ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

DETECTION OF PERSONS AND HEIGHT ESTIMATION IN VIDEO SEQUENCE

Ondrej Kainz*, František Jakab*, Miroslav Michalko*, Peter Feciľak*
* Department of Computers and Informatics, Technical University of Košice, Slovakia

DOI: 10.5281/zenodo.48321

ABSTRACT

The principal goal of this paper is the design and subsequent development of a solution for visual monitoring of specific area. Monitoring includes detection of movement and detection of person in the video sequence. Further additional information is to be extracted, i.e. the number of persons in the area and the height of subjects. Authors of paper propose own solution based on prior comparative analysis of current works and design mobile solution, where the development board handles all the data processing. Intel Galileo development board was selected. Implementation and subsequent testing proves the hardware and software solution to be fully functional.

KEYWORDS: calibrated camera; detection; height estimation; human body; monitoring; video sequence.

INTRODUCTION

Detecting the objects in the image or stream of images is a complex task that has been addressed by many authors. This field of computer science is very important and addresses wide range of techniques and processes. Main goal of the detection process is finding specific object and distinguish it from the background. In order to do that several image processing techniques are being deployed. Utilization of detection provides broad range of use – from the area of human detection to the detection of tumors in the health industry.

The topic we focus on in this paper is detection of human body in the dynamic image, i.e. video sequence. Solution we present may be used in the field of security for the monitoring of premises or simple monitoring of selected area.

Another important area of interest is the height estimation of a person from a visual means. Information about the height may be used in criminology or for person identification – note that it may be also used as a one parameter for recognition of individual.

In this research the main focus is on the utilization of the deployment board and its sole use for both just stated techniques, i.e. detection and height estimation. Solution is to be fully functional and ready to be utilized in the real-world environment.

STATE OF THE ART: MOTION DETECTION

Detection of person in the image is very common and very widely used process that covers numbers of techniques. Most often utilized techniques are background subtraction [1], utilization of optical flow or temporal differences. Efficient and lately used is also technique called histogram of oriented gradients (HOG). In the next paragraphs we present the most recent developments in these fields of research.

Background subtraction was used by [2] to detect moving objects, authors presumed the static background being the input for the system, research was also focused on the detection of human beings, however also other moving objects were considered. Authors in [3] used also this technique to obtain silhouettes, while achieving scale and view invariance together with the fuzzy logic, which is being used to model a robust background. Research states the accuracy of a solution to be 98 percent. Combination of multiple techniques, i.e. HOG, segmentation and background subtraction was carried out by [4], authors use background subtraction just to subtract the person. Subtraction was also used in [5] for detection of foreground objects that were later on processed for human head detection.

Optical flow approach and region segmentation was used by [6], while implementing also geometric flow proved to be successful for human detection process. Authors present this method to be superior over HOG. The analogous technique is used by [7] to reduce the HOG computation time. Optical flow algorithm was used not only for human

detection but also for gesture motion detection, by [8]. Another approach, i.e. anomaly detection was proposed by [9], where authors propose Histograms of optical flow orientation and magnitude to standardize patterns in the scene. Note that optical flow algorithm may be used also for recognition [10].

Temporal differences was used by [11] and focused on detection of moving objects in general, however human detection was also considered, technique was as well used by [12] and [13]. Generally this technique is less popular than above stated.

Last and the widely utilized is HOG. Authors in [14] utilized it to detect human driving a bicycle, as expected outdoor scene provided better detection results of 96.5 percent. In [15] authors also used HOG solely for person detection, approach utilizes also support vector machine (SVM) for human body components identification and thus enhances the detection process. Detection of person using fisheye camera was proposed by [16], authors also used both HOG and SVM with detection rate of 98 percent. Use of skin color was introduced to HOG by [17], where it was used to detection with detection rate of again 98 percent. Another combination of HOG and LogitBoost machine learning algorithm by [18] also provided detection rate to by 98 percent. For the standard HOG detection rates see [19].

DETECTION OF PERSON AND HEIGHT ESTIMATION

In the following part is introduced method for human detection and height estimation in the dynamic video sequence in a real time. Such a detection is carried out by analysis of individual pixels from reference background and further via comparison with the new image, thus the object may be clearly identified. Detection of human body is utilizing proposed novel algorithm that is easy and fairly accurate. Height estimation is based on prior camera calibration. As the hardware the microcontroller Intel Galileo is utilized and the whole calculation is also being handled by this very device.

Hardware Solution

Development board is the crucial part of this solution, specifically was selected Intel Galileo Gen 1, this board allows communication with camera over IP protocol. IP camera is to use power over Ethernet feature – model of IP camera Vivotek FD7131 was selected based on prior testing.

Intel Galileo board uses operating system Linux, which is located on MicroSD card embedded in the development board itself. Traditional networking devices, i.e. router or switch are presumed in the implementation.

In Fig. 1 is shown the proposed hardware implementation of project. IP camera is to be connected in the same networking segment as the board and the board is to carry out the detection processes. Height of person is estimated based on prior camera calibration, further image and additional information of number of persons in the room and their height is feeded to the web interface. Solution was designed to be accessible from the internet.

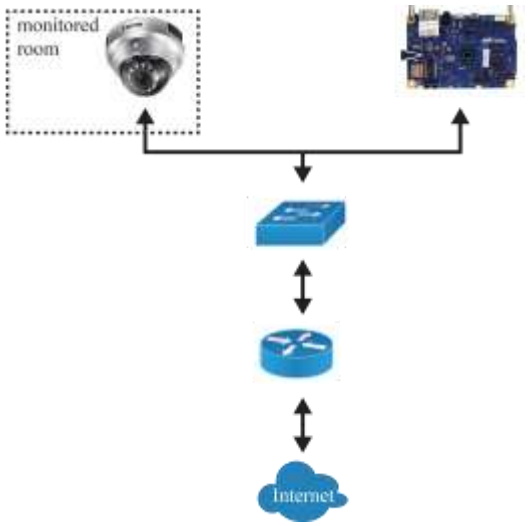

Fig. 1. Scheme of hardware deployment

Software Solution

Analysis was carried out prior to the software implementation, focused on both motion detection and person detection. The solution itself is running on the development board and is as well compiled on this device. The following steps were required for a hardware device to be fully operational and meet criteria set for our solution:

I. Operating system installation for the development board Intel Galileo via MicroSD card.
II. IP camera configuration and further of connection between the development board and IP camera.
III. Initialize the processes within the operating system.
IV. Calibrate the camera and extract the calibration matrix.
V. Implement motion detection within the system, based on comparison of two following frames.
VI. Implement detection of persons based on frame analysis.
VII. Implement height estimation based on prior camera calibration.
VIII. Develop web interface that presents the information as extracted from live stream, i.e. images with detected persons.

Operational Structure of Solution

The overall software solution is consisted from multiple components, the main component are detection of motion, detection of person on the frame and height estimation. All of these were implemented in C and compiled by the very device. Main component is handling the following:

I. Extraction of the first frame containing the background from the IP camera. This frame is considered as a reference background image.
II. Component extracts the frames from camera in cycle, specifically two of them in one moment. All the frames as stored on the development board.
III. Component is carrying out the comparison between the reference image and the extracted frame. Note that all of the pixels are compared.
IV. Once the motion is detected in the previous step, the component initiates the detection process of a person in the frame. If the positive result is obtained the output is stored and later provided to the interface.
V. Before sending the visual data to the interface the information of number of persons and also their height is extracted based on the number of bounding boxes and their sizes.
VI. Component is also providing the output of the detection and estimation in the logged file.

Component runs in cycles, in total three. First downloads the current frame from the IP camera and the latter is used for detection processes. The last one is for estimation of size and also information on the number of the persons in the room. In Fig. 2 is shown the workflow diagram of the component.

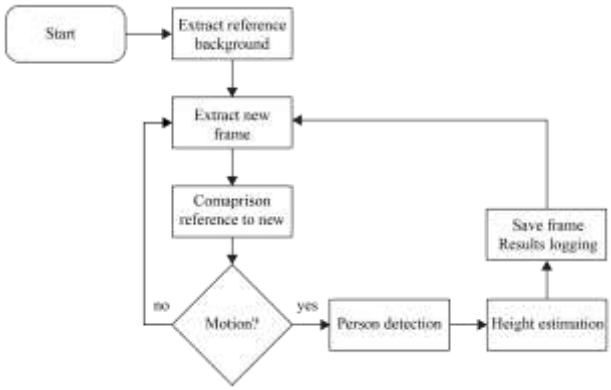

Fig. 2. Operational structure: Component workflow

Motion Detection

The project is to serve for several things, motion detection and subsequent human body detection, and height estimation. Motion detection is crucial step in the whole process. The extracted reference frame is acquired while no human in present in the frame area and thus serves as a static background. Further it is converted to a binary file format for more efficient work with pixels. Then the new frames are being extracted in a cycle and are compared with the

reference frame. Motion detection process is comparing the intensities of the specific pixels, and the threshold may be set within the program. Setting of the proper threshold is highly important and was done after series of tests under the various lightning conditions, e.g. if the value is low it causes too many false positives and vice versa, if the value is high too little true positives are detected. Motion is detected if the specified number of positive pixels is reached.

Person Detection

Another very important step after motion detection is person detection in the frame. This detection step is initiated only if the output of the motion detection is positive, else no frame is considered for human body detection. Detection process in this project is based on the location of detected pixels and also their total number, e.g. detected pixels have to be next to each other for a condition to be successful. Characteristics of human silhouette allow us to define the probable shape of human body and thus enable detection process. Following prior tests the solution is set with the number of pixels to be forming a silhouette. As well, the tool is set to exclude any objects that are not human body silhouettes. One single silhouette is represented by a square of m x n pixels, its size is acquired based on testing in various locations and poses. The outputs of this detection are logged and each bounding box used to mark the person is numbered. In Fig. 3 is depicted the detection of human body in the video sequence.

Fig. 3. Extracted visual output from detection process

Height Estimation

The last step is the height estimation. Camera calibration was utilized to enable this process. Note that height and angle of the camera were known. Camera distortion had to be compensated, since the utilized camera is for survailance of the room and has noticable distortion. Extracted camera parameters allow calculation of the viewing angle and marginal points of the detected human body. Testing was carried out on the number of 30 persons with the standard deviation of approx. 5 centimeters while standing. This may be caused by the type of camera used, mainly the camera resolution and camera placement. In the future research various types of cameras are to be deployed in the system.

DETECTION RESULTS

Two various scene were tested as a part of the implemented project. Testing in the both scenes were tested for various sensitivities of a pixel's intensity. Following the output we may conclude that the detection process is successful. In the Table 1 and 2 are depicted the extracted detection outputs:

Table 1. Results of Human Body Detection: First Scene

First Scene	Low sensitivity	Medium sensitivity	High sensitivity
True positive	71%	60%	58%
False positive	29%	40%	42%

Table 2. Results of Human Body Detection: Second Scene

Second	Low	Medium	High

Scene	sensitivity	sensitivity	sensitivity
True positive	47%	88%	92%
False positive	43%	22%	8%

In both cases were considered inside premises, i.e. scene one and two. For the first case, the area of lecture room was monitored with fair lightning condition. The latter scene was hall with lower lightning conditions. Camera was attached to a ceiling in a first and to a side wall in the second case. Sensitivity in the second scene was better with higher sensitivity due to the dark area characteristics.

CONCLUSION

The principal goal of this paper was the design and deployment of mobile solution for detection of human body and subsequent estimation of its height. The solution proved to be usable in the real environment. The main novelty in this approach is the utilization of solely development board for detection and calculation processes. Utilized board was Intel product – Intel Galileo generation 1 and allowed processing of video stream in a real time from the IP camera. Restrictions of such solution are hardware resources that are limited for this development board and hence the choice of motion detection and human detection algorithms was limited. Bearing this in mind the chosen detection algorithm was combination of multiple techniques used in the area of object detection. Height estimation of person proved to be also possible using our solution and may be further used for monitoring purposes.

In the future research the system is to be deployed in fully real-world environment and is to serve as a basis for development of more complex monitoring system.

ACKNOWLEDGEMENTS

We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, supported by the Research & Development Operational Programme funded by the ERDF.

REFERENCES

[1] A. Elgammal. Background Subtraction:Theory and Practice. Morgan & Claypool, 2014.
[2] S. Suresh, P. Deepak and K. Chitra, "An Efficient Low Cost Background Subtraction Method to Extract Foreground Object During Human Tracking," in International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, 2014, pp. 1432 – 1436.
[3] A. Mahapatra, T.K. Mishra, P.K. Sa, and B. Majhi, "Background Subtraction and Human Detection in Outdoor Videos using Fuzzy Logic," in IEEE International Conference on Fuzzy Systems (FUZZ), Hyderabad, 2013, pp. 1 - 7.
[4] S. Park and J.-H. Yoo, "Human Segmentation Based on GrabCut in Real-time Video Sequences," in IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, 2014, pp. 111-112.
[5] Y. Wang, Z. Zhou, E.K. Teoh and B. Su, "Human Hair Segmentation and Length Detection for Human Appearance Model," in 22nd International Conference on Pattern Recognition (ICPR), Stockholm, 2014, pp. 450-454.
[6] H. Han and M. Tong, "Human Detection Based on Optical Flow and Spare Geometric Flow," in Seventh International Conference on Image and Graphics (ICIG), Qingdao, 2013, pp. 459-464.
[7] D. Kumar and K.V. Suresh, "HOG-PCA Descriptor with Optical Flow based Human Detection and Tracking," in International Conference on Communications and Signal Processing (ICCSP), Melmaruvathur, 2014, pp. 900-904.
[8] Y. Lyu, Y. Yang and J. Ru, "Gesture Motion Detection Algorithm Based on Optical Flow Method," in IEEE International Conference on Computer and Communications, Chendgu, 2015, pp. 128-132.
[9] R.V.H. Colque, C.A. Caetano and W.R. Schwartz, "Histograms of Optical Flow Orientation and Magnitude to Detect Anomalous Events in Videos," in 28th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Salvador, 2015, pp. 126-133.
[10] K. Subramanian, V.B. Radhakrishnan and S. Sundaram, "An Optical Flow Feature and McFIS Based Approach for 3-Dimensional Human Action Recognition," in IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Singapore, 2014, pp. 1-6.
[11] N. Singla, "Motion Detection Based on Frame Difference Method," International Journal of Information &

Computation Technology, vol. 4, no. 15, pp. 1559-1565, 2014.
[12] M.P. Paulus, J.S. Feinstein, S.F. Tapert and T.T. Liu, "Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection," Neuroimage, vol. 21, no. 2, pp. 733-743, 2004.
[13] W. Shuigen, Ch. Zhen and D. Hua, "Motion Detection Based on Temporal Difference Method and Optical Flow field," in Proceeding ISECS '09 Proceedings of the 2009 Second International Symposium on Electronic Commerce and Security, Washington, 2009, pp. 85-88.
[14] H. Jung, J.K. Tan, S. Ishikawa and T. Morie, "Applying HOG Feature to the Detection and Tracking of a Human on a Bicycle," in 11th International Conference on Control, Automation and Systems (ICCAS), Gyeonggi-do, 2011, pp. 1740-1743.
[15] A. Rajaei, H. Shayegh and N.M. Charkari, "Human Detection in Semi-dense Scenes Using HOG descriptor and Mixture of SVMs," in 3th International eConference on Computer and Knowledge Engineering (ICCKE), Mashhad, 2013, pp. 229-234.
[16] A.-T. Chiang and Y. Wang, "Human detection in fish-eye images using HOG-based detectors over rotated windows," in IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Chengdu, 2014, pp. 1-6.
[17] X. Meng, J. Lin and Y. Ding, "An Extended HOG Model: SCHOG for Human Hand Detection," in International Conference on Systems and Informatics (ICSAI), Yantai, 2012, pp. 2593-2596.
[18] H. Ninomiya, H. Ohki, K. Gyohten and N. Sueda, "An Evaluation on Robustness and Brittleness of HOG Features of Human Detection," in 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), Ulsan, 2011, pp. 1-5.
[19] P.-Y. Chen, Ch.-Ch. Huang, Ch.-Y. Lien and Y.-H. Tsai, "An Efficient Hardware Implementation of HOG Feature Extraction for Human Detection," in IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 2, pp. 656-662, October, 2013.

AUTHOR BIBLIOGRAPHY

	**Ondrej Kainz** was born in 1988. In 2013 he graduated (MSc. in Applied informatics) from the Technical university in Kosice, Slovakia. Since the very same year he is a PhD student at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at the Technical university of Kosice. His scientific research interests include computer vision, e-learning, human-computer interfaces, computer graphics, computer networks, biological engineering and body area network.
	**František Jakab** was born in 1959. Graduated from St. Petersburg Electro technical University, PhD degree received from Technical university of Kosice & is associated professor since 2008. He has extensive experience in networking and utilization of ICT in education where he established well known research centre - Computer Networks Laboratory. He has been an coordinator of several large international projects, coordinator of national wide ICT projects and research grants. Involved into Cisco Networking academy program in Slovakia and since 2001 in position of coordinator of Program in Slovakia, 2008-2014 He was awarded as a "IT person of the year 2006" in Slovakia.
	**Miroslav Michalko** received his M.Sc. and Ph.D. degree in Informatics from Technical University in Kosice (Slovakia). For more than 10 years he is member of well recognized research institution - Computer Networks Laboratory at Department of Computers and Informatics at Technical University of Kosice (DCI TUKE). Now he is an Assistant Professor at DCI TUKE and his lectures are focused on Computer Networks, he is more than 10 years active instructor in Cisco Networking Academy. His research includes multimedia content delivery, video streaming services, web and cloud services, multimedia on mobile devices, innovative

	teaching&learning techniques and IoE/IoT solutions.
	**Peter Feciľák** completed doctoral studies in the field of informatics at the Technical University of Kosice. His research activities focus on process optimization in routing and quality of services in computer networks as well as on developments in smart-grid technologies. He is actively participating in various projects (KEGA, VEGA, grant projects in cooperation with commercial partners). Within the Networking Academy program, NetAcad, coordinates the activities of more than 70 networking academies in several countries. As instructors trainer of Networking Academy program he also completed industry certification of Cisco CCNA, CCNP, Security.