Human Surveillance System for Security Application

Mozammel Chowdhury(✉), Junbin Gao, and Rafiqul Islam

School of Computing & Mathematics, Charles Sturt University, Bathurst, Australia
{mochowdhury,jbgao,mislam}@csu.edu.au

Abstract. Human surveillance is an important research activity for security concern. Due to the increasing demand of security in different domains, development of smart and efficient surveillance system has attracted immense interest in recent years. Most of the existing surveillance systems are based on monocular camera and limited by their fixed view angles and hence cannot provide sufficient three-dimensional depth information for person recognition and tracking. This paper proposes an efficient and cost-effective human surveillance system using stereo vision technique. The system uses a multi-view stereo camera pair for image capturing and analyzes the stereoscopic pictures to estimate the 3D depth information for accurate detection and tracking of the human objects. The system can provide automatic warning in case of unrecognized people and entrance in the restricted zones. Experimental results are arranged to demonstrate the robustness and efficiency of our proposed system. Our system is very inexpensive and computationally fast comparable to the existing state-of-the-art surveillance systems.

Keywords: Security · Surveillance system · Access control · Stereo vision · 3D depth information · Person detection and tracking

1 Introduction

Human surveillance is attracting more importance nowadays due to the increasing demand of security and defense in different environments including door access control, border surveillance, immigration control, monitoring employee activities, identifying suspicious people, theft and vandalism deterrence, preventing criminal acts and so on [1, 2].

Several techniques have been developed in last decades for automatic surveillance of people using CCTV cameras and sensors. According to the number of cameras used in these techniques, surveillance systems can be classified into two categories: monocular and multi-camera based system. Most of the conventional surveillance systems widely used for security applications in supermarkets, airports, stations, ATM booths and other public places, employ monocular or single camera [3-7]. They are limited by their fixed view angles, fixed resolutions and limited depth information. These limitations make it complex to estimate and recover the precise 3D information as well as motion behavior of human objects for accurate and robust tracking.

B. Thuraisingham et al. (Eds.): SecureComm 2015, LNICST 164, pp. 711–724, 2015.
DOI: 10.1007/978-3-319-28865-9_47

Due to the recent advancement in vision technology, multiple camera based surveillance systems have attained the superiority for tracking people with different view angles [8-10]. These systems are capable of viewing an object from multi-viewpoints and hence can deal better with occlusions. However, such systems are very expensive and difficult to set up due to the problems of establishing their geometric relationships or synchronizations since they require a large number of cameras [11].

In recent years, researchers have proposed stereo vision based surveillance systems for security applications [12-14]. Stereo vision has the advantage to estimate the 3D position of an object in a given coordinate system from two stereoscopic images [15]. Stereo vision based surveillance systems can easily segment an image into objects to distinguish people from their shadows and provide more accurate location information for their tracking. Most of these systems generally employ a static pan-tilt-zoom (PTZ) camera, whose pose can be fully controlled by pan, tilt and zoom parameters. The PTZ cameras are able to obtain multi-angle views and multi-resolution information However, the main disadvantage of these cameras is that they are unidirectional and their image resolution is poor [16, 20]. The existing stereo based surveillance systems are computationally expensive and hence they are not suitable for real time applications.

To overcome the aforesaid challenges in surveillance systems, we propose an efficient, computationally fast and cost-effective surveillance model for security applications. The system employs a low-cost stereo camera pair for image capturing and recovers the 3D depth information of the human object exploiting a fast stereo vision algorithm. The proposed system includes robust and efficient algorithms for human face identification, stereo correspondence matching, 3D depth extraction, and location estimation of the human objects for security monitoring.

The rest of the paper is structured as follows. Section 2 provides a brief discussion on most related works in this area. In Section 3, we present the architecture of our proposed surveillance system. Experimental results with real time image sequences are reported in Section 4. Finally, Section 5 concludes the paper and gives directions for future work.

2 Related Works

In the last few decades, lots of related research works have been performed by the researchers to develop smart and efficient human surveillance systems. In this section we attempt to review some of them which are more relevant.

Chen *et al.* [16] propose a vision system based on an omnidirectional camera and a PTZ camera. The omnidirectional camera monitors the surveillance object and the PTZ camera captures the image of the target object. These two cameras work in a master-salve mode. This surveillance system is employed mostly in indoor environment. Adorni *et al.* [17] propose a binocular vision system using two omnidirectional cameras which is generally used in robot vision. The system is capable of enlarging view fields. The main disadvantage of these vision systems is that the image resolution of the omnidirectional camera is poor. This drawback has great limitation towards their application in real time.

Munoz-Salinas *et al.* [18] propose an object tracking method which can combine color and depth information using dual static cameras. In another work [19], they use plan-view maps to represent stereo information more efficiently. The main drawbacks associated with these systems are that they cannot obtain multi-resolution and multi-visual-angle information.

Bimbo *et al.* [20] propose a novel framework exploiting two PTZ cameras aiming to relate the feet position of a person in the image of the master camera with the head position in the image of the slave camera. Benjamin *et al.* [21] present a multi camera based surveillance system that can automatically extract useful information from a given scene. It also alerts the user if the tracked object breaks certain defined regulations. Wang [22] illustrates an overview of the recent advances in the field of multi camera video surveillance. It compares the existing solutions and also describes the prevalent technical challenges.

Darrell *et al.* [23] propose a system for tracking people using stereo cameras. The stereo method is used to isolate people from other objects and background. They integrated color and face detection modules in the system. Bahadori and Iocchi [24] propose a semi-automatic surveillance system for museum environment using stereo vision. The system can detect the situations for providing warning messages to the surveillance personnel.

Manap *et al.* [25] propose a system for smart surveillance using stereo imaging. The system uses two smart IP cameras to obtain the position and location of objects. The position and location of the object are automatically extracted from two IP cameras and subsequently transmitted to an ACTi Pan-Tilt-Zoom (PTZ) camera, which then points and zooms to the exact position in space. This work involves video analytics for estimating the location of the object in a 3D environment and transmitting its positional coordinates to the PTZ camera.

Cui and Li [26] propose a surveillance system mainly used for indoor scene monitoring using binocular vision. The system uses two PTZ cameras and employs a rectification-disparity-based method to establish correspondence between two image sequences. The system utilizes depth information to deal with occlusion problem in object tracking.

However, these existing surveillance systems are commercially expensive and require high computation time. With a view to overcome the limitations of the existing systems, we attempt to propose an effective and inexpensive surveillance model using a fast stereo vision technique. The proposed system can measure the precise 3D position of the human object for accurate detection and tracking.

3 Proposed System Architecture

The proposed human surveillance system consists of the following components: (i) Pre-processing for image refinement, (ii) Face detection for person identification, (iii) Stereo correspondence matching for finding disparities in the image sequences,

(iv) Dense depth estimation for recovering the 3D position of the human object, and (v) Person tracking or localization. Fig.1 shows the general architecture of the proposed system.

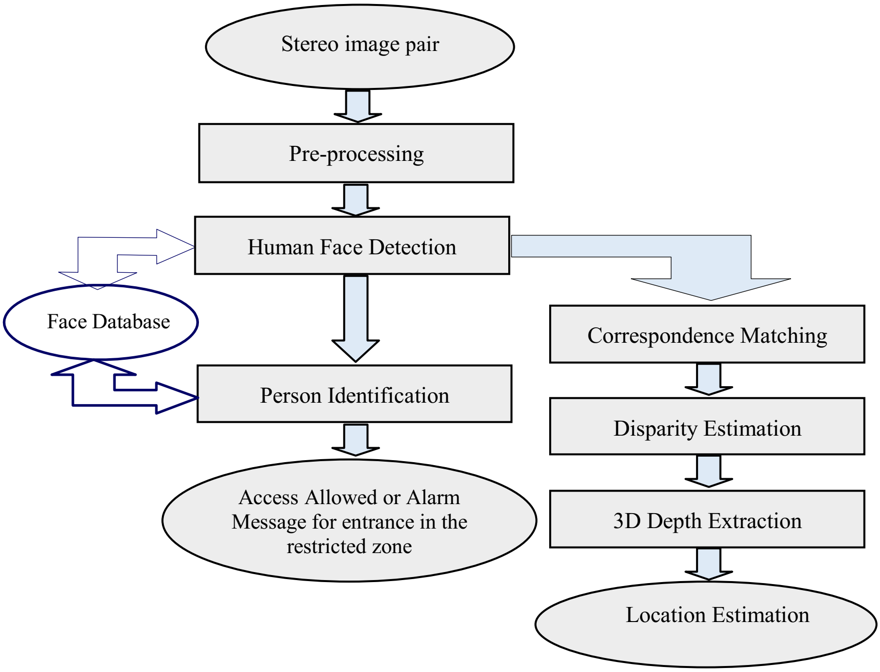


Fig. 1. Architecture of the proposed surveillance system.

3.1 Preprocessing of the Stereo Images

In real time stereo vision systems, there may be significant amount of noise in the captured image pair due to the differences in camera orientation and lighting condition. For this reason, we employ a fuzzy median filtering technique [27] for refining the stereoscopic images corrupted by noise. This filter employs fuzzy rules for deciding the gray level of the pixels within a window in the image. This is a variation of the Median filter and neighbourhood Averaging filter with fuzzy values.

3.2 Face Detection

Human face plays an important role in person recognition in vision-based surveillance system. Face detection is concerned with determining the part of an image which contains face. Different techniques [28-32] have been developed for face detection in last decades, which includes: geometric modeling, genetic approach, neural network, principal component analysis, color analysis and so on.

This paper proposes a fast and robust face detection technique based on skin color segmentation. For detecting the face area, the image is first enhanced using histogram equalization because, the face images may be of very poor contrast because of the limitation of lighting conditions. Then the face skeleton is detected from the largest connected area of the skin color segmented image. The method considers the frontal view of the face in color scale image. The overview of the proposed color histogram based face detection method is shown in Fig. 2.

The efficiency of color segmentation of a human face depends on the color space. While the input colour image is typically in RGB format, the RGB model is not used in the detection process because the RGB colour model is not a reliable model for detecting skin colour [29]. The RGB components are subject to luminance change and hence face detection may fail if the lighting condition changes from image to image. Consequently, we use HSV color model for fast and effective detection process.

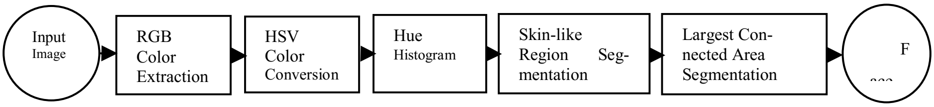


Fig. 2. Block diagram of the face detection method.

In the HSV color model a color is described by three attributes Hue, Saturation and Value. Hue is the attribute of visual sensation that corresponds to color perception associated with the dominant colors, saturation implies the relative purity of the color content and value measures the brightness of a color. The HSV space classifies similar colors under similar hue orientations. The image content is converted from RGB to HSV color space using the following equations:

$$H = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right\} \tag{1}$$

Ranging [0,2π], where $H = H_1$ if $B \leq G$; otherwise $H = 360^\circ - H_1$;

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \tag{2}$$

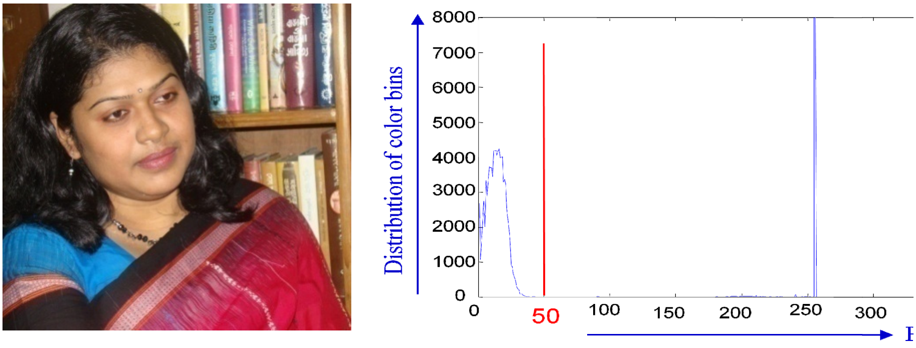
$$V = \frac{\max(R,G,B)}{255} \tag{3}$$

Where R, G, B are the red, green and blue component values which exist in the range [0,255].

Let a color image $I(x, y)$ consists of three color channels $I = (I_R, I_G, I_B)$, at (x, y) of size $M \times N$. First a hue histogram $H(i)$ is obtained by counting the number of pixels, given by the following equation:

$$H(i) = \frac{n(H(I_R, I_G, I_B) = i)}{M \times N} \tag{4}$$

Where n indicates the number of pixels with a hue value $H(I_R, I_G, I_B) = i$ and $M \times N$ is the total number of image locations. Fig. 3 shows a typical color image and its corresponding hue histogram.



(a) A Color image (b) Hue histogram of the color image

Fig. 3. Hue histogram of a typical color image.

3.3 Stereo Correspondence Matching

Stereo matching algorithms are used for correspondence matching of points or blocks between two image sequences to estimate the accurate depth and location of the human object. In stereo vision, two images of the same scene are taken from slightly different viewpoints using two cameras: left and right camera of same focal length and parameters, which are placed in the same lateral plane. Fig. 4 shows the process of capturing an object by two stereo cameras placed in different viewpoints separated by small distance.

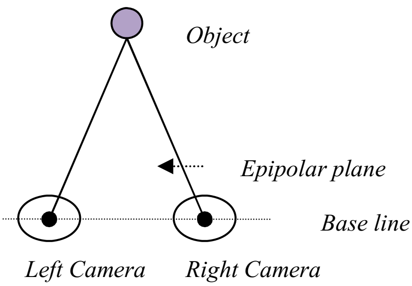


Fig. 4. Stereo vision process: an object is captured by two horizontally aligned cameras.

The main task in stereo vision is finding correspondence matching to estimate the dense disparity. In stereo imaging, for most pixels in the left image there is a corresponding pixel in the right image in the same horizontal line. The difference in the coordinates of the corresponding pixels is known as disparity, which can be expressed by the following equation:

$$d = x_L - x_R \tag{5}$$

Stereo algorithms are mainly classified into two categories: local and global methods [33]. The local algorithms [33-36], also referred to as window-based or area-based algorithms are typically faster and suitable for real time applications rather than global approaches [37- 40]. However, they have less accuracy compared to global methods. The local or window-based stereo algorithms traditionally estimate dense disparity by means of pixel correspondence matching through window cost computation using any one of the following statistical measures: sum of absolute differences (SAD), sum of square differences (SSD), or normalized cross correlation (NCC) [15, 34]. To determine the correspondence of a reference pixel in the left image, the window costs are calculated for all target pixels on the same epipolar line in the right image within a search range. The pixel in the right image that gives the best window cost i.e., the minimum SAD/SSD value or the maximum NCC value indicates the corresponding pixel of the reference pixel in the left image.

In this work, we consider the detected face area in the left image as a reference block and match with another target face block in the right image. The window cost $W_C(x, y, d)$ of a reference pixel at position (x, y) in the left image block with disparity d is computed with the following SAD measure, employing a window centered at position (x, y) in the left image block and another window centered at position $(x+d, y)$ in the corresponding right image block.

$$W_c^{SAD}(x, y, d) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} \left| f_L(x+i, y+j) - f_R(x+i+d, y+j) \right| \tag{6}$$

Where $f_L(x, y)$ and $f_R(x, y)$ are the intensities of the pixels at position (x, y) in the left and right image blocks, respectively. $(2m+1)$ and $(2n+1)$ are the width and height of the rectangular window, respectively.

In this paper, we propose a fast algorithm based on local approach to compute the window costs for correspondence matching. To determine the correspondence of a pixel in the left image block, we just compute the window cost for candidate pixels in the right image block whose intensities are different within a certain threshold value (δ). To achieve a substantial gain in accuracy with less expense of computation time, our algorithm perform correspondence matching only on the diagonal pixels of the square windows rather than employing conventional matching upon all pixels in the windows. Empirically we find that this diagonal matching operation reduces significant computation time compared to the state-of-the-art stereo methods, which is a fruitful contribution towards the development of a fast and effective surveillance system.

The proposed fast correspondence matching or disparity estimation algorithm is depicted as follows:

Algorithm Proposed Stereo matching

1. For each candidate pixel (x,y) in the left image, search the corresponding pixel on same epipolar line in the right image within a search range employing a square diagonal window:

for $d' = -d_{\max}$ to $+d_{\max}$ **do**

if $|f_L(x, y) - f_R(x + d', y)| < \text{threshold}(\delta)$ **then**

Calculate $W_C(x, y, d')$

2. Find d such that, $d = \arg\min W_C(x, y, d')$

3. Repeat steps 1 and 2 to calculate disparities of all pixels in the left image.

3.4 Depth Extraction and Location Estimation

In stereo vision, the depth or 3D information of points in the images can be calculated from the estimated disparity map and the geometry of the camera settings. This process is illustrated in Fig. 5 where, L and R are two pinhole cameras with parallel optical axes; O_L and O_R are two center points of the left and right camera respectively with same focal length f. The baseline, which is the line connecting the two lens centers of the cameras is perpendicular to the optical axes. Let b is the baseline distance and x_L is the x-coordinate of the projected 3D point onto the left camera image plane and x_R is the x-coordinate of the projection onto the right image plane.

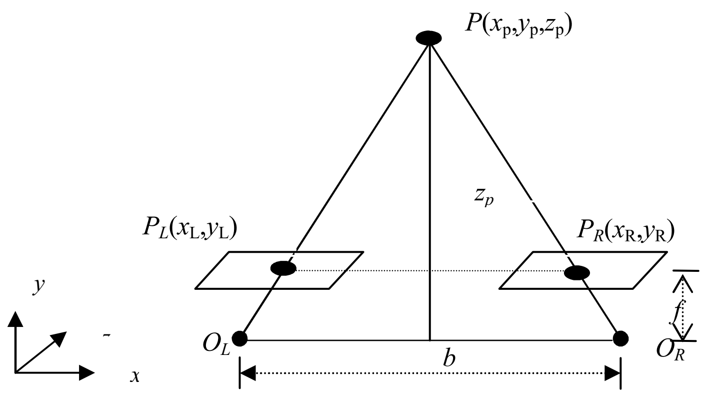


Fig. 5. Dense depth estimation in stereo image pair using triangulation

For a world 3D point $P(x_p, y_p, z_p)$, we can extract the dense depth information from the camera geometry as follows:

$$\frac{z_p}{b} = \frac{z_p - f}{b - (x_L - x_R)} \tag{7}$$

Thus the depth,

$$z_p = \frac{bf}{x_L - x_R} = \frac{bf}{d} \tag{8}$$

This is the distance of the target object from the stereo camera positions.

We can recover the 3D point P from its projections P_L and P_R. Therefore, we have:

$$x_p = \frac{bx_L}{x_L - x_R} = \frac{bx_L}{d} \tag{9}$$

$$y_p = \frac{by_L}{x_L - x_R} = \frac{by_L}{d} \tag{10}$$

Since the depth, z_p indicates a distance value (i.e. in mm or cm), we have to modify the equation (8) for its uniformity because, the parameters (b, f, d) in the equation possess different units. This modification is vital during measuring the distance of objects, otherwise it gives erroneous result. Accordingly, we can reform the equation (8) through converting the unit of the disparity value (d) by dividing it with the pixel size (normally in mm/pixel) of the camera. Thus the depth or distance of the target object becomes,

$$z_p = \frac{bf}{ds} \tag{11}$$

Where, s is the size of a pixel of the stereo camera. Thus, once we can estimate the 3D depth or distance value of the target, we can easily track or localize the human object.

4 Experimental Evaluation

The effectiveness and robustness of this approach is justified using different images captured by the stereo camera pair with different positions, expressions and lighting conditions. Experiments are carried out on a computer with 2.2 GHz Intel Core i5 processor and 4GB RAM. The algorithm has been implemented using Visual C++. We use two SONY VISCA cameras of same focal length and intrinsic parameters for stereo imaging.

The face images are analyzed to demonstrate the feasibility of the proposed detection method. When a complex image is subjected in the input, the face detection result highlights the facial part of the image. The face detection results with our proposed method are depicted in Fig. 6. Images of different persons are taken at different environments both in shiny and gloomy weather. To evaluate our proposed method we consider images with different expressions, pose, orientation, structural components and illumination. The system can also cope with the problem of partial occlusion. Our system demonstrates better performance in case of the frontal face images in simple background while provides worst results for the images in complex background. We perform experiments to compare our proposed algorithm with RGB and YIQ (Luminance, Hue, Saturation) based face detection methods. The results as shown in Table 1 and Table 2, confirms the robustness of our proposed face detection algorithm comparable to others methods.

Fig. 7 represents the stereo matching results with our proposed stereo algorithm. The algorithm computes the disparity values through matching the correspondence pixels within the selected blocks in the left and right image pair. The left and right image sequences are captured through the left and right camera respectively, placed in same epipolar axis. The correspondence matching is accomplished using SAD measures using a window of size 3×3 pixels. The disparities are computed with a search range of –10 to +10 pixels for a threshold level of 25. We estimate these parameters empirically in order to optimize quality of disparity results.

Experientially we find that computational cost increases with the enlargement of the window size. Fig. 8 represents a plot of computational time for different window size, which shows that a window of size 3×3 pixels is a good choice in respect to computational speed. We compare our algorithm with similar stereo methods in terms of execution time, as reported in Table 3. Experimental results confirm that our proposed disparity estimation algorithm outperforms with significant reduction of computation time compared to other existing methods.

(a) Original image (b) Detected face (a) Original image (b) Detected face

Fig. 6. Face detection process: (a) the original image, and (b) detected face image.

We estimate the distance or location of the human objects for five real image pairs using the obtained disparity values and known camera parameters. The focal length of the stereo cameras used in this simulation is 35 mm, pixel size is 0.1165 mm, and the baseline distance between two cameras is 20 cm. Table 4 shows the results of the location estimation process.

(a) Left image (b) Right image (c) Estimated disparity

Fig. 7. Stereo matching process: (a) Left image, (b) Right image, and (c) Disparity map.

Table 1. Performance of different detection methods in terms of detection accuracy (in %)

Type of Face image	No. of test image	Detection Accuracy (%)		
		RGB Color Based	YIQ Color Based	HSV Color Based (Proposed method)
Frontal	50	62.37	78.24	98.18
Tilted	50	58.42	74.65	95.34
Partial Occluded	20	56.72	72.29	91.56
Complex Background	30	50.35	67.17	86.72

Table 2. Performance of different detection methods in terms of computation time

Detection Method	Computation Time (second)
RGB Color Based	1.23
YIQ Color Based	0.87
Proposed Method	0.46

Table 3. Performance of different stereo methods in terms of computation time

Method	Computation Time (second)
Linear Stereo Matching [33]	15
Conventional Area Based [34]	1. 23
Large-scale Stereo Matching [36]	0.96

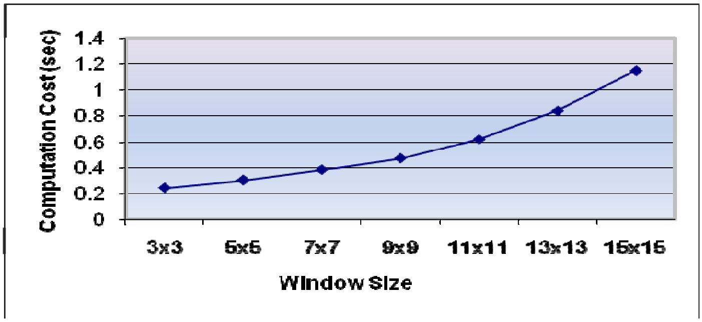


Fig. 8. Computational time versus Window Size.

Table 4. Results of Location estimation

Test	Actual Distance (cm)	Measured Distance (cm)	Error	Error (%)
Test 1	150	148.5821	-1.4179	0.94527
Test 2	200	197.2895	-2.7105	1.35525
Test 3	250	252.6481	2.6481	1.05924
Test 4	300	292.1137	-7.8863	2.62877
Test 5	350	343.7251	-6.2749	1.79283

5 Conclusion

Surveillance systems have gained importance due to the increase of safety and security of people. In this paper we propose an effective and inexpensive human surveillance system consisting of fast and robust algorithms for face detection, stereo correspondence matching, dense depth evaluation and human location estimation. The effectiveness of the proposed algorithms has been justified using real image sequences with complex and simple backgrounds. Experimental evaluation confirms that our algorithms perform superiority comparably to the state-of-the-art existing methods. Our next approach is to extend the algorithm for multi-face detection and location estimation.

References

1. Valera, M., Velastin, S.A.: Intelligent distributed surveillance systems: a review. In: Proceedings of Vision, Image and Signal Processing, pp. 192–204 (2005)
2. Irgan, K., Ünsalan, C., Baydere, S.: Low-cost prioritization of image blocks in wireless sensor networks for border surveillance. Journal of Networks and Computer Applications (JNCA) **38**, 54–64 (2014)

3. Kieran, D., Weir, J., Yan, W.: A framework for an event driven video surveillance system. Journal of Multimedia **6**(1), 3–13 (2012)
4. Danielson, P.: Video surveillance for the rest of us: proliferation, privacy, and ethics education. In: Proc. of International Symposium on Technology and Society, vol. 1, no. 1, pp. 162–167 (2002)
5. Niu, W., Li, G., Tong, E., Yang, X., Chang, L., Shi, Z., Ci, S.: Interaction relationships of caches in agent-based HD video surveillance: Discovery and utilization. Journal of Networks and Computer Applications (JNCA) **37**, 155–169 (2014)
6. Haritaoglu, I., Harwood, D., Davis, L.S.: W4: Real-Time Surveillance of People and Their Activities. IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(8), 809–830 (2000)
7. Pai, C., Tyan, H., Liang, Y., Liao, H., Chen, S.: Pedestrian detection and tracking at crossroads. Pattern Recognition **37**, 1025–1034 (2004)
8. Wang, X.: Intelligent Multi-Camera Video Surveillance: A Review. Pattern Recognition Letters, 1–25 (2012)
9. Bodor, R., Morlok, R., Papanikolopoulos, N.: Dual camera system for multi-level activity recognition. In: Proc. of the IEEE/RJS International Conference on Intelligent Robots and Systems, vol. 1, pp. 643–648 (2004)
10. Yang, T., Li, S.Z., Pan, Q., Li, J.: Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) **1**, 970–975 (2005)
11. Black, J., Ellis, T.: Multi-camera image measurement and correspondence. Measurement **32**, 61–71 (2002)
12. Mittal, A., Davis, L.S., M2Tracker: A multiview approach to segmenting and tracking people in a cluttered scene using region-based stereo. In: Proc. of the 7th European Conf. on Computer Vision (ECCV 2002), pp. 18–36. Springer-Verlag (2002)
13. Darrell, T., Gordon, G., Harville, M., Wood-fill, J.: Integrated person tracking using stereo, color, and pattern detection. International Journal of Computer Vision **37**(2), 175–185 (2000)
14. Ko, J., Lee, J.: Stereo Camera-based Intelligence Surveillance System. Journal of Automation and Control Engineering **3**(3), 253–257 (2015)
15. Chowdhury, M.M., Bhuiyan, M.A.: Fast Window based Stereo Matching for 3D Scene Reconstruction. The International Arab Journal of Information Technology **10**(4) (2013)
16. Chen, C., Yao, Y., Page, D., Abidi, B., Koschan, A., Abidi, M.: Heterogeneous fusion of omnidirectional and PTZ cameras for multiple object tracking. IEEE Transactions on Circuits and Systems for Video Technology **18**(8), 1052–1063 (2008)
17. Adorni, G., Cagnoni, S., Mordonini, M., Sgorbissa, A.: Omnidirectional stereo systems for robot navigation. In: Proc. IEEE Workshop on Omnidirectional Vision and Camera Networks, pp. 79–89 (2003)
18. Munoz-Salinas, R., Aguirre, E., Garcia-Silvestre, M., Gonzalez, A.: A multiple object tracking approach that combines colour and depth information using a confidence measure. Pattern Recognition Letters **29**, 1504–1514 (2008)
19. Munoz-Salinas, R.: A Bayesian plan-view map based approach for multiple-person detection and tracking. Pattern Recognition **41**, 3665–3676 (2008)
20. Bimbo, A.D., Dini, F., Lisanti, G., Pernici, F.: Exploiting distinctive visual landmark maps in pan-tiltzoom camera networks. Computer Vision and Image Understanding **114**, 611–623 (2010)

21. Justin, W.B., Benjamin, L.P., Kyle, K.E., Randy, L.M.: TENTACLE: Multi-Camera Immersive Surveillance System. Small Business Innovative Research (SBIR) Phase I Report, Air Force Research Laboratory (2011)
22. Wang, X.: Intelligent Multi-Camera Video Surveillance: A Review. Pattern Recognition Letters (2012), 1–25 (2012)
23. Darell, T., Gordon, G., Harville, M., Woodfill, J.: Integrated Person Tracking Using Stereo, Color, and Pattern Detection. Computer Vision and Pattern Recognition (1998)
24. Bahadori, S., Iocchi, L.: A stereo vision system for 3d reconstruction and semi-automatic surveillance of museum areas. In: Workshop of the Italian Association for Artificial Intelligence (AI*IA) (2003)
25. Manap, N., Caterina, G., Soraghan, J., Sidharth, V., Yao, H.: Smart surveillance system based on stereo matching algorithms with IP and PTZ cameras. In: 3DTV-CON 2010, pp. 65–68 (2010)
26. Cui, Z., Li, A.: A Novel Binocular Vision System for Surveillance Application. Journal of Multimedia **8**(4), 307–314 (2013)
27. Satter, A.K.M.Z., Chowdhury, M.M.H.: A Fuzzy Algorithm for De-Noising of Corrupted Images. International Journal of Computer Information Systems (IJCSI) **6**(4), 15–17 (2013)
28. Blanz, V., Vetter, T.: Face recognition based on fitting a 3-D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
29. Uddin, J., Mondal, A.M., Chowdhury, M.M.H., Bhuiyan, M.A.: Face detection using genetic algorithm. In: Proceedings of 6th International Conference on Computer and Information Technology, Dhaka, Bangladesh, pp. 41–46, December 2003
30. Rowley, H., Shumeet, H.B., Kanade, T.: Neural Network-Based Face Detection. IEEE Transaction on Pattern Analysis and Machine Intelligence **20**(1), 23–37 (1998)
31. Gopalan, R., Jacobs, D.: Comparing and combining lighting insensitive approaches for face recognition. Computer Vision and Image Understanding **114**(1), 135–145 (2010)
32. Castillo, C.D., Jacobs, D.W.: Using stereo matching with general epipolar geometry for 2-D face recognition across pose. IEEE Trans. Pattern Anal. Mach. Intel. **31**(12), 2298–2304 (2009)
33. De-Maeztu, L., Mattoccia, S., Villanueva, A., Cabeza, R.: Linear stereo matching. In: IEEE International Conference on Computer Vision (ICCV 2011), pp. 1708–1715 (2011)
34. Di Stefano, L., Marchionni, M., Mattoccia, S.: A fast area-based stereo matching algorithm. Image and Vision Computing **22**(12), 983–1005 (2004)
35. Hosni, A., Bleyer, M., Rhemann, C., Gelautz, M., Rother, C.: Real-time local stereo matching using guided image filtering. In: Proc. IEEE-ICME, pp. 1–6 (2011)
36. Geiger, A., Roser, M., Urtasun, R.: Efficient large-scale stereo matching. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part I. LNCS, vol. 6492, pp. 25–38. Springer, Heidelberg (2011)
37. Yang, Q.: Stereo Matching Using Tree Filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence **37**(4), 834–846 (2015)
38. Mei, X., Sun, X., Dong, W., Wang, H., Zhang, X.: Segment-tree based cost aggregation for stereo matching. In: CVPR, pp. 313–320 (2013)
39. Yang, Q.: A non-local cost aggregation method for stereo matching. In: CVPR, pp. 1402–1409, (2012)
40. Tatsunori, T., Yasuyuki, M., Takeshi, N.: Graph cut based continuous stereo matching using locally shared labels. In: CVPR 2014 (2014)