Markerless Laser-based Tracking for Real-Time 3D Gesture Acquisition

Alvaro Cassinelli*
University of Tokyo, Japan.

Stephane Perrin
University of Tokyo, Japan.

Masatoshi Ishikawa
University of Tokyo, Japan.

1 Introduction

We describe here our latest work on a prototype finger-tracking system based on a wide-angle photodetector and an eye-safe laser diode whose beam is steered by a pair of micro-mirrors. 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.

2 System description and Results

Tracking is based on the temporal analysis of the backscattered light measured during a millisecond-long circular laser saccade generated around the current position of the tracked object. This saccade is supposed to remain fully inside the object while tracking; 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 photodetection, both the width and the position of that portion can be precisely determined by a computer, and an accurate translation vector computed and used to re-center the saccade back inside the object. Using this simple closed-loop mechanism, a tracking resolution of about 0.2° (or 3 mm at a distance of one meter) was achieved in a prototype tracking system using off-the-shelf electromagnetic-actuated micro-mirrors (tracking range $\pm 25^\circ$)[Perrin et al. 2003]. 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 modelling the various sources of noise, we estimated and later verified that, without resorting to synchronous photodetection 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. 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 was only limited by the speed of the A/D interface cards [Perrin et al. 2004]. Fig.1 demonstrates real-time tracking of a bouncing ball. Simultaneous tracking of two or more targets without replicating any part of the system was recently demonstrated (see Fig.2), enabling parallel input from multiple users, as well as simultaneous tracking of diverse body parts, manipulation of objects in virtual space, and resizing windows and control information screens, as imagined in Spielberg's film "Minority Report" - but without the need to wear special gloves.

Figure 1: Real-time tracking of a pingpong ball.

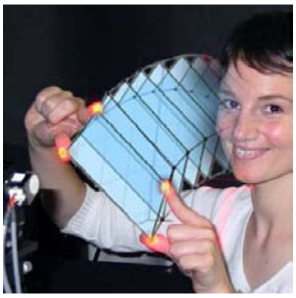

Figure 2: Simultaneous acquisition of multiple unmarked targets.

3 Conclusion

The proposed active-tracking mechanism has some appealing features over other conventional (passive or active) tracking systems: First, it does not require the user to carry any special device (such as active transceivers or reflective optical markers). Second, although depth estimation of distant targets may be difficult (full 3D measurement is limited to about 160 cm in the present configuration), tracking and three-dimensional data acquisition is achieved without the need for a stereoscopic camera - in fact, no imaging device is needed at all. Lastly, its hardware simplicity is such that, using state-of-the-art Micro-Opto-Electro-Mechanical-System (MOEMS) technology, it should be possible to integrate the whole system on a single chip, making a versatile human-machine gestural input interface for use in mobile computing devices. Additional information can be found at www.k2.t.u-tokyo.ac.jp/fusion/LaserActiveTracking/index-e.html.

References

PERRIN, S., CASSINELLI, A., AND ISHIKAWA, M. 2003. Laser-based finger tracking system suitable for MOEMS integration. *Image and Vision Computing conference, New Zealand (IVCNZ 2003), Massey Univ.* (November), 131–136.

PERRIN, S., CASSINELLI, A., AND ISHIKAWA, M. 2004. Gesture recognition using laser-based tracking system. *6th International Conference on Automatic Face and Gesture Recognition 2004 (FG 2004), Seoul, Korea* (17-19 May), 541–546.

*e-mail: alvaro@k2.t.u-tokyo.ac.jp, Web: www.k2.t.u-tokyo.ac.jp