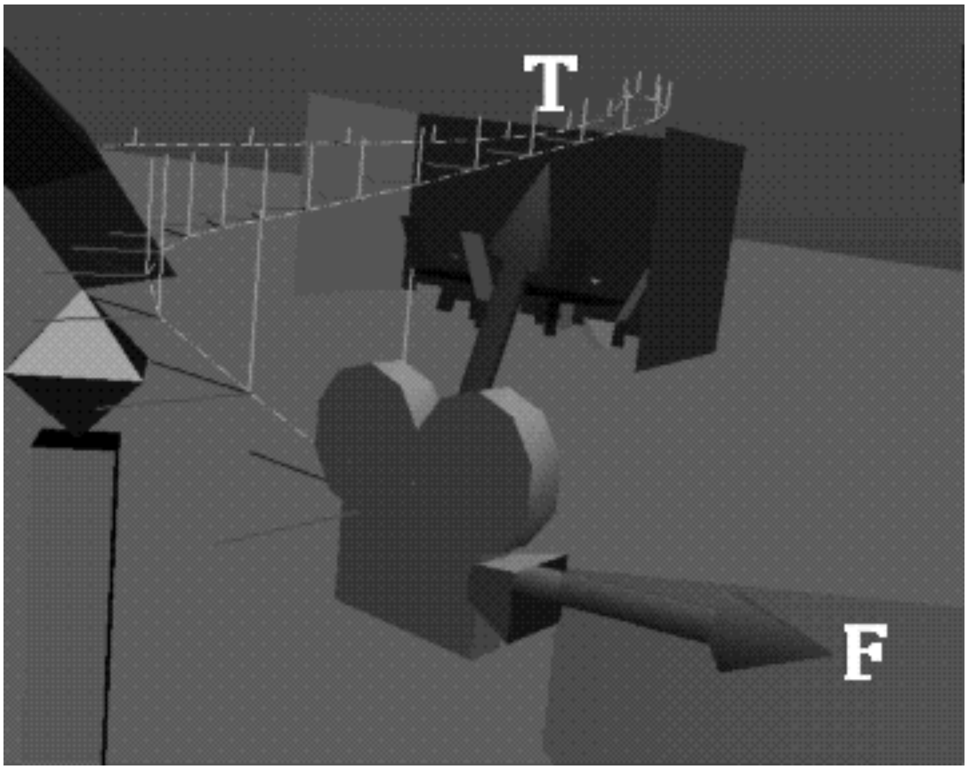

Figure 1. The virtual camera driven by a force and a torque

Linear Motion

The linear motion under a total net force **F** can be computed by solving the differential equation

$$\ddot{\mathbf{X}} = \frac{\mathbf{F}}{m} \tag{1}$$

where m is the mass of the camera and **X** the position vector of its center of mass.

To compute the total net force **F** we have to take into account the driving force $\mathbf{F_d}$ and the forces that are generated by all the various frictions $\mathbf{F_f}$.

When the camera is moving at a speed below a given threshold velocity $\mathbf{v_s}$ and the driving force is smaller than a specified limit $\mathbf{F_s}$, we consider that the camera is in a static situation and that the frictions are caused by the springs that are present in the camera mount. Therefore we have

$$\mathbf{F_f} = k_{vs}\dot{\mathbf{X}} + k_{ss}(\mathbf{X} - \mathbf{X_o}) \tag{2}$$

where $\mathbf{X_0}$ is the position where the camera first entered the static situation, k_{vs} is the damping factor of the springs and k_{ss} the spring constant.

When we are at speeds higher than $\mathbf{v_s}$ or the driving force is bigger than $\mathbf{F_s}$, we consider the dynamic situation, where the frictions are mainly viscous. In that case we have

$$\mathbf{F_f} = k_{vd}\dot{\mathbf{X}} \tag{3}$$

where k_{vd} is the viscous friction coefficient.

It is useful to control the behavior of the camera separately for each of its principal axes. For this reason, the equations will be solved by projecting them onto the body-fixed reference frame, and a different value of the friction parameters will be specified for each one of the local axes.

Angular Motion

The rotational dynamics is expressed in a body-fixed reference system by the equation

$$\mathbf{I}\ddot{\theta} + \dot{\theta} \times \left(\mathbf{I}\dot{\theta}\right) = \mathbf{T_d} \tag{4}$$

where θ is the orientation of the camera expressed using Euler angles, **I** is the moment of inertia tensor of the body and is constant in the body-fixed reference frame, and $\mathbf{T_d}$ is the total net torque applied to the camera.

The modeling of the friction is analogous to the translational case. We can define $\omega_\mathbf{s}$ to represent the threshold angular velocity and $\mathbf{T_s}$ to represent the threshold torque. In a static situation the friction torque $\mathbf{T_f}$ is given by

$$\mathbf{T_f} = k_{vs\text{-}rot}\dot{\theta} + k_{ss\text{-}rot}\left(\theta - \theta_0\right) \tag{5}$$

where θ_0 is the orientation where the camera first entered the static situation, $k_{vs\text{-}rot}$ is the damping factor of the springs and $k_{ss\text{-}rot}$ the spring constant.

In the dynamic case we have

$$\mathbf{T_f} = k_{vd\text{-}rot}\dot{\theta} \tag{6}$$

where $k_{vd\text{-}rot}$ is the viscous friction coefficient.

As for the translational case, it is useful to control the behavior of the camera separately for each of its principal axes, and a different value of the friction parameters will be specified for each one of the local axes.

Specifying the Camera's Behavior

The behavior of the virtual camera in response to user actions is completely specified by the parameters of the physical model of the camera. A different value of the friction parameters can be specified for every dimension in the reference frame attached to the camera. By this means we provide an easy and intuitive way for controlling the behavior of the model separately for each of the six degrees of freedom of the camera and several interesting camera motions or "camera metaphors" can be specified.

Modifying Mass and Inertia Tensor

The camera mass parameter determines the degree of acceleration control by the user. If all other parameters are set to zero, the camera maintains a steady velocity and changes its velocity only in response to user input, resulting in a pure acceleration control metaphor. Higher mass results in a smoother, more continuous motion and a higher degree of acceleration control, although it also makes it more difficult to bring the camera to rest at a given location. A high mass parameter is useful in situations where a smooth camera motion is desired, or where continuous motion is wanted with minimal input from the user, for example, tasks such as moving outside a building or inside large interior spaces.

In an analogous way, the inertia tensor determines the degree of torque control by the user over the camera orientation. A large inertia tensor results in smooth panning and tilting motions. This is often desirable because jerky camera rotation can be disorienting. Without a correspondingly high rotational friction parameter, however, it can be difficult to stop the camera from rotating. This is usually much more disorienting than the analogous translational situation. High inertia tensors are useful for the same sorts of tasks that high mass parameters are useful for: slow steady examination of a large scene. For more precise control up close to objects, a lower inertia tensor is preferable.

Viscous Friction Coefficient

The viscous friction parameter specifies the degree of velocity control. If the other parameters are set to zero, the camera metaphor becomes a pure velocity control one. Velocity control is useful for tasks where quick stopping and changes of direction are necessary, such as avoiding obstacles or inspecting objects up close. A typical architectural walk-through situation where a high viscous friction coefficient would be useful is a small room with furniture. One problem with high viscous friction camera metaphors is that the camera motions are not usually very smooth and the user must give continuous input while the camera is moving. For many interactive tasks, this is not a problem. For other tasks, a suitable combination of velocity and acceleration metaphors can be formed by adjusting various amounts of the mass and viscous friction parameters.

The rotational analog to the viscous friction coefficient is the rotational viscous friction coefficient, which determines the amount of angular velocity control. Angular velocity control is particularly useful because of the need to stop camera rotation quickly. The balance between the moment of inertia and the rotational viscous friction coefficient determines the smoothness of camera panning motions, and many real-world camera mounts have adjustable angular viscous friction controls.

Threshold Force and Velocity

A small static friction parameter establishes a threshold force below which the camera will stay stationary. The threshold velocity determines the speed below which the static friction force comes into effect, bringing the camera to rest more rapidly than the viscous friction forces. Changing these parameters can be useful for tasks in which the user wants to alternate between motion along different degrees of freedom. For example, moving the camera along only one axis at a time is easy as long as the threshold force is not exceeded in the other axes. Likewise, a task that involves hopping to a fixed location, looking around in different directions without moving, then hopping to the next location, is much easier with a small amount of static friction.

By setting the static friction parameter extremely high, its possible to, in effect, lock a particular degree of freedom, resulting in a constrained motion. For example, by locking one of the rotational degrees of freedom, the camera is forced to always maintain the same "up" direction, and by locking one of the translational degrees of freedom, the camera is forced to move in a plane. This can be useful to simulate a walking person's point of view. Locking the vertical and horizontal translational degrees of freedom, so the camera can only move forward and turn, results in a flying-vehicle camera metaphor.

Static Behavior: Damping Factor and Spring Constant

The spring constant and damping factor parameters control the vibrational behavior of the camera mount when it is in the static state. A small amount of damped vibration smooths out the jerkiness in the transition between dynamic and static states. It also provides a small degree of position control feedback while the camera is in the static state. In this way, a small applied force will move the camera slightly, but it will pop back to its rest position. A larger force, above the static friction threshold, will set the camera in motion. This allows the user to get an idea of what direction an applied force will act in before actually moving the camera's position. If the static friction parameter is set extremely high, then the camera becomes locked in the static state and a position control camera metaphor results.

Continuous Variation Between Different Behaviors

It is generally accepted that no one particular type of camera motion or camera control is appropriate for all tasks (Ware and Osborne 1990). In virtual building walkthrough, the user might require acceleration control for moving smoothly as he surveys the overall spatial organization of the building, while the task of inspecting some detail, where the camera viewpoint is close to the point of interest, might require more precise velocity control.

Unfortunately, genuinely useful tasks often require a combination of metaphors or a sequence of alternating metaphors. A common way to deal with this is to allow the user to continually change camera control metaphors by swapping through the different interaction modes. Modes can be changed by selecting menus, striking keys or mouse buttons. There are two problems with this technique. First it is obviously inconvenient and unnatural to continually change modes. The user has to stop performing his task to swap computer and mental modes. After a while, this can be distracting and will often inhibit the user from changing metaphors until the current one becomes really impracticable. Secondly, it is not necessarily true that there are only a finite number of discrete useful metaphors. There are, in fact, a lot of in between situations where no pure metaphor suits the task best (Mackinlay et al 1990).

The parameterized physical camera model provides a solution to this problem by giving us a way to control the camera behavior through its parameter values. For a given task, the user can experiment with the camera metaphor and tune the parameters interactively, through valuators, until a subjectively "best" set of parameters is found. These parameters can then be saved and restored, either automatically or by the user, whenever that particular task is encountered. Alternately, the user can interactively control certain camera parameters while he is performing his task. For example, a mouse button or foot pedal can be set to momentarily increase the amount of viscous friction, acting as a kind of break.

Ideally, however, the user should only be required to use a minimum of input and should not have to be overly aware of the camera metaphor at all. This can be achieved by having the application adjust the camera control parameters algorithmically as a function of position or some aspect of the task at hand. This is potentially the most powerful use of the parametric camera model.

For example, we have created a scalar viscosity field within a scene such that the camera viscous friction parameter increases in the vicinity of objects. This results in a camera metaphor that continuously varies from mainly acceleration control when the camera is far away from objects, to mainly velocity control when the camera is close to an object. In this way, the camera's behavior varies as a function of its distance from to objects.

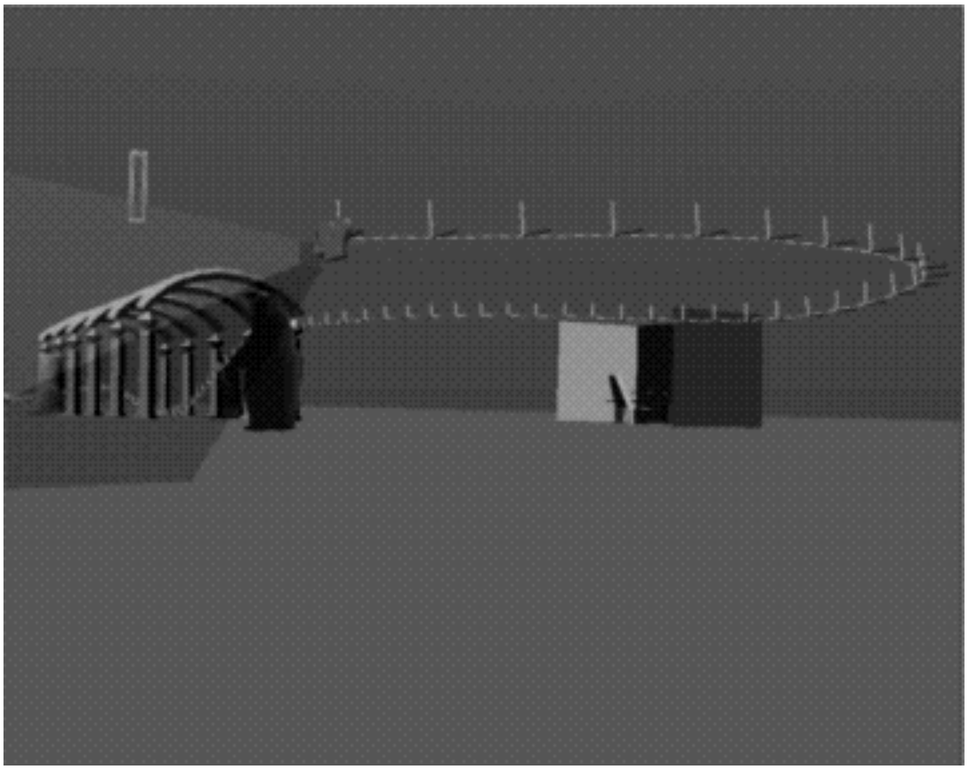

Figure 2. Exploring a scene

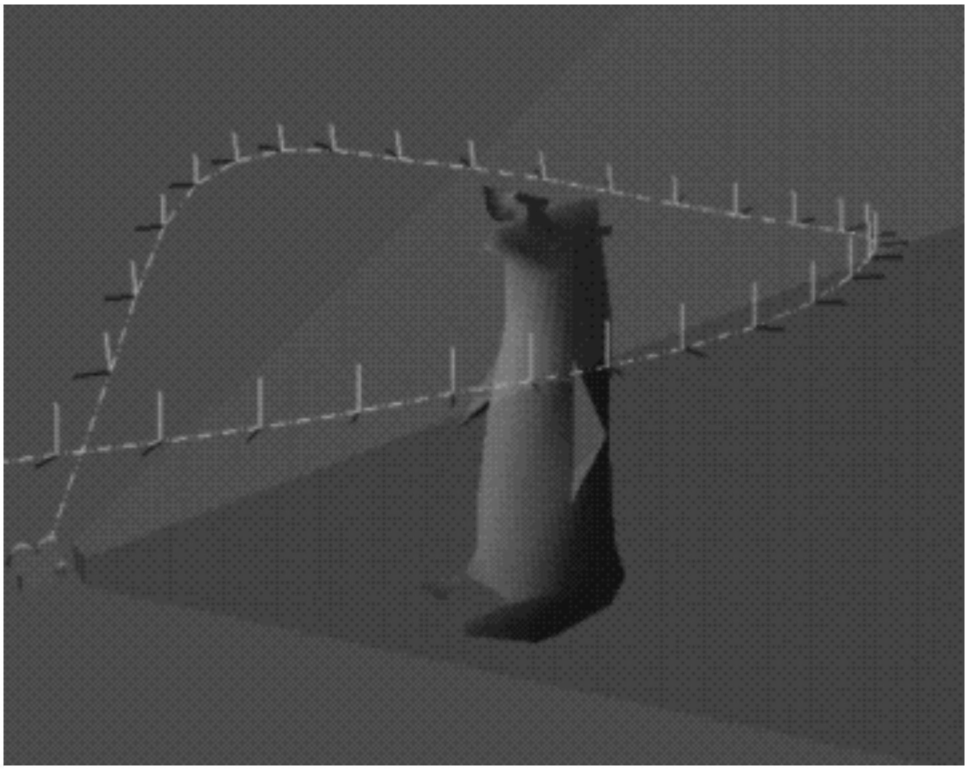

Figure 3. Inspecting an object

Implementation

Our dynamic camera control system is implemented on a Silicon Graphics Iris workstation in C using an object-oriented style of programming, on top of the Fifth Dimension 3D interaction toolkit (Turner et al 1990). In addition to the typical sorts of 3D classes, such as lights, hierarchical models and cameras, the toolkit abstracts every input device as an instance of an input device class. These input device objects communicate their data through a uniform event message protocol.

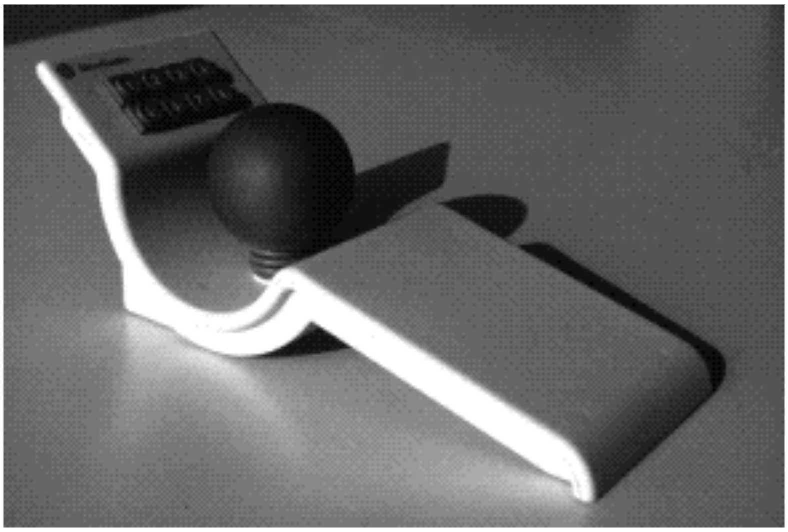

Figure 4. Spaceball six degree of freedom input device

For interactive camera control, we have found the most natural input device to be the Spaceball, produced by Spatial Systems. It is composed of a rigid sphere containing strain gauges mounted on a plastic base in order to offer a comfortable rest position for the operator hand and arm. The position and orientation are specified by rotating or pushing the sphere in a certain direction. Although the sphere moves only slightly, the force and torque applied are measured and an incremental position and orientation is transmitted, at a configurable rate, to the host computer.

The interface between the input device objects and the camera object is implemented by a camera controller object. This object receives events from input devices, interprets the data according to a particular camera metaphor, and updates the camera object's position and orientation accordingly. Since our camera Controller involves a dynamic simulation, it also receives tick events from a clock object at regular time-step intervals.

At each time step, a numerical solution to the equations of motion (1) and (4) is calculated using a second-order approximation for the derivatives of $\mathbf{X}$ and Θ. This explicit procedure evolves the dynamic solution from given initial conditions. The current and the previous value of position and orientation are used to solve for the value at a small time Δt later. No oversampling is necessary, because the precision required is not high (the user interacts with the solver), and Δt represents the time increment between displayed frames. Our current implementation allows us to have an interactive display rate (more than 10 Hz) on a Silicon Graphics Iris 4D/80 with fully shaded scenes containing up to two-thousand polygons. The time spent for the dynamic computations is negligible with respect to the redraw time.

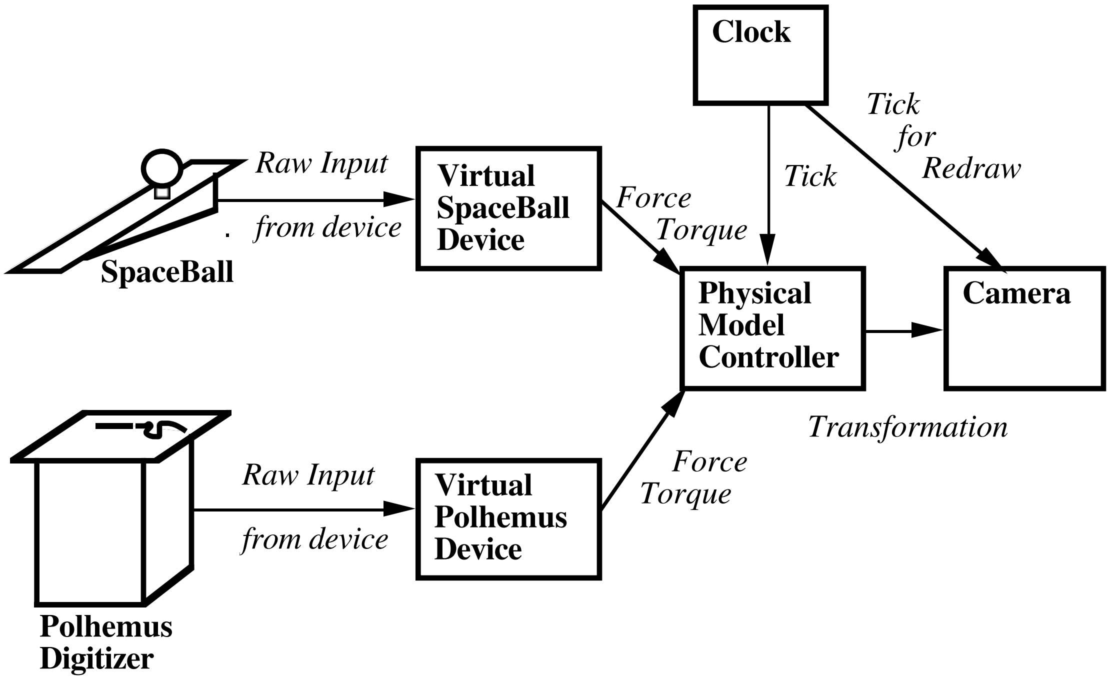

Figure 5. Event communication diagram for the camera controller

One of the advantages of this kind of software architecture is that various other devices, such as the Polhemus 3D, can be interchanged with the Spaceball as input to the Camera Controller. The use of this kind of devices is essential for providing an intuitive way of controlling the dynamic camera. Obviously, pressure-sensitive input devices are usually more appropriate because they provide a passive form of "force-feedback". In our case, the device that gave the best results is the Spaceball. Also, different types of Camera controllers with different behaviors can be swapped in and out, and the same controllers can be used to control Light objects or hierarchical models.

Conclusions and Future Work

We believe that the physically-based camera control model provides a powerful, general-purpose metaphor for controlling virtual cameras in interactive 3D environments. Because it is based on a real camera model, it is natural for the user to control. Its parameters are physically-based and, therefore, easy to understand and intuitive for the user to manipulate. Its generality and control parameters make it configurable to emulate a continuum of camera behaviors ranging from pure position control to pure acceleration control. As it is fully described by its physical parameters, it is possible to construct more sophisticated virtual camera control metaphors by varying the parameters as a function of space, time, application data or other user input. Also, when used with force-calibrated input devices, the camera metaphor can be reproduced exactly on different hardware and software platforms, providing a predictable standard interactive "feel".

We are currently working on extending our model to specify camera paths for computer animation, and on exploring different kinds of algorithmic control of the camera parameters by creating various types of parameter fields within the scene space, and we are planning to use the dynamic model approach for other kinds of interactive tasks such as modeling, assembling, and specifying the animation of different kinds of objects.

Acknowledgements

The authors are grateful to Angelo Mangili for his technical help.
The research was partly sponsored by *Le Fond National pour la Recherche Scientifique*.

References

Baum R., Wingel J.W (1990) Real Time Radiosity Through Parallel Processing and Hardware Acceleration *Proceedings 1990 Workshop on Interactive 3D Graphics* ACM: 67-75

Brooks F.P. Jr (1986) Walkthrough - A Dynamic Graphics System for Simulating Virtual Buildings *Proceedings 1986 Workshop on Interactive 3D Graphics* ACM : 9-22

Feynman R.P., Leighton R.B., Sands M. (1963) The Feynman Lectures on Physics Addison-Wesley Reading, Massachusetts.

Greenberg D.P. (1974) Computer Graphics in Architecture, *Scientific American* 230(5): 98-106.

Hahn J.K. (1988) Realistic Animation of Rigid Bodies *Computer Graphics* 22(4) : 299-308

Mackinlay J.D., Card S.K, Robertson G. (1990) Rapid Controlled Movement Through a Virtual 3D Workspace, *Computer Graphics* 24(4) : 171-176

Magnenat-Thalmann N., Thalmann D. (1986) Special Cinematographic Effects Using Multiple Virtual Movie Camera, *IEEE Computer Graphics & Applications* 6(4): 43-50

Shinogawa Y., Kunii T.L., Nomura Y., Okuno T., Young Y. (1990) Automating View Function Generation for Walkthrough Animation Using a Reeb Graph, *Proceedings Computer Animation 90* Springer, Tokyo: 227-238

Turner R., Gobbetti E., Balaguer F., Mangili A., Thalmann D., Magnenat-Thalmann N. (1990) An Object Oriented Methodology Using Dynamic Variables for Animation and Scientific Visualization *Proceedings Computer Graphics International 90* Springer-Verlag : 317-328

Ware C., Osborne S. (1990) Exploration and Virtual Camera Control in Virtual Three Dimensional Environments *Proceedings 1990 Workshop on Interactive 3D Graphics* ACM : 175-183

Watson V. (1989) A Breakthrough for Experiencing and Understanding Simulated Physics *ACM SIGGRAPH Course Notes on* State of the Art in Data Visualization, IV-26 - IV-32

The Authors

Russell Turner is a researcher at the Computer Graphics Laboratory of the Swiss Federal Institute of Technology in Lausanne, Switzerland. He received his B.S. in Physics and his M.S. in Computer and Information Science from the University of Massachusetts at Amherst. He has also worked as a software engineer for V.I. Corporation of Amherst, Massachusetts. His research interests include computer animation, physical modeling, user-interfaces and object-oriented programming. He is a member of IEEE and ACM.
E-mail: `turner@lig.di.epfl.ch`

Francis Balaguer is a researcher at the Computer Graphics Laboratory of the Swiss Federal Institute of Technology in Lausanne, Switzerland. He received his *diplôme d'ingénieur informaticien* from the Institut National des Sciences Appliquées (INSA) in Lyon, France. His research interests include 3D interaction, computer animation, and object-oriented programming.
E-mail: `balaguer@lig.di.epfl.ch`

Enrico Gobbetti is a researcher at the Computer Graphics Laboratory of the Swiss Federal Institute of Technology in Lausanne, Switzerland. He received his *diplôme d'ingénieur informaticien* from the same institute. His research interests include visualization, computer animation, human-computer interaction and object-oriented programming.
E-mail: `gobbetti@lig.di.epfl.ch`