Intuitive Control of “Bird’s Eye” Overview Images for Navigation in an Enormous Virtual Environment

Shinji Fukatsu  Yoshifumi Kitamura  Toshihiro Masaki  Fumio Kishino

Graduate School of Engineering
Osaka University
2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan

Abstract
We propose a manipulation technique to intuitively control the “bird’s eye” overview display of an entire large-scale virtual environment in a display system that present a user with both overviews (global views) and a life-size virtual environment (local view) simultaneously. It enables efficient navigation even in enormous and complicated environments using both global and local views. The motion of the bird’s eye viewpoint interlocks with the relative motion of the user’s viewpoint and his/her hand, therefore, the user can control the “bird’s eye” viewpoint by intuitive manipulation. Sophisticated display techniques are obtained based on the proposed method by introducing some constraints on the parameters of the “bird’s eye” viewpoint. Experimental results show a combination of the bird’s eye overview image and the life-size local image is displayed to the user by reflecting his/her intuitive manipulation.

1 Introduction
The progress of graphics computations and network capabilities has allowed opportunities in enormous virtual environments. Interactions in large complex virtual environments require various sophisticated techniques and algorithms to handle large-scale geometric databases, for instance, efficient algorithms for displaying complex geometric data to achieve video frame rates [1]. In spite of such efforts, a virtual reality user tends to easily lose his position and direction in an immersed virtual environment without sophisticated assistance. In order to enable a user to obtain accurate spatial perception in an environment, global information on surroundings and local information from a life-size view have to be presented.

Humans construct cognitive maps of their environment for use in navigation [2], and maps assist the user in maintaining knowledge of his current position and direction in a virtual environment [3]. The adequate presentation of environmental overviews from other viewpoints and in different scales assists the user in accurately recognizing the spatial environment. If an overview (global view) is presented on the background view of a life-size virtual environment (local view), and if the way to control the scale, direction and/or viewpoints of the overview is transparent and easily understood by the user, he/she will easily recognize the entire environment.

There is some literature devoted to viewpoint determination of the virtual environment. For example, eyeball-in-hand, scene-in-hand and flying-vehicle-control are metaphors for viewpoint motion control techniques in a virtual environment [4]. Methods of multiple viewpoint and scale (i.e., global view and local view) are also proposed. For example, the “world-in-miniature” duplicates miniature copies of the entire virtual environment [5], and the “virtual GIS” has an overview map inserted in its ordinary display [6]. These examples on immersive types of life-size virtual worlds help the user navigate the user and manipulate virtual objects, however, a method to control the scale, direction and/or viewpoint of the overview has not been discussed. An approach that enables a user to efficiently construct cognitive maps of his environment is to provide an “authority” for intuitively controlling the viewpoint and scale with the overview.

In this paper, a method for intuitively controlling the “bird’s eye” overview display of an entire large-scale virtual environment is proposed. “Interlocked pairs of coordinate systems” is introduced, i.e., the bird’s eye viewpoint interlocks with the user’s viewpoint. As a result, the user can control the “bird’s eye” viewpoint and its orientation by his/her intuitive manipulation. Sophisticated display methods are obtained by introducing some constraints on the parameters of the “bird’s eye” camera.
2 Overview Use in an Enormous Virtual Environment

2.1 The User Interaction

Interactions in an enormous and complex virtual environment require various sophisticated techniques and algorithms to handle the large-scale geometric databases. For example, efficient algorithms that display complex geometric data to achieve video frame rates are essential in enormous virtual environments. Example of literature devoted to this problem are: progressive mesh representations for level-of-detail [7], texture-based object simplification [8, 9], simplification using impostors [10, 11, 12], scene management techniques [13, 14], and hierarchical visibility computations [15, 16].

Despite great effort in the efficient display of complex geometric data, the user can easily lose his/her current position, direction, and path to the desired destination in an enormous virtual environment without sophisticated assistance. The immersive life-size virtual view displayed to the user through a HMD lacks global information about the environment. Therefore, navigation techniques that give the user useful information from a global point of view in an enormous virtual environment are important issues. Several walk-through techniques for controlling the user’s viewpoint, e.g., gaze-directed steering, pointing/gesture steering, are compared in the paper [17]. The eyeball-in-hand, scene-in-hand and flying-vehicle-control are metaphors for virtual camera control in a virtual environment [4]. Scaled-world grab is an automatic scaling manipulation technique [18], however, it may complicate the user’s sense of relative size compared to his body in the immersed virtual environment.

2.2 Overview Use

The adequate presentation of global views and life-size local views can assist the user in obtaining accurate spatial perception of the enormous virtual environment. Two main ideas exist for presenting overviews of the environment. The first idea is to use miniature copies of the entire environment and this is called world-in-miniture (WIM) [5]. It offers multiple viewpoints and scales, and is helpful in navigating and manipulating the virtual environment. The user can change the position and direction of the miniature by using a hand-held 6-D tracker. The second idea is to use “bird’s eye” overview images taken by virtual cameras from certain viewpoints, and “Virtual GIS” is an example that has the overview map inserted in the display [6]. Both overview examples present a global view in addition to a local view, however, a method for controlling the bird’s eye viewpoint has not yet been discussed. A method for intuitively controlling the viewpoint of the bird’s eye to provided overview images is described in the next section.

2.3 Miniature Copies vs. Bird’s eye Overview Images

Two main ideas for presenting overviews are compared in this subsection. The method using miniature copies requires multiple copies of geometry corresponding to multiple different miniatures, therefore, the order of the data size of geometry of the virtual world is proportional to the number of miniatures presented. This drawback implies that miniature copy method is not suitable for enormous and complicated environment such as virtual nature environment.

On the other hand, multiple overview images taken by virtual cameras at multiple viewpoints and scales are put in a virtual environment as texture in the method of using “bird’s eye” overview images. Therefore, the original data size of geometry of the virtual world does not increase even the number of “bird’s eye” overview images increase.

The method proposed in this paper for controlling the viewpoint of overview can be applied to both cases, however, the effectiveness of the proposed method is shown by using the method of “bird’s eye” overview images in this paper.

2.4 Overview Display

Several display configurations are considered to present a user with both overviews (global views) and a life-size virtual environment (local view) simultaneously. The first one is to use multiple windows on a workstation monitor, in which separate windows are used for the different views. Figure 1(a) shows this example. User's eye position is derived from the position of a 6-D tracker attached to a stereo viewing glasses (e.g., LCD shutter glasses). Accordingly, the system present non-distorted images with depth sensations and motion parallax. The display system with large screen can be used instead of the desktop workstation monitor in this configuration, and both tiled- or overlayed-windows are suitable for the different views.

The second configuration consists of a workstation monitor and a see-through HMD as shown in figure 1(b). This configuration lets the user see the life-size local view with the global views superimposed by the function of a optical see-through HMD. It is difficult to present life-size local images with depth sensations because of the difficulty of wearing both a stereo viewing glasses and HMD. The display system with large screen or projection table also can be used in this configuration.

The third configuration uses a closed-view HMD with a 6-D tracker (figure 1(c)). The overview (global views) are presented on the background view of a life-size virtual environment (local view). Because the HMD does not allow any direct view of the real world, a user cannot use the real objects (e.g., frames of the monitor) as the landmark to identify the absolute position/orientation in a 3-D environment. He/she easily lose his position and direction in an immersed virtual environment. Therefore, we start investigating a sophisticated navigation technique for this configuration at first, then we apply this technique to other display configurations.

Several parameters in all above configurations are changeable, e.g., the number of windows, the locations where the global views presented, scales/magnifying power of overviews. The way how to control these parameters must be considered according to the application.

3 Intuitive Control of the Viewpoint for the “Bird’s Eye” Camera

This section describes a method that intuitively control the position and orientation of the “bird’s eye” camera.
a coordinate for a palmtop virtual environment that corresponds to the user’s hand motion. This fourth coordinate is used only to determine the bird’s eye viewpoint, and the bird’s eye overview image is generated according to its position and orientation. The user observes the generated bird’s eye overview image through the window on the projection plane.

![Diagram of coordinate systems](image)

Figure 2: Coordinate system of proposed method for bird’s eye overview images.

### 3.2 Interlocked Motion between Bird’s Eye Viewpoint and User’s Viewpoint

In order to achieve an intuitive control of viewpoint and orientation of the “bird’s eye” camera, the “interlocked motion” between the bird’s eye viewpoint and the user’s viewpoint is proposed. Figure 3 shows the simplified explanation to 2-D. The relative motion (position and orientation) between the bird’s eye coordinate and the world coordinate system for an entire virtual environment are interlocked with the relative motion (position and orientation) between the user’s coordinate system and the coordinate for the palmtop virtual environment. In other words, a pair of coordinate systems (i.e., the user’s coordinate and the coordinate for the palmtop virtual environment) interlocks with the other pair of coordinate systems (i.e., the bird’s eye coordinate and the world coordinate system for entire virtual environment). Figure 4 shows the pairs of interlocking coordinate systems. $X, Y, Z$ correspond to the world coordinate system axes, and $x, y, z$ correspond to the coordinate for palmtop virtual environment axes. The correspondence between the origin of the palmtop virtual environment and the origin of the world coordinate is supposed to be established.

Point $P_u$ is the intersection of a perpendicular line through the horizontal plane of the palmtop virtual environment and the viewpoint of the user, $O_u$. Point $S_b$ is the intersection of the horizontal plane of the palmtop virtual environment and a straight line passing through $O_u$ and the center of the window of the bird’s eye overview image, $C$. The angles $(\psi, \phi, \theta)$ are rotations of the straight line $O_uS_b$ in the coordinate for the palmtop virtual environment around $x, y, z$ axes, respectively. Similarly, point $P_b$ is the intersection of a perpendicular line through the horizontal plane of the entire virtual environment and the bird’s eye viewpoint, $O_b$. Point $S_b$ is the intersection of the horizontal plane of entire virtual environment and the axis line passing through $O_b$. The angles $(\psi_b, \phi_b, \theta_b)$ are rotations of the straight line $O_bS_b$ in the world coordinate $X, Y, Z$ axes, respectively.

### 3.1 Coordinate for a Palmtop Virtual Environment

Six parameters for translation ($X, Y, Z$) and rotation ($\psi, \phi, \theta$) of the bird’s eye viewpoint must be determined to generate an overview image. In order to control these parameters intuitively, “interlock” of the bird’s eye viewpoint and the user’s viewpoint is introduced. Figure 2 shows the coordinate systems used in the proposed method. The world coordinate represents the position and orientation of all objects in the entire virtual environment, and it corresponds to globally-aligned directional representations (north, south, east, and west). The user’s coordinate determines his/her view direction and view volume and corresponds to user-aligned directional representations (front, rear, left, and right). The third coordinate is for the “bird’s eye” overview. A bird’s eye camera that perceives the environment is located at the origin of this coordinate.

In addition to the above three coordinate systems, a coordinate for a palmtop virtual environment that corresponds to the user’s hand motion is introduced. For this purpose, we introduce a new coordinate for a palmtop virtual environment that corresponds to the user’s hand motion.
The interlocked pairs of coordinate systems are used to determine the parameters of the bird's eye camera. \((X_b, Y_b, Z_b)\), the \(XYZ\) coordinate of the bird's eye camera is obtained by magnifying the \(xyz\) coordinate of the user's viewpoint \((x_u, y_u, z_u)\) \(A\) times. Here, \(A\) is the magnifying power given in advance. The angles \((\psi_b, \phi_b, \theta_b)\), orientation components of the straight line \(O_bS_b\) and the world coordinate, are determined by using the orientation components of the straight line \(O_uS_u\) and the coordinate for the palmtop virtual environment.

An example of interlocked motion of coordinate systems is shown in figure 5. If a user rotates his/her hand, the coordinate for palmtop virtual environment is rotated with same angle. At the same time, the position and orientation of bird's eye viewpoint changes due to the change of relative position and direction of user's viewpoint against the coordinate for palmtop virtual environment. By introducing “interlocked pairs of coordinate systems”, the user's viewpoint and orientation reflect a straightforward position and orientation of the bird's eye viewpoint. As a result, the user can control the position and orientation of the bird's eye camera intuitively.

Figure 3: Interlocked pairs of coordinate systems (simplified to 2-D).

Figure 4: Interlocked pairs of coordinate systems.

Figure 5: Interlocked motion between bird's eye viewpoint and user's viewpoint.

4 Constraints on the Attitude of the “Bird's Eye” Camera

Six parameters of position and orientation for bird's eye camera can be determined by the interlocked pairs of coordinate systems as described in the previous section. Based on this idea, sophisticated display methods using both the bird's eye overview image and the life-size local image are obtained by using constraints on parameters to determine the attitude of the bird's eye camera. Figure 6 shows the parameters that determine the attitude of the bird's eye camera. Here, \((X_b, Y_b, Z_b)\) indicate the position of the bird's eye camera in the world coordinate system; \((\psi_b, \phi_b, \theta_b)\) indicate the rotation angle of each axis of the world coordinate \((XYZ)\) that corresponds to pitch, roll, and heading in the bird's eye coordinate.

Constraining one or two of six parameters provides a different useful meaning with display methods using the interlocked pairs of coordinate systems. The user can easily understand the relationship of the global and local views by using each constraint. Five constraint elements on parameters and changes of the bird's eye overview images caused by each constraint are explained using figures 4 and 6 below.

Figure 6: Correspondence of rotation parameters of bird’s eye camera and world coordinate system.
(1) Constraint on X and Y

This constraint is for the parameters X and Y, the 2-D location of the bird’s eye camera (point P in figure 4) in the horizontal plane of the entire virtual environment. If \( X_b \) and \( Y_b \) are fixed to the XY coordinate of a certain point in the horizontal plane of the entire virtual environment, the bird’s eye camera is always located above this point. For example, if the user’s position in the XY plane is selected as the fixation point, the bird’s eye camera is always located above the user’s head even if he/she moves.

An extension of this constraint is to fix the point \( S_b \) (in figure 4), this is the intersection of the central axis of the bird’s eye camera and the horizontal plane of the entire virtual environment. In this case, the user in the virtual environment will always be located at the center of the bird’s eye overview image.

(2) Constraint on Z

By fixing \( Z_b \), the height of the bird’s eye camera, the bird’s eye overview image from a certain fixed height can be obtained.

(3) Constraint on pitch (\( \psi \))

The user can observe the bird’s eye overview image with same viewing angle when \( \psi_b \), the angle of depression of the bird’s eye camera, is equal to \( \psi_u \), the intersection of the horizontal plane of the palmtop virtual environment and a straight line passing through \( O_u \) and the center of the window of the bird’s eye overview image. In another case, if \( \psi_b \) is fixed to a certain angle (e.g. 45 degrees), the angle depression of the bird’s eye camera always becomes this angle.

(4) Constraint on roll (\( \phi \))

The user can observe a bird’s eye overview image that has a parallel horizontal line with a life-size local image when \( \phi_b \) is equal to \( \phi_u \). On the other hand, if \( \phi_b \) is fixed to a certain angle, a stable bird’s eye overview image is obtained regardless of the jagged motion of the user’s head/hand.

(5) Constraint on heading (\( \theta \))

The constraint on \( \theta \) is useful for the user’s understanding of spatial sensation in the virtual environment. Two types of alignments can be established by using this constraint, i.e., the world-aligned setup and the user-aligned setup [3]. The world-alignment setup is established if \( \theta_b \) is aligned to the representative direction in the world coordinate (e.g., north). In this case, a bird’s eye overview image is obtained where north always corresponds to the up side of the window. The user-aligned setup is established if \( \theta_b \) is aligned to the user’s direction (\( \theta_u \)) in the world coordinate. In this case, the bird’s eye overview image is obtained so that the user’s viewing direction in the XY plane always corresponds to the up side of the window of the overview image.

Combination of Constraints

The above constraints can be used with others. In this case, a bird’s eye overview image that has new useful features can be obtained. For example, if both \( Z_b \) and \( \psi_b \) are fixed on a certain constant value, the user can easily perceive the “size” in the virtual environment because the scale factor in the center of the bird’s eye overview image is constant.

After some parameters are constrained or fixed, the other parameters for the attitude of the bird’s eye camera are determined by the interlocked pairs of coordinate systems.

5 Implementation and Experiments

5.1 Implementation

The proposed “bird’s eye” overview display technique is implemented on Onyx2 InfiniteReality™ workstation (Silicon Graphics Inc.) with two R10000 CPUs, 512 Mbyte main memory, 64 Mbyte texture memory, four geometry pipelines and eight display pipelines, using IRIS Performer™ 2.0. The user controls the position and orientation of the palmtop virtual environment by using a hand-held 6-D tracker. The experimental configuration is shown in figure 7. The user’s viewpoint and orientation are also measured by a 6-D tracker attached to his/her head. PerformerTown™ is used for the enormous virtual environment in the experimental system. Figure 8 (a) shows the life-size local image with a bird’s eye overview image presented to a user. Here, a window for the bird’s eye overview image is located almost the center of the user’s view for the easy explanation. Here, parameters such as location, size and scale/magnifying power of overviews windows can be changed with keyboard in our current implementation. Figure 8 (b) shows the perspective view which explains positions and directions of the user and bird’s eye camera in the condition of (a).
5.2 Manipulation of the Bird’s Eye Overview Image

The various kinds of constraint described in section 4 are used to control the position and orientation of the bird’s eye camera with the interlocked pairs of coordinate systems described in 3.2. Here, in the following experiments, two base constraints are used as a standard condition to be compared. The first base is a constraint on \( XY \), i.e., the intersection of the central axis of the bird’s eye camera and the horizontal plane of the entire virtual environment (point \( S_b \) in figure 4) is fixed to the user’s position. The second base is the constraint on heading (\( \theta_b \)) which is fixed to the user’s direction (\( \theta_u \)). By using these constraints, the user-aligned setup is established where the user’s position is always in the center of the bird’s eye overview image as shown in figure 9. Figure 10 shows the positions and directions of the user and the bird’s eye camera on the two base constraints.

From this standard condition, we examine the change of the user’s view by manipulating one of the two base constraints.

Figure 8: Snapshot of user’s view and user and bird’s eye camera positions.

Figure 9: Standard image and condition compared in the following experiments.
5.2.1 Change of User’s View by Manipulating Constraint on X and Y

This experiment manipulates the constraint on XY. At first, the hand-held tracker is moved to the right from its standard condition as shown in figure 11(a). The bird’s eye overview image changes according to the motion of the user’s hand as shown in figure 12(a) in comparison to its standard condition shown in figure 9(a). On the other hand, when the user (with a tracker attached to his/her head) moves to the right from a standard condition without moving his/her hand as shown in figure 11(b), the bird’s eye overview image does not change due to his/her motion except the user’s position as shown in figure 12(b) in comparison to the standard condition shown in figure 9(a).

If the constraints on a standard condition work, the user’s position is always fixed at the center of the bird’s eye overview image even if he/she moves. Figure 13 shows the user’s view caused by movement of the user to the right under a standard condition.

5.2.2 Change of User’s View by Manipulating Constraint on heading

This experiment manipulates the constraint on heading. At first, the hand-held tracker is rotated in the XY plane from its standard condition as shown in figure 11(c). The bird’s eye overview image changes due to the motion of the user’s hand as shown in figure 14(a) in comparison to its standard condition shown in figure 9(a). On the other hand, when the user (with a tracker attached to his/her head) rotate in the XY plane from the standard condition without moving his/her hand as shown in figure 11(d), the bird’s eye overview image does not change due to his/her motion except the position of the user as shown in figure 14(b) in comparison to the standard condition shown in figure 9(a).

As described before, if the constraints for a standard condition work, the user’s position is fixed always at the center of the bird’s eye overview image even if he/she moves. Figure 15 shows the user’s view caused by a rotation of the user under the standard condition.

5.3 User Studies

Experiments have to be conducted to determine how intuitively a user can control the position and direction of bird’s eye camera, and how efficiently he/she construct the cognitive map of the environment by recognizing the position and direction of himself/herself and the bird’s eye camera. It may be necessary to compare the usefulness of our proposed method with world-in-miniature (WIM) [5] in some of these experiments. It is also useful to evaluate the effectiveness of introducing the constraints on the attitude of the bird’s eye camera described in section 4. For example, it is useful to evaluate the followings by changing the combinations of constraints; accuracy of user’s global understanding of the environment and local understanding around the user, perception of size, and recognition of user’s own position and direction in the environment.

Efficiency of the proposed method is evaluated by measuring the completion time of a certain task, e.g., the orienteering to reach the goal by finding several succeeding landmarks using map and compass in a 3-D maze. On the other hand, accuracy is evaluated by comparing the the correctness of the production of a task, e.g., constructed map of the environment, generated movie by following the instructed route, or the bird’s eye overview image taken with an instructed composition. The result of these experiments on user studies will be reported in the next opportunity.

6 Summary and Conclusions

A method which enables a virtual reality user to intuitively control the “bird’s eye” overview display of an entire virtual environment was proposed. The key idea for intuitive manipulation was for the “interlocked pairs of coordinate systems” to interlock the motion of the bird’s
eye viewpoint with the motion of the user’s viewpoint. In addition, some sophisticated display methods were obtained by introducing some constraints on parameters for determining the position and orientation of the “bird’s eye” camera. Experimental results showed that changes in the bird’s eye overview image and the life-size local image were directly caused by the user’s motion. Future work includes more investigation on constraints with multiple bird’s eye overview images, user studies to determine intuitiveness and efficiency of the proposed method, investigation on a proper method of changing the scales/magnifying power of overviews using 3-D interaction devices, and virtual object manipulation using the bird’s eye overview image.

Acknowledgement: This work was supported in part by the research program of the Telecommunications Advancement Organization of Japan.

References
Figure 12: Change of user's view by manipulating constraint on $X$ and $Y$ coordinate.

(a) Translation of attached hand tracker.

(b) User translation.

Figure 13: Change of user's view caused by user's translation under standered condition.
Figure 14: Change of user’s view by manipulating constraint on *heading*.

Figure 15: Change of user’s view caused by user’s rotation under the standered condition.