Design of a Laser Pointer Follower Robot

Facundo García-Cárdenas*, Nelson Soberón*, Cristian Amaya*, and Victor Murray*†‡
*Department of Electrical Engineering, Universidad de Ingenieria y Tecnologia - UTEC, Lima, Peru.
†Department of Electrical and Computer Engineering, University of New Mexico, New Mexico, USA.
‡John A. Paulson School of Engineering and Applied Sciences, Harvard University, Massachusetts, USA.

***Abstract*—We propose a system that performs automatic laser spot detection and tracking. Our laser spot detection approach is based on the HSV color model, thresholding, and center of mass. For tracking, a system of quadrants was used to move a camera depending on the center of mass. Also, the distance to the laser spot is calculated using a stereo vision configuration and geometric relationships. The proposed architecture is composed of an embedded computer that controls a low cost, semi-autonomous sentry robot attached to it. Our results show a robust real-time system that can detect a laser spot and track it, with a distance error lower than 0.80% on average and an execution time of 62 milliseconds per frame, that can be used in different applications.**

***Index Terms*—Stereo vision, laser tracking**

I. Introduction

Due to the recent increase in demand for streaming virtual classes and conference presentations, speakers often require special tools to move the camera towards a specific region of interest. In this context, laser spot tracking has become a useful tool for applications where laser pointers are used to pinpoint a part of a presentation or procedure, an action that the camera can use to center the laser spot in the image. This technology has been successfully deployed for the automatic detection of laser points for remote assistance [1], telemedicine [2], and projector-augmented environments [3]. Nowadays, different methods are proposed to perform this task; for example, in [4], a review of a laser spot detection system is presented. This method is based on thresholding and background subtraction to perform laser tracking. Another methodology was proposed in [5], where the authors used a system based on the Hue-Saturation-Value (HSV) color model and camera settings. The application of these papers was focused on using a laser pointer as a mouse in personal computers. For tracking a laser dot, a mechanical system must be present. In this sense, in [3], a similar application is presented. This system was able to center a laser spot in the camera performing the tracking, using two stepper motors. The laser dot detection was based on intensity-thresholding and computing the center of gravity to detect the laser position.

In this paper, we present an architecture and algorithm to detect and track a laser point. The primary process for the reliable laser spot detection system includes thresholding, background subtraction, the region of interest (ROI) extraction, and coordinate mapping. Also, a distance calculation is computed to perform more accurate tracking. We have tested the algorithm using different laser pointer interactions, particularly with a large screen and external appliances. The focus of this work is to present a whole framework to develop a low-cost semi-autonomous sentry robot, ranging from mechanical design to high-level control. The methodology includes the steps used by the sentry to find its objective, calculate the distance to it, and compute the movements required to center the target on its view.

The organization of this manuscript is arranged as follows. In section II, we introduce the architecture of the laser spot detection system. Next, in section III, we present the main algorithm to detect the laser spot using digital image processing algorithms and how to calculate the distance of depth using the stereo vision. Then, in section IV, we present the results in terms of accuracy and error. Finally, in section V, we present the conclusions and some possible future directions.

II. Architecture of the System

The proposed system requires the ability to track a laser spot in any environment, by automatically segmenting the target from the background and calculating the distance between the object and the sensor. Therefore, this sentry should capture real-time video feed using a stereo camera configuration, a technique that allows the computation of the distance from a set of cameras to an object by imitating how human vision determines depth. Thus, the cameras are arranged horizontally and separated, which generates enough disparity between the images to represent depth. Then, stereo camera parameters are used to relate pixels with distance.

A set of two Fosmon's 6 LED webcams was employed to capture the video. Each of these cameras contains a sensor capable of recording a 640×480 image running at 30 frames per second, with a resolution up to 1.2 megapixels. Even though this camera is capable of achieving that frame rate, since the detection algorithm takes about 60ms to execute, only an average of 16 frames per second was used. For the stereo configuration, each camera sensor was stripped from its module and attached to a closed polylactide (PLA) case, which was 3D-printed to hold both cameras at a distance of 6.2 centimeters from each other. Then, this custom module was assembled on top of a metallic sentry robot with two servo motors, which are used to move the cameras horizontally and vertically to avoid having a dead angle, as shown in Fig. 1.

The architecture of the automated laser spot tracking device is shown in Fig. 2. The system consists of a laser pointer,

978-1-7281-9377-9/20/$31.00 ©2020 IEEE

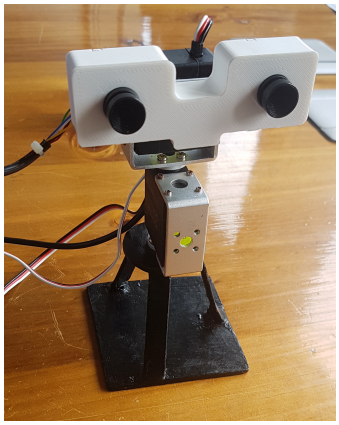

Fig. 1. Sentry robot implemented to track a laser dot.

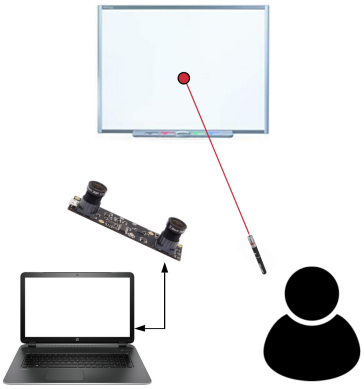

Fig. 2. The architecture of the proposed system. The stereo camera captures the images of the laser dot. These images are processed in the computer, and the coordinates are sent to the camera to track the laser point.

which emits the target laser spot, a sentry robot, which uses a set of cameras to segment the object, a microcontroller, to move the servomotors, and a processing unit; which for these tests was a personal computer. The target is used to select an object of interest, which differs depending on the application of the system. In this context, the sentry provides real-time visual input of the target, calculates the distance to it, and ensures that it is always on sight.

III. Laser Spot Detection

The process of segmenting an object within an image requires pre-processing the input video by extracting the video frames and applying the corresponding filters to each frame. This process varies depending on the characteristics of the object like shape, brightness, and size, as well as the accuracy and speed required for the task. For this application, it is necessary to detect a red, bright, laser spot in real-time. In the next subsections, we describe the different parts of the system.

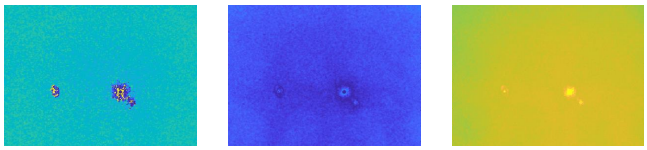

Fig. 3. HSV components, from left to right, of an extracted video frame.

A. Conversion to HSV

The first step consists of extracting some frames from the video input. Each frame is represented using the red, green, and blue (RGB) color model, where each component of the object's color is correlated to the intensity, i.e., the amount of light hitting the object, and, consequently, each other. Moreover, this color model is not aligned to the way the human vision perceives color and, given that the perception of the laser spot depends highly on the surface it is reflected from, it is recommended to use another color model. Therefore, a conversion to the Hue-Saturation-Value (HSV) color model is applied, which facilitates working on scenarios where color description plays an essential role.

The target of this system is defined as a red laser spot. Even though a pure color characterizes this object, it not possible to segment it by filtering the red layer only, as the infrared part of the laser's light affects how it looks on camera. In an ideal scenario, this behavior would not affect the resulting image, as it would only show the information corresponding to the visible spectrum. However, this does not happen in practice, wherein most scenarios the camera sensor spreads the infrared information all across the visible light spectrum. Thus, the spot presents information corresponding to each frequency and is represented as a pinkish white color. This behavior is shown in Fig. 3, where each HSV component of the target is displayed.

The laser dot looks like a bright white dot surrounded by a red shadow because of the reasons stated above. In this example, each HSV layer highlights the laser spot in a different manner. For instance, in the hexagonal model, the hue layer assigns a value depending on the angle formed from the red axis. Thus, as both pure red and white colors correspond to an angle of $0°$, they are assigned a hue value of 0. On the other hand, the value layer corresponds to the range between pure color and white, where the white part of the laser spot would be assigned a value of 1. Lastly, the saturation layer assigns a value depending on the purity of the color. In most scenarios, where the background is composed of mainly gray and brights spots, this layer does not yield any meaningful information. This issue will become more apparent after filtering the image in the following section.

B. Color-space filtering

We apply image binarization using the HSV color model, for each layer, in a way that resembles how humans recognize color. Then, to segment any bright red spots within an image, the upper and lower limits for the hue layer are set to $[0 - 0.056]$ and $[0.944 - 1]$, for the saturation layer to $[0.2 - 0.3]$, and for the value layer to $[0.9 - 1]$. In Fig. 4, we show the

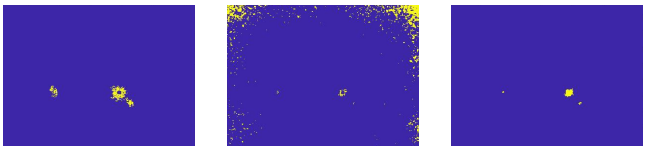

Fig. 4. HSV components, from left to right, of the extracted video frame of Fig. 3, after binarization.

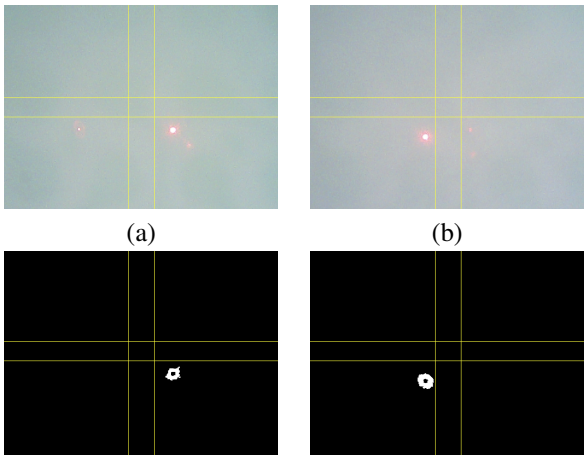

Fig. 5. Segmented objective after the threshold and elimination of reflected points. (a) Original left image. (b) Original right image. (c) Filtered left image. (d) Filtered right image.

results of applying the corresponding thresholds to each HSV layer.

As previously mentioned, the saturation layer presents significantly more noise than its counterparts. Therefore, using this layer could generate false negatives. On the other hand, the hue and value layers present only the target with some added noise, which can be removed by applying an AND operation over these two layers. However, the segmented object would only contain about five pixels on average. Then, a maximum filter of order 121 was applied to smooth and enlarge the target. A filter approach was chosen over a morphological operation due to the computational cost involved, as each iteration must be as fast as possible to guarantee that the sentry can follow the laser spot.

Nonetheless, when working on reflective surfaces, things are more complicated since the light emitted by the laser pointer itself is reflected on the surface. Thus, all 8-connected regions were computed to deal with this issue. Then, a mask was applied to ensure that only the most prominent region, which in most scenarios corresponds to the target, was kept in the image. This method presents a low computational cost, removes most of the noise from the image, and ensures that only the target is segmented. The results of applying these operations to the HSV image are presented in Fig. 5 for a stereo image detecting the laser point. Finally, a set of lines was added to the image to locate the current quadrant more easily.

C. Position computation

Once the target has been successfully segmented, it is necessary to locate the center of the laser spot. To achieve this, it is recommended to compute the number of pixels n that fits the filter and calculate the center of mass of the resulting object as $\mathbf{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$ and $\mathbf{y} = \frac{1}{n}\sum_{k=1}^{n} y_k$, where $\mathbf{x}$ and $\mathbf{y}$ are the coordinates of the center of the target, and (x_k, y_k) are the coordinates of the k-pixel. In other words, this method computes the mean of the pixel coordinates that fitted the filter process.

D. Sentry control

It is necessary to center the laser spot, by moving the sentry accordingly, to ensure that the laser spot is always within the camera view. To achieve this, first, the algorithm determines in which quadrant the target is located. A microcontroller then applies a small increment to the servos to move the sentry and ensure that the target is located inside the center quadrant. For example, if the laser spot is found inside the upper left quadrant, the sentry robot moves to up and left, thereby centering the target.

E. Distance calculation

To calculate the distance from the cameras to the target, the system assumes the same intrinsic parameters for both cameras, which allows finding a geometric relation for computing the distance using the view angle of the cameras ϕ, and their separation B. The proposed method is based on [6] and described in Fig. 6. As we can see, the algorithm assumes that the two cameras are parallel to each other and separated by a distance of B. On the other hand, each camera possesses an angle disparity ϕ_i, a view angle ϕ_0, and produces an image with width x_0. These parameters are used in conjunction with fundamental triangle relations to calculate the distance to the laser spot D as

$$D = \frac{B}{\tan(\phi_1) + \tan(\phi_2)}. \tag{1}$$

However, the angle disparity of each camera is difficult to compute for each iteration. Therefore, the following relationships are applied to substitute ϕ_i in terms of x_i and the view angle ϕ_0

$$\frac{\frac{x_1}{x_0}}{2} = \frac{\tan(\phi_1)}{\tan(\frac{\phi_0}{2})}, \quad \text{and} \tag{2}$$

$$\frac{\frac{-x_2}{x_0}}{2} = \frac{\tan(\phi_2)}{\tan(\frac{\phi_0}{2})}. \tag{3}$$

Thus, the distance to the target D is defined as

$$D = \frac{Bx_0}{2\tan(\frac{\phi_0}{2})(x_1 - x_2)}. \tag{4}$$

It is worth mentioning that the sign convention used in Fig. 6 should be applied for the horizontal disparity, as it ensures that the distance equation works for all cases. On the other hand,

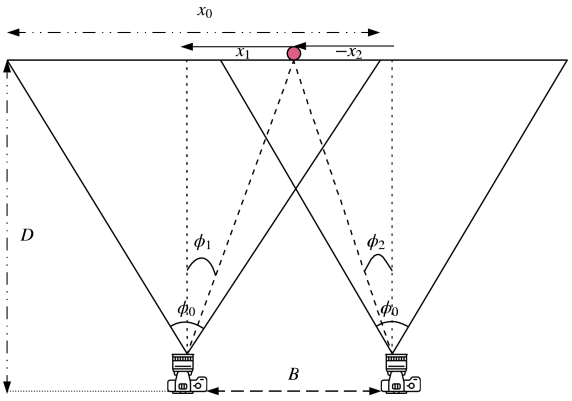

Fig. 6. Diagram for the distance calculation. The method is based on the angle views of the camera and the projection of the points.

the view angle of the cameras was calculated experimentally at about 24^o. Nevertheless, this parameter can be calibrated to obtain better results.

IV. Results and Discussion

We performed 50 experiments, under different environmental conditions, to test the algorithms. The proposed algorithm can detect the laser spot with an accuracy of at least 92% in all the 50 tests. Moreover, this method is capable of dealing with the effects of reflection, working correctly on a colored background, as well as very lit or dark rooms, and discerning between the reflection of the laser pointer itself and the laser spot generated from it. On the other hand, the algorithm was optimized to ensure that it works for real-time scenarios, such as recording a live class or detecting military targets during the day. The algorithm presented a mean computational time of 62ms per frame, using an Intel Core i7-8550U CPU @ 1.80GHz, running on Windows 10. Thus, the resulting image is transmitted at an average of 16 frames per second. Given that the cameras employed are only capable of recording at 30 frames per second, we are using 53.3% of all recorded frames. However, this is enough to ensure that the target is always detected and avoids jittering effects.

The proposed algorithm for the sentry control resulted in an efficient method that can center the spot in an average time of 52ms. However, this result depends highly on the increment value applied to the servomotors. To avoid jitter effects on the camera and to reduce the system's energy consumption, the increment value was set to 0.4. Nonetheless, if more speed was required for the task, it is possible to set the value to the maximum increment tolerated by the servomotor, though this is not recommended.

Finally, to test the distance algorithm, we positioned the structure at several distances from a school board and took five measurements. Since the values did not change significantly, the average was computed for each distance, and the error percentage was calculated. The behavior of the error for different distances can be seen in Fig. 7. On average, the error is lower

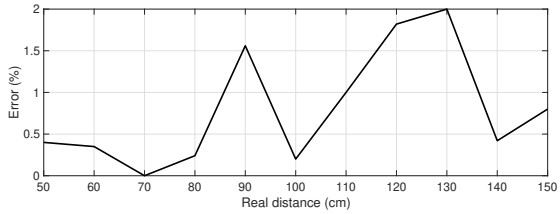

Fig. 7. Percentage of error for distance calculation.

than 0.80%, and we can see that the maximum error was about 2% at a real distance of 130cm. The error increments only at certain distances, unlike a linear relationship; nonetheless, we expect it to increase for more considerable distances. For the sentry robot application, more measurements must be taken to see how the error varies.

V. Conclusions

The proposed method for detecting and follow a laser spot, by applying a threshold in the HSV color model and filtering the 8-connected regions, resulted in a robust implementation capable of detecting the target in very noisy environments, with a high degree of precision. On the other hand, the method for computing the distance from the laser spot to the sentry robot, using fundamental triangle relations within the view angle of the camera and the disparity between the two images, resulted in very precise readings, with a mean error of less than 0.80% for distances lesser than 150cm. The increment-based control system of the sentry robot and the low computational time of the algorithm, allowed to deal with the effects of jittering, which are a common issue in laser tracking systems. It is worth noting that a lower error could be achieved using higher quality cameras, at the expense of incrementing the implementation cost. For future work, this system will be adapted to deal with more considerable distances and detect hidden objects on the field that could be implemented in different applications, e.g., telepresence systems like [7].

References

[1] P. Gurevich, J. Lanir, B. Cohen, and R. Stone, "Teleadvisor: A versatile augmented reality tool for remote assistance," in *SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. Association for Computing Machinery, 2012, p. 619–622.

[2] R. A. Karim, N. F. Zakaria, M. A. Zulkifley, M. M. Mustafa, I. Sagap, and N. H. M. Latar, "Telepointer technology in telemedicine: a review," *Biomedical engineering online*, vol. 12, no. 1, p. 21, 2013.

[3] D. Kurz, F. Hantsch, M. Grosse, A. Schiewe, and O. Bimber, "Laser pointer tracking in projector-augmented architectural environments," in *IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 19–26.

[4] N. F. Zakaria, M. A. Zulkifley, M. M. Mustafa, and R. A. Karim, "A review on laser spot detection system based on image processing techniques," *Journal of Theoretical and Applied Information Technology*, vol. 70, no. 2, pp. 333–344, 2014.

[5] M. Meško and Š. Toth, "Laser spot detection," *Journal of Information, Control and Management Systems*, vol. 11, no. 1, 2013.

[6] J. Mrovlje and D. Vrancic, "Distance measuring based on stereoscopic pictures," in *9th International PhD workshop on systems and control: young Generation Viewpoint*, vol. 2, 2008, pp. 1–6.

[7] J. Avalos, S. Cortez, K. Vasquez, V. Murray, and O. E. Ramos, "Telepresence using the kinect sensor and the nao robot," in *2016 IEEE 7th Latin American Symposium on Circuits Systems (LASCAS)*, 2016, pp. 303–306.