Vision-based Approach for Automated Social Distance Violators Detection

Abdalla Gad*, Gasm ElBary*, Mohammad Alkhedher§, Mohammed Ghazal*, *SMIEEE*
*Department of Electrical and Computer Engineering**
Department of Mechanical Engineering§
Abu Dhabi University
Abu Dhabi, United Arab Emirates
mohammed.ghazal@adu.ac.ae

***Abstract*—Social distancing is a necessary precaution measure taken in order to have more control over the outbreak of infectious diseases such as COVID-19. Most of Social distancing monitoring approaches are based on Bluetooth and mobile phones that require an app to be downloaded on all phones. This paper proposes a different approach to monitor social distancing, using cameras, and combining different computer vision algorithms. The approach utilizes the concept of inverse perspective mapping (IPM) together with the camera's intrinsic information to produce a bird's eye view with real-world coordinates of the frame being processed from a video source. The process starts with image enhancement, foreground detection using Gaussian Mixture Model (GMM) background subtraction, tracking using Kalman filter, computing real-world distance measurements between individuals, and detecting those who have been in less than 2 meters apart as they are considered to be in contact. This tool could assist the efforts of the governments to contain the virus. It can be implemented in closed areas or institutions, monitor the extent of people's commitment, and provide analysis and a faster approach to detect possibly corona suspicion cases. The approach is tested on the task decomposition data set, which included frames of closed areas and the camera's intrinsic parameters. Another data set was created with different scenarios to increase the confidence level of our algorithm. The results showed the success of our approach in detecting the violation in social distancing with accurate measures of the real-world coordinates.**

***Index Terms*—Social Distancing, COVID-19, GMM, Kalman Filter, Inverse Perspective Mapping, Bird's Eye View**

I. Introduction

COVID-19 pandemic has risen as a common enemy for the whole of humanity, killing people and destroying economies. The pandemic effects will remain for years, and the recovery of the communities will take time. While the epidemic has lasted for months until now and everyone is suffering, it's not clear when the virus will be cured that if it will. Economies are being opened again, and everyday life activities are resumed, but this may eventually lead to another wave of the virus if the personnel didn't abide by social distancing and safety regulations. Therefore, being able to measure the extent of the people's commitment in some areas is required. Hence, new means of contact tracing and reporting are needed to be developed to fulfill this objective. This tool will serve a significant role in providing analytical data for the authorities to measure the levels of commitment of people in some locations and provide early warnings to take actions on the uncommitted private and public institutions and markets.

In the World Health Organisation (WHO) definitions, a contact in terms of COVID-19 is defined as either direct contact to a COVID-19 case or being within 1-meter of a case for more than 15 minutes [1]. The process of contact tracing helps control diseases and build a commitment map, which will help in a safer steady recovery. Contact tracing methods were always introduced whenever a pandemic appears, such as the H1N1 and SARS [2], [3]. Similarly, now COVID-19 appeared, which is very contagious, and contact tracing is needed more than before.

Some different technologies and methods are used to achieve contact tracing. A technique used is AI-based, where a video stream is processed, and the algorithm detects the people who violate the social distancing regulations. *Smartvid.io* developed a tool that notifies about social distancing in working sites. *Megvii's* Contactless Screening makes a combination between normal cameras and thermal cameras so detect not only social distancing violations but also report possible cases of COVID-19 [4]. Other methods include mobile apps that adapt short-range wireless technologies like Bluetooth and GPS [5].

However, most of these methods face some challenges like accuracy, cost, privacy, availability, or power consumption. Even though the AI-based approaches are great solutions, they are costly in terms of training and collecting data, and they require high computational units, which will limit its availability. The App-based approaches are more common, but they might lack accuracy due to technical limitations by the user's phone processing unit's performance or the Bluetooth models themselves who have a limited number of devices they can communicate with simultaneously. Privacy issues arise if the collected data are not safely encrypted, as sharing the contact history data will violate the user's privacy to unauthorized parties. Add to that the issue of significant power consumption caused by the GPS and Bluetooth being turned ON all the time. Moreover, not all users may have Bluetooth or GPS on their phones, limiting the uses of these

© IEEE 2021. This article is free to access and download, along with rights for full text and data mining, re-use and analysis.

apps to certain users only or of a specific group of users [6].

However, the approach proposed in this paper is affordable by the different private and governmental institutions as it can be implemented to the current surveillance cameras, unlike the previous methods that are costly or require some authorities that are given to governmental agencies. Moreover, this approach can be used indoors or in places (i.e., factories, hospitals, etc.) where noise affects mobile phones' performance.

This paper uses the IPM model to estimate the distance between objects using a single camera. The inverse perspective mapping approach is described in [7], where the method is used for mapping roads in a bird's eye view. In [8]–[11], IPM-based distance estimation approaches were proposed. However, most of the applications of distance estimation and motion predictions are in the advanced driver assistance systems (ADAS). Some of these applications are proposed in [10], [12].

II. Proposed Method

The first step is to extract the camera's intrinsic information for capturing the scene by using the approach described in [13]. A model of a checkerboard of known square dimensions is used in order to scale the pixels in the image captured by the camera to compute the focal length, principal point, etc. Figure 1 shows the model of the checkerboard used and the successful detection of the points with a low mean error.

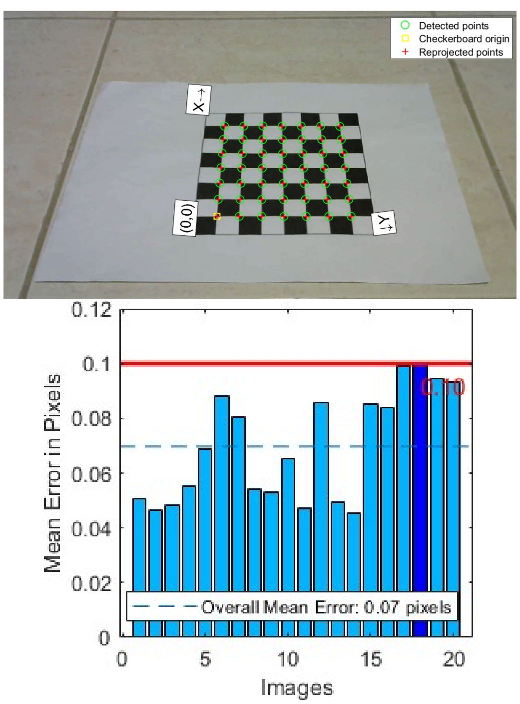


Fig. 1: Detection and projection of checkerboard squares points

These parameters, together with the camera's lens' location in the real world, including the height from the ground level in the scene, the yaw, and the pitch, are used in the inverse perspective mapping process to produce a bird's eye view of the image. Figures 2 and 3 show 2 sample frames and their respective bird's eye view transformation computed by inverse perspective mapping for the 2 data sets used. It can be observed that the centroid of the foreground region (people) is not identifying its location as the person's height is effecting the projection and leads to converging. For this reason, the locations of the foreground regions are specified by the minimum center point towards the lens of the camera as it can be taken at ground level and hence provide more accurate results of the location in real-world coordinates.

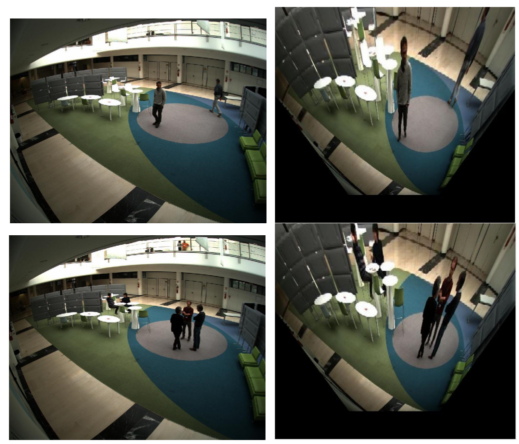


Fig. 2: Sample frames of the task decomposition dataset and their respective bird's eye view transformation

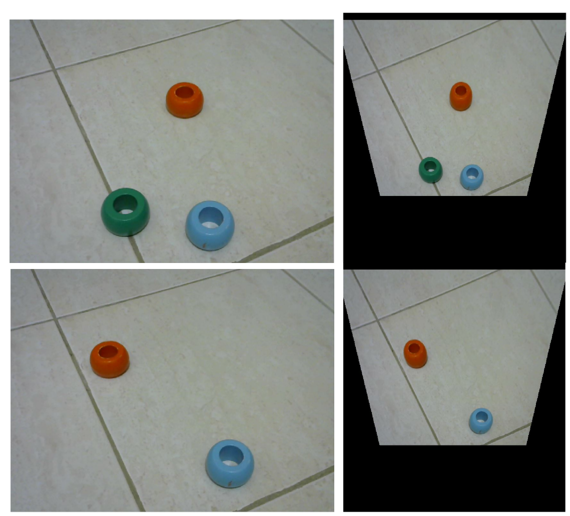
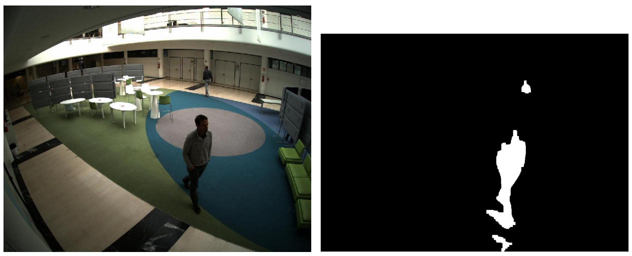
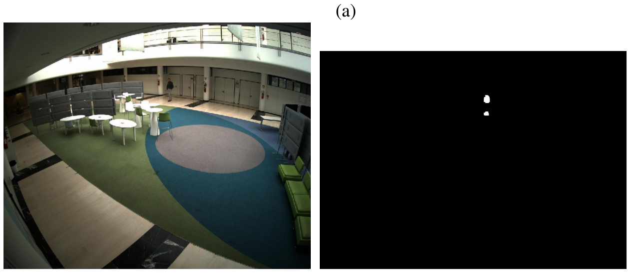


Fig. 3: Sample frames of the custom dataset and their respective bird's eye view transformation

The original input frame is subjected to a Gaussian smoothing filter to reduce the noises in the image. Then GMM background subtraction is applied to detect the foreground objects, which in our case is represented by the individuals moving through the frames captured by the camera. The GMM modeling of the background is introduced in [14], where the statistical distributions of pixels through multiple frames are used to obtain the strongest weighted intensity levels of pixels that have been detected consistently multiple times. These pixels are used to model the background, which is subtracted from the original scene to obtain the difference, and that is the foreground. Figures 4 and 5 show sample frames and the foreground regions using the GMM background subtraction.



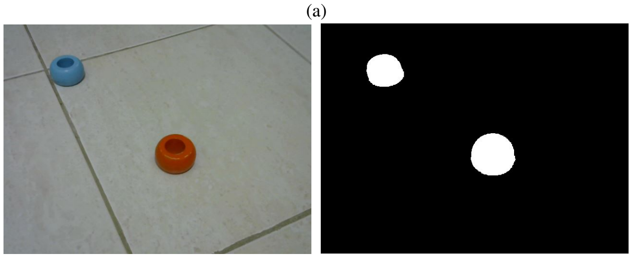
(a)



(b)

Fig. 4: Sample frames of the task decomposition dataset and their foreground segmentation using GMM

(a)



(b)

Fig. 5: Sample frames of the custom dataset and their foreground segmentation using GMM

One drawback of the GMM foreground detection approach is the inability to detect foreground objects that have stopped moving due to the constant updating of the background model, which gives it the advantage of being adaptive to environmental conditions such as illumination variation. To overcome this disadvantage and have unique tracking of different objects in the scene over multiple frames, the approach introduced in [15] is applied where a Kalman filter is used to predict the location of the object in the next scene using constant velocity model and gets updated with every true value obtained till the prediction matches the truth. Figures 6 and 7 show unique tracking of objects in multiple scenes using Kalman filter.

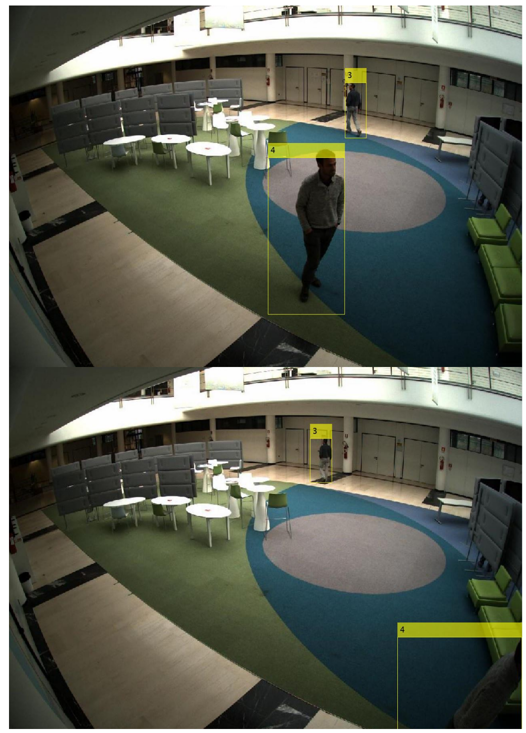


Fig. 6: Unique tracking of objects in multiple frames of the task decomposition dataset

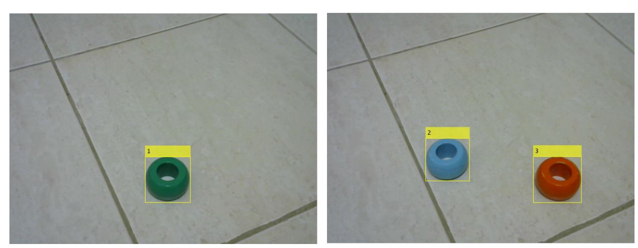


Fig. 7: Unique tracking of objects in multiple frames of the custom dataset

The location in pixels of the objects detected and tracked together with the bird's eye view model obtained previously are used in the projective transformation algorithm to obtain the locations in real-world coordinates [13]. The projective transformation approach utilizes the triangular equations to project the locations of the pixels of the foreground region into their locations in real-world coordinates. Finally, the euclidean distances between the centroids of all the regions projected are computed and compared to a threshold of 2 meters, which specifies that two people have been in contact or not, as shown in figures 8 and 9.

Fig. 8: locations and distances in real world coordinates of the task decomposition dataset

Fig. 9: Locations and distances in real-world coordinates of the custom dataset. Where intentionally, 10cm in real-world is treated as 1m.

To verify that the camera's calculated real-word coordinates and distances are accurate, a comparison between the calculated and measured measurements is performed. The comparison was performed on the distances between the items in 9. The actual distance is obtained using a ruler and compared with the distance calculated through the algorithm. Table I shows the comparison, and since the error was below 5%, retrieving the real-world distances was successful.

TABLE I: Comparison between the calculated and the measured distances between the objects in 9

Distance #	Calculated Distance (Algorithm)	Measured Distance (Ruler)	Error %
Distance 1	12 cm	11.5 cm	3.45 %
Distance 2	9 cm	9.2 cm	2.17 %
Distance 3	18 cm	17.4 cm	1.69 %

Figure 10 shows the flow chart of the overall proposed approach.

III. Experimental Testing and Results

Figure 11 shows the results of the proposed approach in detecting violations in social distancing for individuals who have been in less than 2 meters apart. It can be observed that the inverse perspective mapping approach utilizing the camera intrinsics parameters showed success in accurately identifying the real-world coordinates. It can be observed a limitation in the algorithm in detecting overlapping people and keep track of them. This disadvantage can be overcome by multiple cameras having different views of the same location and coordinating to provide better detection and projection into real-world coordinates.

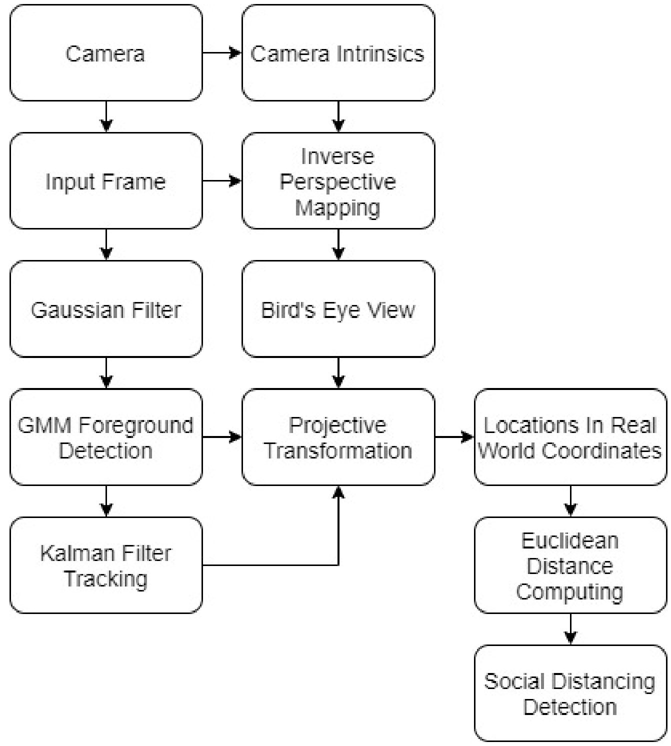


Fig. 10: Flow chart of the proposed method

A. Limitations and Future Work

The proposed algorithm aims not to prevent social distance violations but rather to decrease its excessive occurrence. People would always be in company when spotted in public places, which would raise a lot of warning signs on our system. But the increase of the distance violations in specific regions such as cashier lines should necessitate an action. The proposed approach, together with a face recognition algorithm, can simplify the process of contact tracing. The information about each individual found in public places and the people who have been in contact with are stored on a database and accessed when there has been an infection. This can speed up the process of contact tracing and give better control over the spreading of the disease.

IV. Conclusion

In conclusion, the paper proposed a new approach to detecting social distancing violations using inverse perspective mapping and camera intrinsic parameters. The results showed

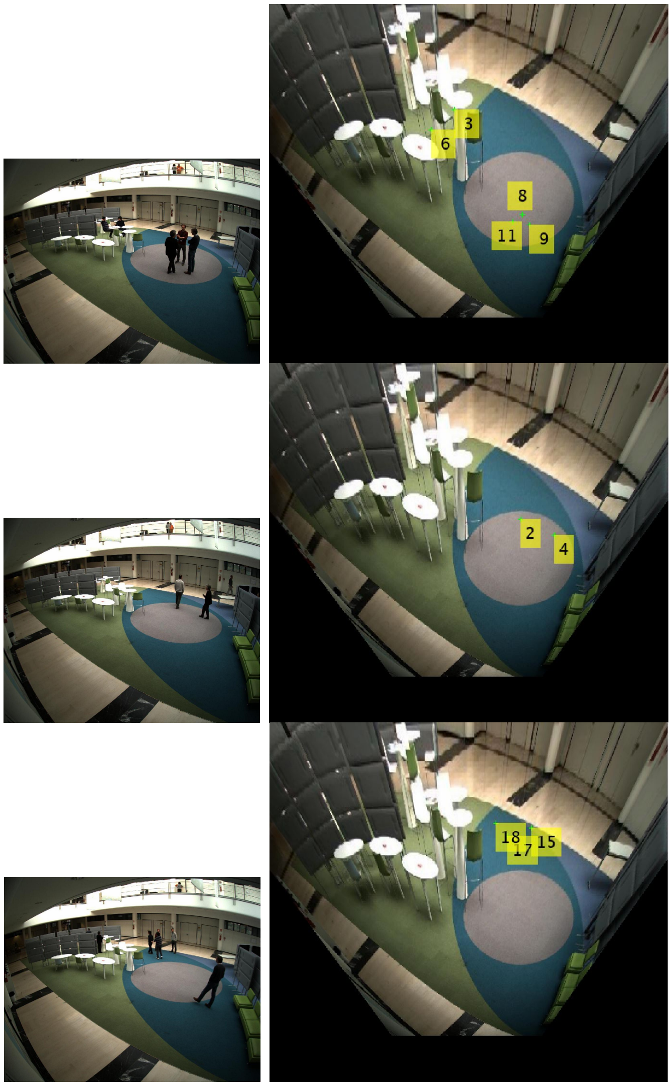


Fig. 11: Detection of social distancing violation

our algorithm's success in obtaining accurate measurements of distances and objects' locations in real-world coordinates. The proposed method can be utilized in closed areas to ensure that social distancing has not been overlooked or violated and hence provide better control over the transmission of infectious diseases.

REFERENCES

[1] W. H. Organization, "Contact tracing in the context of covid-19," May 2020.

[2] V. Lampos and N. Cristianini, "Tracking the flu pandemic by monitoring the social web," in *2010 2nd International Workshop on Cognitive Information Processing*, 2010, pp. 411–416.

[3] K. Leong, Y. Si, R. P. Biuk-Aghai, and S. Fong, "Contact tracing in healthcare digital ecosystems for infectious disease control and quarantine management," in *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies*, 2009, pp. 306–311.

[4] R. Sagar, "How computer vision came in handy for social distancing," https://analyticsindiamag.com/covid-19-computer-vision/ , June 2020.

[5] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the covid-19 pandemic and the role of iot, drones, ai, blockchain, and 5g in managing its impact," *IEEE Access*, vol. 8, pp. 90 225–90 265, 2020.

[6] T. Altuwaiyan, M. Hadian, and X. Liang, "Epic: Efficient privacy-preserving contact tracing for infection detection," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.

[7] J. Jeong and A. Kim, "Adaptive inverse perspective mapping for lane map generation with slam," in *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, 2016, pp. 38–41.

[8] R. Adamshuk, D. Carvalho, J. H. Z. Neme, E. Margraf, S. Okida, A. Tusset, M. M. Santos, R. Amaral, A. Ventura, and S. Carvalho, "On the applicability of inverse perspective mapping for the forward distance estimation based on the hsv colormap," in *2017 IEEE International Conference on Industrial Technology (ICIT)*, 2017, pp. 1036–1041.

[9] A. Bharade, S. Gaopande, and A. G. Keskar, "Statistical approach for distance estimation using inverse perspective mapping on embedded platform," in *2014 Annual IEEE India Conference (INDICON)*, 2014, pp. 1–5.

[10] S. Tuohy, D. O'Cualain, E. Jones, and M. Glavin, "Distance determination for an automobile environment using inverse perspective mapping in opencv," in *IET Irish Signals and Systems Conference (ISSC 2010)*, 2010, pp. 100–105.

[11] P. Wongsaree, S. Sinchai, P. Wardkein, and J. Koseeyaporn, "Distance detection technique using enhancing inverse perspective mapping," in *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*, 2018, pp. 217–221.

[12] A. Awasthi, J. K. Singh, and S. H. Roh, "Monocular vision based distance estimation algorithm for pedestrian collision avoidance systems," in *2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, 2014, pp. 646–650.

[13] A. Bevilacqua, A. Gherardi, and L. Carozza, "Automatic perspective camera calibration based on an incomplete set of chessboard markers," in *2008 Sixth Indian Conference on Computer Vision, Graphics Image Processing*, 2008, pp. 126–133.

[14] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," *Proceedings of IEEE Conf. Computer Vision Patt. Recog, vol. 2*, vol. 2, 01 2007.

[15] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam, and S. S. Aher, "Moving object tracking using kalman filter," in *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)*, 2018, pp. 544–547.