

Research Article

Object Detection and Tracking-Based Camera Calibration for Normalized Human Height Estimation

Jaehoon Jung,¹ Inhye Yoon,^{1,2} Sangkeun Lee,¹ and Joonki Paik¹

¹Department of Image, Chung-Ang University, Seoul 156-756, Republic of Korea

²ADAS Camera Team, LG Electronics, 322 Gyeongmyeong-daero, Seo-gu, Incheon 22744, Republic of Korea

Correspondence should be addressed to Joonki Paik; paikj@cau.ac.kr

Received 6 July 2016; Revised 10 October 2016; Accepted 25 October 2016

Academic Editor: Jesus Corres

Copyright © 2016 Jaehoon Jung et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a normalized human height estimation algorithm using an uncalibrated camera. To estimate the normalized human height, the proposed algorithm detects a moving object and performs tracking-based automatic camera calibration. The proposed method consists of three steps: (i) moving human detection and tracking, (ii) automatic camera calibration, and (iii) human height estimation and error correction. The proposed method automatically calibrates camera by detecting moving humans and estimates the human height using error correction. The proposed method can be applied to object-based video surveillance systems and digital forensic.

1. Introduction

A large-scale video analysis using multiple cameras is gaining attractions in visual surveillance applications. In particular, as the use of the object-based video analysis increases, the demand for extraction of object information is growing up. Since an object information is changed depending on the camera parameters such as a location of the installed camera, viewing angle, and focal length, for this reason, various normalized object feature extraction methods were proposed.

Lao et al. proposed a human motion analysis method for consumer surveillance system [1]. This method estimates human moving trajectories by tracking and recognizing the human motion. Del-Blanco et al. proposed a multiple object detection and tracking framework for the automatic counting of object numbers in a video surveillance application [2]. Lee et al. detected the object region and estimated the depth information using multiple color-filtered apertures (MCA) [3]. Chantara et al. proposed fast object tracking method using adaptive template matching [4]. Chu and Yang detected a moving object using a background model and estimated the object velocity using the object with a previously known length [5]. Maik et al. train the typical poses in both the 2D image and 3D space and represent the located poses as

a silhouette for the human pose estimation [6]. Kang et al. proposed the human gesture detection and tracking method by using the real-time stereo matching [7]. However, this method uses two or more cameras for the depth estimation. In order to estimate the 3D information using single camera, the camera calibration methods [8–11] are proposed. Arfaoui and Thibault used a diffractive virtual grid to estimate camera parameters for a fish-eye lens camera [12]. Neves et al. corrected the fish-eye distortion using parallel lines and then calibrated a static and pan-tilt-zoom (PTZ) cameras using an object height [13]. Zhang et al. tracked homography based on the model plane and then estimated camera parameters using maximum likelihood (ML) approach [14]. Bell et al. used digital display to generate feature points for out-of-focused camera calibration [15]. Kual-Zheng proposed an object height estimation method that extracts feature points and estimates vanishing points using a special pattern such as a cubic box [16]. Gallagher et al. proposed a method to analyze a human age and gender by calibrating a camera and analyzing distances among eyes, nose, and mouth in a human face [17]. Since this method uses a special pattern board for the camera calibration, successful analysis is difficult when the size of a face is small. Shao et al. calibrated a camera using an optical flow method and normalized the object height to estimate

a moving object at the cost of increased computational complexity [18]. Zhao and Hu used a pure translation to calibrate a camera [19], and Li et al. reduced control points given the intrinsic camera parameters to calibrate a pan-tilt camera [20]. Andaló et al. estimated vanishing points by clustering lines in an image and then calculated the object height [21]. However, this method cannot accurately estimate vanishing points when the background does not include sufficient pairs of parallel lines. User input of the human height is another burden of this method.

To solve the abovementioned problems, the proposed method calibrates a camera by detecting and tracking the object region. In addition, a projective matrix, which is a result of the camera calibration, is applied to the proposed human height estimation method, and then estimated human heights are accumulated and corrected using the Random Sample Consensus (RANSAC) algorithm. As a result, the proposed method can estimate the normalized human height using an uncalibrated camera for a visual surveillance system. This paper is organized as follows. Section 2 describes the camera projective model, and Section 3 presents the proposed camera calibration and human height estimation algorithms. Experimental results are shown in Section 4, and Section 5 concludes the paper.

2. Camera Projection Model-Based Calibration: A Review

An object is projected onto a two-dimensional image with different sizes depending on the distance between the object and a camera. In order to estimate the human height using a single camera, the projective relationship between the 3D space information and the 2D image plane is needed. The pin-hole camera projective model [22] is given as

$$sx = A[R \mid t]X, \quad (1)$$

where X represents the coordinate in the 3D space, matrix A contains intrinsic camera parameters, R represents the camera rotation matrix, t represents the camera translation vector, x represents the coordinate in the 2D image plane, and s represents the scale factor. The camera intrinsic parameter is determined by focal length (f_x, f_y), principal point (p_x, p_y), skewness skew, and aspect ratio a as

$$A = \begin{bmatrix} f_x & \text{skew} & p_x \\ 0 & f_y & p_y \\ 0 & 0 & a \end{bmatrix}. \quad (2)$$

To simplify the camera calibration process, the proposed method assuming that $f_x = f_y$, the principal point is the center of the image, skew = 0, and $a = 1$. In the same manner, the camera rotation with regard to z -axis is zero and the translations with regard to x -axis and y -axis are also zero.

Using the vanishing points and lines, Liu et al. [23] compute the camera parameters as

$$\begin{aligned} f &= \sqrt{\left(\frac{a_3}{a_2} - P_y\right)(v_y - P_y)}, \\ \rho &= a \tan\left(-\frac{v_x}{v_y}\right), \\ \theta &= a \tan\left(-\frac{\sqrt{v_x^2 + v_y^2}}{f}\right), \\ h_c &= \frac{h_o}{(1 - d(o_h, v_l) \|o_f - v\| / d(o_f, v_l) \|o_h - v\|)}, \end{aligned} \quad (3)$$

where f represents the focal length, ρ the rolling angle in degree, θ the tilt angle in degree, h_c the camera height, v_l the horizontal vanishing line $a_1x + a_2y + a_3 = 0$, $v = [v_x \ v_y]^T$ the vertical vanishing point, h_o the object height in the world coordinate, o_f the object foot position, o_h the object head position, and $d(A, B)$ the distance measure between two points A and B .

In order to estimate the physical size of an object in the 3D space using the object size in the 2D image, the proposed method detects the moving human to estimate the 3D space information. To estimate the human height, the proposed method assumes the foot position on the flat ground plane. As a result, the foot position in the 2D image plane is inversely projected into the 3D space to obtain the human height information.

3. Normalized Human Height Estimation

The proposed human height estimation algorithm is an extended version of Jung et al.'s work [24] and consists of three steps: (i) moving human detection and tracking, (ii) automatic camera calibration, and (iii) reference object-based human height estimation with error correction. Figure 1 shows the block diagram of the proposed human height estimation method, where I_k represents the k th input frame, R the moving human region, O the human tracking region, P the projective matrix, H_k the k th height estimation result of the human, and H the error corrected height estimation result.

3.1. Moving Human Detection and Tracking. The proposed method first detects a moving human to estimate its height. If the detected human region includes the background region or if the region loses some part of the human body, an accurate estimation of the human height is difficult. For this reason, the proposed method generates a background using the Gaussian mixture model (GMM) [25, 26] and then detects and labels the foreground image. The regions that do not have enough pixels in the foreground image are removed to reduce the noise.

The detected foreground regions include not only a single human region but also a group of human region possibly with

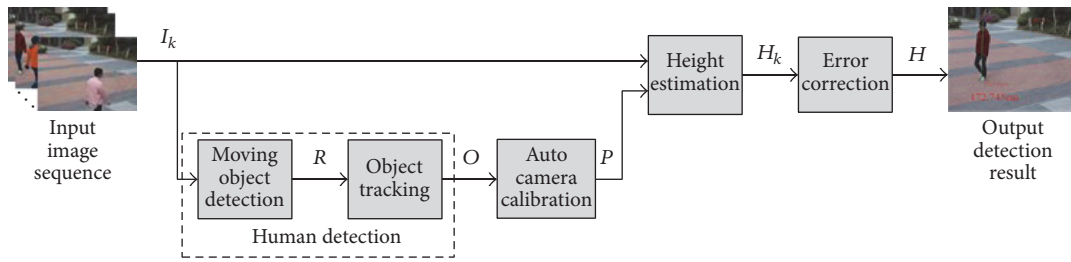


FIGURE 1: Block diagram of the proposed human height estimation method.

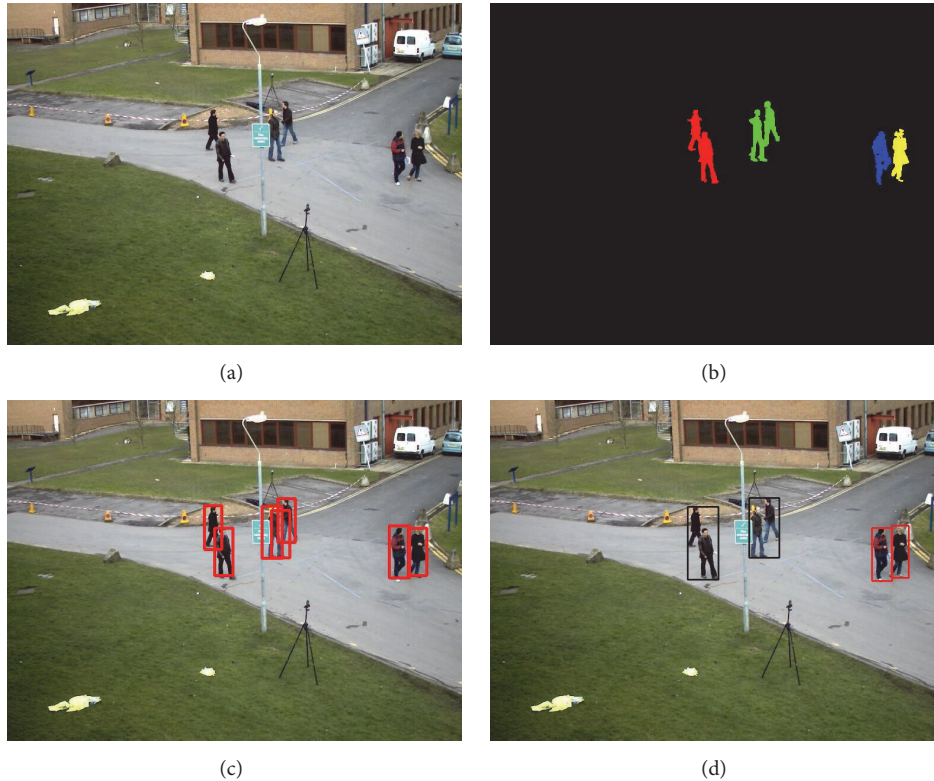


FIGURE 2: Moving human region detection and classification results: (a) an input frame, (b) the detected and labeled foreground regions, (c) the pedestrian detection results, and (d) classified regions.

nonhuman objects, which make human tracking difficult, and as a result, human height estimation error increases. For that reason, the proposed method classifies each region according to whether it is a human region or not. The proposed classification method uses the combined histogram of oriented gradients and local binary pattern (HOG-LBP) and a support vector machine- (SVM-) based human detection method [27]. Using the detected human information, each foreground region is classified into two regions. The first region is a single human region that has only one human object. The second region is a single nonhuman region that has either none or multiple humans. Figure 2 shows the moving human region detection and classification results. Figures 2(b) and 2(c), respectively, show the foreground image and human detection result of a video shown in Figure 2(a). Figure 2(d) shows the region classification result, where single

human and single nonhuman regions are, respectively, represented by red and black boxes.

The proposed method tracks the human and estimates the height in a video using the detected single human region. Although the Kalman filter tracker [28] is a popular stochastic tracking method, it cannot track a nonlinearly moving object. To solve this problem, the proposed method uses a particle filter tracker [29]. In a surveillance input video, human information, such as size and shape, changes while the human is walking. For this reason, the model-based tracking [30] method models the target human using a color histogram to deal with the dynamic characteristics of the moving human. In the proposed method, the HSV color histogram is used to represent the human region to reduce the sensitivity to the illuminance.

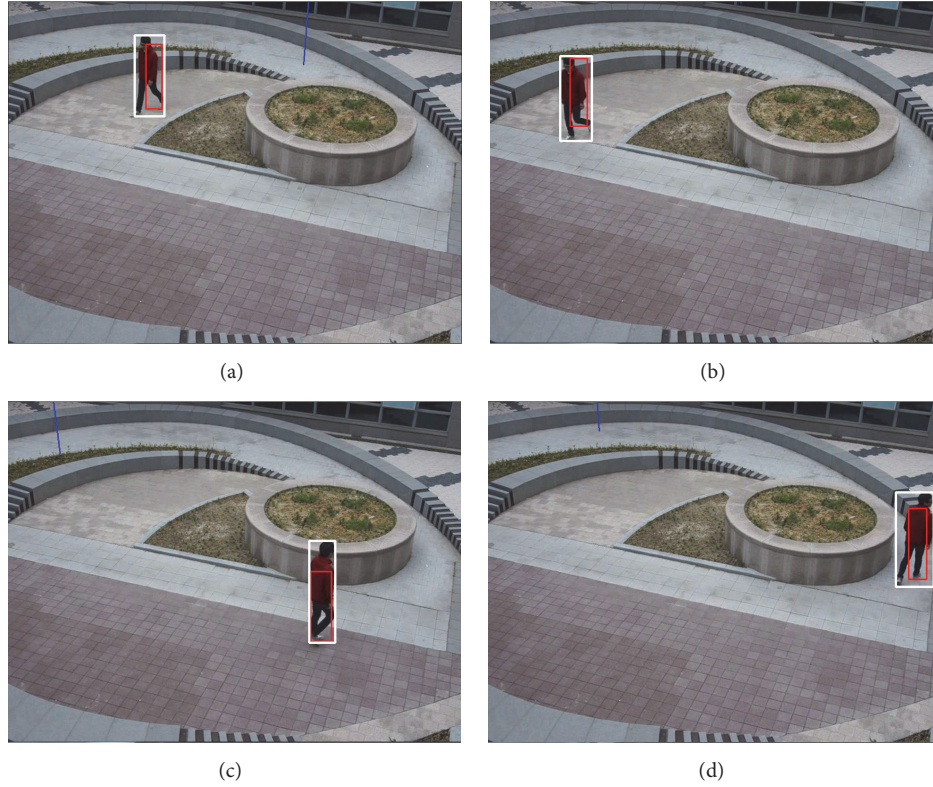


FIGURE 3: Results of human tracking using the proposed method: (a) the 2522th frame, (b) the 2557th frame, (c) the 2743th frame, and (d) the 2819th frame.

The particle filter tracking results may include a probabilistic error and cannot detect the entire human region. Moreover, if the number of particles increases to reduce the tracking error, the time complexity also increases. To solve these problems, the proposed method detects the tracked human region by matching the detected human region with the tracked human regions as

$$O_i = \arg \max_R \left(\frac{1}{n_j} I(R_j, T_i) \right), \quad (4)$$

where O_i represents the i th human region, R_j the j th moving human region that is detected using the background model, n_j the number of pixels in the moving human region R_j , and T_i the tracking region about the i th human. After matching, the proposed method uses additional trackers for unmatched single human regions.

Figure 3 shows the human tracking results using the proposed method. In Figure 3, the red box represents the particle filter tracking result about the moving human, and the white box represents the optimal rectangular region that encloses the detected human region.

3.2. Vanishing Point and Line Estimation Using Human Information. The normalized human height can be estimated in meters in the 3D space by estimating camera parameters. For the automatic camera calibration, the vanishing points and

line should be estimated using the parallel lines in the image. Li et al. estimated the vanishing line and point by extracting the lines from the background structure [31]. However, this method cannot calibrate the camera if the background structure does not have a sufficient number of parallel lines. For automatic calibration without using parallel lines, the proposed calibration method uses the moving human information [23]. More specifically, the proposed method detects both foot and head positions of the human in the 2D image as

$$\begin{aligned} h^{2D} &= \arg \min_o (y_{O_{i,k}}), \\ f^{2D} &= \frac{1}{n_{i,s}} \sum_{k=1}^{n_{i,s}} O_i^s(k), \end{aligned} \quad (5)$$

where h^{2D} represents the human head position in the 2D image, $y_{O_{i,k}}$ the y -axis coordinate of the k th pixel in the i th human region O_i , f^{2D} the foot position in the 2D image, $O_i^s(k)$ the sample foot region that consists of 10% pixels of the i th human region, and $n_{i,s}$ the number of pixels in the sample region.

Both vanishing points and line can be estimated using the detected foot and head positions. The vertical vanishing point can be estimated using the intersection between the foot-to-head lines that include both the foot and head points from the corresponding human region. The horizontal line is estimated using two or more horizontal vanishing points,

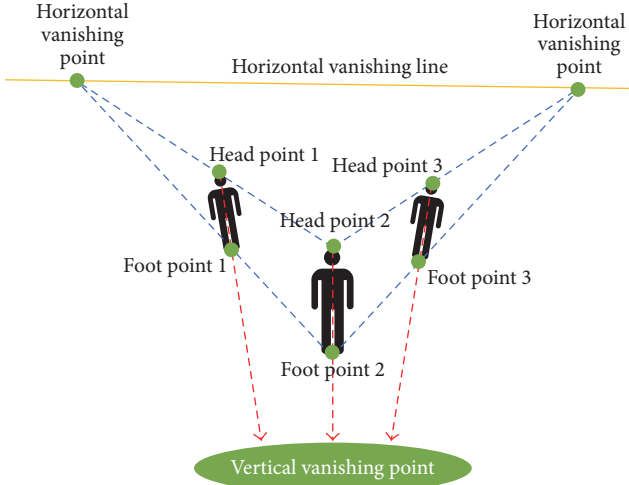


FIGURE 4: Vanishing line estimation process using the human information.

and the horizontal vanishing point is using the intersection between the foot-to-foot and head-to-head lines. The foot-to-foot and head-to-head lines, respectively, include foot and head points. Both vanishing points and line are estimated using the RANSAC algorithm to reduce the estimation error. Figure 4 illustrates the human-based vanishing point and line estimation process.

3.3. Human Height Estimation and Error Correction. The proposed method computes the foot point in the 3D space for the normalized human height estimation using multiple videos acquired by different cameras. To calculate the foot point in the 3D space, the foot point in the 2D image is inversely projected into 3D space. The 3D point is on the line that connects the human foot point in the 3D space with the corresponding image sensor. Since the camera height is estimated based on the ground plane that includes the human foot points, the 3D foot point can be obtained by normalizing the inversely projected point with respect to the z -axis. As a result, the foot point in 3D can be calculated as

$$f^{3D} = \frac{1}{Z} (P^T P)^{-1} P^T f^{2D}, \quad (6)$$

where f^{3D} represents the foot point in the 3D space, f^{2D} the foot point in the 2D image, P the projective matrix, and Z the z -axis coordinate of the point that is inversely projected from the foot point in 2D image.

The reference head point in the 3D space can be estimated by translating the foot point to the vertical direction of the ground plane. Using the reference head point in the 3D space, the corresponding head point in the 2D image is given as

$$h_{\text{ref}}^{2D} = P h_{\text{ref}}^{3D}, \quad (7)$$

where h_{ref}^{2D} represents the reference head point in the 2D image and h_{ref}^{3D} the corresponding head point in the 3D space.

Using the reference head point, the human height can be estimated as

$$H_o = \left| \frac{(y_{\text{ref},h}^{2D} - y_f^{2D})}{(y_h^{2D} - y_f^{2D})} \right| H_{\text{ref}}, \quad (8)$$

where H_o represents the estimated human height, H_{ref} the reference height, $y_{\text{ref},h}^{2D}$ the y -axis coordinate of the reference head point in the 2D image, y_h^{2D} the y -axis coordinate of the human head point in the 2D image, and y_f^{2D} the y -axis coordinate of the human foot point in the 2D image. In this work, the reference height of 1.8 meters was used. Figure 5 shows the human height estimation model using reference height.

The accuracy of human height estimation depends on the detected human region. To reduce the human height estimation error, the proposed method accumulates the estimated human heights in each frame and corrects the errors using the RANSAC algorithm. In the first step of the RANSAC algorithm, n_s sample heights are randomly extracted. Next step computes the sum of squared differences (SSD) between the average height and each estimated human height. The first and second steps repeat n_i times to obtain the error corrected height.

4. Experimental Results

The proposed human height estimation results are shown in this section. The test video was acquired using an uncalibrated camera viewing down the ground plane at the height in between 2.2 and 7.2 meters. Each video sequence has the size 1280×720 and includes moving humans. In addition, the performance evaluation of tracking and surveillance (PETS) 2009 dataset [32] was used to test the proposed algorithm.

Figure 6 shows the result of human height estimation using the proposed method. Although the size of the human in a 2D image looks different by scenes, the resulting normalized height is correctly estimated from all different scenes using a prespecified height of the reference object for the camera calibration.

Figure 7 shows the results of error correction of the estimated height. Figure 7(a) shows the human height estimation error caused by the human pose change. The human height estimation error is corrected using the proposed method as shown in Figure 7(b).

Figure 8(a) shows the human height estimation error caused by an occlusion. The height estimation error of the occluded human is reduced by the proposed method as shown in Figure 8(b).

Figure 9 shows estimation results of multiple human heights. As shown in Figure 9(a), the height of separated human is estimated. In Figures 9(b), 9(c), and 9(d) some humans adjoined and made multiple human regions. As shown in Figures 9(b), 9(c), and 9(d), the proposed method estimates the human height by classifying each region.

Figure 10 shows the estimated human height in video frames, where the ground truth of the object height is 1.75

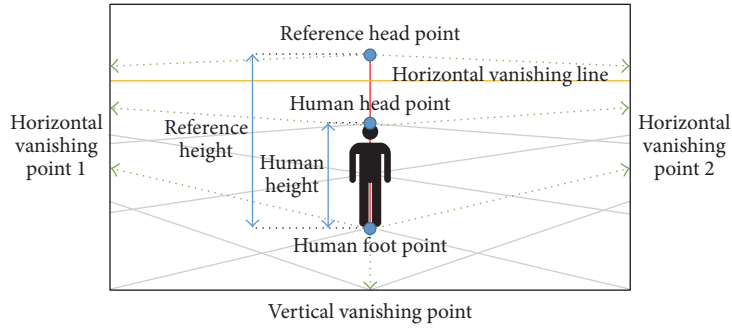


FIGURE 5: Human height estimation model using reference height.



FIGURE 6: Height estimation results of the same person in three different scenes: (a) scene 1, (b) scene 2, and (c) scene 3.

meters shown as the solid curve. The dotted and daggered curves, respectively, show the estimated human height without and with error correction. As shown in Figure 10, the human height estimation error is reduced by 0.027 to 0.012 meters using the proposed error correction method.

5. Conclusions

An automatic calibration method is presented using object detection and tracking by multiple, uncalibrated cameras. As a result, the camera parameters including the camera

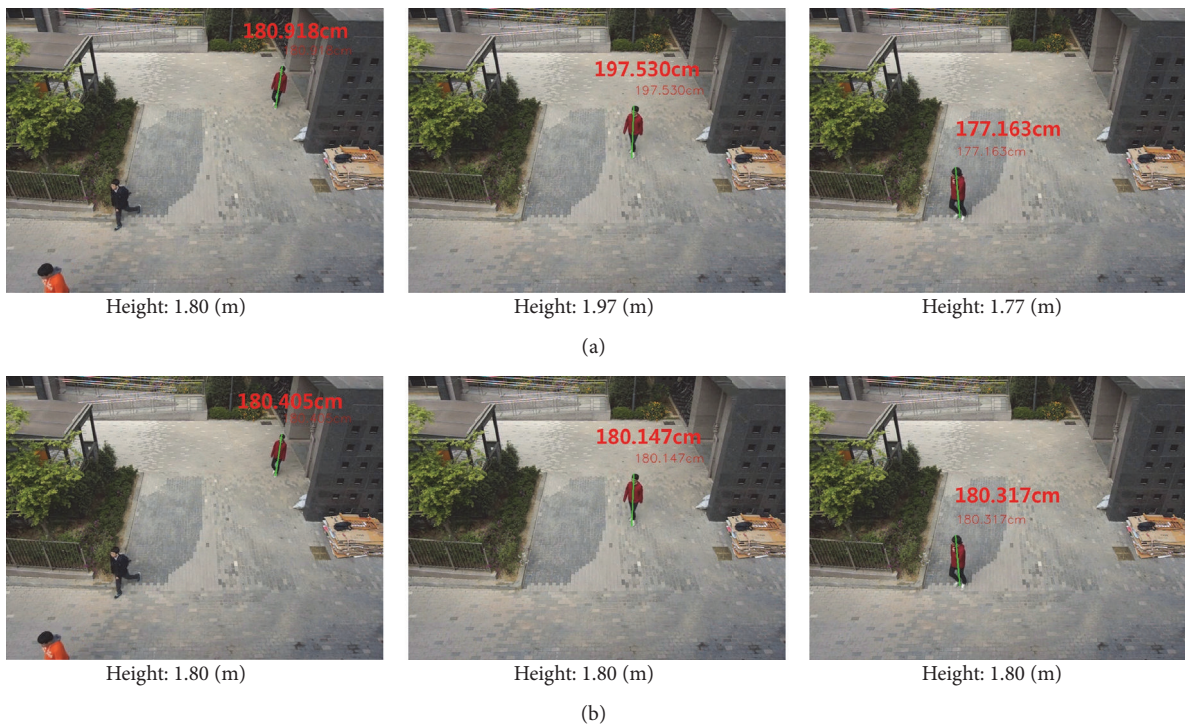


FIGURE 7: Height estimation error correction results of the walking human: (a) human height estimation error caused by the human pose change while the human is walking and (b) error corrected human height estimation results of (a).

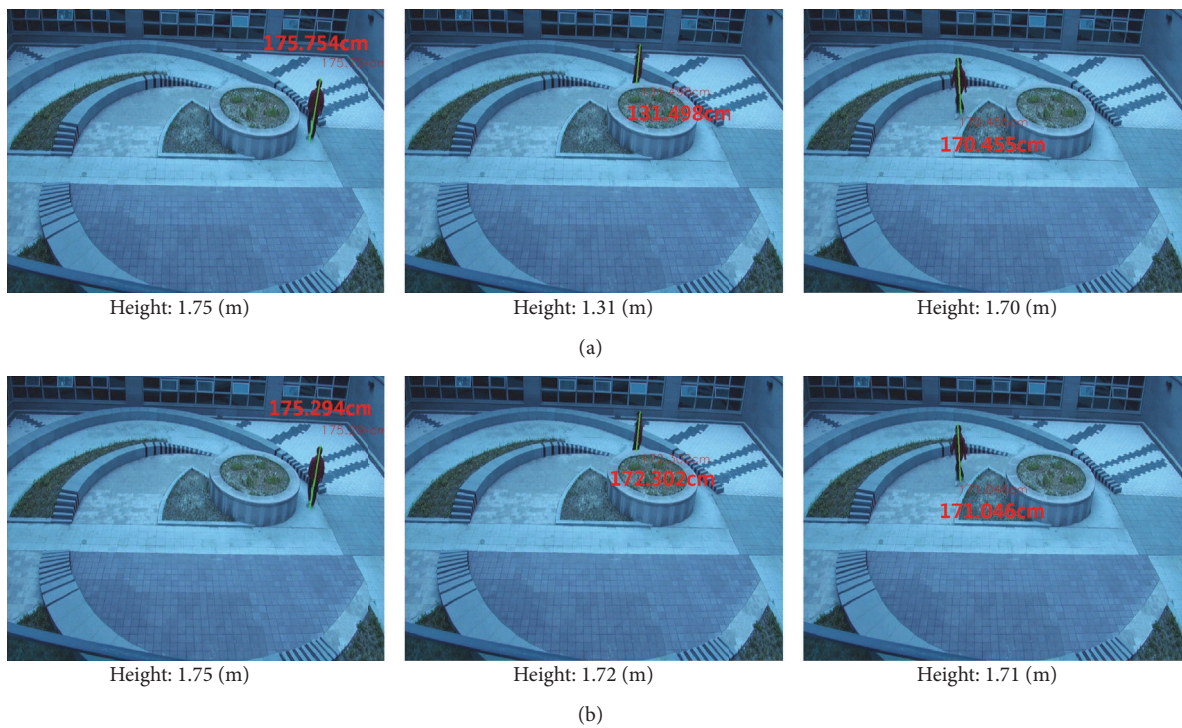


FIGURE 8: Height estimation error correction results of the occluded human: (a) human height estimation error caused by an occlusion and (b) error corrected human height estimation results of (a).

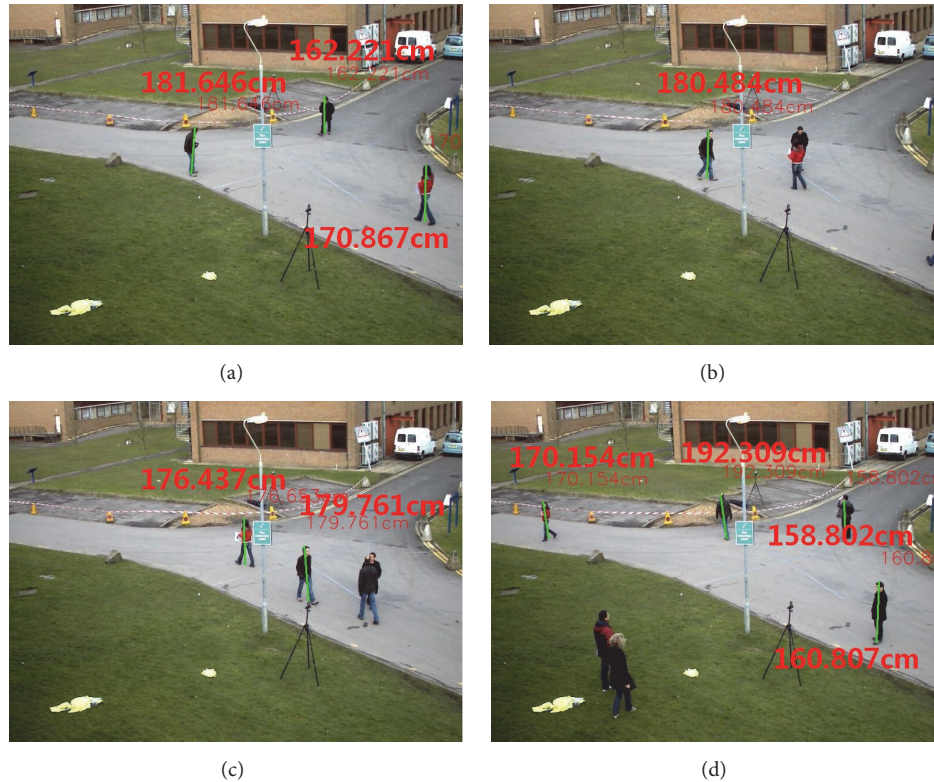


FIGURE 9: Multiple humans height estimation results: (a) 430th frame, (b) 461st frame, (c) 484th frame, and (d) 560th frame.

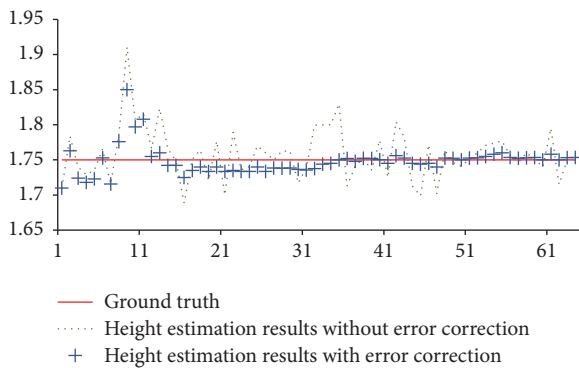


FIGURE 10: Comparison between the estimated height and ground truth.

height are estimated. Moreover, the proposed algorithm can be applied to estimate the normalized human height. As a result, the normalized human height is estimated using multiple uncalibrated cameras. The proposed method can be applied to object tracking and recognition in a very-large area video surveillance system.

Competing Interests

The authors declare that they have no competing interests.

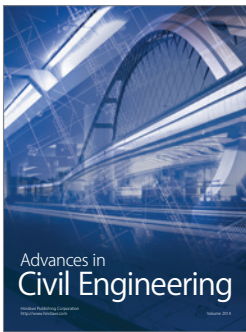
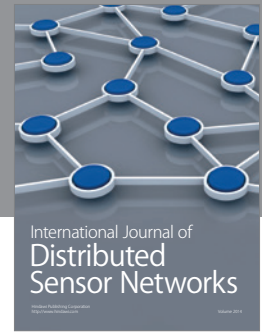
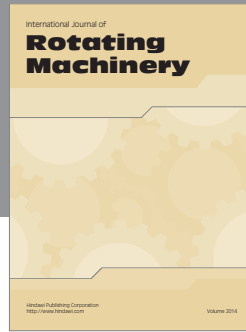
Acknowledgments

This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (B0101-15-0525, Development of Global Multitarget Tracking and Event Prediction Techniques Based on Real-Time Large-Scale Video Analysis), by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2016-H8501-16-1018) supervised by the IITP (Institute for Information & Communications Technology Promotion), and by the Technology Innovation Program (Development of Smart Video/Audio Surveillance SoC & Core Component for Onsite Decision Security System) under Grant 10047788.

References

- [1] W. Lao, J. Han, and P. H. N. de With, "Automatic video-based human motion analyzer for consumer surveillance system," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 591–598, 2009.
- [2] C. R. Del-Blanco, F. Jaureguizar, and N. Garcia, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 857–862, 2012.
- [3] S. Lee, J. Lee, M. H. Hayes, and J. Paik, "Adaptive background generation for automatic detection of initial object region in

- multiple color-filter aperture camera-based surveillance system,” *IEEE Transactions on Consumer Electronics*, vol. 58, no. 1, pp. 104–110, 2012.
- [4] W. Chantara, J. Mun, D. Shin, and Y. Ho, “Object tracking using adaptive template matching,” *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 1, pp. 1–9, 2015.
- [5] H.-C. Chu and H. Yang, “A simple image-based object velocity estimation approach,” in *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control (ICNSC '14)*, pp. 102–107, Miami, Fla, USA, April 2014.
- [6] V. Maik, J. Park, D. Kim, and J. Paik, “Model-based human pose estimation and its analysis using hausdorff matching,” *TECHART: Journal of Arts and Imaging Science*, vol. 2, no. 2, p. 51, 2015.
- [7] S. Kang, A. Roh, C. Eem, and H. Hong, “Using real-time matching for human gesture detection and tracking,” *Techart: Journal of Arts and Imaging Science*, vol. 1, no. 1, pp. 60–66, 2014.
- [8] S. Lee, S. Jeong, H. Yu et al., “Efficient image transformation and camera registration for the multi-projector image calibration,” *TECHART: Journal of Arts and Imaging Science*, vol. 3, no. 1, p. 38, 2016.
- [9] H. Kim, D. Kim, and J. Paik, “Automatic estimation of spatially varying focal length for correcting distortion in fisheye lens images,” *IEEK Transactions on Smart Processing & Computing*, vol. 2, no. 6, pp. 339–344, 2013.
- [10] A. Poulin-Girard, S. Thibault, and D. Laurendeau, “Influence of camera calibration conditions on the accuracy of 3D reconstruction,” *Optics Express*, vol. 24, no. 3, p. 2678, 2016.
- [11] C.-H. Teng, Y.-S. Chen, and W.-H. Hsu, “Camera self-calibration method suitable for variant camera constraints,” *Applied Optics*, vol. 45, no. 4, pp. 688–696, 2006.
- [12] A. Arfaoui and S. Thibault, “Fisheye lens calibration using virtual grid,” *Applied Optics*, vol. 52, no. 12, pp. 2577–2583, 2013.
- [13] J. C. Neves, J. C. Moreno, and H. Proença, “A master-slave calibration algorithm with fish-eye correction,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 427270, 8 pages, 2015.
- [14] Y. Zhang, L. Zhou, H. Liu, and Y. Shang, “A flexible online camera calibration using line segments,” *Journal of Sensors*, vol. 2016, Article ID 2802343, 16 pages, 2016.
- [15] T. Bell, J. Xu, and S. Zhang, “Method for out-of-focus camera calibration,” *Applied Optics*, vol. 55, no. 9, pp. 2346–2352, 2016.
- [16] L. Kual-Zheng, “A simple calibration approach to single view height estimation,” in *Proceedings of the 9th Computer and Robot Vision (CRV '12)*, pp. 28–30, May 2012.
- [17] A. C. Gallagher, A. C. Blose, and T. Chen, “Jointly estimating demographics and height with a calibrated camera,” in *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV '09)*, pp. 1187–1194, Kyoto, Japan, September 2009.
- [18] J. Shao, S. K. Zhou, and R. Chellappa, “Robust height estimation of moving objects from uncalibrated videos,” *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 2221–2232, 2010.
- [19] B. Zhao and Z. Hu, “Camera self-calibration from translation by referring to a known camera,” *Applied Optics*, vol. 54, no. 25, pp. 7789–7798, 2015.
- [20] Y. Li, J. Zhang, and J. Tian, “Method for pan-tilt camera calibration using single control point,” *Journal of the Optical Society of America A*, vol. 22, no. 1, pp. 156–163, 2014.
- [21] F. A. Andaló, G. Taubin, and S. Goldenstein, “Efficient height measurements in single images based on the detection of vanishing points,” *Computer Vision and Image Understanding*, vol. 138, pp. 51–60, 2015.
- [22] Y. Li, J. Zhang, and J. Tian, “Camera calibration from vanishing points in image of architectural scenes,” in *Proceedings of the British Machine Vision Conference*, pp. 382–391, Nottingham, UK, September 1999.
- [23] J. Liu, R. Collins, and Y. Liu, “Surveillance camera autocalibration based on pedestrian height distributions,” in *Proceedings of the British Machine Vision Conference*, pp. 1–11, January 2011.
- [24] J. Jung, H. Kim, I. Yoon, and J. Paik, “Human height analysis using multiple uncalibrated cameras,” in *Proceedings of the 2016 IEEE International Conference on Consumer Electronics (ICCE '16)*, pp. 213–214, Las Vegas, Nev, USA, January 2016.
- [25] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, vol. 2, pp. 28–31, Cambridge, UK, August 2004.
- [26] Y. Kim, S. Jeong, J. Oh, and S. Lee, “Fast MOG (Mixture of Gaussian) algorithm based on predicting model parameters,” *TECHART: Journal of Arts and Imaging Science*, vol. 2, no. 1, pp. 41–45, 2015.
- [27] W.-J. Park, D.-H. Kim, Suryanto, C.-G. Lyuh, T. M. Roh, and S.-J. Ko, “Fast human detection using selective block-based HOG-LBP,” in *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP '12)*, pp. 601–604, Orlando, Fla, USA, September 2012.
- [28] S. S. Pathan, A. Al-Hamadi, and B. Michaelis, “Intelligent feature-guided multi-object tracking using Kalman filter,” in *Proceedings of the 2nd International Conference on Computer, Control and Communication*, pp. 1–6, February 2009.
- [29] T. Zhang, S. Fei, X. Li, and H. Li, “An improved particle filter for tracking color object,” in *Proceedings of the International Conference on Intelligent Computation Technology and Automation (ICICTA '08)*, vol. 2, pp. 109–113, Hunan, China, October 2008.
- [30] E. J. Rhee, J. Park, B. Seo, and J. Park, “Subjective evaluation on perceptual tracking errors from modeling errors in model-based tracking,” *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 6, pp. 407–412, 2015.
- [31] B. Li, K. Peng, X. Ying, and H. Zha, “Vanishing point detection using cascaded 1D Hough transform from single images,” *Pattern Recognition Letters*, vol. 33, no. 1, pp. 1–8, 2012.
- [32] Performance evaluation of tracking and surveillance 2009, <http://www.cvg.reading.ac.uk/PETS2009/a.html>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

