

Laser-Based Finger Tracking System Suitable for MOEMS Integration

Stéphane Perrin, Alvaro Cassinelli, Masatoshi Ishikawa
University of Tokyo, Ishikawa Hashimoto Laboratory
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
alvaro,sperrin,ishikawa@k2.t.u-tokyo.ac.jp

Abstract

This paper describes a proof-of-principle demonstration of a novel and simple active tracking mechanism using a laser diode, steering mirrors, and a single non-imaging photodetector. Tracking is based on the analysis of a temporal signal corresponding to the amount of backscattered light produced during a rapid, local circular scan (or *saccade*) around the presumed object position. The simplicity of the system is such that, using state-of-the-art Micro-Opto-Electro-Mechanical-System (MOEMS) technology, it would be possible to integrate the whole system on a chip, making it an interesting input interface for portable computing devices.

Keywords: Active Tracking, Human Machine Interface, MOEMS

1 Introduction

Vision systems capable of active illumination (i.e. capable of active control of an illumination source) have a number of advantages that may outweigh their drawbacks of their relatively costly or delicate opto-mechanical system. In particular, by exploiting synchronous detection and wavelength filtering to enhance the signal-to-noise ratio, laser-based active illumination vision systems are fairly insensitive to harsh or changing lighting conditions [1], [2]. Moreover, in certain applications active illumination can free the system from extensive image processing, and should be considered whenever real-time response times are needed (millisecond range). Examples include extremely simple magnetic levitation feedback control [3], collision avoidance and ranging mechanisms [4], [5], and more refined yet efficient tracking devices for macroscopic [6], [1] or even microscopic applications [7]. Tracking is a particularly interesting case where a laser-based active-vision system may easily outperform a passive vision system. Since the target is likely to represent a small portion of the total field of view, the illuminating energy can be efficiently concentrated around it, and therefore the system will be extremely power-efficient and work equally well with distant targets. Active illumination can be used to enhance robustness of an otherwise passive tracking system based on extensive image processing (correlation, snakes, etc. [8]). However, in certain cases it may be possible to use a simple non-imaging sensor, thus obviating the need for any image processing [9], [5]. This will allow faster response time, reduced cost, and enhanced robustness and compactness of the system.

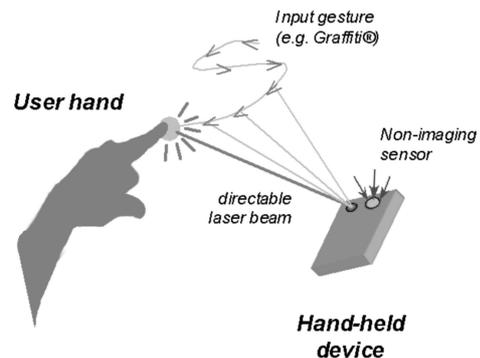


Figure 1: The proposed active tracking system as a human-machine interface for hand-held devices.

Based on these considerations, a very simple active tracking system and algorithm, suitable for integration on a single chip as a Micro-Opto-Electro-Mechanical System (MOEMS) [10], is introduced here. The system depicted in Fig.1 is based on a wide-angle photodetector and a collimated laser beam generated by a laser diode and steered by means of a two-axis micro-mirror. The complexity of the hardware setup is equivalent to that of a portable laser-based barcode reader. It is interesting to note that this tracking system does not require the user to hold any special device. The proposed system can be conceived as a "smart" rangefinder scanner that instead of continuously scanning over the full field of view, would, based on a real-time analysis of the signal, restrict its scanning area over a narrow window around the target.

This paper is organized as follows. Section 2 gives a thorough description of the tracking algorithm, identifying relevant parameters. A demonstrator

setup is introduced in section 3, and its expected performance is compared with the experimental results. Section 4 discusses possible software and hardware improvements and integration issues. In the conclusion section, the most relevant results are summarized and future research directions outlined.

2 Active Tracking

2.1 Principle

Tracking is based on the analysis of a temporal signal corresponding to the amount of backscattered light measured during a laser *saccade*, i.e. a rapid laser scan generated in the neighborhood of the tracked object (see Fig.2). While being tracked, the object continuously backscatters some laser light (Fig.2.a). When the object moves out of the tracking region, the backscattered signal is lost and tracking fails (Fig.2.b). The system then generates a local scanning saccade (Fig.2.c), and re-centers the laser over the new position producing backscattering (Fig.2.d). If this process is repeated rapidly enough, the object will always be within the reach of a small saccade.

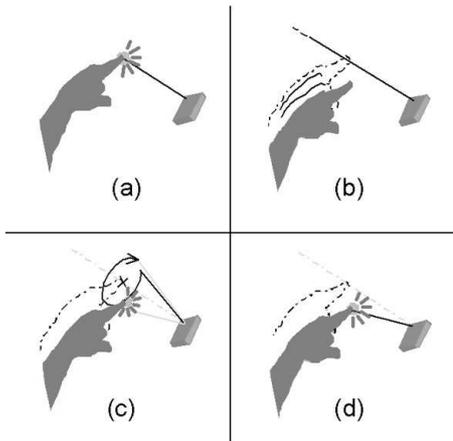


Figure 2: Principle of the non-imaging tracking system based on a laser saccade.

A continuously generated saccade whose trajectory falls fully inside the object surface (Fig.3.a) can be used to obtain a more sensitive tracking: as the object moves, a relatively small portion of the saccade will fall outside the object surface and the backscattered signal will momentarily drop (Fig.3.b). Due to the synchronous operation of the beam-steering mirrors and the photodetection, both the angular width and the relative position of that portion can be determined by the computer. Using such information, an accurate translation vector is derived and used to re-center the saccade back inside the object again, even for small displacements of the object. A circular saccade was selected because this trajectory is easy to generate and has good symmetry properties that translate into reduced algorithm complexity.

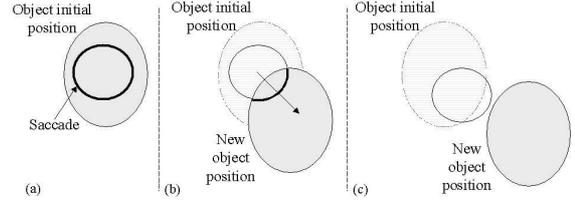


Figure 3: Principle of active tracking based on a continuously generated saccade.

2.2 Algorithm and Parameters

2.2.1 Tracking Method

In our present system the circular saccade is composed of a discrete set of N regularly distributed points. Once the whole saccade has been completed and the signal from the photodetector properly thresholded, a binary signal results that tells, for each point of the saccade, whether or not the tracked object was in the path of the laser beam. Using this binary signal, the new position of the object can be computed and the center of the next saccade moved accordingly.

The configurations considered are all shown in Fig.3. The configuration in Fig.3.a is when the saccade is completely within the object surface, and Fig.3.c when it is completely outside the object. In both these configurations, the next saccade is started at the same position because either the position of the object is known (Fig.3.a), or else the tracking phase has not started or it has failed (Fig.3.c). The configuration in Fig.3.b occurs when the saccade is only partly inside the object surface. The center of the saccade is translated toward the portion of the saccade that is on the object surface. Because in this study tracking is limited to objects of convex shape whose movement is considered uniform and rectilinear during the time needed for the completion of a saccade, in principle, no configurations other than those shown in Fig.3 should occur.

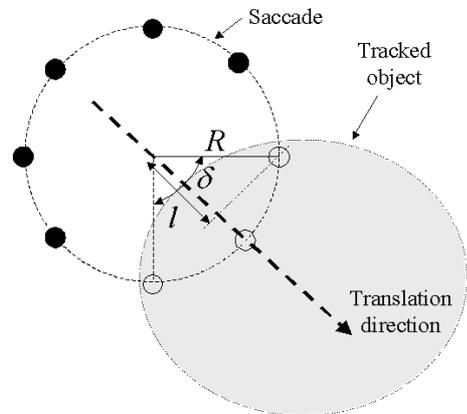


Figure 4: Computation of the translation vector.

The argument (direction) and the norm (distance) of the translation vector are computed as follows. The

direction of translation corresponds to the bisectrix of the portion of the saccade that is inside the object surface. The translation distance depends on the number of points N_L ($1 \leq N_L < N$) composing the portion of the saccade completely inside the object surface. The translation should result in a saccade that will be fully completed on the object surface. The angular width of the portion of the saccade inside the object surface is $\delta = 2\pi \frac{N_L - 1}{N}$ (see Fig.4). If $\delta \leq \pi$, the translation norm d should be equal to the radius of the saccade (indicated by R) plus the distance $l = R \cdot \cos(\frac{\delta}{2})$, that is $d = R + l$ (see Fig.4). If $\delta > \pi$, the translation norm should be $d = R - l$.

2.2.2 Parameters

Two parameters of the tracking algorithm have to be set in order to actually track the object, namely the radius R and the number of points N of the saccade.

The number of points N that constitute the saccade determines the time needed to complete it. Let t_s be the time needed for completing a saccade; then $t_s = N \cdot t_p$, where t_p is the time elapsed between two points. To be tracked by the system, an object should not move more than the saccade diameter ($2R$) during the time elapsed between two successive saccades (t_s). That is, the upper limit for the tracking speed V_l is given by equation 1.

$$V_l = \frac{2R}{t_s} = \frac{2R}{N \cdot t_p} \quad (1)$$

In order to obtain as high a tracking speed as possible, the radius R should be as large as possible or the number of points N should be as small as possible, assuming a constant value for t_p (relation 1). However, the fewer the saccade points, the less accurate the computed direction (see Fig.5). That is, a good compromise must be found between the speed and the precision.

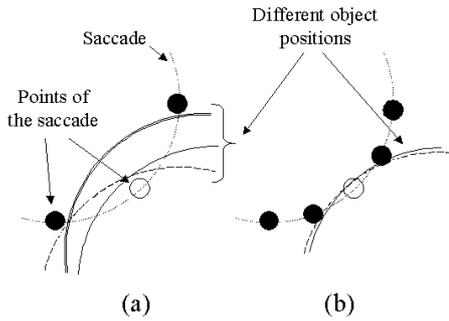


Figure 5: Examples of object positions leading to the same computed translation vector depending on the number of points N .

The radius R of the saccade is fixed and defined on the object surface at a given distance. In order to be completely on the object, the saccade should be smaller than the smallest dimension of the object. Nonetheless the radius should not be too small as this would

result in a decrease in the maximum attainable tracking speed (equation 1), or would allow the object to move freely without actually being tracked. It follows that the optimal size of the saccade should match the smallest dimension of the object. In fact, R should be smaller than its optimal value in order to allow some variation in the distance of the object from the tracking system (which is less constraining for the user).

3 Experimental Demonstration

3.1 System Description

As explained above, the proposed tracking system is composed of a laser source, a wide-angle photodetector, and a two-axis beam-steering mechanism.

3.1.1 Laser Source

The laser source used was a focusable ($20 \text{ mm}/\infty$) diode laser delivering a maximum optical power of 4.5 mW at a peak wavelength of 633 nm (Class IIIa). Visibility of the source greatly simplifies the calibration of the system, but in a final working system using the same optical power, a more eye-safe far infrared ($>1400 \text{ nm}$) diode laser may be preferred. The power delivered by the laser through the beam-steering mirrors was fixed to a maximum of 1 mW, enough to produce a good signal-to-noise ratio for objects as far as about 70 cm (this corresponds to a Class II laser source, similar to that of a barcode scanner, and in normal conditions does not represent a hazard to the eye). As a general rule of thumb, the necessary optical power scales as the square of the working distance, since the backscattered signal is assumed to originate by strong diffuse reflection at the object surface (see below). Therefore, for distances less than 10 cm (typical of portable devices) the required optical power would drop below $30 \mu\text{W}$. To obtain a good angular resolution during the saccade, it is important to properly focus the laser on the object surface. In principle, automatic focusing is possible (as in a pick-up optical head), but a collimated narrow beam may be a more cost-effective solution. This configuration worked well in the experiments. However, by using a beam focused at a narrow distance (less than 10 cm), the eye hazard can be further reduced and thus more optical power can be used to enhance the signal-to-noise ratio.

3.1.2 Light Detection

In the present setup, a wide-angle photodetector placed in the neighborhood of the beam-steering mechanism collects all the light from its surroundings. The backscattering signal was characterized as a function of the distance for several fairly isotropic scattering surfaces, including a piece of white paper, brown cardboard, and the skin of a fingertip (with

measured reflectivities of roughly 80%, 70%, and 60%, respectively). Robust tracking is possible as long as the contrast (defined as the ratio between the backscattered signal from the tracked object and the background signal) remains high. Although a wavelength selective filter is placed in front of the photodetector, in a dimly lit room where illumination fluctuations reach several tens of nanowatts, the working distance seems limited to about 70 cm using a 1 mW laser source. Presently, an object/background discrimination threshold is set manually, depending on the lighting conditions, the tracking distance, and the reflectivity of the selected surface. However, automatic calibration of such a threshold should be straightforward.

3.1.3 Beam-steering Mechanism

The beam-steering mechanism represents the critical part of the system, determining both the ultimate tracking performance and the compactness of the whole setup. In this proof-of-principle experiment, a pair of high-performance closed-loop galvano-mirrors with axes perpendicular to each other (GSI Lumonics, Model V500) are used for both generating the saccade and performing the actual tracking. The galvano-mirrors have a typical clear aperture of 5 mm and a maximum optical scan angle of $\pm 25\%$. In the experiments, the saccade diameter was about one centimeter, and the working distance about 70 cm, which corresponds to a scanning angle of 1° and a sinusoidal input command amplitude of 100 mV. For amplitudes less than two or three hundred millivolts, it was verified that the angular response of the galvano-mirrors remained linear up to at least 1 kHz. For higher frequencies, the output progressively degenerated into a triangular signal, and above 3.5 kHz the system behaved chaotically.

The actual mirror command is not exactly sinusoidal though: since the saccade is composed of a discrete set of sampling points, the galvano-mirrors work in stepped-scan mode, meaning that they move rapidly from one position to another. (The amount of backscattered signal is measured only at the end of each step.) In order to generate a circular saccade, two sampled sinusoidal signals in phase quadrature are sent respectively to the horizontal and vertical galvano-mirrors by the computer at the maximum rate allowable by the program execution and the hardware interface. The program execution speed is negligible (around $1 \mu\text{s}$ on a Pentium III), but the present configuration of the input/output interface boards limits the sampling rate to 5 kHz; in other words, the sample interval t_p is 0.2 ms.

There is a significant delay between the command and the actual angular response, which leads to a systematic error while estimating the argument of the translation vector. This angular shift can be easily corrected by

software. This simple strategy worked well in the experiments. If more precision were needed (or variable delays expected to take place in the system), it is always possible to electronically read the absolute position of the mirrors and then determine the relative translation vector.

3.2 Expected Results

Fig.6 shows the values of the maximum attainable tracking speed (given by equation 1) as a function of the number of points N composing the saccade and for two values of the saccade radius R . It was found that for values of R larger than two thirds of R_{obj} , the tracking performance degraded drastically. It is interesting to note that in the present system the

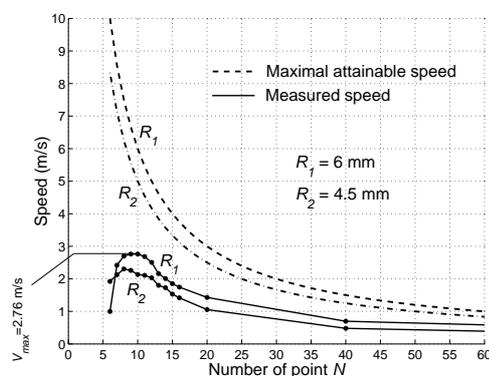


Figure 6: Maximum theoretical and measured speeds as a function of the number of points N for two values of R .

angular speed of the mirrors is not a limiting factor; therefore, for an object of a given size and assuming that the saccade radius is always properly adjusted to it, its distance from the tracking system has no effect on the tracking speed. In fact, the relatively slow interface boards are responsible for the limited speed performance of the present system. If instead, the angular speed of the mirrors were the limiting factor, the maximum tracking speed would decrease as the object gets closer to the tracking system.

3.3 Experimental Results

The evaluation of the performance of the system was done by measuring the maximum speed an object can move without being lost by the tracking system. The tracked object was a circular piece of white paper, $R_{obj} = 18 \text{ mm}$ in diameter, following a circular trajectory at different uniform speeds. The distance between the mirrors and the object remained constant. The absolute measured speed is shown in Fig.6. The maximum experimental tracking speed $V_{max} = 2.76 \text{ m/s}$ was measured for $R = \frac{2}{3}R_{obj}$ and $N = 9$ and $N = 10$.

As expected (equation 1), the measured tracking speed was a decreasing function of the number of points and an increasing function of the radius. But when N is smaller than 8 points ($R = 4.5$ mm) or 9 points ($R = 6$ mm), the speed that the system could reach decreased because the tracking precision became critically low (see Fig.5.a). When the new position was quite different from the real position of the object, the saccade was translated in the wrong direction and the subsequent saccade was eventually completely outside the object (the tracking failed). Presumably, this phenomenon is also the main reason for the significant mismatch (around 50%) between the measured and the maximum attainable theoretical speeds for values of N larger than 10.

The experiments showed another phenomenon. Since the object cannot be expected to stop while the saccade is being performed, the portion of the saccade inside the object surface is slightly different from the one that would have been obtained if the saccade was generated instantaneously. This accounts for a systematic error in the computed translation vector which depends on the speed of the object. Such an error is not noticeable for small speeds, but tends to affect the tracking robustness for larger speeds. Software correction should be possible by estimating the average object speed. The speed of a natural hand gesture was measured to be less than 2.5 m/s. Therefore, the present system was able to track a finger tip (between 1.5 cm and 2 cm in size), as shown in Fig.7.



Figure 7: Tracking of a finger (one second duration)

4 Future Work

The experiments were carried out in a controlled environment (uniform background and stable lighting conditions). However, it is likely that simple algorithmic improvements will bring enough robustness to the tracking system to allow it to function in real environments without having to introduce any modification to the present optical setup. Automatic calibration and memorization of the object-to-background discrimination threshold will

also allow robust tracking when the contrast varies or when multiple bodies are in the field of view. Filtering techniques (e.g. Kalman filters) can be used to estimate the object speed, thus reducing the speed-dependent systematic error on the translation vector as well as treating momentary occlusion of the tracked object. Lastly, modelling kinematic constraints can enhance tracking robustness for real objects.

In order to improve the tracking speed, a first modification to the present hardware would consist of replacing the relatively slow interface boards. Indeed, the galvano-mirror operating frequency can be pushed well above 1 kHz, despite some attenuation and deformation of the output response. If we follow the manufacturer's specifications, saccade frequencies up to 3.5 kHz should be possible before the galvano-mirror response becomes unusable. This would correspond to a tracking speed of about 35 m/s for a finger-sized object at a distance of one meter away.

As shown in the experiments, the tracking robustness depends strongly on the relative size of the saccade with respect to that of the tracked object. As the object moves in front of the tracking system, such relative size may vary and depart from its optimal value. Therefore, estimating the distance from the system to the tracked object may be essential for performing dynamic optimal fitting of the saccade radius. Under well-controlled conditions, the mean intensity of the backscattered signal may be enough to determine the actual distance of the object. However, for distant objects, resorting to more sophisticated telemetry techniques may be necessary. Even simple electronic time-of-flight measurements can give sufficient resolution depth for the present purposes [6], and may also allow the possibility to perform tracking in three dimensions.

The tracking robustness also depends on the signal-to-noise ratio. A simple way to improve the contrast is to perform synchronous detection by modulating the laser source. Cheap commercial range-finders using such techniques are able to function at distances up to several hundreds of meters using Class II lasers. A slightly more sophisticated way is to arrange the optical setup in a "pick-up head" configuration by means of a beamsplitter. That way, the photodetector will only integrate the light coming from a narrow window centered on the target, discarding spurious signals from the rest of the field of view. This configuration would allow for longer tracking distances using safe optical powers, while gaining at the same time robustness in regards to lighting conditions and background characteristics (interesting applications include tracking a limb at a distance of several meters [11], vehicles [4], etc.).

The scaling down of the whole setup is presently limited by the bulky nature of the galvano-mirror pair (roughly $10 \times 10 \times 10$ cm³). One or two-axis

micro-electro-mechanical micromirrors (packaged on a chip 1 cm^2 in size) are already commercially available. Simple resonant MEMS mirrors are ideal for fast saccade scans in the millisecond range and above (see for instance [12]); however, good absolute pointing accuracy can only be achieved with a more refined MEMS design [13]. Therefore, using a pair of micromirrors can be necessary to get the best of a MOEMS based system [14].

Compared with macro-scale systems, MEMS are smaller, lighter and more rugged [15]. The extreme simplicity of the proposed tracking system combined with the recent advances in MEMS makes the proposed system a good candidate for monolithic integration on a silicon chip. Very similar MOEMS have been successfully built in the past [16], and challenges related to the integration of an optical source along with its beam-steering mechanism have been addressed both theoretically and experimentally in fields as diverse as free-space optical switching [13], single-chip scanners [16], and automatically aligned free-space optical links [17].

5 Conclusion

A simple non-imaging tracking system was proposed as a human-machine interface and successfully demonstrated. In the current configuration, the maximum tracking speed was limited to about 3 m/s at a maximum working distance of 70 cm, which is enough to track natural hand gestures.

Further research will be conducted towards a complete MOEMS integration of the tracking system as well as a thorough analysis and improvement of the saccade based tracking algorithm. Development of an appropriate software platform suited to a hand-held integrated version of the tracking system will also be a topic of interest (e.g. Graffiti characters and gesture recognition [18]).

References

- [1] F. Blais et al. The NRC 3d laser tracking system: IIT's contribution to international space station project. *Proc. 2001 Workshop Italy-Canada on 3D Digital Imaging and Modeling Application of : Heritage, Industry, Medicine and Land*, 2001.
- [2] R. L. Allwood and S. Tetlow. Improvements of underwater viewing range by the use of a scanning laser-based illuminator. *IEE Conf. on Electronics in Oceanography*, June 1997.
- [3] J. Boehm et al. Sensors for magnetic bearings. *IEEE Trans. Magnetics*, 29(6):2962–2964, Nov. 1993.
- [4] D. William. Advanced laser-based tracking device for motor vehicle lane position monitoring and steering assistance. *Collision Avoidance and Automated Traffic Management Sensors*, pages 128–137, 1995.
- [5] H.R. Everett. *Sensors for Mobile Robots: Theory and Application*. Wellesley, MA, June 1995.
- [6] J. Strickon and J. Paradiso. Tracking hands above large interactive surfaces with a low-cost scanning laser rangefinder. *ACM CHI'98 Conf.*, Apr. 1998.
- [7] J. Enderlein. Tracking of fluorescent molecules diffusing with membranes. *Appl. Phys.*, B 71:773–777, 2000.
- [8] A. Blake and A. Yuille. *Active Vision*. The MIT Press, Cambridge, MA, 1992.
- [9] T. Marrin and J. Paradiso. The digital baton: A versatile performance instrument. *Proc. 1997 Int'l Computer Music Conf.*, 1997.
- [10] S. J. Walker and D. J. Nagel. Optics and MEMS. *USA Naval Research Lab., Memorandum Report*, May 1999.
- [11] J.A. Paradiso and F. Sparacino. Optical tracking for music and dance performance. *4th Conf. on Optical 3D Measurement Techniques*, Sept. 1997.
- [12] P.R. Patterson et al. A scanning micromirror with angular comb drive actuation. *15th IEEE Int'l Conf. on MEMS, MEMS2002*, Jan. 2002.
- [13] V.A. Aksyuk et al. Lucent microstar micromirror array technology for large optical crossconnects. *Proc. SPIE, MOEMS and Miniaturized Systems*, 4178:320–324, Aug. 2000.
- [14] P.M. Hagelin and O. Solgaard. Optical raster-scanning displays based on surface-micromachined polysilicon mirrors. *IEEE J. Selected Topics in Quantum Electronics*, 5(1):67–74, 1999.
- [15] M.C. Wu et al. Optical MEMS: Huge possibilities for lilliputian-sized devices. *Optics and Photonics News*, pages 25–29, June 1998.
- [16] M.H. Kiang et al. Micromachined polysilicon microscanners for barcode readers. *IEEE Photonics Technology Letters*, 8(12):1707–1709, 1996.
- [17] J. M. Kahn et al. Emerging challenges: Mobile networking for smart dust. *J. Commun. and Networks*, 2(3):188–196, Sept. 2000.
- [18] S. Perrin and M. Ishikawa. Quantized features for gesture recognition using high speed vision camera. *SIBGRAPI 2003, Sao Carlos, Brazil*, Oct. 2003, accepted.