Development and Implementation of a Monitoring System for STRC with the Aid of Head Detection

Gretchel Karen L. Alcantara, Ivan Darren J. Evangelista, Jerome Vincent B. Malinao, Ofelia B. Ong, Reginald Steven DM. Rivera, and Engr. Leonard U. Ambata
Department of Electronics and Communications Engineering,
De La Salle University, Manila.
gretchel_alcantara@dlsu.edu.ph

***Abstract*—The Development and Implementation of a Monitoring System for STRC with the Aid of Head Detection aims to automate the research room operations to increase its efficiency. It consists of a head detection system, website managed database and a mobile application controlled room access. The website and mobile application cater to two types of users namely, the administrator and the students. The website exclusively features announcements and management of equipment inventory both handled by the administrator. The mobile application features an IN/OUT button which manages the users access on the five (5) research rooms in the building. The head detection system detects, tracks and counts each person that enters a research room. The count noted by the system is then compared to the count of students that used the mobile application for room access. The e-mail and SMS notification system observes the comparison of the two counts and sends the appropriate notification to the administrator when there is a violation in the conditions set for the two systems. The database serves as the central point as it holds every information needed by each element in order to communicate with the whole system.**

***Index Terms*—Digital Signal Processing; Head Detection; Monitoring System.**

I. Introduction

The Science and Technology Research Center (STRC) is primarily utilized for research, and experimentation in the university. The rooms that the group specifically conducted its study on, are the research rooms for the Electronics and Communications Engineering (ECE) department. These five (5) research rooms are managed and monitored by the Electronics and Communications Engineering Coordinator.

The process of reserving research rooms in the Science and Technology Research Center (STRC) building is by passing reservation forms along with the other required documents to the ECE coordinator. There is no automated method of checking the request status, and access to the room depends on whether or not someone is available to open it. There is no other accessible method to monitor the students for the ECE coordinator. The only way is by monitoring the Closed Circuit Television (CCTV) feed in one's office. Having said this, the current process creates an issue of inconvenience and unreliability. Inconvenient in a sense, of having no assurance whether or not one will be able to access its reserved room for academic purposes, and unreliable, meaning that the facilitator has no other means of strict assurance and control of the people accessing the room. These issues can be resolved by creating an automated system, which caters to the needs of both the faculty, and the students.

The project was to develop and implement a monitoring system in the five (5) research rooms that utilized the existing CCTV cameras wherein a head detection scheme was incorporated. The head detection scheme was used to aid in the monitoring and limiting of entry of students. Figure 1, shows an overview of how the monitoring system works. A mobile application and website is provided for the administrator and students to utilize the functions provided for each user, as well as access to the research rooms. An SMS and email notification was utilize to alert the administrator once violations are done.

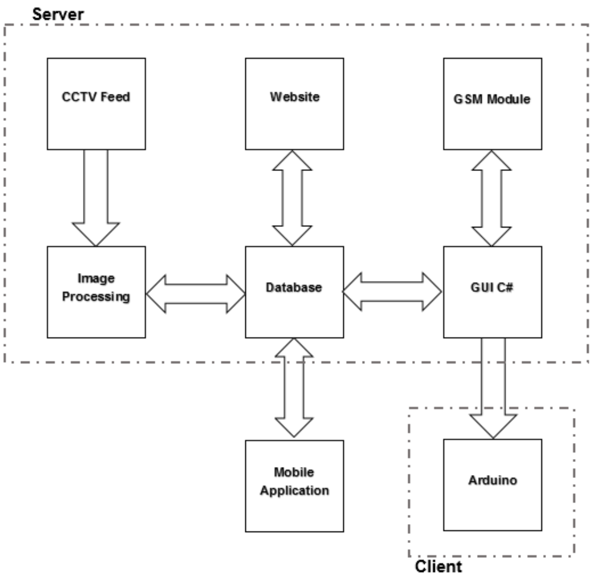

Figure 1: Block Diagram of the Monitoring System

II. Image Processing

The whole process for conducting the detection consists of three major parts; detection, tracking, and update, as shown in Figure 2.

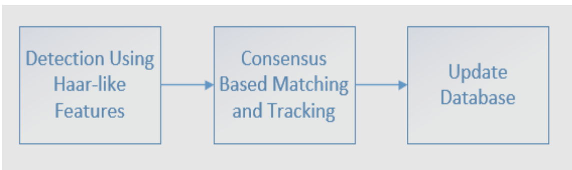

Figure 2: General Flow for Image Processing

Figure 3: Head Detection Process

First, the group utilized the Viola-Jones framework to be able to detect those accessing the research rooms. This framework focuses on Haar-like features. Instead of pixel-based computing, it clusters and categorizes features into adjacent rectangular regions. For example, the eye region, which is brighter and the nose and cheek region, which is darker are considered as two rectangular regions. These are extracted to create an integral image which allows faster computing since it adapts the sliding window approach thanks to its feature categorization. These will produce an adamant amount of features hence the need to select relevant features to speed up the computing, and improve the predicting power of the supposed classifier. In order to do so, one must conduct Adaptive Boost training or AdaBoost training. AdaBoost is a learning algorithm that trains the weak classifiers, and combines them into a cascaded classifier. Adaptive because the training process increments in stages, and no longer repeats the rejected images by the previous weak classifiers. They are weak in a sense that these classifiers cannot detect the object of interest on its own. This will then produce a cascaded classifier that can be used to detect the object of interest. In the group's case, the heads of the accessing students. The count of the detected will be fed to the tracking code.

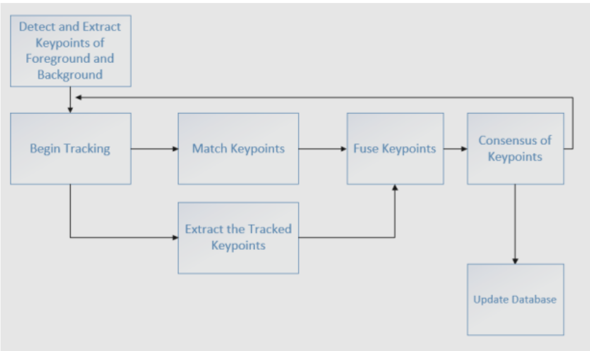

Figure 4: Clustering of Static-Adaptive Correspondences for Deformable Object Tracking

Second, the data taken from the detection will be fed into the tracking code, Clustering of Static-Adaptive Correspondences for Deformable Object Tracking (CMT) or formerly known as Consensus-based Matching and Tracking of Keypoints for Object Tracking. It is an algorithm for tracking that uses key points for tracking. Initially, it breaks down the frame into key points separating those of the object of interest and the background. It makes use of both static correspondences from the initial object of interest and the adaptive correspondences from the previous frame. The adaptive correspondences functions in such a way that it estimates the optical flow of the object by using the Lucas-Kanade method. The static correspondences match the current frame and the initial frame by method of key point description and matching. Both methods will be compared and fused together to get the current key point of the object, which will be used for determining both, the changed position of the object and the consensus. It applies hierarchical agglomerative clustering on each of the current key points in finding the largest consensus of the object of interest. The angle of rotation and the center of the object will be used for the next frame for estimating the optical flow. Lastly, this data will be then used to update the count on the database.

The Open Source Computer Vision Library, or the OpenCV library offers thousands of algorithms that caters to image processing. It is also advantageous in terms of computational efficiency, and real-time applications hence, was used as the study's general library.

III. Mobile Application

A user-friendly mobile application was also provided by the group for both the students and the administrator. The mobile application was made through Android Studio with support for devices that run on the Android operating system up to the latest firmware, Android Marshmallow. The main feature of the application is to be the key for the entry and exit to the research rooms. A student is first required to register to the system. He/She would then have to upload the group's requirements through the requirements page of the application. This eliminates the need to pass documents to the administrator which lessens the paper trail and the need to always go back and check whether the group has already been granted access. During registration, the MAC address of the smartphone/ device is also acquired automatically. This is to ensure security and a one to one account-device correspondence. The students are only required to enter their ID numbers during login because the system automatically checks the MAC address if it corresponds to the registered account. Figure 5 shows a screenshot of the mobile application's login page.

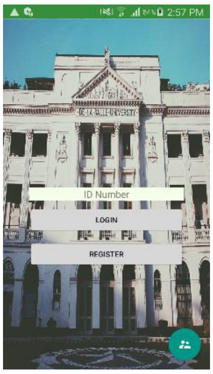

Figure 5: Mobile Application Login Page

The door access provides users two buttons, mainly the Entry and Exit buttons which will be used to unlock the doors of the research rooms. A student is only allowed inside their group's assigned room and thus the door only unlocks when the assigned room is picked through the drop down list. Once the user picks the room, the system checks the database if the right room was picked or not. If correct, then the door unlocks. If the wrong room was picked, the user is prompted about it and the door does not unlock. This ensures that the number of people entering the room will be in line with the limit provided by the administrator.

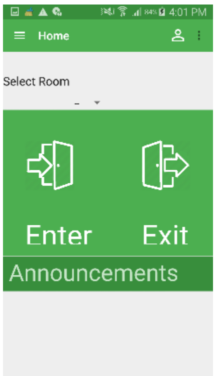

Figure 6: Mobile Application Door Access

Other features are also present inside the mobile application such as the viewing of the CCTV footage, an announcements list and a list of the equipment borrowed by the researching students.

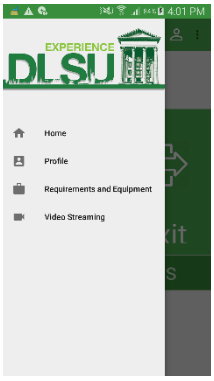

Figure 7: Mobile Application Menu

The mobile application is not only for the students' use but also for the administrator. A separate button may be seen on the bottom right hand corner of the login page for the administrator access. Once this is pressed, the mobile application system checks different parameters to see whether the button press was prompted by an administrator or not. Once the administrator logs in successfully, a list of all the students registered on the database will be provided. The administrator would then be able to change the students' permission status to access the rooms and update other parameters if the need arises. The administrator would also have access to the rest of the mobile application.

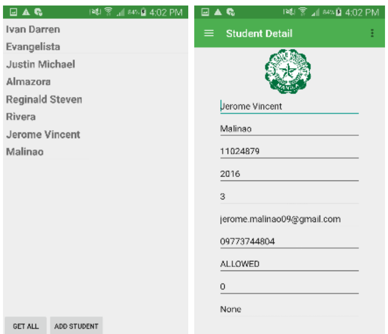

Figure 8: Administrative pages

IV. Website and Database

The database serves as the central point for each section of the system, letting each of these sections pass around information. The database is created with PostgreSQL, containing tables that would contain the information of students who submitted their room reservation request through the mobile application, an inventory of the equipment available for borrowing, administrator announcements, messages received and sent by the administrator through SMS, and the count of students permitted through the mobile application and the count of students detected by the CCTV with the aid of head detection.

The group created a website with Visual Studio Express for Web, utilizing C# as the language that would serve as an interface letting students access and manage student profiles that were submitted along with their room reservation and view the status of their room reservation and equipment they have borrowed. Figure 9 shows the page a logged-in student user would see when updating his or her information submitted initially.

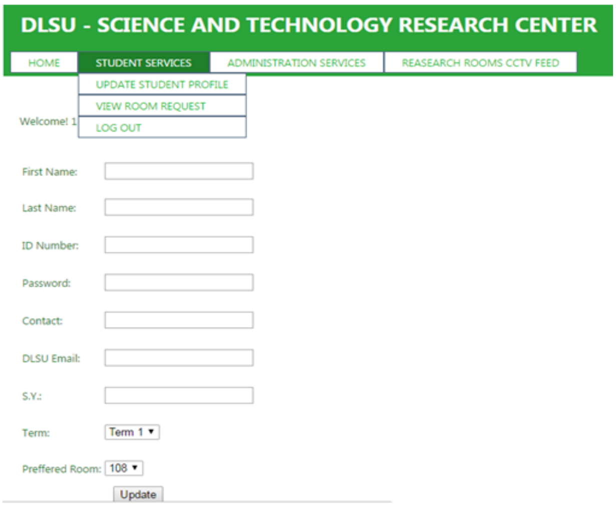

Figure 9: Update Student Profile Page

Through the website, the administrator may also access and manage all student profiles and equipment record contained in the database.

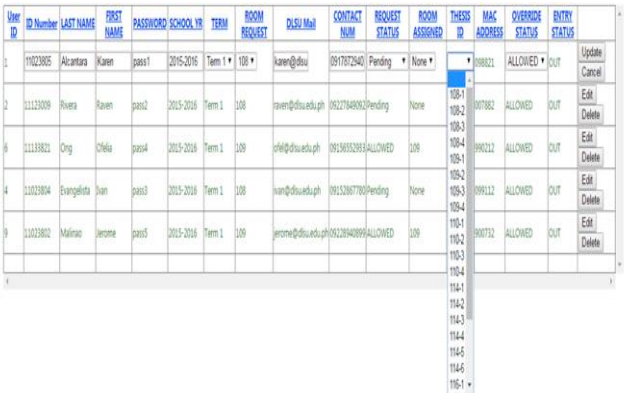

Figure 10: Profile and Requests Page

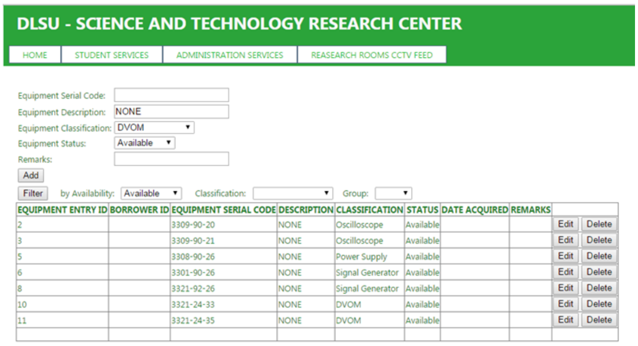

Figure 11: Equipment Page

The administrator may also monitor the number of students within a research room through an e-mail notification system which checks the equality of the count of students within a room allowed by the mobile application versus the count of students detected by the head detection system. Figure 11 showcases an example e-mail notification that the administrator would receive when there is a disparity in the number of students for each room.

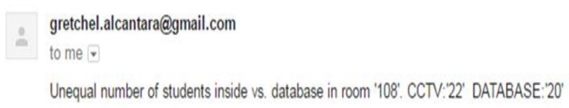

Figure 11: Room Monitoring E-mail Notification

The services available for students and the administrator are all password restricted, only letting users who have accounts registered within the database have access. The website also contains announcements regarding the (5) research rooms at STRC and general DLSU events which is available for general viewing.

V. ARDUINO AND CONTROLLER

The system consist of image processing for the CCTV, mobile application for door access, website and database for administrator purposes which falls on the software part of the system. The GUI controller was made in order to provide a transition of the software part of the system to the hardware part which is the remote Arduinos found at each room.

The main function of the GUI controller is to allow communication between the software and the hardware part of the system and to provide the necessary decision making for the system. The GUI controller monitors the database every five (5) seconds to determine if there is a request from the mobile application. This will be then evaluated if the said request has chosen the same room as the assigned room. Another status that was being checked is if the student's access to the room and the override status system is both "ALLOWED". Once these conditions were satisfied, the magnetic locks at the room will be opened else it will remain closed.

Another function of the GUI controller is keeping track of the number of students that are being detected by the image processing and the number of request made by the students. If there is unequal number of students inside the room and the request made by the students, the GUI controller will trigger the SMS module to send a text message to the administrator for notification. The SMS module can also receive SMS from the administrator that would allow the usage of a certain room for the day. It requires a text format in order to change the override status of the room.

VI. DATA AND RESULTS

The group established a testing environment for its head detection and tracking. As for the head detection, a fixed region of interest (ROI) was used as the focus for detection. The tests were conducted in the five (5) different research rooms for the ECE department with two (2) different types of footages namely, CCTV footage and a webcam footage mounted on the wall similar to the location of the CCTV camera. There was no extracted CCTV footage taken from the room 116, because it was not made available for access. The tests that were conducted utilized utmost 20 samples per room done in different time periods.

Table 1
Head Detection Accuracy

Room	HoG	HAAR (CCTV)	HAAR (Webcam)
108	60%	90%	100%
109	70%	80%	100%
110	66.7%	90%	100%
114	50%	71.4%	95%
116	No footage available	No footage available	90%

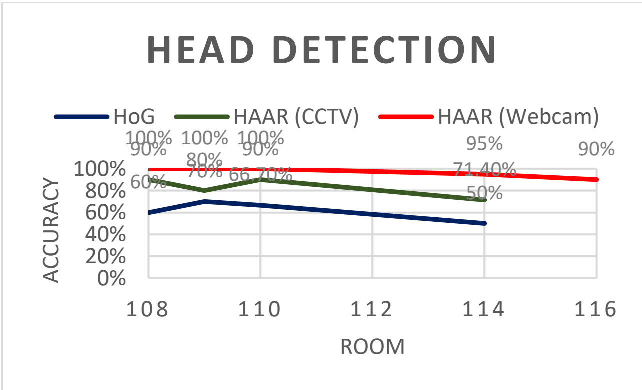

Figure 12: Comparison of Results

Table 1 and Figure 12 show the comparison between the different algorithms used by the group for its head detection system. The Histogram of Gradients (HoG) had an over-all accuracy of 62%. The haar-like feature detection for CCTV, and webcam footages had an accuracy rate of 83% and 97%, respectively. Similar to Haar detection, HoG detection focuses more on features rather than pixel computation. Haar focuses on facial features, while HoG focuses on the whole body of the person; this results to a longer computation period and a less accurate predictive power. The webcam footage yields better results because, the extracted CCTV footages from the DVR are compressed. Figure 13 shows the detection performance of the HAAR-like feature detection for both the CCTV and webcam footages. It can be observed that the CCTV footages contain false-negatives.

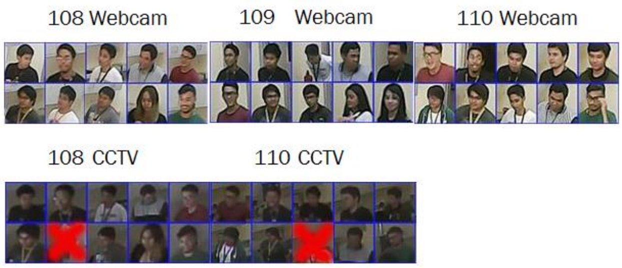

Figure 13: Head Detection Results

The group did not consider detections with false-negatives and false-positives as successful. In order to further improve the performance; the group conducted a test solely for false-negatives. Objects such as: bottles, a pillow with a face, a fire extinguisher, a jar, and other objects that resemble facial features and a neck. Based on the results, the objects were not falsely detected. Although the test samples were not detected, based from the images below one can observe that there were in fact, false-positives.

Figure 14: False-positives Test

Table 2
Tracking Time

Sample	TLD (s)	CMT (s)
A	7	5
B	5	1
C	8	5
D	7	6
E	5	1

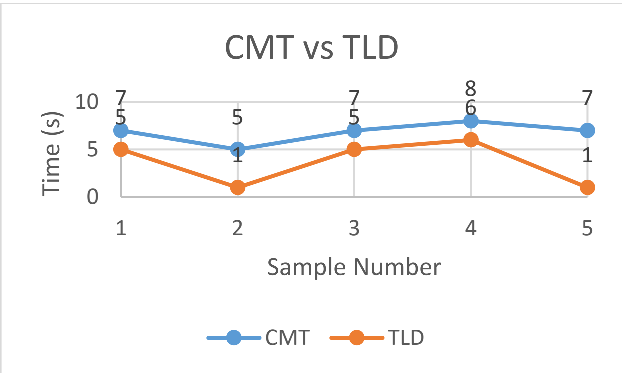

Figure 15: Comparison of Results

Table 2 and Figure 15 shows the comparison of the time duration of the two tracking algorithms on different samples. It can be observed that the Track-Learn-Detect (TLD) algorithm produced a low accuracy rate of only 36%. As for the Clustering of Static-Adaptive Correspondences for Deformable Object Tracking (CMT) algorithm, it produced an accuracy rate of 68%. The accuracy was computed by dividing the length of time it tracks with respect to ten (10) seconds, which is the duration of the time it took in the region of interest. Although the CMT algorithm is more accurate in real time tracking, it did not track that well with the CCTV footage. One reason could be the compressed pixels of the footage. Figure 16 shows the CMT result for sample A.

Figure 16: CMT for Sample A

The errors are due to certain factors such as the resolution of the camera, illumination in the environment, as well as the positioning of the camera.

For the door access test, the group performed this test by getting the success rate of each room per entry and exit. The mobile application's "in/out" function served as the key for each room by manipulating the user's status in the database which will then trigger the corresponding response of the magnetic locks. A pie graph was constructed for each room, representing the success vs. failure rates of the corresponding magnetic lock's response.

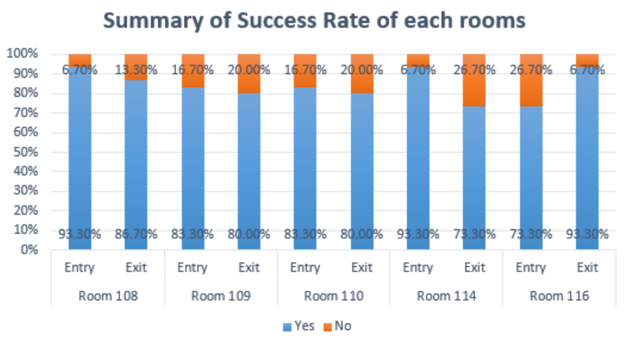

Figure 17: Summary of Success Rate of each rooms

Table 3
Summary of Success Rate of each room

	Room 108		Room 109		Room 110		Room 114		Room 116	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
Yes	93.30%	86.70%	83.30%	80.00%	83.30%	80.00%	93.30%	73.30%	73.30%	93.30%
No	6.70%	13.30%	16.70%	20.00%	16.70%	20.00%	6.70%	26.70%	26.70%	6.70%

In Table 3 and Figure 17, the room which has the most number of unsuccessful outcome is the entry to Room 116 and the exit to Room 114. While on the other hand, the entry to Rooms 108, 114 and the exit to Room 116 have the least number of unsuccessful outcome. The main reason why the number of errors is much lesser in Room 108 than Room 116 is that the router where the WiFi modules are connected is much closest to Room 108. Each room were tested having 30 samples for every entry and exit.

Table 4
Reliability of Each Room

	Room 108	Room 109	Room 110	Room 114	Room 116
	Entry - Exit	Entry - Exit	Entry - Exit	Entry - Exit	Entry - Exit
Reliability	80.89%	66.64%	66.64%	68.39%	68.39%
Complement	19.11%	33.36%	33.36%	31.61%	31.61%

In Table 4, it is shown that the reliability of each room wherein the success rate for entry and exit were multiplied together. It is multiplied together since for every entry there should be an exit, this implies that the door access for each room is a series system. Which is why the reliability for Room 109 and Room 110 is the lowest since both entry and exit success rate is 83.3% and 80% respectively. The reliability for rooms 114 and 116 is a little greater than the reliability for rooms 109 and 110 because of the success rate of 93.3%.

VII. CONCLUSION

The study consists of a head detection system, database managed website, and a mobile application controlled room access. The overall accuracy for the head detection implemented with HAAR-like features provided an accuracy of 97% for webcam footages, and 83% for the CCTV footages; the tracking accuracy rate of 68.75%. This entailed the group to have a more accurate and stable system as compared to the other algorithms. As for the door access test, the overall success rate of the door entry is 85%, while the total success rate of the door exit is 82%. The reason for the difference in the success rates of the entry and exit is due to the delay and the unstable network connection. A lot of factors have played a vital part in the result of the study. Some of which are the fixed conditions in the testing environment, such as the illumination in the rooms, CCTV camera positions, resolution of the cameras, and the network connection in the building. In Spite of the factors stated, the group was able to deliver an operational monitoring system for the research rooms inside the Science and Technology Research Center

REFERENCES

[1] S. H. Son, R. C. Beckinger and D. A. Baker, "DRDB: a distributed real-time database server for high-assurance time-critical applications," in Computer Software and Applications Conference, 1997. COMPSAC '97. Proceedings., The Twenty-First Annual International, Washington, DC, 1997.

[2] [2] T. D. Hui, "MICROCONTROLLER-BASED LOCK USING COLOUR SECURITY CODE," 7 June 2013. [Online]. Available: http://portal.fke.utm.my/fkelibrary/files/thendaohui/2013/682_THENDAOHUI2013.pdf.

[3] [3] S. R. Khan, "Development of Low Cost Private Office Access Control System," International Journal of Embedded Systems and Applications, pp. 1-7, June 2012.

[4] [4] T. O'Leary, "Electromagnetic Locks: Approved, Listed, Recognized & Certified," October 2010. [Online]. Available: http://www.securitron.com/Other/Securitron/Documents/Electromagnetic_Locks_Article.pdf.

[5] [5] M. Piccardi, "Background Subtraction Techniques: A Review," April 2004. [Online]. Available: http://www.cs.nccu.edu.tw/~whliao/es2008/BackgroundSubtractionReview-Piccardi.pdf.

[6] [6] Z. Kalal, K. Mikolajczyk and J. Matas, "Forward-Backward Error: Automatic Detection of Tracking Failures," in International Conference on Pattern Recognition, Istambul, Turkey, 2010.

[7] [7] D. G. Lowe, "Distinctive Image Features," in International Journal of Computer Vision, 2004.

[8] [8] Campus Component, "GSM Interfacing Board," [Online]. Available: http://www.campuscomponent.com/media/download/GSM%20Module.pdf.

[9] [9] Engineers Garage, "Introduction to Image Processing," [Online]. Available: http://www.engineersgarage.com/articles/image-processing-tutorial-applications.

[10] [10] Gregor, "Principle Of Work Of Magnetic Door Locks And Their Advantages," The Techno Phobe, [Online]. Available: http://thetechnophobe.com/magnetic_door_locks.html.

[11] [11] D. N. Hassan, "Introduction to Computer Vision and Image Processing," [Online]. Available: http://www.uotechnology.edu.iq/dep-cs/mypdf/subjects/4sw/4ip.pdf.

[12] [12] "Background Subtraction," [Online]. Available: http://www.ics.uci.edu/~dramanan/teaching/cs117_spring13/lec/bg.pdf.

[13] [13] "Background Subtraction for Detection of Moving Objects," [Online]. Available: https://computation.llnl.gov/casc/sapphire/background/background.html.

[14] [14] G. a. P. R. Nebehay, "Consensus-based Matching and Tracking of Keypoints for Object Tracking," in Winter Conference on Applications of Computer Vision, 2014.

[15] [15] G. a. P. R. Nebehay, "Clustering of (Static-Adaptive) Correspondences for Deformable Object Tracking," in Computer Vision and Pattern Recognition, 2015.

[16] [16] D. Müllner, Fastcluster: Fast Hierarchical Agglomerative Clustering Routines for R and Python, Standford University.

[17] [17] S. a. C. M. a. S. R. Y. Leutenegger, "BRISK: Binary Robust Invariant Scalable Keypoints," in Proceedings of the 2011 International Conference on Computer Vision, 2011.

[18] [18] D. G. Viswanathan, "Features from Accelerated Segment Test (FAST)," 2006.