Camera-Based Log System for Human Physical Distance Tracking in Classroom

Somrudee Deepaisarn[a,1,*], Angkoon Angkoonsawaengsuk[a,2], Charn Arunkit[a,3], Chayud Srisumarnk[a,4], Krongkan Nimmanwatthana [a,5], Nanmanas Linphrachaya[a,6], Nattapol Chiewnawintawat[a,7], Rinrada Tanthanathewin[a,8], Sivakorn Seinglek[a,9], Suphachok Buaruk [a,10], Virach Sornlertlamvanich [b,c,11]
[a]*Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, Thailand*
[b]*Faculty of Engineering, Thammasat University, Pathum Thani, Thailand*
[c]*Faculty of Data Science, Musashino University, Tokyo, Japan*
[1]somrudee@siit.tu.ac.th, [2]6222771725@g.siit.tu.ac.th, [3]6222770669@g.siit.tu.ac.th, [4]6222781922@g.siit.tu.ac.th, [5]6222770966@g.siit.tu.ac.th, [6]6222780536@g.siit.tu.ac.th, [7]6222781914@g.siit.tu.ac.th, [8]6222782037@g.siit.tu.ac.th, [9]6222780460@g.siit.tu.ac.th,[10]d6522300067@g.siit.tu.ac.th, [11]virach@gmail.com
* Corresponding Author

***Abstract*—In the pandemic of COVID-19, the indoor physical distancing protocol has been one of the recommendations for people to avoid close contact with each other in order to prevent contagious clusters. This paper proposes an end-to-end camera-based human physical distancing recording system for an indoor environment, specifically, a classroom. The recording system aims to automatically trace the locations of persons and the directions of their movements in a classroom, also with respect to the on- and off-seat activities. No identity of persons is kept in the recording log system, but locations of individual persons at each timestamp are obtained; hence, the spatial and temporal distribution can be studied further. In this paper, we illustrate the overview workflow of the human and seat detection as well as the log system storing human physical distancing actions.**

I. Introduction

The pandemic of a novel coronavirus began in Wuhan, Hubei, China, in late December of 2019. It quickly spread worldwide in a short period of time [1], where Thailand announced its first COVID-19 case as early as the first month of 2020. Thailand has been suffering ever since then. There were subsequent state quarantines due to the constant outbreak within Thailand [2]. At the peak of the pandemic, Thailand had almost thirty thousand confirmed cases of COVID-19 in one day, and the number has drastically gone down, but there are still cases all around in the country [3]. COVID-19 has left devastating wounds and resulted in turnarounds in many sectors of Thailand. The economy has significantly shrunk and slowed down to the point that it can cause severe recessions as after-effects of the pandemic due to decreased social interactions such as travelling. People have lost their jobs and incomes, which in turn cause a chain effect in the decrease of household consumption [4]. Moreover, in the educational sector of Thailand, workers and students have never prepared for such situations before. This requires many adaptations for all academic staff and institutes to manage learning during the event of a new normal [5].

The way people contact others has changed since the wake of the COVID-19 pandemic. This was first incurred by the global and nationwide lockdowns of most countries around the globe, people were forced to migrate to online space [6], especially for workers who have to remotely work from home. The characteristic of the new social culture is the so-called ”New Normal” [7]. The governmental strategic policy includes nationwide lock-downs in many countries, giving general guidelines for self-care and home isolation, and promoting physical distancing schemes for people to follow [8]. Physical distancing requires that the maintaining distance between persons must be at least 1.8 meters or 6 feet apart and the interactions should not include any kind of physical contact [9]. Similarly, there should not be a large number of people in a public space, especially in tight or closed-area [10]. However, in practice, it is difficult for people to stay in proper physical distancing at all times, especially in public spaces. Even though it is a set of non-pharmaceutical and public health interventions intended to prevent the aerosol spread of the novel Coronavirus , the activities between people and behaviors have been closely integrated and required people to physically engage with each other at a close distance. Although the practice of physical distancing is extremely crucial [11], people still neglect its importance and omit to give information to medical personnel. Regardless, the concerns of public monitoring come with manual labor and privacy concerns in order to constantly watch the monitoring feedback. Therefore, we proposed a certain end-to-end system to detect and save information for medical staff to utilize for further prevention and detection.

In this paper, the concept of the physical distancing system will be illustrated in the purposed topic of Camera-based Classroom Seat and Person Physical Distancing Log System Application. The application covers from the start of detecting the seat matrix location coordinate within the classroom. The seat detection system is proposed by the application of a perspective transformation along with the extraction of contour color on the top part of the lecture table. On top of that, the person and referring to the person with his or her location and whereabouts in the physical distance calculation

for determining whether he or she is in the proximity of the virus spreading by applying the purposed ratio calculation between window frame size and referenced room dimensions. The physical distancing calculation system was modified using YOLOv3, a real-time object detection system framework [12] to accurately detect activity inside the classroom environment as well as the bird-eye view interface that we used in the further development in order to satisfy the indoor physical distancing monitoring purposes.

Regarding the implementation after the detection work which is directly related to the storage management as well as privacy concern on video camera [13], we proposed the application of dictionary data structure in order to store the data in a more efficient method in terms of storage space, privacy concern and physical distancing protocol comparing to directly storing raw footage video data [14]. The developed dictionary recording system will be stored the distancing activity as a log file which can be used and manipulated for further analysis in future work.

II. Related work

A. Perspective Transformation

Perspective Transform, in the context of image processing, is the methodology that is normally implemented to alter one's perspectives to a different perspective dimension in space, e.g., from a camera viewing angle into a bird-eye viewing angle [15]. The method deals with the conversion of perspectives of an image and it allows a perception of the image from a different field of view. For our case, perspective transformation disregards the problem of people blocking the camera and it grants a better viewing angle for us to process the image detection [16]. The perspective transformation presents an alternate way of satisfying the need for depth estimation and approximating the distance between people as well as allowing the top-view angle for determining precise locations [17]. In this paper, this is done using the function provided by OpenCV open-source tool [18].

B. Person Detection

In this case, we utilized the "You Only Look Once" Version 3 algorithm (YOLOv3) for real-time person detection. It is an algorithm that uses features learned by a fully-connected convolutional neural network to detect an object. Objects detection or person detection has been crucial for many real-world applications and YOLOv3 can perform greatly on the task since it runs faster and more effectively compared to other detection methods. This can be referenced from many applications such as surveillance and security or even counting objects based on recognition and detection [19]. For example, a pedestrian detection system that utilizes the algorithm of RT-YOLOv3 or Real-Time YOLOv3 [20]. Most of these applications have been applied to detecting human and object in outdoor environments.

C. Contour Extraction

OpenCV is a Python open-source computer vision and machine learning software library. It was built with the main purpose of functioning in real-time computer vision processing [21]. In the library provided by OpenCV, there are many functions available for the users to utilize and implement. Contour detection grants us the ability to detect the borders of objects, and localize them easily in an image [22]. This could be one of the first pre-processing steps for analysing images. We, therefore, adopted the use of the contour detection technique in our pipeline, allowing efficient and fast debugging when scanning all seats in a classroom scenario.

III. System Pipeline

Our proposed system is an end-to-end camera-based human physical distancing recording system. The system itself takes a video file as input and stores the physical location of the seat and people in every fixed timestamp. The system will keep the coordinates of seat and person locations in a log file in the local storage, instead of keeping the video data to reduce the storage capacity [23].

A. Seat Detection

To perform the detection of each seat in a room, we applied the principle of contour color extraction [24] on the top part of the seat/table as in Fig. 1 which shows the robust feature extracted as its unique contour shape. The result of the contour detection was proven to work well for the first rows of the seat. Meanwhile, the rest of the rows still could not be detected that well.

The perspective transformation helps transform the projection of the original image into a new visual plane that can be suitable for our requirement [15]. The equation of the perspective transformation is defined in Equation (1).

$$\left[x', y', w'\right] = \left[u, v, w\right] \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{1}$$

The original image is represented by u and v and is also corresponding to the perspective transformation of the image coordinates x and y, where $x = x'/w'$ and $y = y'/w'$.

We, therefore, applied the perspective transformation by selecting the region of interest (ROI) which, for this work, is the top seat area of every seat in the room before performing the seat contour detection to enhance the detection ability of the system.

B. Person Detection

Detection of a person, in our case, is designed to perform using the well-known YOLOv3 since it has the ability to detect objects based on the trained weights and data set [25], [19]. We implemented the algorithm with the weighted and pre-trained model of which we specifically selected the function that can be utilized to perform human body detection as in Fig. 2. In the standard pre-trained YOLOv3, the bottom center of the detection box on the human body was defined to be

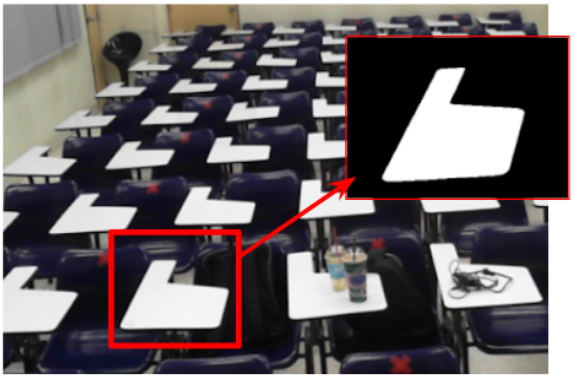

Fig. 1. Detection of the white top-seat table contour.

a reference point of the person's location. In this work, we improved the body reference point to suit our environment by firstly calculating the height-to-width ratio of the box rounded to the detected human object. The ratio is denoted as $Q_{\text{Box}} = h_{\text{Box}}/w_{\text{Box}}$. Where h_{Box} and w_{Box} represent the height and the width of the box, respectively. As a result of our experiments, the value for Q_{Box} that is greater than or equal to 1.6, refers to a standing person in the box; otherwise, the person is sitting down (on a seat). After we had determined the constant threshold for classifying standing and sitting persons based on the dimensions of the detected box, we then defined the two criteria for adjusting the new reference point of the box accordingly.

Therefore, we modified the algorithm that is used to identify the location of the person in each frame. Instead of using the conventional point at the bottom-middle point of the detected human body, we use the appropriate reference point c which can take values of either 0.775 or 0.675 in the case of a person standing or sitting, respectively. That is the value for the vertical proportion from the top-middle point of the box, where the reference point is located. The standing and sitting scenarios are illustrated in Figs. 3(b) and 3(c) respectively, compared to Fig 3(a) where the conventional algorithm is applied to determine the reference location of a person. Such reference points for standing and sitting cases were conditionally given by Equation (2).

$$c = \begin{cases} 0.775; & \text{if } Q_{\text{Box}} \geq 1.6 \\ 0.675; & \text{otherwise} \end{cases} \tag{2}$$

Therefore, this proposed condition and calculation approach was then used to improve the accuracy of the persons' locations in the classroom environment.

C. Physical Coordinate Calculation

After determining the seat and person reference point coordinates, we performed the calculations to turn the general image coordinates of the detected object to calculate physical coordinates. The physical coordinate calculations were performed by taking the two required reference parameters quoted from room physical dimensions. These reference parameters

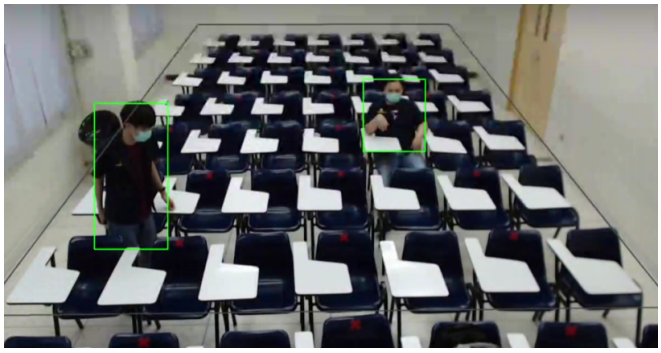

Fig. 2. Human body detected using YOLOv3 in green boxes. The image reference frame is drawn with a black line

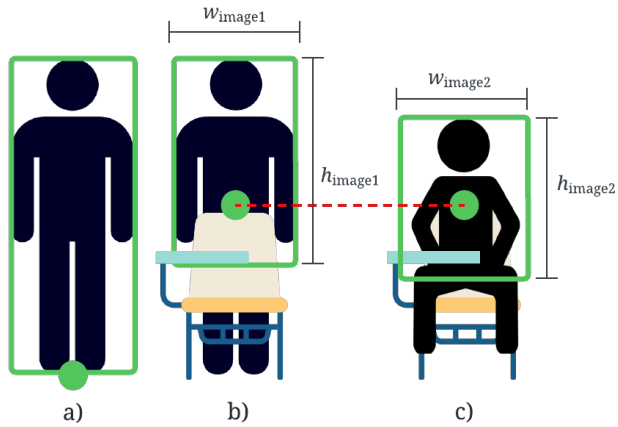

Fig. 3. Illustration of human body detection reference point (green) for (a) original approach (at the bottom-middle point of the green box), (b) a person standing between seats(77.5% of vertical proportion from the top-middle point as $h_{image_1}/w_{image_1} \geq 1.6$), and (c) sitting on the seat(67.5% of vertical proportion from the top-middle point as $h_{image_2}/w_{image_2} < 1.6$).

were used to calculate the ratio between the real size of the room and the pixel size of the region of interest on the image frame. The ratio of image and room size is calculated as in Equation (3).

$$Q_{\text{Physical}} = \frac{ImageSize}{RoomSize} \tag{3}$$

This Q_{Physical} is a ratio used to determine the physical coordinate (actual location) of the seat and persons. This was calculated using the window (screen) scale and the object pixel coordinate as in the following Equation (4).

$$C_{\text{Physical}} = \frac{(C_{\text{Image}} \times S_{\text{Image}})}{Q_{\text{Physical}}} \tag{4}$$

The calculation to calibrate the size of the image to the real-world size in both$x$ and y axes were performed separately for each of the room's dimensions using the Equations (3) and (4). The physical coordinates denoted as C_{Physical} representing the actual location of the object that appear in the image at the pixel coordinate C_{Image} with the scale of S_{Image} on the screen, which is then calibrated by the physical ratio Q_{Physical}. Finally,

the calculated physical location C_{Physical} for x and y axes will be used to project the physical coordinate value on the bird-eye view interface as well as to be kept in the log file for future physical distancing protocol uses.

D. Log File

This paper proposes the approach that stores seat and human physical distancing activities as dictionary data, .JSON, log-file instead of the traditional way of keeping the whole .mp4 or .mov video file type [26]. This novel application of a log file system is proposed in our work more efficiently in terms of manageable ways for storing raw video files itself and letting the person monitor the distancing activity all the time. Our proposed method takes every relevant object and its physical coordinates detected from the surveillance camera video input and then stored them in a form of .JSON formatted log file. The log file consists of three major parameters that can influence the physical distancing protocol.

IV. Implementation

Initially, we recorded the video data in a classroom using a webcam Logitech C920 HD Pro.

The video frame of the empty room was used to initialize the seat coordinate detection and all video frames were used for human location identification. Seat locations were extracted by applying color contour detection on the seat's unique area. Before applying color contour detection, we preprocessed the first frame sliced to enhance the ability of seat color contour detection by perspective transforming the image. The perspective transformation gives the model the ability to better detect every seat in the room compared to applying the model directly to the original image in Fig. 4. The seat matrix coordinate and person coordinates are calibrated on the same reference frame in the bird-eye view. We modified the pre-trained model's person detection parameters and calculate the physical location. This was done by firstly adjusting the confidence value of the model which consequently made the model more responsive to the detection of people at further distances. Then we calculated the physical location of the person with respect to the person coordinate in the image. This information about the seat and human coordinates is recorded in the log file for their displacement measures.

These data are all kept in form of a dictionary of numerical lists with respect to date and timestamps, which is convenient for any future observation needed. The overall diagram of the system workflow is demonstrated in Fig. 5.

V. Performance Evaluation

We investigated the uncertainty of determining the seat physical locations by computing the difference between the calculated value and the actual measurement value for every seat position (relative to the reference frame) as illustrated in Fig. 6.

The test footage was recorded at a resolution of 1920x1080. The program was run on a system using an Nvidia RTX 3070 GPU. In terms of the computational costs, we measured the average time taken to process a single frame from the video source to be 0.08 seconds.

Fig. 4. Classroom seat detection before applying the perspective transformation.

VI. Result and Discussion

A. Detection Enhancement

Our developed system initiates seat positioning in a classroom using seat contour detection. In the classroom environment, considering the top section of the seat, the constraints that are resulted from further seats cause the contour detection to be less precise than in the front-row seats. We then proposed the use of perspective transform for enhancing the detection ability for the seats at further distance [27]. The perspective transformation gives a higher ability for the seat detection algorithm to detect the seats in the back rows as seen in the images before and after applying the perspective transformation shown in Fig. 4 and Fig. 5 step d), respectively.

B. Physical Distancing

We used the upper left-hand corner of the detected seat bounding area after the perspective transformation had been applied in Fig. 5 step d) as the reference point for the system. This reference point was used as the origin of the person and seat coordinates that can then be used for the calculation of the physical locations (see Fig. 6). Perspective transformation of the image space into a bird-eye view allowed physical coordinates to be calculated [28]. Actual distance in real-world units can be estimated using the system. However, there are still some errors from the transformation, which can generally be found in a number of related researches [29], [30], [31]. Paying particular attention to selecting the area of interest and the proper algorithm to calibrate the image-physical distance would help gain a more precise image-based distance measurement.

We have implemented the system for the location estimation of the persons and seats using the camera-based system in order to work out the actual human and seat locations in the classroom. As a result of the performance evaluation compared against the accepted value (i.e. the actual physical distance measured in the classroom), the average error of seat location computed using the proposed system is ± 5.25 cm with a standard deviation of 4.64 cm. This suggests the overall performance of our camera-based system in determining

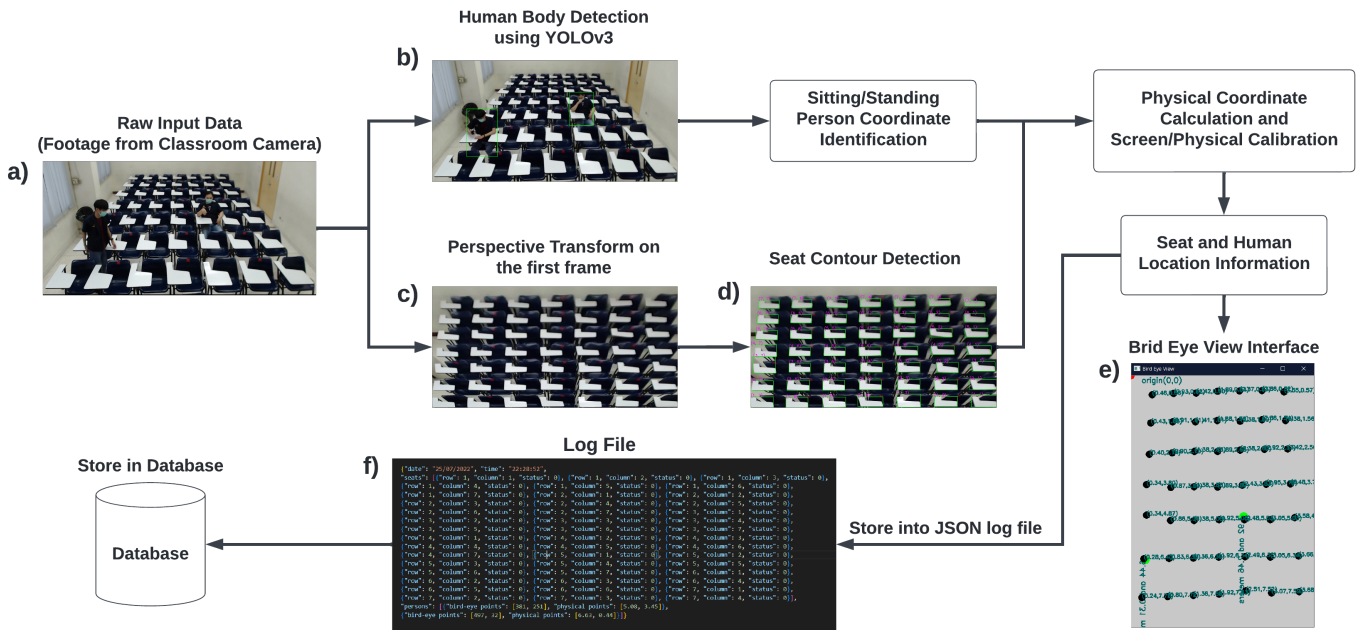

Fig. 5. The system workflow for camera-based indoor physical distancing log recording system.

locations in the real world. This scale of the error shows a slight effect on the physical distancing protocol in the indoor environment, specifically, the classroom when compared scale-wise. Therefore, the system can be generalized to obtain the distances between persons. In particular, the system provides the accurate determination of whether or not the person is at an appropriate distance apart from other persons, following the official distancing recommendation for COVID-19 prevention i.e. 180 cm [10]. Where the uncertainty of 5.25±4.64 cm causes a tiny significance to the scale of the distancing; therefore, the system is reliable enough in terms of giving a proper physical distancing suggestion.

C. Seating Behavior

Sitting and standing persons are classified based using the condition defined in Equation (2). That is we consider the height-to-width ratio of the area surrounding a detected person as the criteria for classifying whether a person is on- or off-seat. According to the on- or off-seat classification, an appropriate reference point to represent the spatial location of the person in the classroom space was made as seen in Fig. 3. This proposed method partly solved the problem of the hidden part of the human body behind the seat, hence more precise locations were achieved in our indoor environment when compared to the traditional approach in Fig. 3(a). The improvement of the body reference point shows the ability to develop a more precise location identification system by giving the adaptable reference point based on the various environment in the future. In addition, the effect of body movements when sitting on a seat was observed. The physical locations of persons sitting on particular seats were calculated. Taking into account the usual movements, the maximum distance of any person away from his/her seat was determined. We found that such maximum sitting boundary is 46×87 cm with respect to the seat centroid, which is reasonable according to the actual dimensions of the seat. Hence, the distance error determined in the previous section is neglected compared to the movement of people. This helps define a sitting area of a person and suggests whether a particular seat is occupied or available, which is worth recording in the log file for reference and analysis purposes.

VII. Conclusion

We developed the person and seat log file recording system for physical distancing observation from video data from a classroom with limitations on human detection. Human body reference point determination techniques and physical coordinate calculations were performed to create positioning strategies that can identify the locations of persons and seats physically. The proposed strategies were evaluated and analysed for the effectiveness of the system usage. The system also proved to be possible future development for a more variate environment. The output from the system can be recorded as a log file for analysis afterwards.

Acknowledgment

The authors gratefully acknowledge the financial support provided by the Thammasat University Research fund under the TSRI, Contract No. TUFF19/2564 and TUFF24/2565, for the project of “AI Ready City Networking in RUN”, based on the RUN Digital Cluster collaboration scheme, and the Center of Excellence in Intelligent Informatics, Speech and Language Technology and Service Innovation, Thammasat University, Thailand.

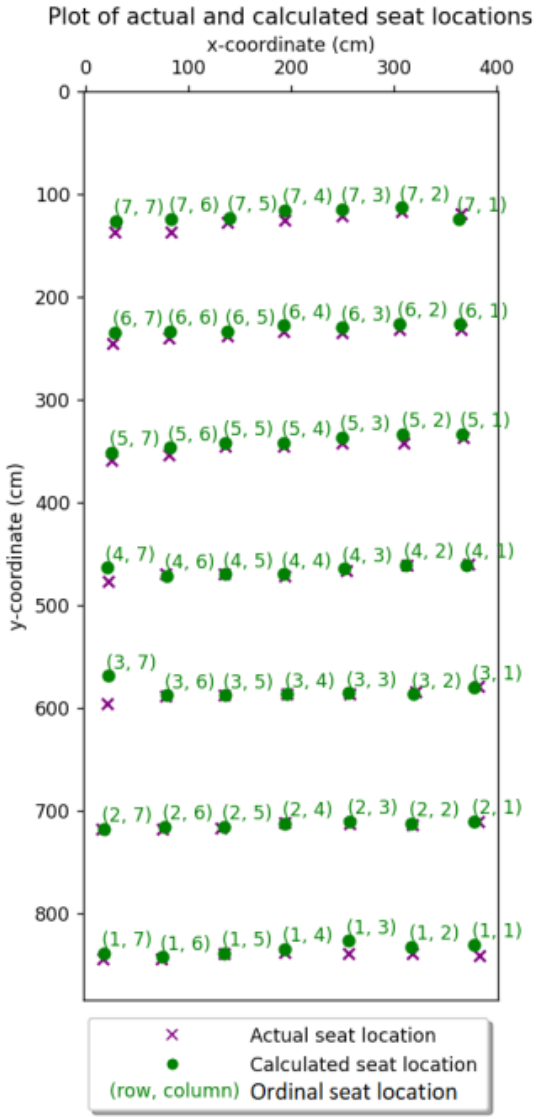

Fig. 6. Plot of error between the actual (measured) locations and the calculated locations (using the perspective transform technique) of the seats in the classroom with the specified ordinal seat locations as recorded in the log file.

References

[1] Y.-C. Wu, C.-S. Chen, and Y.-J. Chan, "The outbreak of covid-19: An overview," *Journal of the Chinese medical association*, vol. 83, no. 3, p. 217, 2020.

[2] S. Jobsri, "Ministry of public health's communication in the situation of coronavirus disease (covid-19) pandemic," *Journal of communication art, Sukhothai Thammathirat Open University.*, vol. 11, no. 1, pp. 13–28, 2021.

[3] World Health Organization, "WHO coronavirus (covid-19) dashboard: Thailand." https://covid19.who.int/region/searo/country/th, 2022. visited on 2022-08-07.

[4] P. Katewongsa, D. A. Widyastari, P. Saonuam, N. Haemathulin, and N. Wongsingha, "The effects of the covid-19 pandemic on the physical activity of the thai population: Evidence from thailand's surveillance on physical activity 2020," *Journal of sport and health science*, vol. 10, no. 3, pp. 341–348, 2021.

[5] T. Intharawiset, T. Jareoan-sa, and P. Yuang-soi, "Reflection on thai education after covid 2019," *Journal of Legal Entity Management and Local Innovation*, vol. 7, no. 4, pp. 323–333, 2021.

[6] Y. Xiao, B. Becerik-Gerber, G. Lucas, and S. C. Roll, "Impacts of working from home during covid-19 pandemic on physical and mental well-being of office workstation users," *Journal of Occupational and Environmental Medicine*, vol. 63, no. 3, p. 181, 2021.

[7] A. Bozkurt, *Handbook of Research on Emerging Pedagogies for the Future of Education: Trauma-Informed, Care, and Pandemic Pedagogy: Trauma-Informed, Care, and Pandemic Pedagogy*. IGI Global, 2021.

[8] W. Marome and R. Shaw, "Covid-19 response in thailand and its implications on future preparedness," *International journal of environmental research and public health*, vol. 18, no. 3, p. 1089, 2021.

[9] K. Pearce, "What is social distancing and how can it slow the spread of covid-19," *Johns Hopkins University*, pp. 1–5, 2020.

[10] A. Venkatesh and S. Edirappuli, "Social distancing in covid-19: what are the mental health implications?," *Bmj*, vol. 369, 2020.

[11] E. M. Aquino, I. H. Silveira, J. M. Pescarini, R. Aquino, J. A. d. Souza-Filho, A. d. S. Rocha, A. Ferreira, A. Victor, C. Teixeira, D. B. Machado, *et al.*, "Social distancing measures to control the covid-19 pandemic: potential impacts and challenges in brazil," *Ciencia & saude coletiva*, vol. 25, pp. 2423–2446, 2020.

[12] L. Zhao and S. Li, "Object detection algorithm based on improved yolov3," *Electronics*, vol. 9, no. 3, p. 537, 2020.

[13] J. R. Reidenberg, "Privacy in public," *U. Miami L. Rev.*, vol. 69, p. 141, 2014.

[14] Q. N. T. Thi, T. K. Dang, H. L. Van, and H. X. Son, "Using json to specify privacy preserving-enabled attribute-based access control policies," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, pp. 561–570, Springer, 2017.

[15] X. Li, S. Li, W. Bai, X. Cui, G. Yang, H. Zhou, and C. Zhang, "Method for rectifying image deviation based on perspective transformation," in *IOP Conference Series: Materials Science and Engineering*, vol. 231, p. 012029, IOP Publishing, 2017.

[16] E. Koufogiannis, N. Sgouros, and M. Sangriotis, "Robust integral image rectification framework using perspective transformation supported by statistical line segment clustering," *Applied Optics*, vol. 50, no. 34, pp. H265–H277, 2011.

[17] N. Khatri, A. Dasgupta, Y. Shen, X. Zhong, and F. Shih, "Perspective transformation layer," *arXiv preprint arXiv:2201.05706*, 2022.

[18] J. Zhang, J. Zhang, B. Chen, J. Gao, S. Ji, X. Zhang, and Z. Wang, "A perspective transformation method based on computer vision," in *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, pp. 765–768, IEEE, 2020.

[19] P. Gupta, V. Sharma, and S. Varma, "People detection and counting using yolov3 and ssd models," *Materials Today: Proceedings*, 2021.

[20] J. Luo, Y. Wang, and Y. Wang, "Real-time pedestrian detection method based on improved yolov3," in *Journal of Physics: Conference Series*, vol. 1453, p. 012149, IOP Publishing, 2020.

[21] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Real-time computer vision with opencv," *Communications of the ACM*, vol. 55, no. 6, pp. 61–69, 2012.

[22] S.-L. Wang, W. H. Lau, and S. H. Leung, "Automatic lip contour extraction from color images," *Pattern Recognition*, vol. 37, no. 12, pp. 2375–2387, 2004.

[23] M. Ajmal, M. H. Ashraf, M. Shakir, Y. Abbas, and F. A. Shah, "Video summarization: techniques and classification," in *International Conference on Computer Vision and Graphics*, pp. 1–13, Springer, 2012.

[24] S.-W. Hong and L. Choi, "Automatic recognition of flowers through color and edge based contour detection," in *2012 3rd International conference on image processing theory, tools and applications (IPTA)*, pp. 141–146, IEEE, 2012.

[25] N. I. Hassan, N. M. Tahir, F. H. K. Zaman, and H. Hashim, "People detection system using yolov3 algorithm," in *2020 10th IEEE international conference on control system, computing and engineering (ICCSCE)*, pp. 131–136, IEEE, 2020.

[26] J. van Rest, "Surveillance and video analytics: factors influencing the performance," 2015.

[27] I. Ansari, Y. Lee, Y. Jeong, and J. Shim, "Recognition of car manufacturers using faster r-cnn and perspective transformation," *Journal of Korea Multimedia Society*, vol. 21, no. 8, pp. 888–896, 2018.

[28] J. C. Marutotamtama and I. Setyawan, "Physical distancing detection using yolo v3 and bird's eye view transform," in *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, pp. 50–56, IEEE, 2021.

[29] S.-F. Lin, J.-Y. Chen, and H.-X. Chao, "Estimation of number of people in crowded scenes using perspective transformation," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 31, no. 6, pp. 645–654, 2001.

[30] J. Yu, N. Gao, Z. Meng, and Z. Zhang, "High-accuracy projector calibration method for fringe projection profilometry considering perspective transformation," *Optics Express*, vol. 29, no. 10, pp. 15053–15066, 2021.

[31] V. Kocur and M. Ftáčnik, "Detection of 3d bounding boxes of vehicles using perspective transformation for accurate speed measurement," *Machine Vision and Applications*, vol. 31, no. 7, pp. 1–15, 2020.