

Smart Laser-Scanner for 3D Human-Machine Interface

Álvaro Cassinelli

Alvaro_Cassinelli@ipc.i.u-tokyo.ac.jp

Stéphane Perrin

Stephane_Perrin@ipc.i.u-tokyo.ac.jp

Masatoshi Ishikawa

Masatoshi_Ishikawa@ipc.i.u-tokyo.ac.jp

University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Abstract

The problem of tracking hands and fingers on natural scenes has received much attention using *passive* acquisition vision systems and computationally intense image processing. We are currently studying a simple *active* tracking system using a laser diode, steering mirrors, and a single non-imaging detector, which is capable of acquiring three dimensional coordinates in real time without the need of any image processing at all. Essentially, it is a smart rangefinder scanner that instead of continuously scanning over the full field of view restricts its scanning area, on the basis of a real-time analysis of the backscattered signal, to a very narrow window precisely the size of the target. The complexity of the whole setup is equivalent to that of a portable laser-based barcode reader, making the system compatible with wearable computers.

Categories & Subject Descriptors: H.5.2 [Information interfaces and presentation]: User Interfaces – *Input Devices and Strategies, Interaction Styles*; H.1.2 [Models and Principles]: User/Machine Systems – *Human Factors*.

General Terms: Performance, Algorithms.

Keywords: Active Tracking, 3D Acquisition, Smart Laser Scanner, Device-less Computer Interaction.

INTRODUCTION

Alternatives to the keyboard/mouse are necessary on handheld computers [1]. On PDAs, input of text is often done through a touch sensitive screen and a stylus using a prescribed input method, such as Graffiti™. Input and viewing space are thus merged, allowing a substantial saving of physical space. The next logical step would be to remove the need for any (dedicated or merged) input space, as well as the need for any additional input device (stylus, data-gloves [3], etc). This would allow inputting data by just executing bare-handed gestures in front of a portable device - that could then be embodied in a keyboard-less wrist-watch for example. Ultimately, coupling such interface with a projective (screen-less) display would completely remove the need of any physical interaction area (for viewing or inputting), allowing further downsizing of mobile computing devices. The system

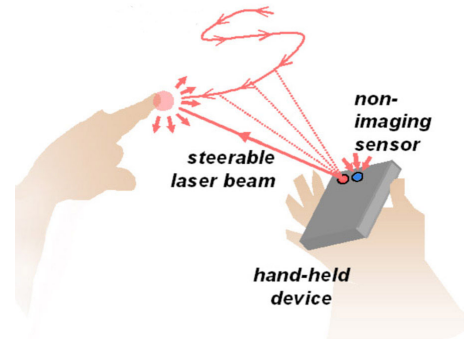


Figure 1. Active tracking using “smart” scanning.

under study is depicted in Figure 1; it is based on a wide-angle photo-detector and a collimated laser beam generated by a laser diode and steered by means of a two-axis micro-mirror.

TRACKING PRINCIPLE

Tracking is based on the temporal analysis of the backscattered light measured during a rapid circular laser “saccade” generated around the presumed position of the tracked object. Figure 2 shows a simplified, step-by-step tracking sequence: (a) The saccade is supposed to remain fully inside the object while tracking; (b) As the object moves, a small portion of the saccade may fall outside the object and the backscattered signal will momentarily drop. Due to the synchronous operation of the beam-steering mirrors and the photo-detection, an accurate re-centering vector (white arrow) is computed; (c) The center of the saccade is updated accordingly. These three steps are repeated continuously (each circular scan takes about two milliseconds). Adaptive saccade shapes are also under study (these would change if required so as to match important features of the tracked object).

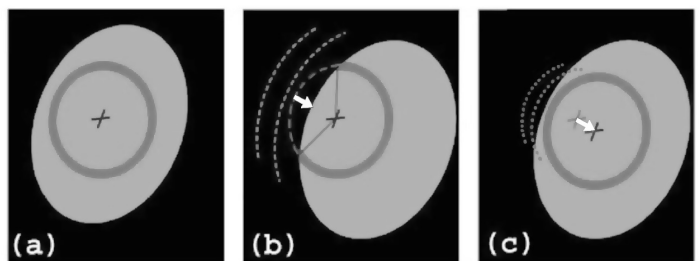


Figure 2. Tracking principle

SYSTEM PERFORMANCE

Using this simple closed-loop mechanism, a tracking resolution of about 0.2° (i.e. 3 mm at a distance of one meter) was achieved in a prototype using off-the-shelf electromagnetic-actuated micro-mirrors producing a field of view of $\pm 25^\circ$ [2]. Depth is computed at the end of each saccade from the averaged intensity of the backscattered signal - which evolves roughly as the inverse square of the distance. Discrimination between the tracked object (e.g. fingertip) and the background (including other parts of the hand) is possible as long as the signals present enough contrast. By modeling the various sources of noise, we estimated and later verified that, without resorting to synchronous photo-detection nor any signal post-processing, the prototype is already capable of sub-centimeter depth discrimination at a distance of up to 30 cm (with 90% confidence), and has a peak resolution of less than 4 mm at a distance of 5 cm from the system. Although such performance would satisfy the requirements of a 3D interface for hand-held devices, we recently integrated a lock-in amplifier that enables the system to work under normal illumination conditions, while at the same time extending its operation range up to several meters. Maximum tracking speed was measured to be around 3 m/s at a distance of about one meter from the mirrors (the typical speed of a finger performing gestures is < 2.5 m/s), and is only limited by the relatively slow A/D interface cards. Figure 3 shows the real-time tracking of a ping-pong bouncing ball.

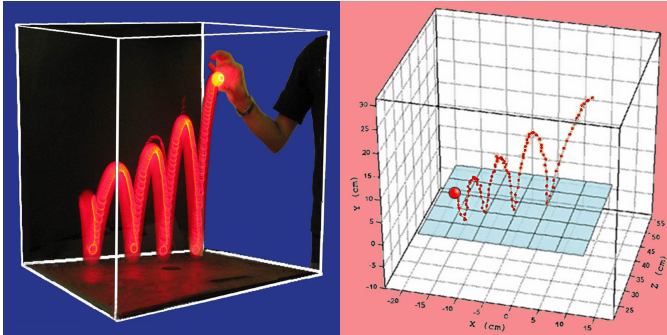


Figure 3. Real-time tracking and 3D acquisition.

THE SYSTEM AT WORK

Simultaneous tracking of two or more fingers without replicating any part of the system is also possible (see Figure 4). Despite the lack of tactile feed-back, virtual reality applications for this marker-less, multi-target tracking system are countless: the system would allow, for instance, multiple users to interact on the same virtual space; or a single user to control several virtual tools at the same time, resize windows and control information screens as imagined in Spielberg's "Minority Report" film, but without the need to wear special gloves nor markers. Several proof-of principle applications are being developed

for a hands-on demonstration, such as zooming and scrolling on a map using real-time 3D tracking of a single finger, drawing, and manipulation of virtual 3D models using simultaneous tracking of two or more fingertips.

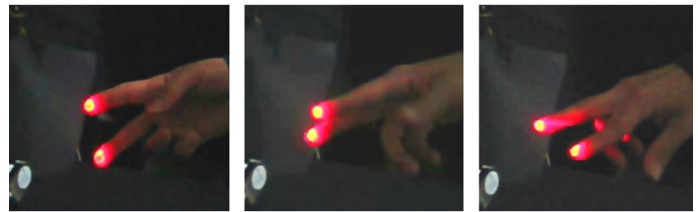


Fig. 4: simultaneous tracking of (bare!) fingertips.

CONCLUSION

An active tracking system capable of tracking hands and fingertips without the need for reflective markers is proposed and demonstrated. The system has some appealing features over other conventional (passive or active) tracking systems, namely:

- **Few constraints on the user/environment.** The system does not require the user to carry any special device (such as active transceivers or reflective optical markers); it does not require any special projection area (other than the hand/finger itself); it is an active-illumination system, so it does not impose stringent illumination conditions.
- **Real-time, non-imaging full 3D acquisition.** Tracking and three-dimensional data acquisition is achieved without stereoscopic cameras. No image processing required.
- **System on-chip integration.** Its hardware simplicity is such that, using state-of-the-art Micro-Opto-Electro-Mechanical-System (MOEMS), it should be possible to integrate the whole system on a single chip, making a versatile human-machine input interface for use in mobile computing devices.

Additional information can be found online [4].

REFERENCES

1. Bowman, D. A. et al. *3D User interfaces: theory and practice.*, Addison-Wesley (2005).
2. Perrin, S., et al, "Laser-Based Finger Tracking System Suitable for MOEMS Integration", Proc. Image and Vision Comput. (IVCNZ), pp.131-136, (2003).
3. Sturman, D. J., et al, "A Survey of Glove-based input", IEEE Computer Graphics and Applications, Vol.14, No.1, pp.30-39, (1994).
4. <www.k2.t.u-tokyo.ac.jp/fusion/LaserActiveTracking>