# MobiVR - A Novel User Interface Concept for Mobile Computing

Ismo Rakkolainen
*Institute of Signal Processing, Tampere University of Technology,*
*P.O. BOX 553, 33101 Tampere, Finland*
*Ismo.Rakkolainen@tut.fi*

## Abstract

*As mobile computing and wireless Internet is emerging, current windows-based user interfaces with mouse and keyboard are becoming impractical. New kinds of user interfaces are needed for hand-held devices.*

*We present a pragmatic gestural user interface approach for hand-held devices. The basic idea is to combine any tracking method, a near-eye microdisplay, and wireless communication into a hand-held device in a manner that enables intuitive finger or stylus pointing, and use the device as a universal information appliance.*

*An early prototype employing a computer vision-based tracking method is implemented. The prototype enables a normal windows-based, full-resolution screen view in a hand-held device. The user's finger can replace a mouse.*

*The presented MobiVR concept seems to be useful and easy to use, and can be applied to numerous future hand-held devices. It could also facilitate e.g., mobile augmented and virtual reality applications.*

## 1. Introduction

What is in Figure 1? A computer? No, they are only the present computer peripherals for input and output (I/O). The microprocessor itself is very small and can be embedded to many devices. The I/O peripherals could also look very different.

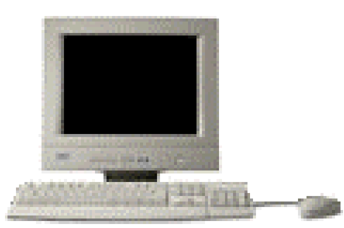

**Figure 1. The present computer I/O peripherals.**

The user interfaces (UI) are becoming increasingly important in computing. The fundamental properties of computers have changed radically over the years, but human-computer interaction (HCI) has not changed much for two decades. Communications and computing are converging, and mobile digital devices have become small. Current second generation (2G) mobile phones and wireless networks are sufficient for mobile voice communication. The emerging 2½G and 3G networks are mostly aimed for data communications like multimedia messages, mobile Internet and email. The dominant paradigm of computing seems to be changing towards mobile computing and wireless Internet.

The user interfaces must follow this major context of computing. The current UI approaches like Graphical User Interfaces (GUI) with traditional I/O devices like mice and keyboards are mostly inappropriate for the future hand-held computers, which will come in diverse form factors and functions like personal digital assistants (PDAs), hand-held games, mobile phones, digital audio players, smart cameras, etc.

The pen-based UI of a PDA is good for some purposes, but the small, low-resolution screen will not be enough for many mobile tasks like web browsing. The need for bigger displays and better user interfaces for hand-held devices is paramount.

A good user interface should be so intuitive and easy-to-use that it would not require any substantial training. The quest for better user interfaces for mobile devices is taking place, although no universally best UI for all tasks, users, contexts, devices, and environments exist.

Having a decent sized display and an easy way of inputting data and checking information on a pocket-sized device is difficult to obtain. Current high-resolution LCD displays on laptop and tablet PCs are rather large for small hand-held appliances. Some solutions to have a large display on a small appliance are e.g., flexible roll-up displays, as proposed in [7] and recently in the epyrus prototype by Siemens at CeBit 2003.

Near-eye microdisplays would provide truly lightweight and high-resolution virtual screens for wearable computers [2] and hand-held devices. There are some PocketPC PDAs with high-resolution near-eye microdisplays, like Mobintech's Mobile Internet Platform, Interactive Imaging Systems' Ice communicator, Hitachi's Wearable Internet Appliance and the discontinued Inviso's eCase. Also prototype mobile phones with near-eye microdisplays have been presented. McLaughlin Consulting Group predicts that over 10 million microdisplays will be embedded to some mobile phones, PDAs, and Internet appliances (http://mcgweb.com/presentations/USDC.htm) in 2005, with the number growing very fast.

The hand-held devices should be small, yet their displays and interfaces should be big and convenient to use. A microdisplay can solve the display issue, but the UI becomes then a problem. Cursor movement could be handled by e.g., a touchpad or a tiny joystick, but they may be slow and clumsy to operate with the same hand that is holding the hand-held device. Hand gestures could be a more natural and intuitive way of pointing.

We present a novel, multimodal approach for the user interface of mobile devices [8]. The MobiVR concept combines several well-known technologies and elements in a new way, which is an invention by definition. The novelty lies in the combination of these elements. The implementation of the concept is still in a very early phase.

MobiVR consists of a near-eye microdisplay, a tracking system that facilitates finger pointing or similar gestures, and a wireless network connection. It can interact with any local or remote applications, information stores, or devices, which export their interfaces to it.

MobiVR enables to use a traditional WIMP user interface in a hand-held device. Preferably a computer vision-tracked finger is used for pointing, but other pointing devices and tracking methods can also be used to reduce the required computation of the tracking. The index finger is an intuitive, convenient, and universal pointing tool, which is always on and available, and is used since very early childhood and across all cultures.

The structure of this paper is as follows. Chapter 2 discusses the prior art. Chapter 3 presents the basic idea of MobiVR user interface, and Chapter 4 describes the early implementation of it. Future work is discussed in Chapter 5 and conclusions are given in Chapter 6.

## 2. Prior art

Hundreds of I/O devices and novel UIs have been presented to improve Windows, Icons, Menus, and Pointing (WIMP) GUI and also to add more natural, perceptive, and compelling capabilities to HCI. One candidate for some tasks is finger pointing, which is an intuitive and well-known method for HCI.

There are several variations to finger tracking. Often video cameras and computer vision methods are used for tracking. The cameras are scattered around a desktop PC or a room, or sometimes they have been attached to a head-mounted display (HMD). The tracking system can employ infrared (IR) emitters and IR detectors [e.g., 9, 13] and can also be finger-mounted [10].

In the following, some earlier works, which are relatively near and relevant to the MobiVR concept, are presented. European patent EP-A-0679984 describes a system employing a HMD with cameras to track the user's finger for input.

Mann [4] has presented a finger mouse for wearable computing. It enables to control a cursor with a finger, allowing the user to draw outlines of objects or to control the local wearable computer's pull-down menus. It requires a HMD with an attached camera. The drawing of real objects requires, however, very precise registration with the real world, which is a major problem.

MIT Media Lab's augmented reality UI for wearable computing employs finger tracking [12]. It allows the user to replace the mouse with a finger. The MIT finger tracking is used to control the operating system, or digitize an image and virtually annotate it. It augments the traditional computer interface.

Hand Mouse [3] is a wearable computing input interface, which also uses finger pointing to replace mouse. The Hand Mouse is not hand-held, but requires head-mounted cameras, displays, and other sensors and devices.

Marisil [6] is a hands free UI for future mobile media phone. It uses a see-through HMD with attached miniature cameras, and tracks the user's fingers. The HMD augments the user's palm with virtual buttons, which the user can point with a finger. The user needs no hand-held devices or artifacts to make a phone call, but a see-through HMD with cameras is required. Real-time registration is essential for the method, which is computationally heavy and difficult to implement precisely. The palm and the pointing finger are usually of the same color, and are thus not easy to distinguish from each other.

The MobiVR concept has several differences with the abovementioned UIs. Although it is a little similar to some wearable computing and virtual reality (VR) user interfaces, the main contribution is that MobiVR adjusts the user interface for mobile computing by giving it a new form and a relevant context. The device is hand-held, not head-mounted, and thus does not occlude the real world all the time, as many wearable and VR user interfaces. This is usually convenient for most consumers and real usage situations.

MobiVR is conceptually a different thing from HMD-based systems. The idea of being hand-held affects the context, purpose, possibilities, usability, mobility, and convenience of use, and thus the very essence of the device. MobiVR is differently constructed and is suitable for partly different tasks.

In addition, current HMDs may require adjustments, are fragile, sizeable, and may be inconvenient for short tasks and for mobile users out on the streets. HMDs may not become widely accepted consumer products in mobile use, even if they reduce in size and cost. HMDs are socially unacceptable for many situations and people, as they for example block visibility to real world, isolate people into their virtual worlds, and remove eye contact. Also, MobiVR requires no computationally heavy registration with the real world, as is the case with augmented reality (AR), although AR and VR are also possible with the MobiVR.

Also, the earlier works do not use IR light for tracking, and were designed to help the operation of the local computer, not to interact with any remote devices. MobiVR also adds wireless communications and ubiquitous computing [14] aspect to finger or stylus pointing. The wireless transmission channel (e.g., Bluetooth, GPRS, WLAN, or IrDA) enables the user to access any remote device, appliance, or computer that exports its UI to MobiVR. The exported UI format can be e.g., XHTML or other web standards.

Siemens has recently introduced a stylus- and camera-based pointing concept for hand-helds [11], which is a little similar to MobiVR. It employs a camera embedded into the back of a PDA or a mobile phone, and tracks a pen with computer vision. The user looks onto the ordinary LCD display and points the stylus behind the device. One major difference lies in the small-resolution display, which still does not provide a sufficient resolution and/or small size for many tasks. As a conclusion, we have found no hand-held devices similar to MobiVR.

## 3. The basic idea of the MobiVR concept

The basic components of MobiVR are a wireless communication channel, a near-eye microdisplay, and a tracking system to track a finger or some pointing device (see Figure 2).

Finger tracking based on camera and computer vision is a convenient and intuitive method of pointing, as no hand-held or finger-mounted artifacts are needed. In the basic form, the field of views of the camera and the display align.

Arm strain and fatigue can be avoided by tilting the camera or by adding an extra camera to the bottom of the device (see Figure 2) so that the arm can remain low or even on top of a desk.

The device provides a "virtual touch screen" for the user and couples the virtual interface of a wireless handheld appliance to the hand-eye coordination of the user. The user can point to the items on the virtual display intuitively with a finger (see Figure 3), and thus replace a mouse.

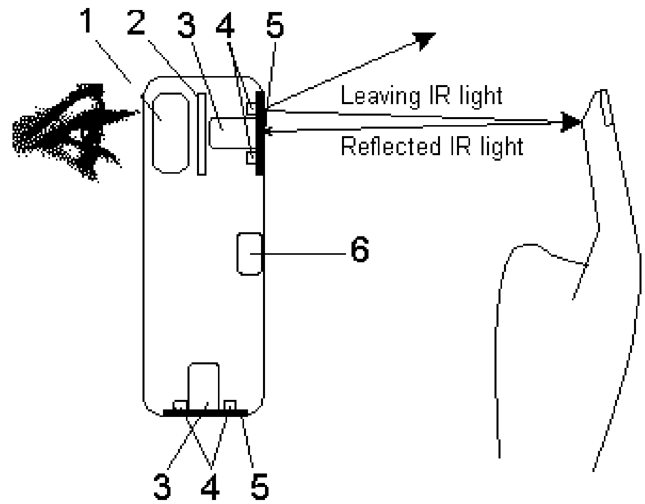


**Figure 2. The basic components of the MobiVR user interface device. The optics (1) and the microdisplay (2) show the UI of a remote appliance. The camera (3) tracks the user's finger. Optional IR light sources (4) and IR pass filter (5) can simplify the tracking, and there can be several of them. Also e.g., inertial sensors (6) can be used.**

Objects can be selected ("clicking the mouse") e.g., by a gesture like pressing the icon or a thumb movement, or the device or a pointing stylus may incorporate a conventional button for selection. Also other methods like speech recognition can be used.

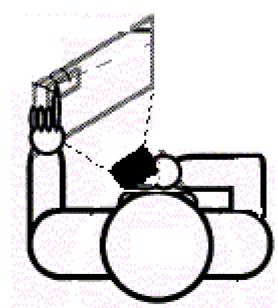

**Figure 3. The user can point a finger to the objects on the virtual screen.**

Current GUIs depend upon click-and-point paradigm, i.e., entering an object and selecting it. However, precise pointing may be time-consuming and especially double clicking may be difficult, as cursor may move accidentally. Interesting future possibilities are crossing-based interfaces [1], which may be faster to use in general, and also in gestural interface devices like MobiVR. Only crossing the object boundary is then enough to select an object.

The device can be described, thought of, and applied in many ways. It could be seen as a hand-held web browser, an integrated UI component of a hand-held appliance (multimedia mobile phone, PDA, wristwatch, camera, etc.), a universal, ubiquitous remote control, etc. Its design can resemble e.g., a camera, a stick, a remote control, or a binocular.

Computer vision-based finger tracking can be made robust and real-time. An IR camera, IR light sources near the cameras, and IR pass filter with visible light cut-off (see Figure 2) can significantly simplify the finger tracking process, which is a very well known idea. Instead of, or complementing IR tracking, also finger-mounted reflectors or structured light can be used. Since there is immediate visual feedback on the finger tracking and an opportunity for correction, the system should be useable without achieving 100 % accuracy on tracking. It is also easy to orient the camera towards a suitable, plain background, e.g., a tabletop.

If several cameras are used, the distance of the hand or finger is relatively easy to attain. Also e.g., inertial sensors can be employed. These options enable many possibilities for the user interface like manipulating objects in three dimensions, creating immersive 3D desktop views, or applications like hand- or finger-controlled games.

Also other tracking methods can be employed. They require some hand-held sensors or transmitters, which can be embedded into a small pointing stylus. As computer vision-based tracking is computationally heavy, the other generally known tracking methods can significantly simplify the tracking task.

Further, the pointing accuracy of the finger is limited by its size, and possibly by other reasons like a disability. More accurate pointing is possible, if a pointing stylus or some other tracked object is used. The tracked object can be hand-held, hand- or finger-mounted, or integrated or attached to a glove, wristwatch, ring, etc.

## 4. The implemented prototype

An early prototype of the UI device was implemented with modified low-cost, off-the-shelf components. The basic components were a PC, a Sony PLM-S700E HMD, and NaturalPoint™ [5] Smart-Nav™ AT hands free mouse, which was attached onto the HMD.

The near-eye microdisplay could be non-see-through or semitransparent, when augmented reality applications become possible. We used the HMD as a non-see-through display in our prototype.

The visual tracking of Smart-Nav uses four IR LEDs located behind the visually opaque front lens to illuminate the provided tiny IR reflectors (adhesive dots and reflective rings). A CMOS camera inside the Smart-Nav picks up the reflection and a chip processes the image data. Information is passed on to a computer via a USB port where it is translated by the Smart-Nav software into cursor movement. Non-relevant IR sources like reflections from sunlight are discarded.

The Smart-Nav system is meant to replace a mouse in a desktop PC environment and the camera is attached on top of a monitor. However, in our application, the camera is looking the opposite way, i.e., from the user to the world. Smart-Nav has settings for various camera positions, but the required mirror setting was not available.

We solved the problem by attaching a mirror and the Smart-Nav components suitably onto the HMD and using proper Smart-Nav settings to make it work properly for this application and to align field of views of the camera and the HMD. The field of views of the HMD and the camera in the Smart-Nav happen to be the same.

Figure 4 shows an image of the implemented early prototype without the protective, IR-transparent plastic covering. It is still much bigger than what it could be with better components. It employs a binocular display, but a monocular display would reduce the size significantly. The weight of the device is 135 grams.

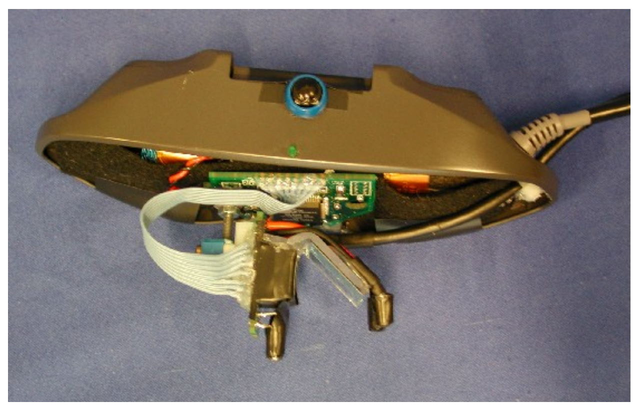

**Figure 4. An early prototype of MobiVR. A mirror and the reconfigured SmartNav components are in front of the HMD.**

The composed system works relatively well and responds fast to finger movements. Sometimes Smart-Nav lost the finger, especially if the reflectors were blocked or they exited from the field of view of the camera. Computer vision software of our own could have given better control over the system, but at this point it was not yet implemented.

The AT version of Smart-Nav features also dwell clicking, which enables to left click, right click, double click, and drag by simply holding the cursor still, or “dwelling” on an icon or button. The dwell click time interval can be set from 0.2 to 2 seconds. The feature could remove the need to click items, and thus can prevent accidentally moving the cursor point while clicking.

Figure 5 shows MobiVR in use. The implemented prototype is still a little clumsy, as it uses standard components not designed for the device. No real usability studies have yet been conducted with the prototype. We will put a video to http://www.cs.tut.fi/~ira/MobiVR.html in the near future to demonstrate the early prototype.

**Figure 5. MobiVR in use. The ring reflector can be seen on the pointing finger.**

Informal initial usability observations affirm that our system provides a means of interacting with a hand-held device in an intuitive and easy-to-use manner. It worked well and gave a good taste of the potential of the concept. It seems to fit to mobile or relatively short tasks like accessing data and giving commands. The ubiquitous computing aspect is not yet well presented in the early prototype.

## 5. Future work

A more advanced prototype will be made soon. With a little work the whole system could be built much smaller and truly mobile. The ubiquitous computing aspect needs further work.

It is possible to add another display to the other end of the device, as we did in the prototype. This, together with inertial sensors would enable stereoscopic views, novel user interfaces, and virtual reality applications in a hand-held appliance.

We could also try alternative tracking methods like ultrasonic tracking, inertial or acceleration sensors, and other options. For example, two cameras would provide the distance of the pointer, enabling 3D manipulation.

One electromechanical tracking method would be to employ three thin wires attached from the display unit to the pointing device, and to measure their outstretching. The method would provide the screen position of the pointing stylus, and also the distance. This is not as inconvenient as it may sound at first. The method effectively prevents the loss of the pointing stylus, and is cheap and easy to construct. In conjunction with a camera, one wire would give the distance for the position observed by a camera and could also provide power for the IR light on a stylus.

Finger tracking is needed only when the eye is near the display and the finger is visible for the camera. A sensor can be placed near the display's ocular to notice the eye proximity for switching the device on and off. This enables considerable power saving.

Also usability studies for the system would be very important and interesting to find out about task performance, user acceptance, and possible technical and usability shortfalls.

## 6. Conclusions

The mobile and ubiquitous trend in computing can fundamentally change the paradigms of the present user interfaces and the ways we interact with computers. In our work we have peeked into the mobile future by implementing an early prototype of a novel user interface. The presented UI concept for mobile computing derives largely from VR, but gives it a new form and context, which is more suitable for mobile computing.

It is interesting to note that cameras are already being embedded into mobile phones and soon also near-eye microdisplays are emerging. The physical infrastructure for the MobiVR concept will soon become a part of many appliances, imposing no extra hardware, weight, or cost for using MobiVR.

The implementation of the MobiVR concept is in an early stage, but the idea seems to be very interesting and relevant for mobile devices. It may be a viable solution to the real issue on how to have a decent sized display and an easy way of inputting data and checking information on a pocket-sized device.

The invention is fully compatible with WIMP, featuring also a smooth transition path to many post-WIMP user interfaces like augmented reality and virtual reality, and for special applications like hygiene virtual keyboards in hospital appliances etc. The MobiVR concept seems to be a natural, fast, and easy-to-use tool for mobile, hand-held computing.

## 7. Acknowledgments

Thanks to Mikko Asunta for skillfully constructing the prototype. The work has patents pending.

## 8. References

[1] Accot, J., Zhai, S., More than dotting the i's - Foundations for crossing-based interfaces. In *Proceedings of CHI 2002,* (Minneapolis, MI, April 2002), 73-80.

[2] Billinghurst, M., Bowskill, J., Dyer, N., Morphett, J., Spatial Information Displays on a Wearable Computer. IEEE Computer Graphics and Applications, Vol. 18, No. 6, Nov.-Dec. 1998, 24 –31.

[3] Kurata, T., Okuma, T., Kourogi, M., Sakaue, K., The Hand-mouse: A Human Interface Suitable for Augmented Reality Environments Enabled by Visual Wearables, In *Proceedings of ISMR'2001*, (Yokohama, Japan, March 2001), 188-189.

[4] Mann, S., Wearable computing: a first step toward personal imaging. *IEEE Computer* Vol. 30, No. 2, (Feb. '97), 25-32.

[5] NaturalPoint hands free mouse, http://www.naturalpoint.com/.

[6] Pulli, P., Antoniac, P., MARISIL - User Interface Framework for Future Mobile MediaPhone, In *Proceedings of Mobile Multimedia Systems & Applications* (Delft, Holland, November 2000), 1-5.

[7] Rakkolainen, I., Timmerheid, J., Vainio, T., A 3D City Info for Mobile Users. Proceedings of the 3rd International Workshop on Intelligent Interactive Assistance and Mobile Multimedia Computing (IMC'2000), November 9-10, 2000, Rostock, Germany. 115-121.

[8] Rakkolainen, I., Finnish pat. appl. 20011321, filed on 6/21/2001.

[9] Rekimoto, J., Matsushita, N., Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display. In *Workshop on PUI'97* (Banff, Alberta, Canada, October 1997), ACM Press, 30-32.

[10] Sibert, J., Gokturk, M., A finger-mounted, direct pointing device for mobile computing. In *Proceedings of UIST'97*, (Banff, Alberta, Canada, October 1997), ACM Press, 41-42.

[11] Siemens Virtual Mouse at CeBit 2003. http://www.my-siemens.com/MySiemens/CDA/Index/0,2730,HQ_en_0_cms%253A9270,FF.html.

[12] Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., Picard, R., Pentland, A., Augmented Reality Through Wearable Computing. Presence, special issue on Augmented Reality, Vol 6, No. 4, (Fall 1997), 386-398.

[13] Starner, T., Auxier, J., Ashbrook, D., Gandy, M., The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. In *Proceedings of ISWC 2000* (Atlanta, GA, October 2000), 87-94.

[14] Weiser, M., The Computer for the 21st Century. Scientific American, Vol. 265, No. 3, 94–104.