Highly Reliable Time-of-Flight L-Shaped Gesture Recognition System

Haris Kudić and Dejan Jokić

***Abstract* — Gesture recognition is a field of study that involves recognizing human movements and gestures through sensors. In this paper, a basic gesture recognition system is proposed that uses three Time-of-Flight (TOF) VL53L0X distance sensors positioned in an L-shape able to recognize gesture through cover glass and up to 40 cm of distance. The system is capable of recognizing four basic gestures: swipe right, swipe left, swipe up and swipe down. This system can be applied in various fields such as Human-Computer Interaction (HCI), Gaming, Virtual Reality (VR) and Robotics, this paper will focus on the implementation and evaluation of the proposed system. The inspiration for the system is to simplify interaction with medical panel PCs and monitors while improving the hygienic aspect of the same, while taking into consideration data privacy.**

***Keywords* — Gesture Recognition, HCI, Time-of-Flight, VL53L0X.**

I. Introduction

Gesture recognition technology has been widely used in various applications, such as human-computer interaction, gaming, and virtual reality. There are different types of sensors that can be used for gesture recognition, including camera-based systems, ultrasonic sensors, infrared sensors etc. However, TOF distance sensors have become increasingly popular in recent years due to their low cost, high accuracy, and small size [1], [2].

The choice of a proper gesture recognition device is extremely important, it is to be assumed that these devices will be, more often than not, covered with some kind of protective glass cover. Therefore, a gesture control system needs to overcome this. Also, the technology on which it is based needs to be chosen as to not be affected by the environment conditions. Most of the gesture sensors on the market are based on the infrared light and therefore their accuracy can greatly be influenced by external IR light, which represents big issues while using these devices in the daylight conditions. Some of these sensors work on a principle of four directional IR photodiodes, which presents no issue regarding data protection, while others take pictures using IR cameras, in order to recognize the motion. Under certain regulations and conditions this may represent an issue [3]-[5].

Nevertheless, the biggest flaw of the readily available sensors has been the maximum recognition distance and influence that the environment conditions can have on the performance. Therefore, this paper demonstrates an affordable, custom built gesture recognition system that overcomes mentioned issues offering cover glass support, no or low external IR light influence, data protection and long recognition distance.

Camera based gesture recognition systems represent one of the best choices due to their fairly easy implementation nowadays and broad range of gestures to be recognized. Even so, they do have certain disadvantages in some applications, one of the biggest ones being data privacy. Camera based systems are required to capture and analyze private data, which can represent an issue in certain situations and for certain users. Other than privacy, for implementation of simple gesture recognition, camera-based recognition system together with the necessary computation power unit, may represent a lot costlier solution [6].

Time-of-Flight (TOF) sensors are non-contact sensors that measure the distance to an object by measuring the time it takes for a light pulse to travel to an object and back. TOF sensors work based on the principle of triangulation, meaning they measure the time that it takes for a light pulse to travel to an object and back, which can then be used to calculate the distance of an object based on the speed of light [7]. TOF sensors are widely used in various applications such as automation, automotive, robotics, human-computer interaction etc. [8].

One of the main advantages of using TOF sensors for gesture recognition is their high accuracy. TOF sensors can measure the distance to an object with an accuracy of a few millimeters, which makes them more than suitable for gesture recognition on short distances where change of distance is more important than the actual distance. Other than that, TOF sensors are small in size which makes them perfect candidate for embedding and integrating into broad range of devices [7], [8]. This paper aims to provide an example of a cost effective and simple, yet reliable gesture recognition system comprised of TOF sensors, able to provide good accuracy and ease of use even in non-optimal conditions. The system should be able to accommodate for the cover glass, take into consideration data-protection and privacy.

II. Experiment

The gesture recognition system described in this paper is meant to be used in applications where space is not a limiting factor like medical grade monitors, specialized devices etc. Since it is designed to be controlling bigger, rather than smaller devices, the gestures performed should be performed with the whole hand rather than with a finger.

Submitted on February 14, 2023.
Published on April 12, 2023.
H. Kudić, International Burch University, Bosnia and Herzegovina.
(e-mail: haris.kudic@stu.ibu.edu.ba)

D. Jokić, International Burch University, Bosnia and Herzegovina.
(e-mail: dejan.jokic@ibu.edu.ba)

This ensures better accuracy, and it is more natural for larger devices.

A. Hardware Setup

Hardware component of the gesture recognition system consists of the following parts:

- Microcontroller used for recognizing gestures;
- Three VL53L0X Time-Of-Flight sensors;
- The system (host) for interacting with the gestures.

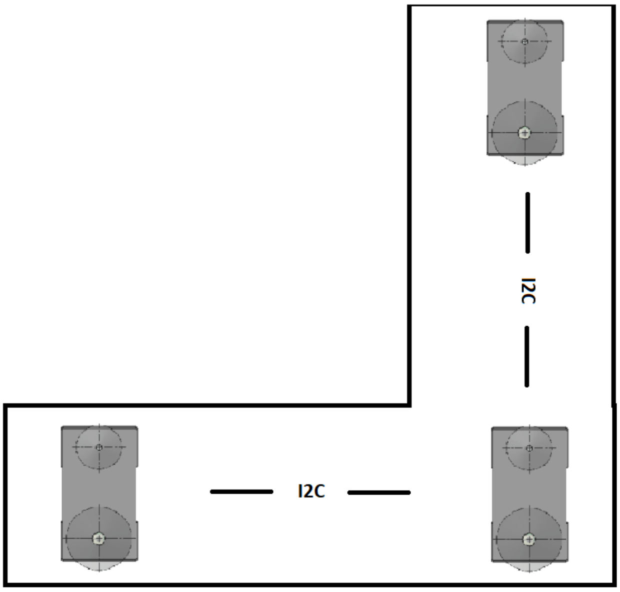

Fig. 1. 'L' - shaped TOG Gesture Recognition System.

This gesture recognition system is built on top of three time-of-flight ST Electronics VL53L0X ranging sensors with effective measuring of up to 2 m. These sensors have been chosen because of their excellent performance during daylight and also their ability to compensate for the cover glass. This will ensure that the device is usable even when it is to be built in behind a glass, which makes it a great candidate for incorporating this system in monitors and similar devices. In order to detect four types of motion, swipe left, swipe right, swipe up, swipe down, the sensors are positioned to be in 'L' shape as shown in Fig. 1. This will enable to detect the order in which the sensors are reading changes in their respective measurements of distance when gesture is being performed and thus successfully detect the gesture. Another reason an ‘L’ shape and three TOF sensors were chosen is so that the system could be built in the corner of a larger device, meaning that the size of the gesture recognition system can be kept small for a more natural way of performing gestures, whereas the actual device being controlled can be a lot bigger. While using four gestures would add to the accuracy of the system, it would also limit its usage due to larger area that it would occupy and for most of today’s devices this would be troublesome [9].

Mutual displacement of the TOF sensors is extremely important because of their effective field of views of the collector. This is the main constraining part in the size of the gesture system overall. The FOV for each of these sensors, taken from the datasheet, is 25° [9] as shown in Fig. 2. Placing these sensors closer to each other will decrease the distance where their respective FOVs overlap, causing a decrease in the maximum recognition distance. Proportionally, increasing the displacement will increase the effective maximum recognition distance.

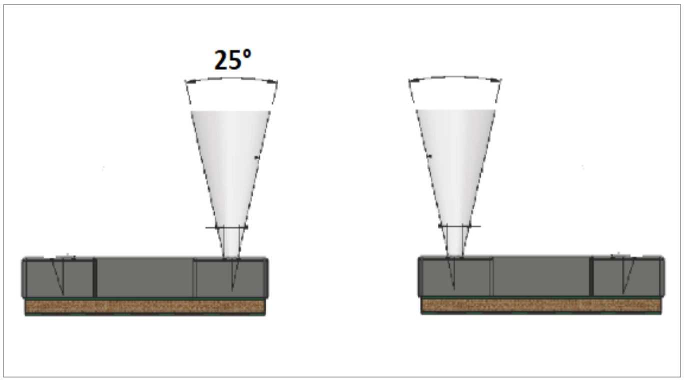

Fig. 2. Collector Field of View.

For the recognition system a maximum recognition of 40 cm has been chosen and the displacement of the sensors has been calculated according to this distance. It is worth noting that since the gesture recognition system is based on the order in which the sensors are being activated, it is not necessary to completely isolate respective FOVs, but rather to have a region in which these are not overlapping at the maximum distance. This region here has been chosen as half of the FOV and calculations have been done accordingly. The FOV range edge of the first sensor at 40 cm should lay in the middle of the FOV of the second sensor.

From Fig. 2 is clearly visible that each FOV represents a right circle cone with its half-angle being 12.5°. In order to calculate the necessary distance between two sensors, it is first necessary to calculate its base radius. This can be easily calculated utilizing (1).

$$r = h * \tan(\theta) \quad (1)$$

where;
r = radius of the base of the cone
h = height of the cone (40 cm)
θ = half-angle of the cone (12.5°)

Plugging in the values, the cone for its base at 40 cm will make a circle with radius of 8.8678 cm. Having calculated the radius of the base, Fig. 3 shows the circular bases of the two neighboring sensors and it is quite clear that the radius is also the distance the two sensors need to be apart considering that one’s FOV ends at the center of the other one.

VL53L0X sensors use I2C bus as a communication protocol which makes interconnection of the sensors fairly simple, as they can all be connected in parallel using only two wires, aside from power lines. As seen from data sheet, the sensors have the same default hardware I2C address, but this can be programmed by keeping the XSHUT pin low so that is possible to connect multiple sensors in the system.

For the purpose of this experiment, Arduino Mega 2560 for reading TOF sensors values and recognizing gestures has been used, while as a host a Windows PC has been used to gather information about performed gestures. TOF Sensors have been soldered onto a bare PCB and connected to the Arduino via I2C lines.

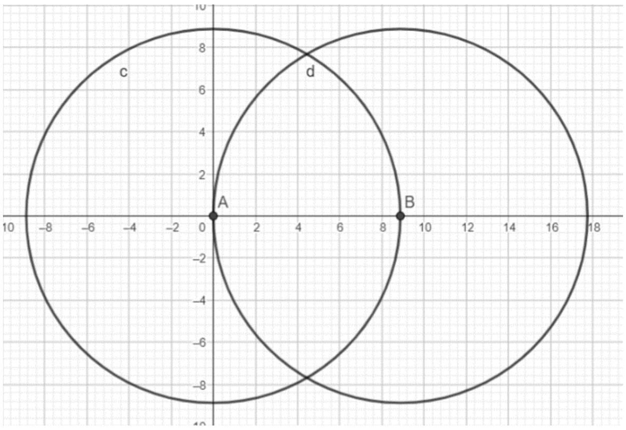

Fig. 3. Circular bases of two neighboring sensors.

B. Software Setup

The software part relies on reading the values of all three sensors continuously to notice any difference in the measurements. In order to demonstrate this, a VL530L C library has been used. Library supports all necessary functions to setup sensors, as well as to read the values and if necessary to calibrate them.

In order to setup the sensors, first it is necessary to program their I2C address, as previously said. Failure to do so will not enable the use the sensors as there will be a conflict on the I2C bus. It is worth mentioning that the same process needs to be done whenever the sensor loses the power as the changes are not persistent. In this step it is necessary to know which I2C address corresponds to which physical sensor, since this will be used later on to actually determine the gesture. Once the I2C addresses are programmed, it is necessary to initialize the sensors which is greatly simplified by the library.

Sensor initialization programs the registers with default values, if necessary, the sensors can be further calibrated for distance in order to compensate for the cover glass. Since the gesture recognition logic is based on a threshold rather than exact values, the calibration is not necessary as it is simpler to just add an offset to the readings of the sensors. This will compensate for different readings when the cover glass is present. If for some reason the sensors would provide rather high differences in readings for the same distance, then all sensors would need to be calibrated.

Once the sensor setup is done, it is possible to write the function for the actual gesture recognition. As previously said, gesture recognition is based on reading the measurements of all sensors and identifying when the sensor measurement is within the specified range, in this case from 5-40 cm (with no offset applied for cover glass). The measurements of all three sensors are taken in a continuous loop while also logging the sequence in which the sensors enter the specified range. The sequence is determined after the specified time elapse which can be used to control the speed of the gesture recognition, smaller time frame will indicate that the user must swipe faster in order to achieve the gesture.

It is worth noting here that the gesture recognition function will halt when two sensors are within the specified range at the same time, this is to further limit the false gesture readings. Effectively this would mean that the sensors readings will be accepted only in the allowed regions for each sensor individually (not intersected parts of FOVs in Fig. 3). Once the reading has been done and the gesture has been recognized, the microcontroller can either broadcast a message of a successfully performed gesture to the host or perform a specified command.

In certain situations, when it is necessary to execute certain gesture a number of times, in order to alleviate the effort of the user, it is possible to perform the same gesture by holding the hand over two bottom sensors immediately after the gesture. This needs to be done in a timely manner so that the system does not read the movement as a gesture but as a repeated gesture state.

III. Discussion and Results

In order to test the accuracy of the gesture recognition mechanism, each of the four gestures has been performed 100 times indoors under natural lighting and room temperature. All gestures have been performed by hand movements at different speeds at distance 15-30 cm from the sensors. The average time between repeating gestures was kept between 1-2 seconds.

The proposed system demonstrates that three TOF distance sensors positioned in an L-shape can be used to recognize basic gestures with high accuracy. The system was able to recognize the four basic gestures: swipe right, swipe left, swipe up, and swipe down with an average accuracy of 97.25% (Table I). Somewhat different accuracy for different gestures could be due to order in which the TOF sensors are being read which could influence in some gestures failing to recognize more often than other.

TABLE I: Gesture Recognition Accuracy

Gesture	Accuracy
Swipe Left	98%
Swipe Right	98%
Swipe Down	96%
Swipe Up	97%
Avg.	97.25%

The results of the experiment show that the proposed system is suitable for real-world applications provided its robustness to variations in the participants' hand movements, which means that it can recognize the gestures in a natural way at different speeds and different heights. TOF sensors enable usage of the gesture recognition system even in the daylight due to their compatibility to overcome IR influence from the environment. This makes the system suitable for applications such as human-computer interaction, gaming and virtual reality. The system proposes itself as an especially good solution where simple, touchless interaction with the device is needed.

However, the proposed system has some limitations. The system is only capable of recognizing four basic gestures: swipe left, swipe right, swipe up and swipe down. In order to add more complexity to the gesture recognition, extra programming would need to be done. The system could also have difficulties with different type of cover glass so this would need to be examined on a per case basis.

Another limitation could be the physical size of the system, as the sensors require some minimum distance between them in order to be able to perform gestures at certain height so certain devices (monitors, panel PCs etc.) could have too thin bezels for incorporating this system into their design.

IV. Conclusion

In this paper, a simple, custom gesture recognition system has been proposed using three Time-Of-Flight VL53L0X distance sensors positioned in an L-shape. The system is capable of recognizing four basic gestures: swipe right, swipe left, swipe up and swipe down with an average accuracy of 97.25%. The proposed system could be further improved by adding more gestures and improving the overall gesture performance. System is suitable for real-world applications such as human-computer interaction, gaming and virtual reality where space is not a constrain due to somewhat larger size.

The inspiration for the system design was an easier and more hygienic control of medical monitors and panel PCs which would lower the need for touching the devices and thus transmitting bacteria and viruses, while also taking into consideration data protection and privacy by omitting usage of any type of sensitive information (camera-based gesture system). The proposed system can be implemented in various fields to simplify human-computer interaction and automation process in all types of industries where a simple human interaction is needed.

Conflict of Interest

Authors declare that they do not have any conflict of interest.

References

[1] Li J, Chen Y, Liu Y. A survey of gesture recognition using depth sensors. *Multimed Tools Appl.,* 2018 Nov 15;77(23):27757-27787.
[2] Kao SK, Lin CY. A review of gesture recognition using time-of-flight sensors. *Sensors*, 2018 Jul;18(7):2037.
[3] Kim K, Kim J, Kim HK. A review of infrared gesture recognition technology. *Measurement*; 2018 Oct 1;141:37-47.
[4] Chen L, Liu Y, Li J. A review of gesture recognition using stereo vision. *Measurement*, 2018 Nov 1;142:63.
[5] Weber AM. The role of time-of-flight cameras in human-computer interaction. In *2013 IEEE International Conference on Robotics and Automation*, 2013 May 6-10; Karlsruhe, Germany. IEEE; 2013, p. 69-76.
[6] Naseer AM, Almagrabi AH. Pros and Cons of Camera-Based Gesture Recognition: A Review. In: *2021 IEEE 5th Middle East and North Africa Conference on Biomedical Engineering (MENA-BME)*, 2021 Mar 29-30; Cairo, Egypt. IEEE; 2021, p. 1-6.
[7] Al-Osaimi AM, Imran MA, Seah WKG. A survey of gesture recognition techniques. *IEEE Commun Surv Tutor*, 2014 Jan 1;16(1):414-54.
[8] Corso JJ. Time-of-flight sensors for gesture recognition. *IEEE Trans Pattern Anal Mach Intell,* 2013 May;35(5):1205-18.
[9] ST Microelectronics. *VL53L0X datasheet*. [Internet]. 2018. Available from: https://www.st.com/resource/en/datasheet/vl53l0x.pdf.