Recognition of Image in Different Cameras using an Improved Algorithm in Viola-Jones

Washington Garcia - Quilachamin[1]
Student PhD, UNMSM, Lima – Peru
Professor Faculty of Engineer
ULEAM, Manta - Ecuador

Luzmila Pro - Concepción[2]
Professor Faculty of Engineer
UNMSM
Lima – Peru

Abstract—**Technological evolution through computer tools has given rise to tasks of impossible recognition for an ordinary man, but at the same time favorable for the safety of people. Deep learning is considered a tool that uses images and video to detect and interpret real-world scenes. Therefore, it is necessary to validate the application of an algorithm with different cameras for the recognition of people, being a contribution to surveillance in domestic environments and of companies. In this research, an algorithm is presented that, through a camera, allows to detect the image of a person. The objective of this research is to validate the process in the recognition of the image with four cameras through the application of the improved algorithm Viola-Jones. The validation was carried out through a mathematical analysis, which allowed us to base the recognition of the image using four different cameras. As a result of the study, an effective and functional validation was obtained, about the results achieved with the application of the algorithm, using the four cameras and effective in the speed-based recognition concerning the different tests performed on the capture and recognition of each image, reducing the recognition time and optimizing the software and hardware used.**

Keywords—Algorithm; video surveillance; cameras; image of a person

I. INTRODUCTION

With technological evolution, the application of algorithms with cameras for the recognition of objects, people or events are part of the studies of artificial vision and deep learning, which is of much interest research, in the innovation of real-world scenes. Therefore, recognition tasks are considered complex for a computer, according to [1], [2]. The authors [3] consider that from the study of the recognition of objects, innovation in facial and people recognition is started.

The authors in [4] proposes that every process, algorithm, model, etc., is based on the effective recognition of human activity (HAR) which consists of understanding what people are doing from their position, figure, movement or other space-time information derived from their video sequences. according to [5], [6] and [7], Some factors incentive this research with the recognition of a person's image is based on the recognition of different types of images that relate to human movements such as running, limping, jumping, etc.

So the authors in [8], establish that the Viola-Jones algorithm is currently one of the most used to solve search problems considering the face of a person. According to [9], [10] pattern recognition systems are developed as a part of

continuous monitoring of human behavior in the area of assisted living, rehabilitation, and entertainment.

The authors in [11], [12] and [13] mention that applications may vary in object recognition systems in a video surveillance system located at an intersection of roads, restaurants, education centers, shopping centers or any other public meeting place.

For what is considered the field of security as a topic of broad action in offices, homes, and smart cities. Some authors [14], express that the security of a people and family is very important by considering that human pattern recognition systems can be connected to through device automatic.

The purpose of this research is to validate the process in the recognition of image applying the improved algorithm. The analysis was carried out through mathematical equations considering a definite integral from a zero point, which allowed us to justify the research carried out about the recognition of image of a person through four cameras, considering as parameters the time and distance from the camera, to the person's position inside the area.

As a result of the study in using an improved algorithm in viola - jones, for the recognition of image, using four cameras, data were obtained relating to the time and distance in which the image is detected with each camera. These data were used in the analysis, through the application of equations and definite integral, considering as reference the author [15]. With the effective use of the algorithm in a distance of 3m to 9m, the image capture and recognition as true positive, false positive, false negative is determined, considering that, at a greater distance of an image concerning the camera, the angle is smaller in recognition of image, otherwise its result related to recognition would be false negative.

This paper is structured as follows:

Section II shows the related work, which explains a previous review of the main models of image recognition over time, details the methodology used, the framework used to construct the proposed algorithm, improved algorithm, criteria, and the mathematical analysis applied.

Section III the results of the application of improved the algorithm are shown, and the analysis in the recognition of image of a person with the four cameras. Finally, the conclusions reached in the present study are detailed, and the future work.

II. METHOD AND RELATED WORKS

The authors in [16], [17], [18], manifest that there are several studies on base at the recognition of an image using algorithms.

Convolutional Neural Network (dCNN) [19], Hidden Markov (HMM) [20], [21],[22] and Support Vector Machines (SVM) [23], as the authors argue, are algorithms models used to classify images, the analyze data, with computational methods for the patterns recognition.

The models described are based on algorithms that can identify images differently, as they consider it [7], [24], [25], , who mention studies related to the recognition of faces, people, traffic signs, tumors and many other aspects of visual data in different activities that the human performs.

Of all the models, the SVM was chosen to be fully implemented in Matlab, because in real-time it detects people not included in an upright position. Also, this model is related to classification and regression problems; this will be used in our research through the algorithms described.

The SVM model is used to perform recognition of people in a real application [23], and with the criteria of the authors of [8], the implementation of the Viola-Jones algorithm allowed us to obtain the capture of the image of a person. Which is represented with a frame, and that these pixels minimizing the amount and time of necessary calculations.

In [8] and [26], the authors state that the Viola-Jones algorithm is based on a series of classifiers and employs a method of approximation based on appearance. It is divided into two stages: a first stage of learning the classifier based on a large number of positive examples and negative examples, and a recognition phase by applying this classifier in relation to the images not known.

These features were used in our research and that allowed us to apply efficiently from an integral image in time real considering a person for this study.

A. Methodology

To analyze the algorithm's execution based on the recognition of people, our research applies an exploratory methodology, considering the data collection of four video surveillance cameras, where 840 was obtained captures, corresponding to 210 images per camera. This study is carried out to determine the effectiveness in the execution of the algorithm viola jones based on the recognition of image.

B. Framework of Work

The purpose of this research is to understand and validate the process to recognize an image based on the using an improved algorithm in Viola-jones, using the image acquisition tool, computer vision system, and deep learning. the procedure is described through the framework shown in Fig. 1.

The improved algorithm will be used to capture a pattern referring to the image of a person, considering that the objective is to validate the recognition of image in real-time through the four cameras hilook, toshiba, max, and turbo. the

calculation of the data about the integral image is expressed in [8].

With the information obtained concerning parameters time and distance, in the recognition of the image, the respective calculations and analysis were developed. These are fundamental in the validation of the algorithm, considering the authors [15] and [27].

C. Algorithm Applied

The algorithm described below was improved, based on Viola-Jones and has been used in the recognition of image in real-time.

```
Data: image, recobj, imqrec, imshow, frame, bboxes, r
image=vision. recobj ( );
 recog=imaq.VideoDevice ( );
set (recog);
    while (true)
        frame=step(recog);
        bboxes=step (facerecognition,frame);
        imqrec=insertObjectAnnotation(frame, ,bboxes,);
        imshow(imqrec)
        exit of imshow
        r=findobj();
    end
end
```

The data obtained that refers to the recognition are imported and saved in the routing path.

D. Viola-Jones

The purpose of this research is to determine an algorithm based on Viola-Jones that allows us to recognize an image and validate your process through the mathematical equation (1). Whereas this work is motivated in the recognition of faces on real-time video, according to [28].

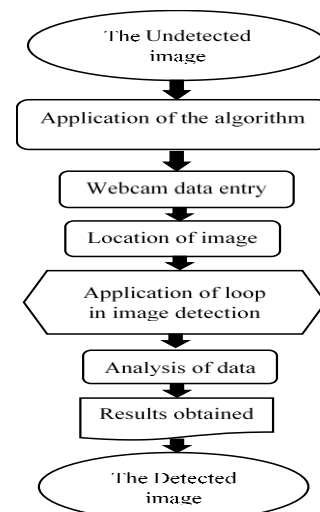$$I(x,y) = \sum_{xi<x,yi<y} i(x',y') \qquad (1)$$

Fig. 1. Framework for Image Recognition.

Where: (x, y) is the integral image calculated in pixels, (x', y') is the original image. Using the integral image, any sum of a rectangular area ABCD can be calculated efficiently, (2):

$$\sum_{(x,y) \in ABCD} i(x', y') = I(D) + I(A) - I(B) - I(C) \qquad (2)$$

E. Tools used

For this study, the improved algorithm in the recognition of image is used and four different types of cameras. These were selected by their size and class.

Among the cameras selected and which were used for the proposed objective, is the computer's webcam and cameras for infrared security type Dome and Bullet. The software that was used for testing and algorithm improvement was Matlab 2017b.

According to [29], the authors describe the parameters of cameras considered in this study and shown in Table I.

640 x 480 pixels were considered for the image resolution of the Toshiba camera because it does not support its configuration, while the image resolution on the Hilook, Max and Turbo cameras is assigned by default when used with the Easy Cap device, which allows the conversion of the image.

F. Mathematical Procedure

The mathematical procedures are based, considering [29], and for the development of this process in our research, it is considered the operation of vectors, matrices, and definite integrals, the equation of the slope (1), was applied the equation mathematical concerning the function f (x) (2) and its derivative as a function of time (3).

$$m = \frac{y_2 - y_1}{x_2 - x_1} \qquad (2)$$

$$f(x) = \frac{y_2 - y_1}{x_2 - x_1} = \frac{dy}{dx} \, dt \qquad (3)$$

$$f(x) = \frac{dy}{dx} \, dt \qquad (4)$$

Where "dy", represents recognition angle, "dx" distance and "dt", concerning the time, which is considered as (x, y) respectively.

The equation of definite integral (4) was considered, which was applied to evaluate the values obtained concerning the time and distance parameters of each camera used for the recognition. Starting from a zero point.

$$\int_0^d \frac{1}{2}(x/t).dx \qquad (4)$$

where "d" will be the distance value and "t" is the time value obtained from each camera, respectively.

TABLE. I.   CHARACTERISTICS OF CAMERAS

Camera	Night Vision	Image format	Image resolution	Iris opening
Toshiba	Negative	YUY2	640 x 480	Permanent
Hilook	Positive	YUY2	720 x 576	Automatic
Max	Positive	YUY2	720 x 576	Automatic
Turbo	Positive	YUY2	720 x 576	Automatic

III. RESULTS AND ANALYSIS

A. Results of the Recognition by Cameras

In Table II, the person's height, recognition angle, camera height, and distance are described, which are the parameters used and the data obtained concerning the recognition of an image, in this case, concerning a person's.

Table III shows the data obtained by the four cameras, considering the time in seconds when recognition the image of a person concerning distance.

The data of Table III were applied in Fig. 2, which are the result of the execution of the algorithm when using the four different cameras in recognition of the image of a person. Lines shown in a different layout, are related to time and distance, concerning each camera used in this research when detecting the image.

TABLE. II.   MEASUREMENT PARAMETERS

Person stature (m)	Recognition angle	Camera height (m)	Distance (m)
1.78	16°	0.88	3.00
1.78	14°	0.88	4.00
1.78	12°	0.88	5.00
1.78	10°	0.88	6.00
1.78	8°	0.88	7.00
1.78	6°	0.88	8.00
1.78	4°	0.88	9.00

TABLE. III.   RECOGNITION TIME BY CAMERA

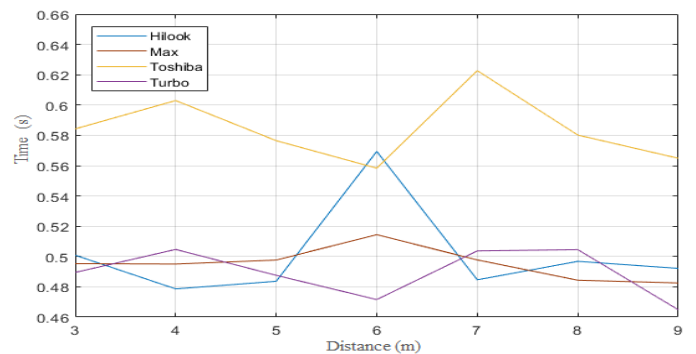| Distance (m)	Types of cameras			
	Hilook (s)	Max (s)	Toshiba (s)	Turbo (s)
3.00	0,5008767	0,4952986	0,5842529	0,4894874
4.00	0,4786243	0,4950434	0,6029704	0,5046813
5.00	0,4836451	0,4976678	0,5764817	0,4875867
6.00	0,5693954	0,5144184	0,5583289	0,4715800
7.00	0,4846173	0,4977600	0,6227089	0,5037189
8.00	0,4968127	0,4843582	0,5801745	0,5045335
9.00	0,4921652	0,4825438	0,5648531	0,4648238

Fig. 2.   Time-Distance Variations.

Fig. 2, shows the line concerning the Hilook camera, which presents variations, and the showing the largest peak of all the lines at 6m and low speeds in the intervals of 4m and 5m., for what is considered that this camera is less efficient in the distance range of 3m to 9m. It also shows the stroke that represents the Max camera, this stroke represents the most continuous speeds of all data and It has a slight drop peak in the 6m interval and represents the second-fastest.

The strokes, of Toshiba Webcam camera, are progressively scaling, presenting a slight peak in 4m., And the second peak of fall, this time lighter in 7m. Even so, the speed values in the last intervals in this camera are slower, but it maintains its continuity. The stroke concerning the Turbo camera is the fastest in most intervals except in 4m, which shows a slight peak of lower speed, also in the interval of 7m and 8m, the same problem occurs. It is slightly slower. Therefore, this camera is the fastest compared to the strokes of the previous cameras.

The results obtained allowed to demonstrate the performance in the speed of the different cameras in the recognition of image, considering that the proposal is not to evaluate the cameras, but the effectiveness using the improved algorithm in viola-jones. Considering this result, the algorithm works, regardless of the type and brand of the cameras.

Table IV shows the image recognition framed in a frame (yellow contour), in this case, concerning a person's, it is considered the distance of a person regarding the camera. The images were captured using four different cameras through the application of the algorithm at different distances and positions in real-time, of course, the values vary according to their specifications.

The application of the algorithm in the recognition of image it allows us to validate its effectiveness in the four cameras. Table V shows the results of the recognition at different focal angles.

B. Analysis in the Recognition of the Image

In this case, concerning a person's. The results obtained in the recognition of the image in the four cameras, considering a distance of 3, 4, 5, 6, 7, 8 to 9m, and with the application of the algorithm configured in Matlab. It was determined that the Turbo camera detects positively the image in the established range.

The Toshiba Webcam camera, detects the image positively in the range of distance of 3m to 6m and false positive at a distance of 7m, positively detecting again in the range of 8m to 9m.

With the Max camera, the recognition of image is false negative at a distance of 3m, considering that in the range of 4m to 9m the recognition is positive.

Finally, with the Hilook camera, image recognition is positive at a distance of 3m, 5m and in a range of 7m to 9m. In this recognition there is a false negative at a distance of 4m and a false positive at 6m. The statistical table of Fig. 2 is also considered, concerning the distance with the time in image recognition, which allows establishing the reliable Turbo HD camera in the application.

C. Results with the Mathematical Equation

The speed is the result in the recognition of the image, concerning distance and time. Table VI shows the results obtained, through the developed calculations considering 0m as a lower limit, up to an upper limit of 3m up to 9m.

TABLE. IV.   IMAGE RECOGNITION

Camera type	Distance						
	3m	4m	5m	6m	7m	8m	9m
Turbo							
Toshiba							
Max							
Hilook							

TABLE. V.   FOCAL ANGLE AND IMAGE RECOGNITION

Camera	Focal angle		Recognition
	Inclination	Rotation	
Turbo	0° to 90°	0° to 360°	Positive
Toshiba	0° to 180°	0°	Positive / False-Positive
Max	0° to 180°	0° to 360°	False-Negative / Positive
Hilook	0° to 180°	0° to 360°	Positive / False-Negative False-Positive

TABLE. VI.   CAPTURE SPEED

Distance (m)	Types of cameras			
	Hilook (m/s)	Max (m/s)	Toshiba (m/s)	Turbo (m/s)
3.00	8,98424702	9,08542847	7,70214405	9,19329078
4.00	16,71457132	16,16019929	13,26764962	15,85158792
5.00	25,84539779	25,11715646	21,68325551	25,63646629
6.00	31,61247878	34,99097233	32,23906196	38,16955766
7.00	50,55535574	49,22050788	39,34422649	48,63823851
8.00	64,41059176	66,06680758	55,15581950	63,42492619
9.00	82,28944265	83,93020489	71,70005794	87,12978983

Fig. 3.   Speed – Distance.

Fig. 3 shows the strokes obtained between the ratio of the speed and the distance, of the four different cameras used for the recognition of the image of a person.
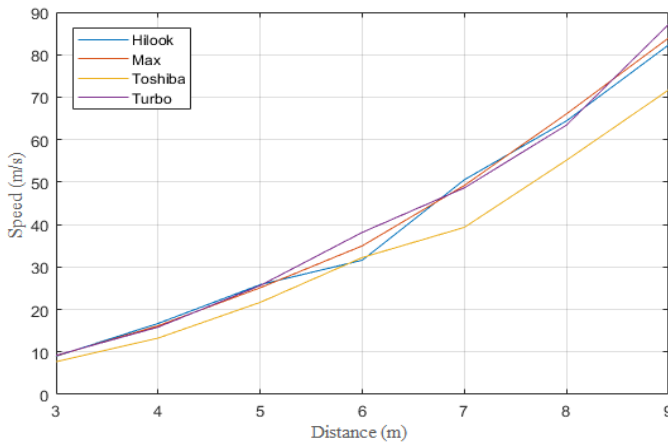
D. Analysis with the Applied Mathematical Equation

Based on the results obtained in Table VI, the speed of the recognition of the image of a person, in a distance range of 3m to 9m. it was determined: that the average speed in recognition by the Hilook camera is 40.058m/s. The Max camera detects the image of the person at an average speed of 40.653m/s. With the Toshiba camera, recognition average speed was of 34.441m/s, and finally, with the Turbo camera, the recognition of the image in a range of 3m to 9m was with an average speed of 41.149m/s.

Fig. 3, shows the data the Table VI, in this figure shows the lines concerning the speed and distance parameters of the four cameras. Its lines present variations to the speed, in relation to

the recognition of the image in a range of 3m to 9m. Of course, the values vary according to their specifications and this determines the Toshiba camera to be the most optimal because of its average speed of 34.441m/s, but it must be considered that this camera works with the computer processor.

IV. CONCLUSION

The use of the improved algorithm in viola-jones, allowed to obtain an effective and functional evaluation concerning 840 captures, corresponding to 210 images per camera, achieved in the recognition of the image of a person in the four cameras. With the data of time, distance and speed applied through the mathematical equation, it allowed determining that one of the four cameras used in this research to show a higher speed than the others, demonstrating that this is the fastest at the time of capturing and recognition of the image of a person, but this camera works with the computer processor.

So that, by applying the algorithm in the cameras, it was possible to validate its efficiency and speed in the recognition of image, considering that this recognition is reduced concerning time, optimizing the software and hardware used.

In this research, the hilook camera is considered the most suitable with an optimal displacement based on the speed of 40.058m / s. This also allows validating the effective in the recognition of the image of a person in the distance range of 3m to 9m, and the results obtained and analyzed strengthen the research developed concerning the using improved algorithm in viola-jones applied in the recognition of the image of a person.

Its validation and subsequent implementation of video surveillance cameras with the improved algorithm will serve as a contribution to the security and surveillance of people in domestic and business areas.

V. FUTURE WORK

This study is considered in the future to link concerning improving energy efficiency, video surveillance and reducing the CO_2 that is affecting climate change.

REFERENCES

[1] X. J. Lin, Q. X. Wu, X. Wang, Z. Q. Zhuo, and G. R. Zhang, "People recognition in multi-cameras using the visual color processing mechanism," Neurocomputing, vol. 188, pp. 71–81, 2016.

[2] Mathworks, "Image Category Classification Using Bag of Features," 2016.

[3] K. Sinhal, "Object Detection using Deep Learning for advanced users," 2017.

[4] S. Sun, J. Zhao, and Q. Gao, "Modeling and recognizing human trajectories with beta process hidden Markov models," Pattern Recognit., vol. 48, no. 8, pp. 2407–2417, 2015.

[5] C. Benabdelkader, R. Cutler, H. Nanda, and L. Davis, "EigenGait : Motion-Based Recognition of People Using Image Self-Similarity," pp. 284–294, 2001.

[6] J. Matai, A. Irturk, and R. Kastner, "Design and implementation of an FPGA-based real-time face recognition system," Proc. - IEEE Int. Symp. Field-Programmable Cust. Comput. Mach. FCCM 2011, pp. 97–100, 2011.

[7] H. Xu, L. Li, M. Fang, and F. Zhang, "Movement human actions recognition based on machine learning," Int. J. Online Eng., vol. 14, no. 4, pp. 193–210, 2018.

[8] M. V. Alyushin and A. A. Lyubshov, "The Viola-Jones algorithm performance enhancement for a person's face recognition task in the long-wave infrared radiation range," Proc. 2018 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. ElConRus 2018, vol. 2018-Janua, pp. 1813–1816, 2018.

[9] M. Y. Santos et al., "A Big Data Analytics Architecture for Industry 4.0," WorldCIST 2017 Recent Adv. Inf. Syst. Technol., vol. 4, 2017.

[10] H. F. Nweke, Y. W. Teh, M. A. Al-garadi, and U. R. Alo, "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges," Expert Syst. Appl., vol. 105, pp. 233–261, 2018.

[11] X. Jiang, J. Sun, H. Ding, and C. Li, "A silhouette based novel algorithm for object detection and tracking using information fusion of video frames," Cluster Comput., vol. 0, 2018.

[12] F. H. K. Zaman, M. H. Ali, A. A. Shafie, and Z. I. Rizman, "Efficient human motion detection with adaptive background for vision-based security system," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 7, no. 3, pp. 1026–1031, 2017.

[13] A. E. Papadakis, E. Tsalera, and M. Samarakou, "Survey on sound and video analysis methods for monitoring face-to-face module delivery," Int. J. Emerg. Technol. Learn., vol. 14, no. 8, pp. 229–240, 2019.

[14] Ilhan Aydin and Nashwan Adnan Othman, "A New IoT Combined Face Detection of People by Using Computer Vision for Security Application," Int. Artif. Intell. Data Process. Symp. (IDAP), Mal., pp. 0–6, 2017.

[15] G. D. Dean, Advanced engineering mathematics with matlab, 4th Editio. 2016.

[16] F. Kamaruzaman and A. A. Shafie, "Recognizing faces with normalized local Gabor features and Spiking Neuron Patterns," Pattern Recognit., vol. 53, pp. 102–115, 2016.

[17] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu, "Silhouette analysis-based gait recognition for human identification," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1505–1518, 2003.

[18] X. J. Wang, "A human body gait recognition system based on fourier transform and quartile difference extraction," Int. J. Online Eng., vol. 13, no. 7, pp. 129–139, 2017.

[19] C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian, "Multi-type attributes driven multi-camera person re-identification," Pattern Recognit., vol. 75, pp. 77–89, 2018.

[20] A. V. Nefian and M. H. Hayes, "Face detection and recognition using hidden Markov models," pp. 141–145, 2002.

[21] A. Mahapatra, T. K. Mishra, P. K. Sa, and B. Majhi, "Human recognition system for outdoor videos using Hidden Markov model," AEU - Int. J. Electron. Commun., vol. 68, no. 3, pp. 227–236, 2014.

[22] L. Speech, Processing, Daniel, H. James, and Martin, "Hidden Markov Models ch9," no. Chapter 20, 2017.

[23] Y. Lu, K. Boukharouba, J. Boonært, A. Fleury, and S. Lecœuche, "Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features," Neurocomputing, vol. 126, pp. 132–140, 2014.

[24] A. Baca, Methods for Recognition and Classification of Human Motion Patterns – A Prerequisite for Intelligent Devices Assisting in Sports Activities, vol. 45, no. 2. IFAC, 2012.

[25] M. Kleinsmith, "Zero to Hero: Guide to Object Detection using Deep Learning: Faster R-CNN,YOLO,SSD," 2016. [Online]. Available: http://cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd/.

[26] P. Viola and M. Jones, "Robust real-time face detection," Proc. Eighth IEEE Int. Conf. Comput. Vision. ICCV 2001, vol. 2, pp. 747–747, 2001.

[27] L. Yolanda, M. Martín, L. Gutiérrez Mendoza, L. Mary, and A. Nieves, "Guidelines to design of virtual learning objects (VLO). Application to the Teaching-Learning Process of Area under the Integral Calculus Curve Lignes directrices pour la conception d' objets d'apprentissage virtuels (OAV). Application au processus d' ense," Rev. Científica Gen. José María Córdova, vol. 14, pp. 127–147, 2016.

[28] R. Klette, Concise Computer Vision. 2014.

[29] Garcia-Quilachamin W. Concepción L.P. Herrera-Tapia J. Salazar R.J.Toala-Mero W., "Validation of an algorithm for the detection of the image of a person using multiple cameras," in Applied Technologies. ICAT 2019. Communications in Computer and Information Science, vol 1194. Springer, Cham, 2020, pp. 486–501.