Handymap: A Selection Interface for Cluttered VR Environments Using a Tracked Hand-held Touch Device

Mores Prachyabrued[1], David L. Ducrest[2], and Christoph W. Borst[1]

[1] University of Louisiana at Lafayette, Lafayette, Louisiana, USA
[2] Louisiana Immersive Technologies Enterprise, Lafayette, Louisiana, USA

Abstract. We present Handymap, a novel selection interface for virtual environments with dense datasets. The approach was motivated by shortcomings of standard ray-casting methods in highly cluttered views such as in our visualization application for coalbed methane well logs. Handymap uses a secondary 2D overview of a scene that allows selection of a target when it is occluded in the main view, and that reduces required pointing precision. Reduced sensitivity to pointing precision is especially useful for consumer-level VR systems due to their modest tracking precision and display sizes. The overview is presented on a tracked touch device (iPod Touch) that is also usable as a general VR wand. Objects are selected by a tap or touch-move-release action on the touch surface. Optionally, redundant visual feedback and highlighting on the main display can allow a user to keep focus on the main display and may be useful with standard wand interfaces. Initial user feedback suggests Handymap can be a useful selection interface for cluttered environments but may require some learning.

1 Introduction

We present Handymap, a novel selection interface that uses a tracked hand-held touch device to address occlusions in highly cluttered views and that does not hinge on ray pointing precision. We are developing a VR-based visualization system for geological interpreters to interpret well log data (spontaneous potential and resistivity curves) from wells situated in Northern Louisiana. Fig. 1 (left) shows a scene from the application. The database contains several hundred well logs that the application can display, creating cluttered views even when smaller subsets are displayed. This causes selection problems with ray-casting interfaces [1]. Standard ray casting uses a virtual ray extended from a hand or controller to select the first intersected object. In a cluttered view, it can be difficult to select a target due to occlusions. A standard ray interface requires navigation to resolve difficult occlusions, which may increase selection time or disturb the view context. In a less extreme case, occlusions reduce the selection target area, making ray-casting slower and less accurate [2] due to increased required pointing precision. The problem appears especially when a user selects distant targets, which occurs in our application in which the user overviews a

This is an author-formatted version. The original publication is available at www.springerlink.com. ISVC 2011, Part II, LNCS 6939, pp. 45-54.

large collection of well logs. A small hand movement becomes a large distant ray movement, reducing pointing precision. This problem is also notably increased in consumer-level VR setups with modest tracking precision and display sizes.

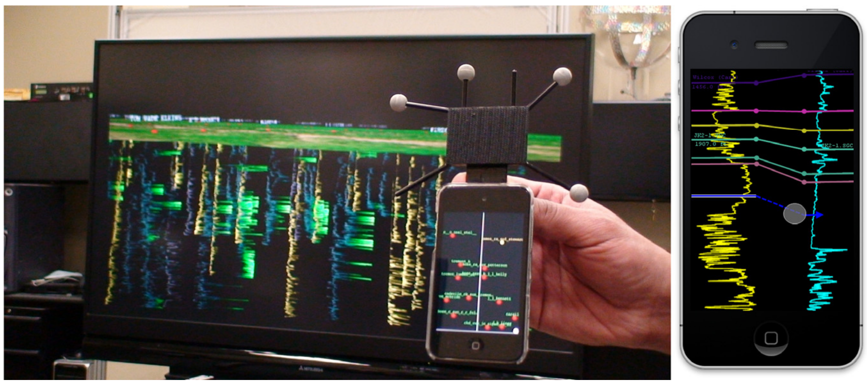

Fig. 1. *Left:* Low-cost well log visualization system (Mitsubishi 3D DLP TV and iPod Touch with markers for OptiTrack camera-based tracking) showing well logs (curves) hanging underneath a terrain generated from SRTM (Shuttle Radar Topography Mission) data. The iPod Touch presents an overview of the well log scene that resolves occlusion in the main view and supports rapid touch-based selection. A middle vertical line represents a virtual ray in the main view, which is locked during a selection step. *Right:* (Constructed conceptual illustration) Well log "picks" illustrated as horizontal lines with associated depth and text annotation. A highlighted pick on the left log is being associated with a pick on the right log by a drag gesture

Various selection techniques address occlusions or pointing precision problems (see Section 2). However, they may not be adequate for our well log application, so we developed Handymap selection. It exploits the scene structure that has a well log dataset distributed on a terrain surface (well log curves hanging underneath) by using a secondary 2D overview of the scene as shown in Fig. 1 (left). It presents the overview on a tracked touch device (iPod Touch) that is also usable for conventional ray interactions. The overview represents well logs with circles and labels surrounding a virtual ray to resolve occlusion in the main view. Although a virtual ray extends from the iPod Touch similar to standard ray-casting interface, it is not used for conventional intersection-based selection. Instead, users touch the handheld display with a thumb to select a well log from the overview.

Handymap visuals can also be used with a standard controller (e.g., Logitech gamepad or InterSense Wand) by presenting the overview on a main display and requiring joystick or pointing interactions. However, the iPod Touch interface allows direct touch selection that may be faster, supports intuitive gesture-based overview adjustments (zooming and panning), and reduces clutter on the main display.

Additionally, the touch interface can further aid interpretation, e.g., to improve management of well log "picks", which are depth levels selected on logs for geologic relevance. The touch surface provides direct depth selection for picks as well as efficient text annotation via its virtual keyboard that allows faster selection of characters than standard VR wand techniques. Geological interpreters have requested to manipulate multiple well logs on the iPod Touch as it may help relate picks between well logs (Fig. 1, right). The interpreters could use the related picks to generate a coarse subterranean surface representation of underground composition.

After describing the Handymap interface, we report initial user feedback. It suggests that Handymap will improve selection in our well log application, and suggests important features and considerations for the interface.

2 Related Work

2.1 Selection Techniques Addressing Occlusion or Precision Problems

We summarize relevant ray-based approaches due to ray interaction dominance and because studies [3, 4, 5] have shown it has better selection performance than techniques like virtual hand-based selection [6]. Olwal and Feiner [7] presented Flexible Pointer, which allows users to bend a virtual ray to point to fully or partially obscured objects. Wyss, Blach, and Bues [8] presented iSith, a technique that addresses occlusions by using an intersection of two rays to define a target. Grossman and Balakrishnan [5] presented Depth Ray, Lock Ray, Flower Ray, and Smart Ray techniques that include mechanisms to disambiguate a target from multiple intersected objects along the ray. All those techniques require the ray(s) to intersect with the target and suffer from limited pointing precision at long distances and with tracker jitter. Also, Flexible Pointer and iSith require additional tracked input devices.

Selection techniques such as flashlight (Liang and Green [9]) or aperture (Forsberg, Herndon, and Zeleznik [10]) lower required pointing precision by replacing the virtual ray with a conic selection volume. Frees, Kessler, and Kay [11] presented the PRISM enhanced version of ray-casting that increases user pointing precision by dynamically adjusting the control/display ratio between hand and ray motions. All those techniques do not work well in highly cluttered views or do not address the case of a fully occluded target. Kopper, Bacim, and Bowman [12] recently presented Sphere-casting refined by QUAD-menu (SQUAD) that addresses occlusions and does not require high pointing precision. It uses sphere volume to define an initial set of selectable objects, and it progressively refines the set using QUAD-menu selection until the set contains only the target. However, evaluation showed that it may not work well with highly cluttered environments due to the required number of refinements. Also, its selection process does not preserve spatial information, while we want a technique that shows some spatial relations.

2.2 Handheld Device Interfaces for Virtual Environments

Aspin and Le [13] compared a tracked tablet PC to a tracked gamepad in immersive projection display environment. They found that using the tablet PC created a greater sense of immersion. Users developed a stronger relationship with the virtual environment because of the interactive visuals and tactile sensation of the tablet. Olwal and Feiner [14] leveraged the high visual and input resolution of a touch display on a tracked mobile device for improved interaction on a large touch display (zooming and selection of small targets). Their user study showed overall higher user preference for this approach over direct touch interaction on the large display. Katzakis and Hori [15] evaluated use of accelerometers and magnetometer on a mobile phone for a 3D rotation task. Their results showed it to be faster than mouse and tablet interactions.

Kim et al. [16] presented a navigation technique called Finger Walking in Place (FWIP) using finger motions resembling leg walking motions on a multi-touch device. This was later adapted to iPhone/iPod Touch for navigation in a CAVE [17].

Song et al. [18] presented volume data interactions using a multi-touch wall display and iPod Touch. In addition to using multi-touch gestures on the iPod Touch, slicing plane position (on the wall display) could be controlled by sliding the iPod Touch on the wall display, with orientation of the slicing plane controlled by tilt sensing on the iPod Touch. Slices could then be annotated on the iPod touch.

3 Handymap Design

3.1 Map Overview

Handymap presents a 2D overview of the virtual environment. We consider different perspectives for the overview, based on projections along a world-up axis, terrain-up axis, or controller-up axis. In any case, the overview represents a 3D region in the environment, where position and orientation of the region change with controller pose and inputs. Well logs in this region are represented by labeled icons on the overview. The overview can be zoomed and panned by scaling and translating the region. To address hand instability and tracker jitter, Handymap incorporates a ray-locking behavior where the overview becomes static, i.e., independent of additional virtual ray (controller) motion during selection and overview adjustment.

3.2 Handymap Interaction

We consider two main interaction types: overview and scene. Overview gestures control prominent features within the overview: well log selection, overview zooming, and overview panning. Overview gestures were our primary focus, but we also incorporate scene-specific gestures to manipulate the scene, e.g., world-grab, view azimuth/elevation, scene panning, and terrain scaling.

Overview gestures (Fig. 2) rely on the user's primary hand and especially the thumb. Well log selection uses a touch-refine-release approach. The user touches the iPod display to initiate the interaction, tentatively indicating the well log closest to the touch point (Fig. 2a). The user can change (refine) the indication by moving the thumb closer to another icon while maintaining touch. Finally, the user releases the touch to select an indicated well log. During this interaction, an indicated well log is highlighted both on the iPod and on the main display. The user can additionally pan the overview region during selection refinement by dragging the thumb to any edge of the display (Fig. 2c). This is a temporary pan and is forgotten when the touch ends. To cancel selection, the user releases the touch at any edge of the display (in a panning zone). To zoom the overview, the user touches the display with the primary thumb and forms a pinch gesture with the secondary hand (Fig. 2b).

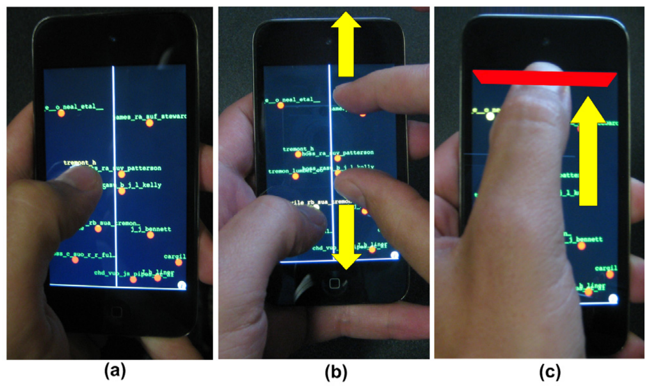

Fig. 2. Overview gestures for well log selection and overview adjustment: (a) Touch-refine-release to highlight and select a well log. (b) Pinch gestures for zooming the overview. (c) Drag gestures for panning the overview (entering the red area pans the overview forward)

Ray-locking can be enabled as a system parameter. If enabled, the overview region is independent of ray (iPod) motion during overview gestures. The best default behavior depends on the VR system type and user characteristics, e.g., consumer-level VR setups with notable tracker jitter may require ray-locking enabled.

Prototype scene gestures typically use the secondary hand. The user paws the iPod display with two fingers next to each other to pan the scene along the world floor plane. Pawing with two fingers separated adjusts view elevation, and rotating one finger about the other finger adjusts azimuth. To uniformly scale the scene, the user pinches two fingers. For grab-the-world type scene manipulation, the user taps once with the primary thumb, then taps and holds to clutch (grab), then moves the (tracked) iPod Touch in 3D space, and finally releases the touch to end the grab.

We use a state machine to prevent distinct gestures from overlapping. A refine gesture (both target indication and overview panning) can transition to any other gesture except world-grab. Overview zooming can transition to refining but not to scene gestures. A fast single tap will not result in selection but is used to detect world-grab. A scene gesture must end (no finger on the surface) before another gesture is detected.

3.3 Handymap Perspective and Overview Calculation

Handymap perspective determines how the scene is projected for the overview. It affects occlusions in the calculated 2D overview and consistency between object layout in the overview and in the main view. We considered three perspectives:

1. World-based: The overview is displayed like a view down from the top of the main display, i.e., parallel to the real world floor (Fig. 3a).
2. Terrain-based: The overview is displayed as though it is parallel to the terrain, i.e., view direction normal to the terrain's principal plane (Fig. 3b).
3. Controller-based: The overview is displayed as though it is parallel to the controller face, i.e., view direction defined "controller-up" axis (Fig. 3c).

In all cases, the overview still rotates according to orientation of the controller (iPod) with respect to an axis parallel to the projection direction.

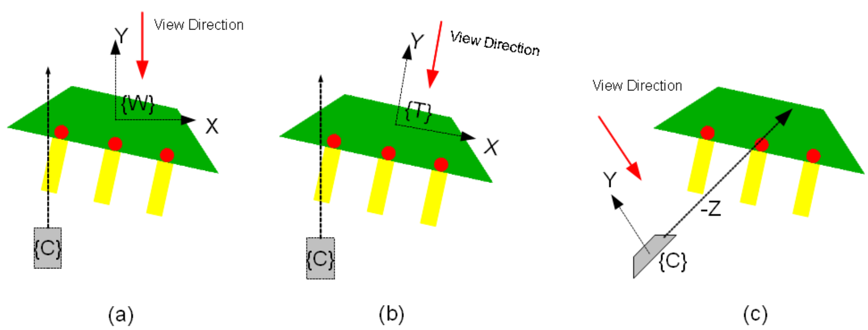

Fig. 3. Three Handymap perspectives considered: (a) World-based perspective. (b) Terrain-based perspective. (c) Controller-based perspective. The figure represents the controller (e.g., iPod Touch) and views from the main display. {W}, {T}, {C} refer to fixed world (main display), terrain, and controller coordinate frames, respectively

3.3.1 Overview Calculation

Given a Handymap perspective, a Handymap coordinate frame is calculated as detailed in the following subsections. The 3D region mapping to the overview is centered and aligned on this coordinate frame. Mapping well log positions for Handymap icons is done by transforming positions to this coordinate frame (reference positions near the terrain surface). The overview shows only well logs whose Handymap coordinates fall within the mapped region based on current scale.

3.3.2 World-Based Perspective

The world-based perspective provides a consistent object layout between the overview and the main view, e.g., objects to the left of the virtual ray in the main view (from user's usual perspective, independent of controller rotations around the ray axis) are represented on the left side of the virtual ray representation in the overview. The world-based perspective has occlusions in the overview when the terrain tilts significantly away from horizontal (with respect to the world).

The Handymap coordinate frame origin is computed as a fixed point on the virtual ray projected onto the world floor or horizontal (XZ) plane. We chose this fixed point by considering a user's typical interaction depth (i.e., typical distance between user and dataset) so that the overview region falls on the terrain where objects of interest reside. The Handymap up (Y) axis matches world up (Y) axis. The Handymap forward (-Z) axis is computed as the virtual ray direction vector projected to the world floor plane. The Handymap left (-X) axis is found by axes cross product.

We chose world-based perspective as the default perspective because it provides a consistent layout and is not limited by our well log data. Since our terrain is nearly planar, it is uncommon and unnecessary to rotate the terrain far from horizontal. With reasonable scale for the overview region, our well log application has no occlusion in the overview with world-based perspective.

3.3.3 Terrain-Based Perspective

The terrain-based perspective better preserves object spacing and eliminates occlusion in the overview when objects are distributed on the terrain's surface (assuming reasonable overview scale). However, the object layout in the overview may be inconsistent with the main view, e.g., objects to the left of the virtual ray in the main view (defined as before) could be on the right side of the overview if the terrain is flipped up-side down in the world.

The Handymap coordinate frame for terrain-based perspective is computed similarly to world-based perspective (Sect. 3.3.2) except that the calculation uses terrain floor (XZ) plane and terrain up (Y) axis in place of world floor and world up.

With normal constrained terrain rotation in our well log application, there is no layout consistency problem. Since there is also no occlusion in the overview, the terrain-based perspective works about as well as the world-based perspective.

3.3.4 Controller-Based Perspective

In the controller-based perspective, the Handymap coordinate frame is simply the controller frame translated to the fixed point (Sect. 3.3.2) on the virtual ray.

The controller-based perspective suffers from both occlusion (in the overview) and layout consistency problems depending on controller orientation. However, it provides the user with the most full and direct control of the overview. The user is free to adjust the overview to avoid these problems. The controller-based perspective may be a good option for 3D data inspected from more angles, but it does not provide notable benefit in our well log application.

3.3.5 Zooming and Panning the Overview

Zooming the overview is accomplished by scaling the overview region. We chose a default region size intended for good distribution of well log icons in the overview. Panning the overview is accomplished by translating the Handymap frame origin on its local view plane axes.

4 User Feedback

We solicited feedback from 5 users about their expectations and suggestions for Handymap based on 30-45 minute sessions. The users were three geosciences *domain experts* (one with VR experience and previous experience with the application) and two *VR experts* with some prior exposure to the application. We used the equipment and scene shown in Figure 4.

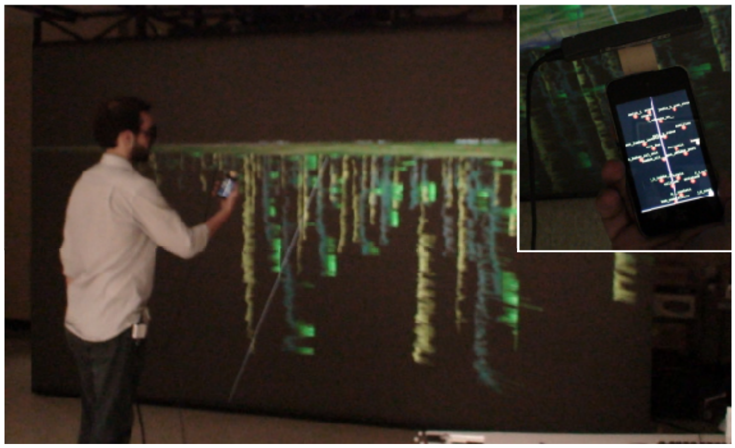

Fig. 4. Environment on Visbox HD13 with Intersense IS-900 tracking of controller and head

We asked each user to compare Handymap to standard ray-casting for well log selection in a cluttered view (Fig 4). All users believe Handymap improves selection (easier and faster) because the overview resolves occlusions, allowing selection without navigation. This is especially appreciated by the domain experts since it does not disturb their interpretation context (e.g., when they want to select two well logs from the same view for comparison). One domain expert and one VR expert stated that ray-casting may be better when a target is close and unobstructed. Two domain experts indicate that Handymap requires learning. One domain expert estimated that it took 10 minutes for proficiency but still expressed a preference for Handymap.

We asked for feedback during both selection of specified targets and free exploration. One domain expert and one VR expert stated that clear presentation of log labels in the main view is important to allow them to find matching target on the iPod touch confidently. One VR expert stated that they could locate a target in the overview easily by relating it to the virtual ray. Two domain experts and one VR expert commented that overview zooming is useful, since it allows them to use Handymap for a larger region and allows finer interaction. One domain expert and one VR expert commented that additional terrain representation on Handymap can be useful, but should be optional. Geologists usually consider topography irrelevant to these interpretations.

Two domain experts commented that seeing a target in the main view when absent in the overview was confusing, demonstrating the importance of reasonable overview scale. One VR expert commented that fingers interfere with text reading on iPod Touch. All users commented that a focus shift between the main display and iPod Touch was a drawback but still express preference for Handymap. One domain expert suggested that tilting the touch surface to the user's eyes during interaction would reduce focus shift. Another domain expert suggested selection should not be cancelled when releasing touch in a panning zone. Two VR experts suggested that additional representations of overview region and touch point in the main view may be helpful.

We also asked each user to test display alternatives for Handymap visuals. One case used main display visuals instead of iPod visuals, placing the overview at the bottom center and aligned with the main display surface. One VR expert stated that the overview cluttered the display and was confusing, further stating that the overview did not feel like a top-down view due to the alignment. They suggested that tilting the overview may help. The other users liked the reduced focus shift, with two domain experts stating that it helps mental focus. A domain expert stated a large overview is helpful. The other two domain experts stated that a single display helped them relate overview to main view. One VR expert stated it avoids finger interference with labels.

Another approach was to omit visual overview and mainly use the Handymap for touch input. In this case, the visual cue, on the main display, was to dynamically highlight the log corresponding to thumb position. One domain expert stated that selecting from the overview was easier. The other users liked the reduced focus shift. One VR expert stated that it related interaction to the main view. One domain expert and one VR expert commented that overview panning is useful since it allows continuous interaction even without looking at the iPod. One VR expert suggested that visual cues in the main view for panning would help, or to limit panning range.

Based on responses, we believe that the touch input aspect of the iPod was more important than its visual display, and extending visual feedback associated with Handymap on the main display is a good next step. We expect the touch display surface to further be useful for other tasks in our application, as suggested in the introduction.

5 Conclusion and Future Work

We summarized the occlusion and pointing precision problems with standard ray-casting in cluttered virtual environments. We described the Handymap selection interface to address these problems in a well log visualization application. User feedback suggests Handymap can be a useful interface for cluttered environments, but that it may require some practice. Easy association of a target in the main view with the corresponding representation in the overview, touch input surface, redundant feedback in the main view, and overview zooming and panning features are important.

Future work should include formal evaluation of Handymap with comparison to other techniques and understanding of design tradeoffs. For example, we want to evaluate the iPod visual display for Handymap to see if it impacts performance over presenting visuals on the main display, considering the focus shift between the touch display and the main display. We will consider extensions to Handymap, e.g., additional 3D representations of overview region and touch point in the main view, or by investigating auto-scaling of the overview region. Finally, we will continue to extend our iPod Touch interface for well log interpretation.

References

1. Mine, M. R.: Virtual Environment Interaction Techniques. Technical Report, University of North Carolina at Chapel Hill. (1995)
2. Steed, A. and Parker, C.: 3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays. 8th International Immersive Projection Technology Workshop. (2004)
3. Poupyrev, I., Weghorst, S., Billinghurst, M., and Ichikawa, T.: Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. Computer Graphics Forum, 17(3). (1998) 41-52
4. Bowman, D. A., Johnson, D. B., and Hodges, L. F.: Testbed Evaluation of Virtual Environment Interaction Techniques. Proceedings of ACM Symposium on Virtual Reality Software and Technology (VRST). (1999) 26-33
5. Grossman, T. and Balakrishnan, R.: The Design and Evaluation of Selection Techniques for 3D Volumetric Displays. Proceedings of ACM Symposium on User Interface Software and Technology (UIST). (2006) 3-12
6. Bowman, D. A., Kruijff, E., LaViola, J. J., and Poupyrev, I.: 3D User Interfaces: Theory and Practice. Addison-Wesley. (2004)
7. Olwal, A. and Feiner, S.: The Flexible Pointer: An Interaction Technique for Augmented and Virtual Reality. Conference Supplement of ACM Symposium on User Interface Software and Technology (UIST). (2003) 81-82
8. Wyss, H. P., Blach, R., and Bues, M.: iSith – Intersection-based Spatial Interaction for Two Hands. Proceedings of IEEE Symposium on 3D User Interfaces (3DUI). (2006) 59-61
9. Liang, J. and Green, M.: JDCAD: A Highly Interactive 3D Modeling System. Computers and Graphics, 18(4). (1994) 499-506
10. ForsBerg, A. Herndon, K., and Zeleznik, R.: Aperture Based Selection for Immersive Virtual Environments. Proceedings of ACM Symposium on User Interface Software and Technology. (1996) 95-96
11. Frees, S., Kessler, G. D., and Kay, E.: PRISM Interaction for Enhancing Control in Immersive Virtual Environments. ACM Transactions on Computer-Human Interaction, 14(1):2. (2007)
12. Kopper, R., Bacim, F., and Bowman, D. A.: Rapid and Accurate 3D Selection by Progressive Refinement. Proceedings of IEEE Symposium on 3D User Interfaces (3DUI). (2011) 67-74
13. Aspin, R. and Le, K. H.: Augmenting the CAVE: An Initial Study into Close Focused, Inward Looking, Exploration in IPT Systems. Proceedings of IEEE Symposium on Distributed Simulation and Real-Time Applications. (2007) 217-224.
14. Olwal, A. and Feiner, S.: Spatially Aware Handhelds for High-Precision Tangible Interaction with Large Displays. Proceedings of International Conference on Tangible and Embedded Interaction (TEI). (2009) 181-188
15. Katzakis, N. and Hori, M.: Mobile Devices as Multi-DOF Controllers. Proceedings of IEEE Symposium on 3D User Interfaces (3DUI). (2010) 139-140
16. Kim, J., Gračanin, D., Matković, K., Quek, F.: Finger Walking in Place (FWIP): A Traveling Technique in Virtual Environments. Proceedings of International Symposium on Smart Graphics. (2008) 58-69
17. Kim, J., Gračanin, D., Matković, K., Quek, F.: iPhone/iPod Touch as Input Devices for Navigation in Immersive Virtual Environments. Proceedings of IEEE Conference on Virtual Reality (VR). (2009) 261-262
18. Song, P., Goh, W. B., Fu, C., Meng, Q., and Heng, P.: WYSIWYF: Exploring and Annotating Volume Data with a Tangible Handheld Device. Proceedings of ACM Annual Conference on Human Factors in Computing Systems (CHI). (2011) 1333-1342