Design of a Robotic System to Control the Actions of the People According to Covid-19 Regulations in a Highly Populated Indoor Area

Atakan Durmaz, *Student Member and Chairman, IEEE METU Robotics and Automation Society*
Junior Student in METU Electrical & Electronics Engineering Department

***Abstract*—This paper presents the work done under METU EEE STAR Program 2020 by Atakan Durmaz. Project is primarily based on having a robot which controls the actions of the students by considering COVID-19 regulations,i.e. whether they wear a mask or not, together with if they are keeping the social distance in a overcrowded and such a place where it is hard to keep social distance all the time like the A building of METU EEE Department. Concepts that will be discussed are based mostly on the implementation of the current technological information to a robotic platform by considering performance and having a cost effective solution. In the theoretical part, we will discuss the operation principles of some popular and expensive sensors, such as LIDAR and IMU together with the design and manufacturing of these sensor systems as well as the overall robot design. Another point will be discussed is about the computer vision applications by using an RGB and a thermal camera, which will form the data. While discussing about these vision and data parts, we will also discuss the computational efficiency. At the end we want to obtain solutions containing filtering and sensor data fusion.**

***Index Terms*—Sensor operation, LIDAR sensor design, thermal video processing, RGB-D video processing, sensor fusion, filtering**

I. Introduction

THE current applications to prevent the spread of the Covid-19 virus are checking body temperature of the people at the entrance of the buildings and also, applying tests to the ones that get any close interaction with the people getting positive test results. These precautions were thought by mainly health workers and they are aimed to prevent the spread of the virus , but the limited work force in the health industry and the careless of the people brought the current situation to far worse than the start of the pandemic. Indoor areas have the high probability of infection and the social distance with wearing protective equipment have a significant effect to decrease this risk [1]. To achieve this low probability of infection, we came up with a solution of using a mobile robot which basically takes the role of a patrol for the Covid-19 regulations, namely social distancing and wearing a mask properly. The next thing was the redearchşng the current implementations for that purpose. Generally, currently used systems are immobile and unable to check every people in a building at the same time. For example, at the entrance of the buildings a disinfection cabinets are being currently used to spray disinfectant to the people at the entrance and measuring their body temperatures to letting them in or not [2].

Our solution provides the real time control of everyone whether they are following all the Covid-19 regulations or not in a crowded building such as the A building of the METU EEE Department. In our solution we preferred to use thermal and an RGB-D camera with a 2D LIDAR system. By using thermal camera, humans can be detected since the body temperature and the room temperature can be separable with an optimal thermal camera. Also, by using these data and the depth data from the RGB-D camera, humans can be detected seperately and the social distancing will be checked. If a group of people do not follow the rules, robot will get close to them and give a warning to keep social distance. Also, from the RGB-D camera, it can be checked that if people wear their masks or not, and whether the usage of the mask is appropriate or not. In determination of the mask condition of the people, we will present an algorithm to make our solution efficient and possible to work on the real time system. Basically we will take our image from the camera and on the same image we will select our region of interest as the area around the mouth and the nose to make the computation more efficient. Also, a 2D LIDAR system will be used for a SLAM application. By using a SLAM algorithm, the robot is planned to cover the whole floor and to check everybody whether they follow the regulations.

This regulation check procedure can also be performed by using already existing CCTV camera structure and the data base. However, immobile camera systems have high probability of getting manipulated or unable to check every point on the same area, i.e. having blind spots, or if the detection algorithm cannot decide whether a mask is properly worn, there is nothing to do in such a case. In our solution, we designed our system to operate on a mobile robotic platform and by doing so, if the algorithm cannot decide whether a person wears a mask or not, the robot can easily orbit around the people to get a clear view of that person to make an accurate decision. This is the main advantage of having a mobile platform.

In this paper, the detailed explanations of our custom and cost effective solutions will be discussed. Also, some other requirements for our platform will be explained also, such as power regulator design.

After this point, constructing the hardware and mechanical parts will be handled, then the implementation of the developed algorithms to our robot and optimization will be performed. Since the sensor data fusion of thermal and RGB-

D camera together with a LIDAR system is planned, after the implementation part, and taking data samples sensor fusion will be done and our system can be used in daily operation.

II. Power Regulator Choice

On the robotic platform, motors, microprocessors, many sensors, such as a camera, distance sensors and a LIDAR sensor are needed to be used, the varying power input of these systems must be provided in the robot from the battery. To provide associated input voltage to these systems, power regulators are needed to be used. In this part, a power regulator system designed by using a step-down regulator, namely LM2576 will be examined. This regulator, is planned to be used for the supply of most of the electronics boards namely, Jetson Nano, stereo camera, and distance sensors. As it can be realized that by this regulator, supplying the input voltages of the motors are not planned since the current drawn by the motors may damage this regulator, which has a current limit of 6.9 A, under the room temperature [3].

The first reason why LM2576 step-down voltage regulator is chosen is that, the input voltage of this regulator can be take any value between 7V and the 40V, which gives us flexibility to choose a battery from a wide range. Secondly, as mentioned before, this regulator has a 6.9 A current limit under room temperature. Going over this limit will damage the regulator, in order to prevent going over 6.9 A limit due to fluctuations in the current, the drawn current ideally from the regulator should not be chosen close to this boundary. To achieve this, multiple regulator circuits are needed to supply current to different parts of the system so that the total current passing through the regulator should not exceed the current limit of the regulator.In Fig.1 below, you can find the circuit design for the LM2576 regulator. In the circuit design, there is a cartridge fuse connected at the output terminal with a current limit of 4 A to ensure that regulator operates in the safe current range.

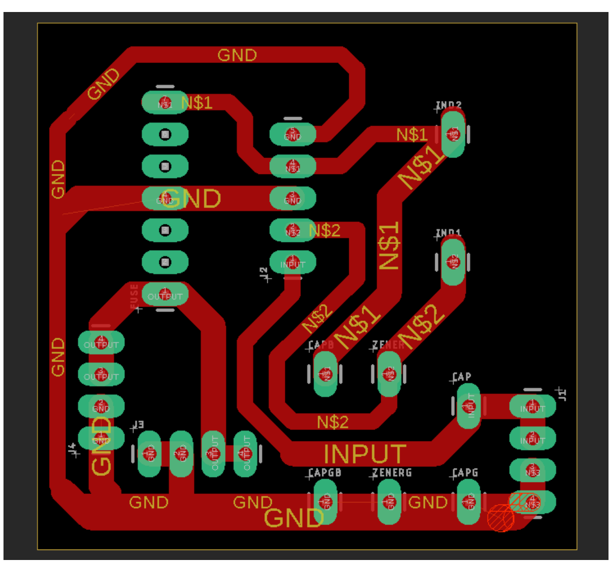

Fig. 1. Circuit design for the power regulator

The second reason of LM2576 may be a good choice as the regulator is that the output voltage is 5 V. That means, it takes input from a broad range of 7 to 40 V and gives constant output of 5 V. While doing this, the regulator does not do this operation by using a resistor to give up the most of the power as linear regulators, instead it is a switching regulator. This reduces the heat produced during operation and the possibility of having any heat related to damages. Also, since there is a step-down operation casing heat dissipation in any way, the power dissipation on the regulator is needed to be checked whether a heat sink is needed or not. The power dissipation equation and the necessary calculations can be seen as follows,

$$P_D = V_{IN} * I_Q + (V_O/V_{IN}) * I_{LOAD} * V_{SAT} \tag{1}$$

,where

$$I_Q = 12mA$$

$$V_{SAT} = 2A$$

$$V_{IN} = 12V$$

$$V_O = 5V$$

$$I_{LOAD} = 3A(average)$$

substituted in the Equation 1, we get the dissipated power as 2.64 W. When we have the copper with area 0.5 square inch, we get

$$Q_{JA} = 50^oC/W$$

with the ambient temperature,

$$T_A = 0^oC$$

$$\Delta T_J = P_D * \theta_{JA} \tag{2}$$

According to the Equation 3 below, since the actual operation temperature exceeds the maximum allowed junction temperature which is 125 ^oC, a heat sink is needed. But as you might realize this is the case for 3 A load current, with less current values satisfying these conditions, there is no need to use a heat sink.

$$T_J = \Delta T_J + T_A = 132^oC \tag{3}$$

This result shows that, there is a requirement for a heat sink to protect equipment from over heating and other damages. This calculation was for the SMD type of LM2576, so with the non-SMD model which includes heat sink, no external heat sink will be required.

III. LIDAR Design

For the mapping and pose estimation of a robotic system, light detection and ranging (LIDAR) sensors are being currently used. These sensors work like a sonar system which performs scanning of the area by using light. The distance is calculated by keeping track of the duration between the pulse and the return. The former represents the time instant when the light beam is sent from the sensor and the latter refers to the time instant when that beam returns back to the sensor.

This duration will be represented as the travel time of the light beam, t_{travel}. Since t_{travel} includes both going and returning of the signal, half of it is required to obtain distance data. So, distance,x can be obtained by the following equation;

$$x = \frac{t_{travel} * c}{2} \qquad , where\ c = speed\ of\ light \qquad (4)$$

Also, since required LIDAR sensors are costly for the particular application, a unique solution is required. To achieve that, we can build a LIDAR sensor, formed as the combination of time of flight(ToF) sensors. as it can be seen in the Fig. 3, when we have n number of ToF sensors represented by boxes, basically we can obtain a partial environmental information. By increasing n, resolution can be improved and the result will converge to a continuous 360 degree information.

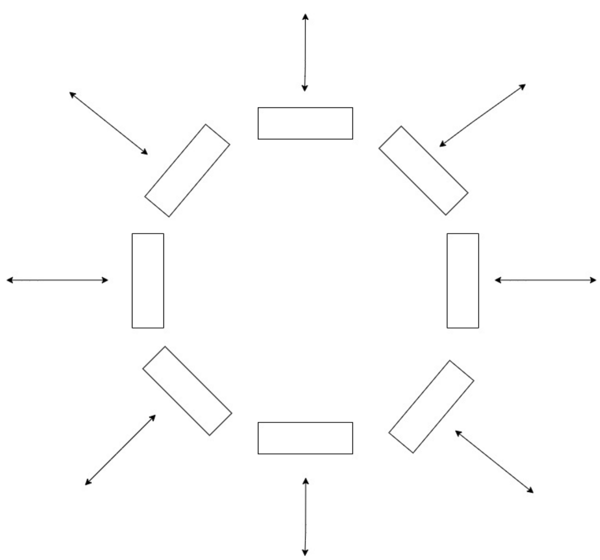

Fig. 2. Positioning of the ToF sensors

However, as one can easily recognize, when n is increased to get a more precise data, cost is also increasing and goes to infinity. On the other hand, the aim of doing this was decreasing the cost and finding a cost efficient solution. Therefore, we need to use something different than increasing n, to improve the precision. To achive high precision for an optimal cost, a step motor is needed. This motor will rotate the ToF sensors with the layout seen in the Fig. 3 above. This ensures that, we will have a continuous data from our sensor rather than discrete data as in the stationary system.Therefore, a rotation action is required.

Unlike the ideal case, in a practical approach we need to think about the errors in the rotation angle and the errors and the distortions in the LIDAR data. To minimize these errors, a feedback information is required. This feedback can be obtained from an inertial measurement unit (IMU), which gives the angular velocity as the output and the angular position is obtained by integrating that over time. By considering also the precision of the IMU sensor and our application, BNO055 IMU sensor will be an appropriate choice.

As mentioned before, we need a rotating ToF sensors and an IMU sensor. The design of the circuit and the 3D design can be seen from the Fig**??** below. Since we have a rotating system and having minimum connection with the outside will be good for the maintainability of the system, also by considering the fact that BNO055 IMU sensor uses I2C protocol, using VL530X ToF sensor, which is also using I2C protol and requires total of 4 connections, namely V_{in}, GND, SCL and SDA, will be a smart choice for our application. These connection are represented as blue circles in the Fig. 3 below. And the other header connections are for the ToF sensors and the IMU sensor. Since the motor will be at the bottom as in Fig. 4 and not to overload system with creating high inertia at the rounding plate, battery is not included to the circuit schematic in Fig. 3.

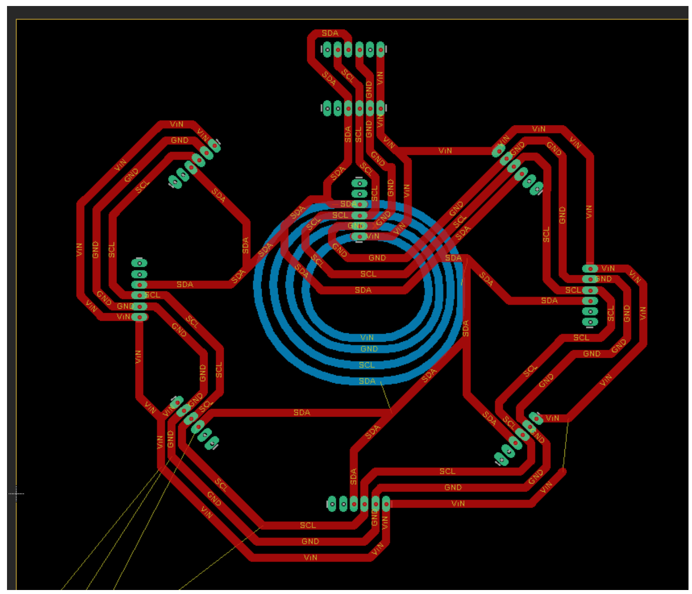

Fig. 3. Circuit design for the LIDAR system

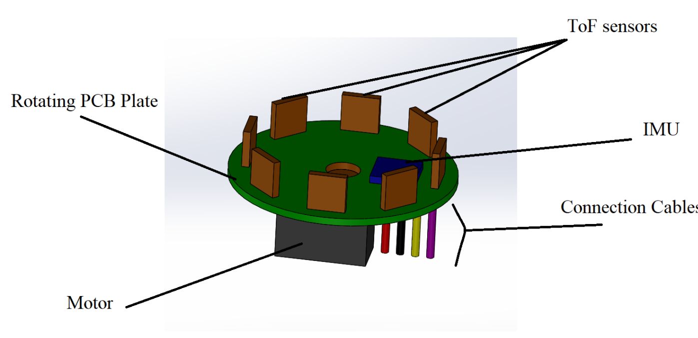

Fig. 4. 3D design for the LIDAR system

This system will give us an output in the form of a 1D array from the ToF sensors containing the all distance measurements from all ToFs and a 1D array from IMU, in the form of d_i's, where i is ranging from 1 to the number of ToF sensors. However, from the outputs of the IMU, only yaw angle will be used, so we will use only that value. Number of ToF sensors is important for determination of the precision, since we have a rotating sensor system for a given period T, time span between the updates of the data, t_u, can be found by the following;

$$t_u = \frac{T}{number\ of\ ToF\ sensors} \tag{5}$$

In this particular application number of sensors are chosen as 8, which gives us an update time $t_u = T/8$. The reason of this update time is that, while rotating sensors, some time passes after a measurement in any direction, arbitrarily name this direction as $+\hat{a}_x$. After k^{th} sensor takes measurement in this direction, t_u time passes bwtween this measurement and the next mesuremnt will be done by $(k+1)^{th}+1)$ sensor. In t_u time $(k+1)^{th}+1)$ sensor gets aligned with $+\hat{a}_x$ and takes a new measurement. Therefore, as t_u decreases, which means an increase in the number of ToF sensors, our somehow blind time decreases. Therefore, if t_u is shorter we can know sudden changes in the environment instantly and get necessary action according to that. The data obtained from this LIDAR system will be send to a ROS package; therefore, it will be easy to work on our controller which runs the ROS environment on Ubuntu Mate 16.04 operating system [4].

IV. Video Processing and Interpretation

To regulate Covid-19 regulations, namely wearing a mask and keeping social distance, a visual input and processing of this input is required. Also, to measure the body temperatures of the people at the environment, we are going to use a thermal camera.

Firstly, for the control of social distancing, again the LIDAR data can be used, but it may be difficult to detect humans in such a dynamical environment such as the A Building of the EEE Department at METU. Even if to use a filter, again there will be a requirement of a visual data, which can be obtained from a camera only [5]. While detecting social distance, we can use a thermal camera. This requirement is due to another issue about the body temperature check for every person. Since we need to check body temperatures of the people we have to use a thermal camera. By using this thermal data we can also determine the social distancing. This can be done by the usage of a simple threshold,i.e. a simple filtering to the thermal data. Thermal cameras can give output as the form of a gray scale image and this image is a matrix containing values from 0 to 255. Depend on the specs of the camera and the distance of the object from the camera, output value may differ [6]. To determine a human we can use the body temperature as a reference. Since the room temperature is around 298 K and the human body at its regular conditions is around 309 K, we can use this difference to detect humans.During this process, there may be some possible errors due to clothing, or more importantly a person can hold a cup of hot coffee or tea, which is around 335 K. If a simple filter [7] is applied to the system, we may get false negatives and false positives with high probability, which is undesired. To minimize these, we can use a filter which performs the weighted sum of this temperature data over a human area on the image. Thus, the false alarm probability will be less compared to a simple filter. Therefore, we need to check the average of this weighted sum over the human image an if the value exceeds 311 K, to be sure about the well being of the person, the robot can get closer to the person and take another measurement just from the forehead this time. Depending on the result, the robot can take action. On the other hand, we have discussed the human detection only and have not mention on the social distancing. The social distancing value may be obtained also from, this thermal camera by using monocular or stereo image depth cues or also using a Kalman filter to sensor data fusion with the thermal image and the depth info that will be obtained from either an RGB-D camera or a stereo camera [8] [9] .

Secondly, by detecting faces, we are aim to check whether people are wearing a mask or not. There are some data sets used for pedestrian detection used in the autonomous drive area[10]. However, in our study, we chose to use a previously prepared face and human detection data sets. Since these are highly worked and very efficient systems to work on a simple development board like Jetson Nano by using OpenCv library [11], we choose to directly implement these data sets. Now, to check mask, if we use the YOLO algorithm, it will bring a lot of computational complexity and it is a very computatonal heavy solution, which is not appropriate on our control board[12].

To be able to have a computational efficient solution, we can use these previously prepared face detection data set. Since face detection is much easier than detecting masks, first we need to use face detection algorithm. After that, on a 2D image we have the region of interest, which contains the face. Then, we now have a smaller image and less data, which is easy to work on. After obtaining this region of interest, containing a face, again we need to use a previously prepared eye detection data set. Again, since this is a highly worked problem, the solution is efficient. However, if we run these two separately, we have lags in our output, which means we have an inefficient solution. To make it efficient, we can run the eye detection only on the region of interest containing face. This comes from the Bayes' Rule. Since we know that we should only detect eyes on a face rather than floating on the around, it is sufficient to check face regions on an image to find eyes.

After that, we now have face region and two eye regions. To detect a mask, we need to look at the mouth region, which is under the eyes. By using this simple idea, we can again create a new region of interest from the face region by taking the region between the bottom of the face region and the bottom of the eye regions. By doing so, we decreased the amount of data needed to be processed, which gives us a far more efficient solution rather than searching in the whole image. A result of a trial can be seen as below in Fig. 5.

In the Fig. 5, as you can see the face and eye areas are shown by red and green boxes, respectively. Also, there is another yellow box, showing the area, which is the area between the bottom of the eyes and the bottom of the face region of interest. This yellow box points the area where whether a mask should be for an appropriate usage of a mask. This image is just a photo of a trial and according the obseravtions, the algorithm is quite efficient up to this point and making it possible to use in a real time system.

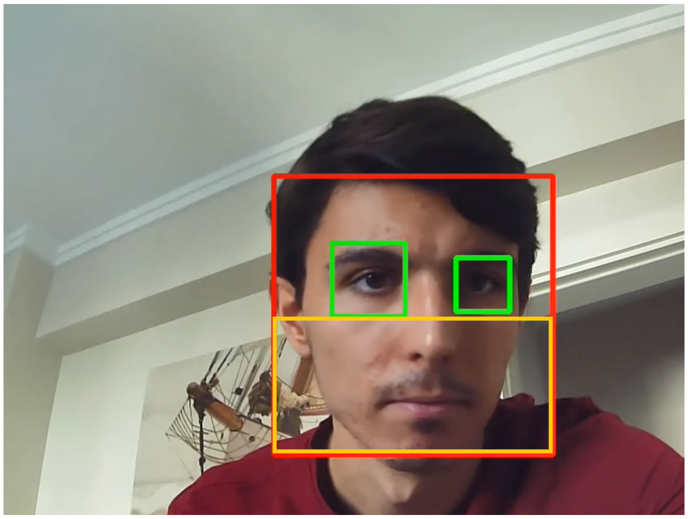

Fig. 5. Face and eye detection with the mask area determination trial result

V. Conclusion

In summary, this paper addresses the issue of controlling actions of the people according to Covid-19 regulations in a highly populated area by using a robotic system, namely EEE Department A Building in METU. The foreseen problems in this processes are due to the limitations of the thermal data as explained in the Video Processing and Interpretation part detailedly, namely false readings caused by hot objects and some heat absorbers. Then, we come up with a solution of filtering. Next, we have discussed about detecting social distancing by using thermal image and depth info by using a sensor data fusion. After that, we have talked about the efficiency of mask detection algorithm and how to optimize this algorithm to operate on a real time system.

On the other hand, we have discussed on the power regulator choice for our system according to our requirements. Finally, we have designed a LIDAR sensor system, which may be a cost efficient solution for our system.

VI. Future Work

The first thing that will be handled is building our hardware and mechanical structure. After that, implementing these algorithms and systems on this physical system is planned to be achieved and the behavior of the robot in a real environment will be observed. On the other hand, by taking data samples and performing our sensor data fusion, the control algorithm will be improved to make our robot interactive with people and to move in a dynamical environment without giving any error. Then, according to our test results together with measurements, other possible custom solutions or applications can be applied to optimize the overall system.

Acknowledgment

The author would like to thank Assist. Prof M. Mert Ankaralı and Prof. İlkay Ulusoy at the Electrical and Electronics Engineering Department in METU for the academic guidance.

References

[1] C. Sun and Z. Zhai, “The efficacy of social distance and ventilation effectiveness in preventing covid-19 transmission,” *Sustainable cities and society*, vol. 62, p. 102390, 2020.

[2] K. P. Rane, “Design and development of low cost humanoid robot with thermal temperature scanner for covid-19 virus preliminary identification,” *International Journal*, vol. 9, no. 3, 2020.

[3] T. Instrument, ““lm2576/lm2576hv series simple switcher® 3 a step-down voltage regulator,” *LM2576 datasheet, Jun*, 1999.

[4] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, “Ros: an open-source robot operating system,” in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[5] G. Welch, G. Bishop, *et al.*, “An introduction to the kalman filter,” 1995.

[6] Y. Abdelrahman, P. Wozniak, P. Knierim, N. Henze, and A. Schmidt, “Exploration of alternative vision modes using depth and thermal cameras,” in *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, 2018, pp. 245–252.

[7] P. S. Diniz *et al.*, *Adaptive filtering*. Springer, 1997.

[8] S. Lee, D. Har, and D. Kum, “Drone-assisted disaster management: Finding victims via infrared camera and lidar sensor fusion,” in *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, 2016, pp. 84–89.

[9] G. K. Humphrey and P. Jolicoeur, “An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification,” *The Quarterly Journal of Experimental Psychology Section A*, vol. 46, no. 1, pp. 137–159, 1993.

[10] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, “Adversarial sensor attack on lidar-based perception in autonomous driving,” in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 2267–2281. [Online]. Available: https://doi.org/10.1145/3319535.3339815

[11] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. ” O’Reilly Media, Inc.”, 2008.

[12] R. Huang, J. Pedoeem, and C. Chen, “Yolo-lite: a real-time object detection algorithm optimized for non-gpu computers,” in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 2503–2510.