LinLED: Low latency and accurate contactless gesture interaction

Stephane Viollet, Martin Chauvet, Jean-Marc Ingargiola

LinLED: Low latency and accurate contactless gesture interaction

Stéphane Viollet**
Aix-Marseille Univ., CNRS, ISM
Marseille, France
stephane.viollet@cnrs.fr

Jean-Marc Ingargiola
Aix-Marseille Univ., CNRS, ISM
Marseille, France
jean-marc.ingargiola@univ-amu.fr

Martin Chauvet
Aix-Marseille Univ., CNRS, ISM
Marseille, France
martin.chauvet@orange.fr

Figure 1: The patented LinLED interface composed of 32 near infrared photodiodes and 32 LEDs arranged along a 1D array. A finger or a hand moving in the air above LinLED at a height up to 40 cm can be located with an an accuracy as small as 1 mm and a latency as fast as 1 ms.

ABSTRACT

An innovative gesture interface called LinLED is described. Traditional interfaces often suffer from limited tracking range and significant latency when locating fingers or hands. However, LinLED presents a game-changing solution. LinLED comprises a 1D-array of photodiodes and offers great precision in locating hands or fingers moving laterally in the air. Its accuracy is as small as 1 mm, which is ten times smaller than the photodiode pitch. Moreover, the latency is as fast as 1 ms, all thanks to its pure analog processing. LinLED can accurately detect any object reflecting infrared within a range of approximately 30 cm by 40 cm. Additionally, LinLED is sensitive to vertical motion (Z-axis motion), enabling the detection of three different types of gestures: selection, swipe, and graded smooth lateral movement. LinLED represents a significant advancement in gesture interfaces, offering high precision, minimal latency, and the ability to detect various gestures effectively.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; Robotics.

KEYWORDS

gesture, optical, fusion, infrared, interface

ACM Reference Format:
Stéphane Viollet, Jean-Marc Ingargiola, and Martin Chauvet. 2023. LinLED: Low latency and accurate contactless gesture interaction. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23 Companion), October 9–13, 2023, Paris, France*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610661.3617164

1 INTRODUCTION

In today's technological landscape, user interfaces have undergone significant development to keep pace with advancements in computer hardware and software. A notable direction in these advancements involves gesture detection, where users can interact with machines through direct interpretation of hand movements without physical constraints [7–9, 11–13, 15]. The LinLED interface represents an alternative solution that enables the detection of objects like the user's finger, hand, or other body parts, all without the need for any physical device [5], surface contact, or being within a camera's field of view [1]. This interface achieves absolute position detection with great precision, surpassing pixel pitch limitations, and operating within a large tracking space. The inspiration for developing the LinLED interface stemmed from studying the electric field signal processing in electric fish [4]. Despite the relatively coarse resolution of its receptors along the body, these animals feature hyperacuity, a phenomenon also observed in visual

*SV, MC, J-MI made the hardware design. MC fabricated the LinLED and developed the software. MC performed the experiments. SV made the model and wrote the paper.

ICMI '23 Companion, October 9–13, 2023, Paris, France

ACM ISBN 979-8-4007-0321-8/23/10…$15.00
https://doi.org/10.1145/3610661.3617164

systems [14]. This hyperacuity in LinLED is achieved through the use of a 1D array of near-infrared LEDs and the weighted sum of Gaussian overlapping curves. As demonstrated by previous research [2], this analog signal processing technique results in highly linear and smooth responses. Overall, the LinLED interface represents a cutting-edge innovation in gesture detection, offering an unparalleled level of accuracy and fluidity without the need for any physical constraints, making it a promising solution for various applications, especially in the automotive field (e.g., multimedia interface).

2 THE LINLED INTEFACE

2.1 A hyperacute optical position sensing device

As depicted in figure 1, the LinLed interface allows to locate a finger translating laterally over an 1D array composed of infrared LEDs and photodiodes. The principle relies on the implementation of the weighted sum of each photodiode's output signal. By assuming that the angular sensitivity of each photodiode fits with a bell-shaped function (see 2), the weighted sum W_S can be written as follows:

$$W_S = \sum_{i=1}^{n} i * Ph_i \tag{1}$$

with Ph_i the output signal of a photodiode. We assume the photodiodes to be located at the i^{th} position on a segment ranging from 1 to n. It can be shown that the W_S function can feature intriguing properties [2]. In particular, it can be shown that for sufficiently large value of the the bell-shaped curves (i.e., an overlap of at least 15% between two adjacent photodiodes), W_S becomes monotone increasing and almost linear. It is worth noting that this property can explain the hyperacute phenomena. Like hyperacuity defined for a visual system [14], the hyperacute LinLed device can locate a finger or a hand with an accuracy much better than the pitch between two adjacent photodiodes.

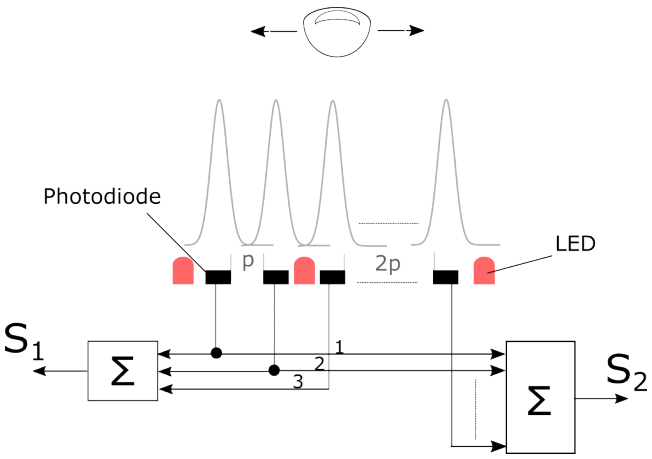

Figure 2: The LinLED interface can be seen as a 1D array of photodiodes featuring Gaussian angular sensitivities. The two summations S_1 and S_2 are: the weighted sum of each photodiode's output signal and the sum over the whole array, respectively.

2.2 Model of finger localization

As illustrated in figures 2 and 3, the output signal $Ph(\Psi_f)$ of a photodiode can be calculated as follows:

$$Ph(\Psi_f) = I_2 \int_{-\infty}^{\Psi_f - \frac{L_1}{2}} s(\Psi)d\Psi + I_1 \int_{\Psi_f - \frac{L_1}{2}}^{\Psi_f + \frac{L_2}{2}} s(\Psi)d\Psi + I_2 \int_{\Psi_f + \frac{L_2}{2}}^{+\infty} s(\Psi)d\Psi \tag{2}$$

with Ψ_f the angular position of the finger. We assume here that the finger can be considered as a bar with a contrast intensity I_1 placed in front of a background of intensity I_2. The angular position Ψ_f can be calculated as follows:

$$\Psi_f = \arctan 2\left(\frac{D}{X_b - X_p}\right) \tag{3}$$

with D the distance to the finger, X_p the photodiode's linear position along the 1D array and X_b the position of the center of the finger. The angular sensitivity $s(\Psi)$ is assumed to follow a Gaussian curve defined as follows [6]:

$$s(\Psi) = \frac{2\sqrt{\pi \ln(2)}}{\pi \Delta\rho} e^{-4\ln(2)\frac{\Psi^2}{\Delta\rho^2}} \tag{4}$$

with $\Delta\rho$ the angle at half-with of the Gaussian.
Thus, by combining equation 4 and (2), we have:

$$Ph = \frac{I_1 - I_2}{2}\left(\text{erf}\left(\frac{2\sqrt{\ln(2)}}{\Delta\rho} * L_1\right) - \text{erf}\left(\frac{2\sqrt{\ln(2)}}{\Delta\rho} * L_2\right)\right) + I_2 \tag{5}$$

with *erf* the well-known error function and the angles L_1 and L_2 defined as follows:

$$L_1 = \arctan 2\left(\frac{D}{X_b - X_p - 0.5\Delta_f}\right) \text{ and } L_2 = \arctan 2\left(\frac{D}{X_b - X_p + 0.5\Delta_f}\right) \tag{6}$$

with Δ_f the finger's width. Therefore, it can be seen that the model of the output signal of one photodiode results from the difference of two error functions, which is also bell-shaped curve.

2.3 Hardware implementation

The analog processing of the LinLed device relies on a modulated infrared signal (30kHz) emitted by the LEDs, making LinLED robust to variations introduced by artificial lighting (neon, LEDs...) and even sunlight. A lock-in amplifier connected to each photodiode was used to demodulate the signal reflected by fingers or hands. Then, the demodulated signal of each photodiode is sent to a summing amplifier by means of a classical operational amplifier. By choosing correctly the set of resistance, the weighted sum (noted S_1 in figure 2) described in section 2 can be directly implemented with analog electronics, making the processing time of this operation extremely fast. We implemented an additional summing amplifier (noted S_2 in figure 2), the output of which is used to calculate the normalized weighted sum described in section 3.2. Figure 4 depicts the hardware of LinLED.

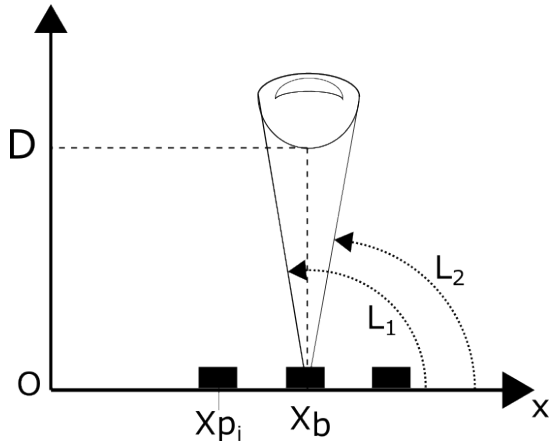

Figure 3: The finger is assumed here to be a bar placed at a distance D from the photodiode and subtending an angle equal to $L_1 - L_2$.

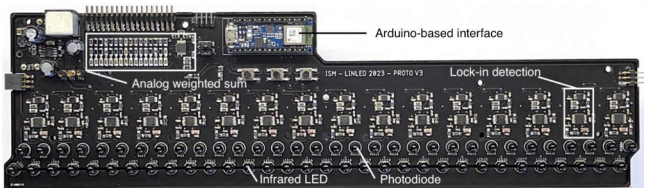

Figure 4: The LinLED hardware.

3 PERFORMANCES IN LOCALIZATION

3.1 Test bench

As described in figure 5, we built a test bench composed of an actuated linear stage used to translate a 3D-printed hand with 3 fingers (3 cm width), moving above LinLed at a distance of 8 cm. A lead screw actuated by a stepper motor was used to displace the hand by step of 0.8 mm (micro-step mode) for a course of 393 mm. For each step, the analog WS and the sum were acquired by the internal arduino board (see figure 4) and sent to a computer through a USB interface. The stepper motor was controlled by means of an external arduino board. For sake of repeatability, we used an artificial hand but real fingers, hands or any object reflecting infrared can be used with LinLED (see figure 9).

3.2 Characterization of the hand localization

Figure 6 shows the 32 channels of LinLED corresponding to the 32 demodulated output signals. As expected from the model described in section 2.2, each signal follows bell-shaped function with respect to the linear position of the hand. In addition, it is clearly shown that the amplitude of the signals is not uniform. This is due to a disparity in the sensitivity of the photodiodes. To compensate for this disparity that can affect drastically the linearity in the localization, we calculated the normalized weighted sum S_{norm} defined as follows:

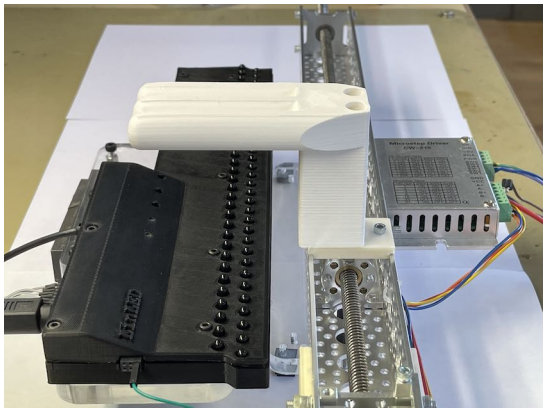

Figure 5: Side view of the 3D printed hand composed of three fingers translating above LinLED by means of linear stage actuated by a stepper motor.

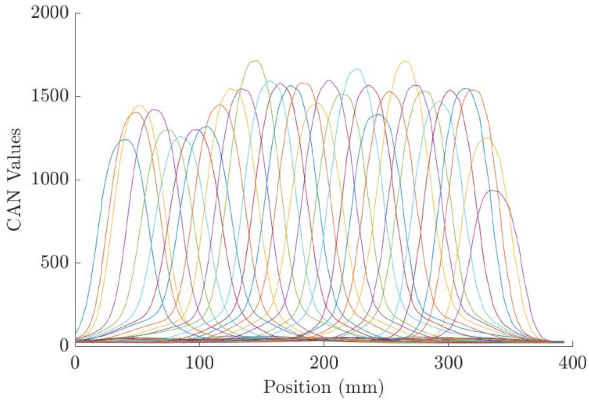

Figure 6: The 32 demodulated output signals of the 32 photodiodes in response to a displacement of the hand along a course of about 400 mm.

$$S_{norm} = \frac{S_1}{S_2} = \frac{\sum_{i=1}^{n} i * Ph_j}{\frac{\sum_{i=1}^{n} Ph_i}{n}} \tag{7}$$

with S_1 and S_2 the weighted sum and the sum of the photodiode's output signals, respectively (see figure 2). S_{norm} is calculated in real time on the arduino board implemented in LinLED.

Figure 7 shows the LinLED response versus the hand displacement. The response is very linear and smooth. In addition, it is clearly shown that the resolution is here of 1 mm, which is 10-time smaller than the resolution of 1 cm imposed by the pitch between two adjacent photodiodes. Figure 8 shows the normalized weighted sum S_{norm} for various distance D ranging from 45 mm to 130 mm. The main interest of the normalized weighted sum is precisely to make LinLED robust to distance variations. The slope of the curves shown in figure 8 remains constant and linear for a finger traveling over 200 mm.

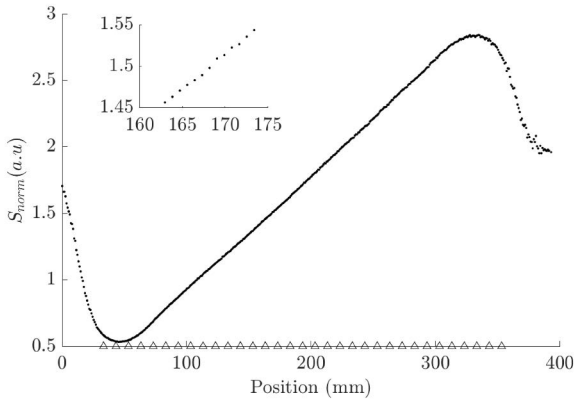

Figure 7: Normalized weighted sum versus linear position of the hand. The inset (magnified view) shows that the resolution of LinLED is equal to 1 mm.

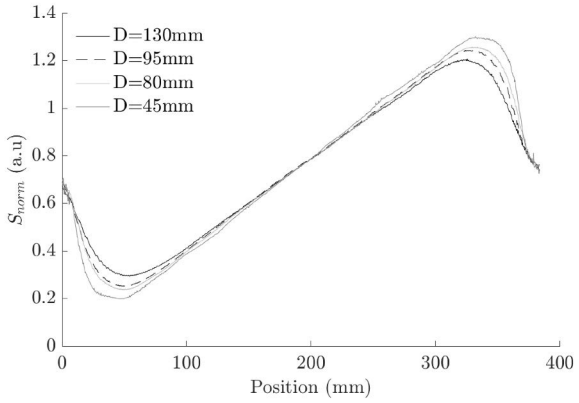

Figure 8: Normalized weighted sum versus position of the hand for four distances of the hand ranging from 45mm to 130mm.

Figure 9 shows the LinLED response versus the hand displacement for three real hands of three various skin tones. The hands were placed at about 9cm above LinLED. The response is very linear and smooth regardless the tone and width of the hand.

3.3 A low latency interface

To characterize the latency of the signal processing, we used a mechanical chopper composed of a rotating blade placed over a pair of LED-photodiode. The chopper emulated a very fast displacement of an object moving swiftly over LinLED. As shown in figure 10, the latency of the analog demodulation is about 0.4ms. The latency of the analog summing is to small to be measured. The total latency including the analog to digital conversion and processing time (32-bit Arduino board) is about 1ms.

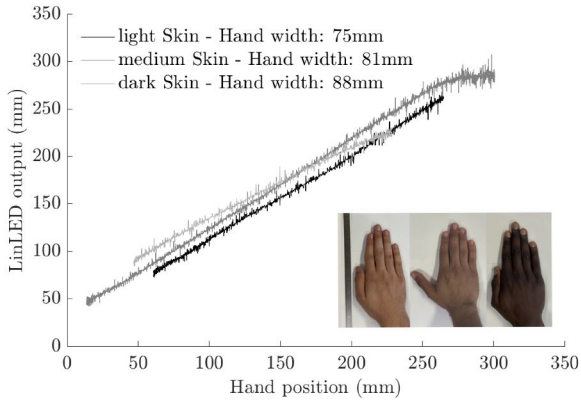

Figure 9: LinLED's output signal versus real hand's lateral position for three various hands placed at about 9cm above LinLED.

4 CONCLUSION

We present an interface designed for gesture interaction named LinLED. This contactless, high-speed, and precise system operates along two orthogonal axes (X and Z) with virtually unlimited tracking space, constrained only by the number of addressable photodiodes. Notably, LinLED's resolution of 1 mm, dictated by the test bench, likely surpasses its true resolution capabilities, promising even finer precision. By measuring linear position, LinLED enables effortless calculation of speed and motion direction. When connected to a computer via a USB interface, it seamlessly functions as a human interface device (HID). LinLED was specifically developed to complement mid-air ultrasonic interfaces, where great accuracy and ultra-low latency are essential for projecting ultrasounds onto the user's hand [3]. In this context, LinLED serves as a compelling alternative to commercial gesture interfaces such as LeapMotion [10]. The tracking space could also be improved by increasing the light intensity and the amplifier gains. To promote accessibility and user-friendliness, we have developed a Windows Software Development Kit (SDK) in C and Python, allowing seamless integration and exploration of LinLED's capabilities. The SDK will soon be available for open access, encouraging innovation and collaboration. Inspired by the works of [1] and [7], we are actively investigating the potential of convolutional neural networks to further enhance LinLED's gesture recognition capabilities by analyzing the temporal patterns of the photodiode signals. LinLED represents a valuable gesture interface solution, characterized by its speed, accuracy, and seamless integration possibilities. Its potential for various applications, combined with a user-friendly SDK and ongoing research into gesture recognition, positions LinLED at the forefront of gesture interaction technology.

5 ACKNOWLEDGMENTS

We acknowledge the support received from the Centre National de la Recherche Scientifique (CNRS), Aix-Marseille University and the french Tech Transfer acceleration Company (SATT Sud-Est).

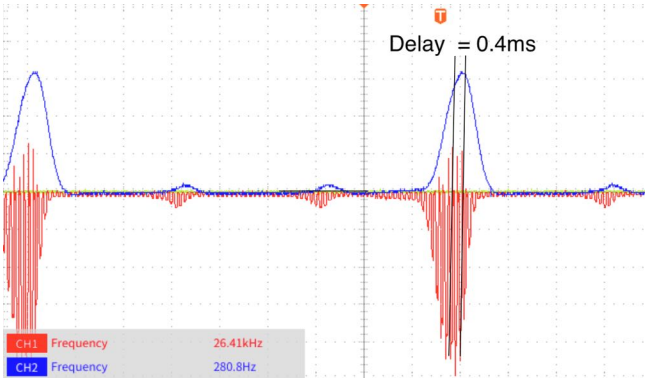

Figure 10: Delay between the raw modulated signal provided by a phtodiode of LinLED and the demodulated signal in response to a rotating blade moving very quickly over the interface.

REFERENCES

[1] Fahmid Al Farid, Noramiza Hashim, Junaidi Abdullah, Md Roman Bhuiyan, Wan Noor Shahida Mohd Isa, Jia Uddin, Mohammad Ahsanul Haque, and Mohd Nizam Husen. 2022. A structured and methodological review on vision-based hand gesture recognition system. *Journal of Imaging* 8, 6 (2022), 153.

[2] Pierre Baldi and W Heiligenberg. 1988. How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. *Biological cybernetics* 59, 4-5 (1988), 313–318.

[3] Kyle Harrington, David R Large, Gary Burnett, and Orestis Georgiou. 2018. Exploring the use of mid-air ultrasonic feedback to enhance automotive user interfaces. In *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*. 11–20.

[4] Walter Heiligenberg. 1987. Central processing of sensory information in electric fish. *Journal of Comparative Physiology A* 161, 4 (1987), 621–631.

[5] Shuo Jiang, Peiqi Kang, Xinyu Song, Benny PL Lo, and Peter B Shull. 2021. Emerging wearable interfaces and algorithms for hand gesture recognition: A survey. *IEEE Reviews in Biomedical Engineering* 15 (2021), 85–102.

[6] Lubin Kerhuel, Stéphane Viollet, and Nicolas Franceschini. 2011. The VODKA sensor: A bio-inspired hyperacute optical position sensing device. *IEEE Sensors Journal* 12, 2 (2011), 315–324.

[7] Kyun Kyu Kim, Min Kim, Kyungrok Pyun, Jin Kim, Jinki Min, Seunghun Koh, Samuel E. Root, Jaewon Kim, Bao-Nguyen T. Nguyen, Yuya Nishio, Seonggeun Han, Joonhwa Choi, C.-Yoon Kim, Jeffrey B.-H. Tok, Sungho Jo, Seung Hwan Ko, and Zhenan Bao. 2023. A substrate-less nanomesh receptor with meta-learning for rapid hand task recognition. *Nature Electronics* 6, 1 (2023), 64–75. https://doi.org/10.1038/s41928-022-00888-7

[8] Nibhrat Lohia, Raunak Mundada, Arya D McCarthy, and Eric C Larson. 2021. AirWare: Utilizing Embedded Audio and Infrared Signals for In-Air Hand-Gesture Recognition. *arXiv preprint arXiv:2101.10245* (2021).

[9] Andualem Maereg, Yang Lou, Emanuele Secco, and Raymond King. 2020. Hand Gesture Recognition based on Near-infrared Sensing Wristband. https://doi.org/10.5220/0008909401100117

[10] Tomás Mantecón, Carlos R del Blanco, Fernando Jaureguizar, and Narciso García. 2016. Hand gesture recognition using infrared imagery provided by leap motion controller. In *Advanced Concepts for Intelligent Vision Systems: 17th International Conference, ACIVS 2016, Lecce, Italy, October 24-27, 2016, Proceedings 17*. Springer, 47–57.

[11] Tomás Mantecón, Carlos R Del-Blanco, Fernando Jaureguizar, and Narciso García. 2019. A real-time gesture recognition system using near-infrared imagery. *PloS one* 14, 10 (2019), e0223320.

[12] Keunwoo Park, Sunbum Kim, Youngwoo Yoon, Tae-Kyun Kim, and Geehyuk Lee. 2020. DeepFisheye. https://doi.org/10.1145/3379337.3415818

[13] Shigeyuki Tateno, Yiwei Zhu, and Fanxing Meng. 2019. Hand Gesture Recognition System for In-car Device Control Based on Infrared Array Sensor. https://doi.org/10.23919/sice.2019.8859832

[14] Gerald Westheimer and Suzanne P. McKee. 1977. Spatial configurations for visual hyperacuity. 17 (1977), 941–947. https://doi.org/10.1016/0042-6989(77)90069-4

[15] Yuki Yamato, Yutaro Suzuki, Kodai Sekimori, Buntarou Shizuki, and Shin Takahashi. 2020. Hand Gesture Interaction with a Low-Resolution Infrared Image Sensor on an Inner Wrist. https://doi.org/10.1145/3399715.3399858