

# Camera Handoff for Multicamera Multiobject Tracking

Po-Tong Wang,<sup>1</sup> Jia-Shing Sheu,<sup>2\*</sup> and Jing-Han Lai<sup>2</sup>

<sup>1</sup>Department of Electrical Engineering, Lunghwa University of Science and Technology,  
No. 300, Sec. 1, Wanshou Rd., Guishan District, Taoyuan, 333326, Taiwan

<sup>2</sup>Department of Computer Science, National Taipei University of Education,  
No. 134, Sec. 2, He-Ping East Road, Da-an District, Taipei, 106, Taiwan

(Received September 3, 2021; accepted January 6, 2022)

**Keywords:** camera handoff, background subtraction, human tracking

In recent years, the unmanned economy has begun to develop rapidly worldwide. Unmanned shops use technology to replace human personnel and increase efficiency. This technology has increased the demand for accurate and efficient camera surveillance systems. Camera handoff is a crucial step in the tasks of continuous target tracking and the maintenance of consistent cross-camera target marking in a multicamera surveillance system. In this study, we proposed a method of indoor multicamera handoff. To ensure continuous target tracking, the minimum number of required frames was maintained and at least one camera tracked each target. We employed the background subtraction approach to detect the target. Next, we used three trackability measures to evaluate the tracking object and trigger camera handoff accordingly, selecting the optimal camera for target tracking. The three measures considered were the resolution, occlusion, and distance to the edge of the camera's field of view. When one of these reached a preset threshold with an increasing trend, the system triggered the camera handoff.

## 1. Introduction

In recent years, the topic of the unmanned economy has attracted considerable attention. Many businesses have begun to launch unmanned stores as pilot outlets. Compared with traditional stores, unmanned stores provide consumers with a faster, more convenient, and more innovative shopping experience. However, although people enjoy the advantages of the unmanned economy, the need for comprehensive coverage has increased, as has the complexity of associated monitoring systems. Given that a single camera, with its limited field of view (FOV), cannot meet the surveillance requirements of most unmanned stores, multicamera networks have emerged. Using multiple cameras increases not only the coverage area but also flexibility. In the past, multiple images were presented to human supervisors for analysis. However, focusing on multiple images simultaneously is challenging for humans. Therefore, a camera mechanism that can automatically track objects of interest from an optimal angle is essential.

The term camera handoff refers to the decision-making process associated with transferring the tracking of a target of interest from one camera to another. Consistent labeling of the target

---

\*Corresponding author: e-mail: [jjashing@tea.ntue.edu.tw](mailto:jjashing@tea.ntue.edu.tw)  
<https://doi.org/10.18494/SAM3616>

allows that target to be identifiable from among the inputs of multiple cameras. The camera-swapping mechanism defines when and to which camera the target tracking should be transferred, thereby optimizing efficiency by using minimal resources without losing the target. In this study, Raspberry Pi-based application modules were used to set up cameras at various angles. The background subtraction method was employed to locate moving objects in the shooting frame, and the target outline was obtained through multilayer image processing. With the target contour results, we calculated the camera's imaging quality score  $Q$  using three trackability measures, namely, the resolution, occlusion, and distance between the target and the FOV boundary. We then compared the  $Q$  values of different cameras in relation to a single target. The  $Q$  value was used to trigger the handoff; the camera with the highest  $Q$  value was used for target tracking. This system can be used not only in unmanned stores but also in sites housing multiple monitoring systems, such as airports, subways, train stations, banks, and schools. It can protect the safety of citizens and facilitate the transition toward smart cities.

In this paper, Sect. 2 provides a review of the literature and methodology. The structure of the proposed system is discussed in Sect. 3. The experimental design and setup are presented in Sect. 4. Finally, the paper is concluded in Sect. 5.

## 2. Literature Review and Methodology

The efficacy of surveillance systems is easily compromised by factors such as lighting, equipment, visual effects, and image clarity. Image enhancement is necessary to enhance image clarity and thereby improve human activity recognition.<sup>(1)</sup> Grayscale conversion, which is applied using a highly effective image enhancement algorithm, can improve image brightness and clarity. Through linear conversion, the gray values of each pixel can be modified, thereby allowing the contrast to be widened, the edges to be sharpened, and the features to be rendered more vividly.<sup>(2)</sup> Retouching is a common image editing operation that can adjust the image content globally or locally. Blurring, the most commonly used retouching function, aims to eliminate noise or introduce special effects.<sup>(3)</sup> Among all image segmentation methods, thresholding is the simplest and most efficient approach for distinguishing a foreground from a background. Specifically, it can be applied to identify the optimal threshold value. Thresholding is divided into two types: bilevel and multilevel. Bilevel thresholding creates two collections of objects, whereas multilevel thresholding clusters the pixels into many similar groups.<sup>(4)</sup> Dilation and erosion are the two most basic morphological operations. Erosion followed by dilation is called an opening operation, whereas dilation followed by erosion is called a closing operation. Multiple morphological image processing methods can be combined on the basis of dilation, erosion, opening operations, and closing operations.<sup>(5)</sup> Edge detection is fundamental to basic image features and segmentation; moreover, it constitutes a key information source for texture features and is central to shape quality analysis. The classical edge detection algorithms are local-window gradient operators, such as the Roberts, Sobel, Prewitt, Kirsch, and Laplacian operators.<sup>(6)</sup> However, owing to their noise sensitivity, these operators generate unsatisfactory results with actual image input. The Canny edge detector has three features: 1) a low signal-to-noise ratio (SNR), which reduces the likelihood of edge detection errors; 2) favorable positioning

performance, in that the detected edge point should ideally be equivalent to the center of the real edge; and 3) a single-edge response, in that false edge detection should be minimized. On account of its satisfactory performance, the Canny edge detector quickly became a standard against which other edge detection methods are measured.<sup>(7)</sup> Foreground extraction is a central task in image processing. Background subtraction is a mature method for moving target extraction. Assuming that a camera is fixed, an empty background without any targets is first stored, and then the current frame is subtracted from the reference background image to subtract the same part in the background and obtain details concerning the moving object.<sup>(8,9)</sup> However, background subtraction is susceptible to various constraints that inhibit its tracking performance, such as local motion in the background and the presence of light and shadow. To alleviate these problems, Kumar and Yadav proposed a background subtraction algorithm that correctly initializes and updates the background module to extract the relevant foreground blob, thereby improving the tracking accuracy.<sup>(10)</sup> Dharamadhat *et al.* employed multiresolution critical point filters for color object tracking on the basis of background subtraction and target matching.<sup>(11)</sup>

Tracking moving objects has always been a key concern in surveillance systems. Zoom cameras are now widely used in large-scale environmental surveillance systems. Camera handoff was developed to manage camera switching in a manner conducive to seamless target tracking.<sup>(12,13)</sup> However, almost all handoff techniques rely on robust camera trackers. State-of-the-art techniques for evaluating camera handoff performance use either annotated videos or simulated data and operate in conjunction with a tracker.<sup>(14)</sup> Most camera handoff algorithms assign priority to the camera that the target is approaching, but this simple rule is usually inadequate and may prevent handoff from occurring where necessary. To select the optimal camera and minimize unnecessary handoff requests, multiple criteria should be considered.<sup>(15)</sup> To maximize the scope of surveillance activities, Khan and Shah introduced a system that tracks people in images obtained from multiple angles, effectively managing camera switching when tracking the same target by determining the FOV limits of other cameras.<sup>(16)</sup> Kim and Kim calculated the proximity probability according to the number of foreground modules and the angular distance between the camera and the object. Their system then selected the camera with the highest proximity probability, with experiments indicating that the selected camera used two ratios to make judgments and was more accurate and stable than those employing a single foreground module ratio.<sup>(17)</sup> Li and Bhanu defined a game theory-based method that used global utility, overall satisfaction with tracking performance, and camera utility to evaluate various competing cameras and thereby achieve seamless handoff.<sup>(14)</sup> Kim *et al.* considered three techniques for camera handoff: 1) purely geometric computation, 2) 2D image matching, and 3) 3D stereo range matching for target tracking during Mars expeditions.<sup>(13)</sup> Limprasert *et al.* used a combination of position and speed information of the tracking target as the basis of a camera switching mechanism, proposing a method for real-time, multitarget tracking in a multicamera airport domain. They employed multiple object tracking accuracy and processing speed indicators to evaluate their algorithms on standard data sets. A comparison of the results with those obtained through established methods demonstrated that significant improvements were attained.<sup>(18)</sup> Tsagkatakis and Savakis proposed the use of camera handoff to reduce memory requirements and optimize camera lenses with limited capabilities.<sup>(19)</sup> Camera handoff can be

performed under various conditions. The present study proposes a system that mainly originates from the system proposed by Yao *et al.*, who used resolution, occlusion, and the distance between the target's centroid and the FOV boundary as conditions and evaluated the camera's effectiveness in tracking targets.<sup>(15)</sup> Purnama *et al.* optimized target tracking in a multicamera airport domain by combining the Kalman filter and FOV line method, which were respectively used as the tracker engine and to identify the cameras to which the target was visible.<sup>(20)</sup>

Despite having been extensively discussed, the subject of object tracking remains disputed because of the complexity associated with the external environment, for example, with regard to deformation and lighting. In response to the occlusion problem, Zhao *et al.* presented a discriminant model and a generative model.<sup>(21)</sup> Wu used subregion classifiers and compressive tracking to improve robustness in the tracking of occluded targets.<sup>(22)</sup> Aslan *et al.* compared the human-like detection effect of CNNs with that of HOG-SVMs in the case of complete occlusion.<sup>(23)</sup> Heimbach *et al.* used the Kalman filter, which has a fast convergence rate and does not require pretraining, to increase the accuracy of feature detection.<sup>(24)</sup> Whoang *et al.* used maximum color distance and shape density to track targets that appear deformed, warped, or overzoomed.<sup>(25)</sup> Li and Lu proposed a robust global learning-based target tracking algorithm to mitigate problems related to occlusion, motion blur, and rotation.<sup>(26)</sup> Hsia *et al.* proposed the use of machine learning to make intelligent recommendations for facial skincare products.<sup>(27)</sup> Pai *et al.* employed a nonlinear support vector machine (SVM) to recognize gesture changes and establish interactivity between a vehicle and its user.<sup>(28)</sup> Wu used the YOLO model to detect multiple people and track their movement.<sup>(29)</sup>

### 3. System Architecture

In this section, the multicamera handoff mechanism is explained in detail. Figure 1 illustrates the overall system structure, which uses IDEF0.<sup>(30,31)</sup> The system is divided into two submodules: people detection and the camera handoff algorithm. The camera handoff algorithm is the main focus of this paper.

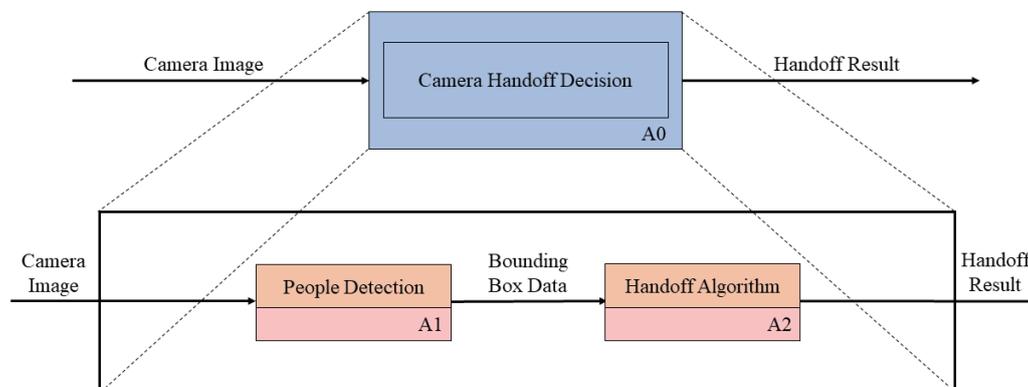


Fig. 1. (Color online) Architecture of the proposed system.

The first submodule involves people detection, as shown in Fig. 2. We first input the source image through grayscale conversion, Gaussian blur, background subtraction, thresholding, Canny edge detection, dilation, and area filtering. Finally, we output the bounding box result. A grayscale image comprises only colors that are shades of gray having equal intensity in the RGB space. We convert an RGB image to a grayscale image using

$$Y = R \times 0.299 + G \times 0.587 + B \times 0.144. \quad (1)$$

Gaussian blur is a data smoothing technique that can reduce random variation and the level of detail in images. Using Eq. (2) for weighting, we take the average of the surrounding points as the center value.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2)$$

Foreground extraction is a primary task in the field of image processing. The background subtraction method is a mature moving target extraction technology. With the camera's fixed preset, the targetless empty background is saved, and then the current frame is subtracted from the reference background image. A moving object mask is obtained by replacing the same part in the background. Thresholding is the simplest image segmentation technique. From a grayscale image, thresholding can be used to create binary images by Otsu's threshold method. The built-in Canny edge detection from the OpenCV library is used to mark the edges of the binary image. Through repeated dilation and erosion, broken edges are connected to obtain as complete a target edge as possible. Applying area filtering once helps to prevent target misjudgment caused by environmental noise. The tracking effect may be poor when the target is overly small or overly far away from the camera.

After people detection, the bounding box result is obtained and can be used as the input for the camera handoff stage. In the camera handoff algorithm, as shown in Fig. 3, we calculate the

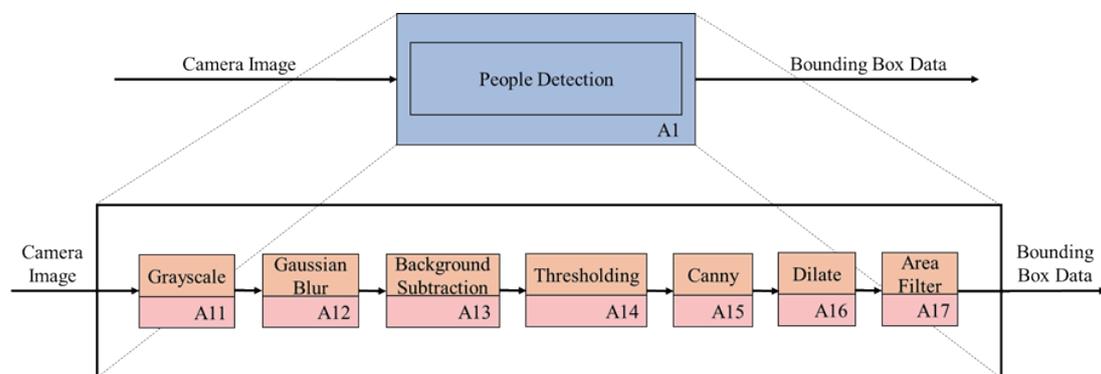


Fig. 2. (Color online) Architecture of the people detection submodule.

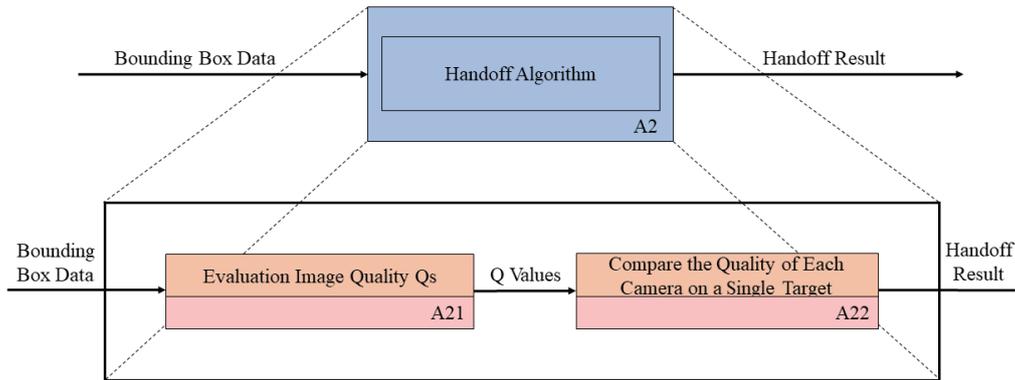


Fig. 3. (Color online) Architecture of camera handoff algorithm.

image quality and express it as a  $Q$  value. By comparing the  $Q$  value of each camera in relation to a single target, we can determine which camera exhibits the optimal tracking effect for a given target.

To maintain continuous object tracking, a handoff request is triggered before the object of interest becomes untraceable or unidentifiable for the current observation camera. The object of interest may become untraceable or unidentifiable because 1) the object is being occluded by other objects, 2) the object is leaving the camera's FOV, or 3) the object's resolution is decreasing. Three trackability measures are defined to determine when to trigger a handoff request: 1) occlusion ( $M_o$ ), 2) distance to the edge of the camera's FOV ( $M_d$ ), and 3) resolution ( $M_s$ ). Occlusion is frequent in indoor environments. In situations involving occlusion, the current frame's bounding box area can simply be compared with the average bounding box area. If the ratio of the two decreases suddenly, this indicates the occurrence of occlusion.  $M_o$  denotes the occlusion and is defined by

$$M_o = \frac{A}{\text{Mean}(A_p)}, \quad (3)$$

where  $A$  represents the bounding box area of the current frame and  $\text{Mean}(A_p)$  is the average of the bounding box area. The distance to the edge of the camera's FOV is crucial for two reasons. First, the target approaching the edge of the observation camera's FOV may indicate that the target is about to pass beyond that FOV. Second, to reserve a sufficient amount of computation time for the execution of camera handoff, the object should remain at a certain distance from the boundaries of the camera's FOV.<sup>(4)</sup>  $M_d$  denotes this distance, which is given by

$$M_d = \frac{D}{D_{max}}, \quad (4)$$

where  $D$  denotes the minimum distance between the target and the FOV vertex and  $D_{max}$  is the distance from the center of the FOV to the edge, as shown in Fig. 4.

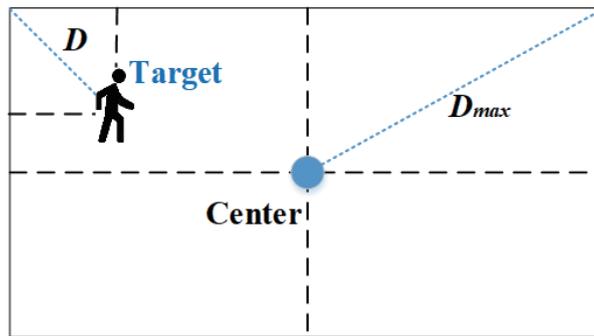


Fig. 4. (Color online) Schematic of  $M_d$ .

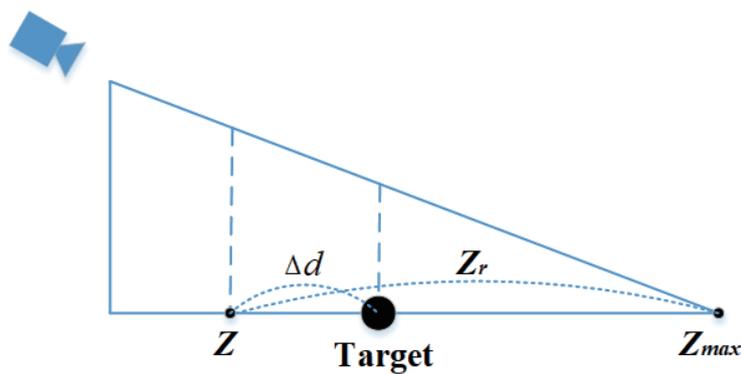


Fig. 5. (Color online) Schematic of  $M_s$ .

The distance between the target and the camera affects the size of the bounding box area, which may reduce the resolution and cause tracking errors.  $M_s$  denotes the resolution, which is defined as

$$M_s = 1 - \frac{\Delta d}{Z_r}, \tag{5}$$

where  $Z$  denotes the distance between the camera and target for the optimal tracking effect,  $\Delta d$  is the distance between the target and  $Z$ , and  $Z_r$  is the distance between  $Z$  and  $Z_{max}$ , as shown in Fig. 5. Here,  $Z_{max}$  is the maximum distance between the target and the camera at which the target can be completely observed,

$Q$  is the quality of images captured by a camera in relation to a target and is obtained as

$$Q = w_s M_s + w_d M_d + w_o M_o. \tag{6}$$

where each  $w$  is an experimental weight obtained by trial and error that depends on the experimental situation, with  $w_s + w_d + w_o = 1$ . The larger the  $Q$  value, the more accurate the camera tracking in relation to the target. The  $Q$  value is used to trigger the camera handoff. By comparing the  $Q$  value in relation to a single target, the system can determine which camera

possesses the optimal tracking ability and whether to switch the observation camera. By inputting the bounding box results,  $M_s$ ,  $M_d$ , and  $M_o$  are calculated. Finally, Eq. (6) is employed to calculate and output the image quality  $Q$ .

#### 4. Experimental Results

The experimental environment of the program comprised a Raspberry Pi board. As shown in Fig. 6, we used a Raspberry Pi 4 model B and Raspberry Pi camera module V2 as experimental cameras. Raspberry Pi is a single-board computer developed by the nonprofit Raspberry Pi Foundation to promote students' understanding of computer science. We used Docker software in the Debian operating system environment, as recommended by the Raspberry Pi Foundation, to produce an image file, as well as languages such as C++ and Python3 for research activities. The Raspberry Pi camera module supports the Model A and B versions of Raspberry Pi. We employed the OV5647 lens produced by OmniVision, which could be directly connected to the Plum expansion board by means of a flexible cable. The five-megapixel lens can record movies in 1080p at 30 frames per second. Here, we used a mobile power bank (connected using a USB-C cable) as the power source, providing a maximum output of 5 V/2.1 A.

We placed two identical cameras (hereafter denoted as Camera 1 and Camera 2) at different positions. As shown in Fig. 7, a schematic of the entire field, the camera's FOV is such that the target entered Camera 1's FOV from the leftmost side of the experimental field and passed through the central area to the right side of the field, where it finally left Camera 1's FOV. During the experiment, the target repeatedly moved back and forth along this path.

Figure 8 depicts the people detection steps mentioned in the previous section. Here, we display the output of this process, the end goal of which was to separate the moving target from the background and mark it on the screen with a bounding box. We used the targetless background, current frame, and previous frame to perform three-way background subtraction and thereby achieve the optimal tracking results.

A substantial amount of debris obscured the experimental field. Thus, adjusting  $w$  achieved the optimal tracking effect. After several comparisons, the optimal  $Q$  value formula for this experimental environment was determined as



Fig. 6. (Color online) Experimental system combining the Raspberry Pi and camera.

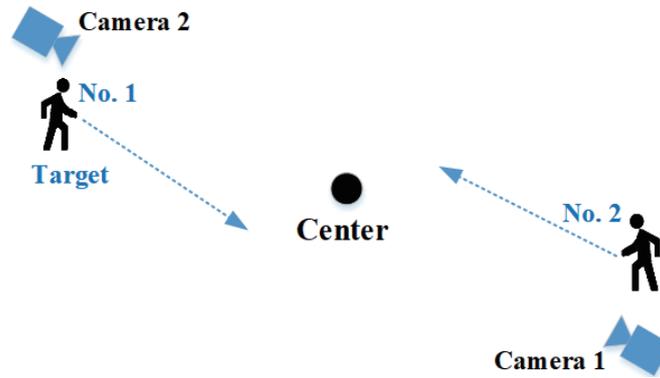


Fig. 7. (Color online) Target path for tracking two people.

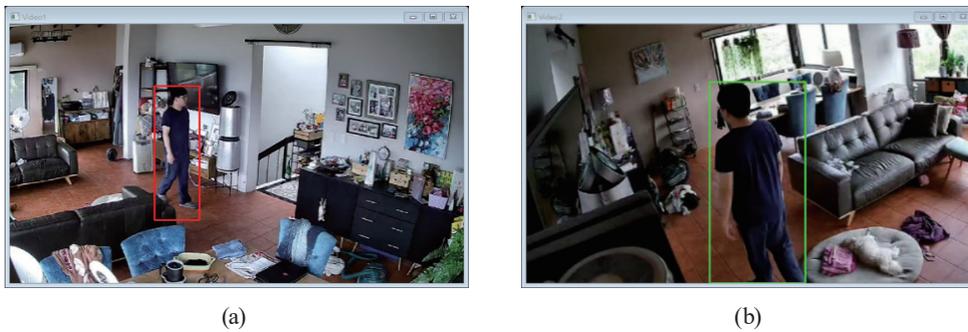


Fig. 8. (Color online) People detection by the two cameras. (a) Camera 1. (b) Camera 2.

$$Q = 0.3 M_s + 0.4 M_d + 0.3 M_o. \quad (7)$$

The color of the bounding box indicates whether the camera was tracking the target. As shown in Fig. 9, Camera 1 had a larger  $Q$  value than Camera 2 and was therefore the current tracking camera. Table 1 presents a summary of a real-world situation wherein the target walked between the two cameras using the path depicted in Fig. 7. At frame  $\mathbf{f}_n$ , the target walked into Camera 1's FOV from the edge of the experimental field and then entered Camera 2's FOV. During the period from frame  $\mathbf{f}_n$  to frame  $\mathbf{f}_{n+20}$ , the target was overly close to Camera 2, complicating target recognition. Thus, a switch was made to Camera 1. At frame  $\mathbf{f}_{n+50}$ , because the target was occluded in Camera 1's FOV, handoff was executed and Camera 2 now tracked the target. At frame  $\mathbf{f}_{n+110}$ , the target walked to the corner of Camera 1's FOV, complicating target recognition. Therefore, a switch was made to Camera 2. During the interval from frame  $\mathbf{f}_{n+110}$  to frame  $\mathbf{f}_{n+202}$ , the target approached Camera 2. Because the target was overly close to Camera 2, handoff was again executed and the target was now tracked by Camera 1. Here, we found that two handoffs occurred: one at frame  $\mathbf{f}_{n+110}$  from Camera 1 to Camera 2, and the other at frame  $\mathbf{f}_{n+178}$  from Camera 2 to Camera 1.

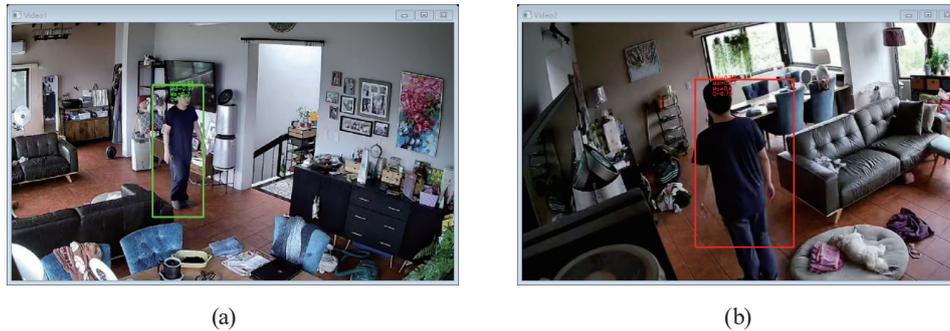


Fig. 9. (Color online)  $Q$  values of different cameras at the same time point. (a) Camera 1 for  $Q = 0.77$ . (b) Camera 2 for  $Q = 0.75$ .

Table 1  
Summary of camera handoff procedures performed in the experiment.

Frame		$f_n$	$f_{n+20}$	$f_{n+50}$	$f_{n+110}$	$f_{n+178}$	$f_{n+202}$
Camera 1	$Q$	0.81	0.77	0.74	0.35	0.58	0.51
	$M_o$	0.87	0.85	0.53	0.22	0.37	0.34
	$M_d$	0.42	0.67	0.8	0.25	0.66	0.49
	$M_s$	0.59	0.7	0.67	0.61	0.68	0.71
Camera 2	$Q$	0.59	0.75	0.73	0.46	0.49	0.27
	$M_o$	1	0.75	1	0.18	0.14	0.16
	$M_d$	0.18	0.87	0.74	0.53	0.6	0.22
	$M_s$	0.73	0.6	0.44	0.63	0.63	0.15

## 5. Conclusions

We developed a system to monitor and assign target tracking tasks to a camera with high tracking ability in a surveillance space. Three trackability measures were employed to calculate the image quality  $Q$ . The value of  $Q$  was evaluated to determine whether camera handoff should be performed. The proposed system possesses two advantages over traditional image processing methods. First, it is simpler and more intuitive; second, AI and deep learning reduce the amount of calculation required, which is conducive to lightweight system development. Furthermore, the experimental results confirmed that the three trackability parameters can indicate the target tracking status. By adjusting the experimental weights, an acceptable  $Q$  value can be calculated and used for camera handoff management. The proposed system is convenient, fully automated, and can achieve superior tracking results in unmanned shops or other indoor spaces requiring multicamera monitoring.

## References

- 1 W. M. Ke, C. R. Chen, and C. T. Chiu: IEEE Trans. Circuits Syst. Video Technol. **21** (2011) 360. <https://doi.org/10.1109/TCSVT.2010.2087475>
- 2 H. Zhu, N. Lin, H. Leung, R. Leung, and S. Theodidis: IEEE Sens. Lett. **4** (2020) 1. <https://doi.org/10.1109/LSSENS.2020.2995060>

- 3 F. Ma, X. Jing, X. Zhu, Z. Tang, and Z. Peng: IEEE Trans. Inf. Foren. Secur. **15** (2020) 115. <https://doi.org/10.1109/TIFS.2019.2917160>
- 4 A. A. Ewees, M. A. Elaziz, M. A. A. Al-Qaness, H. A. Khalil, and S. Kim: IEEE Access **8** (2020) 26304. <https://doi.org/10.1109/ACCESS.2020.2971249>
- 5 M. Abdalla and B. Nagy: IEEE Access **6** (2018) 23108. <https://doi.org/10.1109/ACCESS.2018.2827566>
- 6 P. Bao, L. Zhang, and X. Wu: IEEE Trans. Pattern Anal. Mach. Intell. **27** (2005) 1485. <https://doi.org/10.1109/TPAMI.2005.173>
- 7 Q. Xu, S. Varadarajan, C. Chakrabarti, and L. J. Karam: IEEE Trans. Image Process **23** (2014) 2944. <https://doi.org/10.1109/TIP.2014.2311656>
- 8 W. Kim and C. Kim: IEEE Signal Process Lett. **19** (2012) 127. <https://doi.org/10.1109/LSP.2011.2182648>
- 9 B. Han and L. S. Davis: IEEE Trans. Pattern Anal. Mach. Intell. **34** (2012) 1017. <https://doi.org/10.1109/TPAMI.2011.243>
- 10 S. Kumar and J. S. Yadav: Perspect. Sci. **8** (2016) 317. <https://doi.org/10.1016/j.pisc.2016.04.064>
- 11 T. Dharamadhat, K. Thanasontornlerk, and P. Kanongchaiyos: Proc. 2008 IEEE Int. Conf. Robot Biomimetic (IEEE, 2008) 1255–1260. <https://doi.org/10.1109/ROBIO.2009.4913180>
- 12 C. Lee, S. Lin, C. Lee, and C. Yang: 2010 Int. Conf. Mach. Learn. Cybern. (2010) 3015. <https://doi.org/10.1109/ICMLC.2010.5580739>
- 13 W. S. Kim, A. Diaz-Calderon, S. F. Peters, J. L. Carsten, and C. Leger: IEEE Trans. Cybern. **44** (2014) 2109. <https://doi.org/10.1109/TCYB.2014.2301442>
- 14 Y. Li and B. Bhanu: IEEE Sens. J. **11** (2011) 676. <https://doi.org/10.1109/JSEN.2010.2051148>
- 15 Y. Yao, C. Chen, B. Abidi, D. Page, A. Koschan, and M. Abidi: IEEE Trans. Syst. Man Cybern. C Cybern. **40** (2010) 101. <https://doi.org/10.1109/TSMCB.2009.201750>
- 16 S. Khan and M. Shah: IEEE Trans. Pattern Anal. Mach. Intell. **25** (2003) 1355. <https://doi.org/10.1109/TPAMI.2003.1233912>
- 17 J. Kim and D. Kim: 2008 ACM/IEEE Int. Conf. Distributed Smart Cameras (2008) 1. <https://doi.org/10.1109/ICDSC.2008.4635705>
- 18 W. Limprasert, A. Wallace, and G. Michaelson: IEEE J. Emerg. Sel. Top Circuits Syst. **3** (2013) 263. <https://doi.org/10.1109/JETCAS.2013.2256820>
- 19 G. Tsagkatakis and A. Savakis: IEEE Trans. Circuits Syst. Video Technol. **21** (2011) 1810. <https://doi.org/10.1109/TCSVT.2011.2133970>
- 20 B. Purnama, B. Erfianto, and Y. Hafidz: 2014 Int. Conf. Info. Comm. Technol. (2014) 297. <https://doi.org/10.1109/ICoICT.2014.6914082>
- 21 Q. Zhao, W. Tian, Q. Zhang, and J. Wei: 2016 Int. Conf. Soft Computing Mach. Intell. (2016) 143. <https://doi.org/10.1109/ISCMI.2016.41>
- 22 Z. Wu, J. Yang, H. Liu, Z. Guo, and Q. Zhang: 2015 IEEE Int. Conf. Consumer Electronics (2015) 298. <https://doi.org/10.1109/ICCE-Berlin.2015.7391263>
- 23 M. F. Aslan, A. Durdu, K. Sabanci, and M. A. Mutluer: Electrical and Electronics Engineering, Karamanoglu Mehmetbey University (2020).
- 24 M. Heimbach, K. Ebadi, and S. Wood: 2018 Conf. Signals Syst. Computers (2018) 1499. <https://doi.org/10.1109/ACSSC.2018.8645175>
- 25 I. Whoang, F. Akram, J. H. Kim, and K. N. Choi: 2013 Int. Conf. Image Graphics (2013) 369. <https://doi.org/10.1109/ICIG.2013.80>
- 26 X. Li and H. Lu: 2012 IEEE Int. Conf. Image Process (2012) 413. <https://doi.org/10.1109/ICIP.2012.6466883>
- 27 C. H. Hsia, T. Y. Lin, J. L. Lin, H. Prasetyo, S. L. Chen, and H. W. Tseng: Sens. Mater. **32** (2020) 3235. <https://doi.org/10.18494/SAM.2020.2862>
- 28 N. S. Pai, Y. H. Zhou, P. Y. Chen, W. L. Chen, and S. A. Chen: Sens. Mater. **31** (2019) 2245. <https://doi.org/10.18494/SAM.2019.2211>
- 29 Y. C. Wu, C. H. Chen, Y. T. Chiu, and P. W. Chen: Electronics **10** (2021) 1780. <https://doi.org/10.3390/electronics10151780>
- 30 J. S. Sheu and C. Y. Han: Advan. Technol. Inno. **5** (2019) 10. <https://doi.org/10.46604/aiti.2020.4284>
- 31 J. S. Sheu and C. W. Chen: Sens. Mater. **32** (2020) 3209. <https://doi.org/10.18494/SAM.2020.2860>

## About the Authors



**Po-Tong Wang** received his M.S. degree in computer science from National Taipei University of Education, Taiwan, in 2015 and his Ph.D. degree from the Bio-Industrial Mechatronics Engineering Department at NTU, Taiwan in 2020. He is currently an assistant professor in the Department of Electrical Engineering, Lunghwa University of Science and Technology, Taiwan. His primary research interests include image sensors, color space transformation algorithms, PLCopen control for advanced industrial automation, and industrial applications of image vision and automation in Taiwan. (neojwang@gmail.com)



**Jia-Shing Sheu** received his M.S. and Ph.D. degrees from the Department of Electrical Engineering at National Cheng Kung University, Tainan, Taiwan, in 1995 and 2002, respectively. He is currently a professor in the Department of Computer Science at National Taipei University of Education, Taipei, Taiwan. His research interests include pattern recognition and image processing, especially focusing on real-time face recognition, and embedded systems. (jiashing@tea.ntue.edu.tw)



**Jing-Han Lai** received her M.S. and B.S. degrees from the Department of Computer Science, National Taipei University of Education, Taiwan. Her main interests are in the fields of image processing and embedded systems. (g110832005@grad.ntue.edu.tw)