

Research Article

Automatic Person Identification in Camera Video by Motion Correlation

Dingbo Duan,¹ Guangyu Gao,² Chi Harold Liu,² and Jian Ma¹

¹ Beijing University of Posts and Telecommunications, Beijing 100876, China

² Beijing Institute of Technology, Beijing 100081, China

Correspondence should be addressed to Chi Harold Liu; liuchi02@gmail.com

Received 23 February 2014; Revised 13 May 2014; Accepted 13 May 2014; Published 3 June 2014

Academic Editor: Eugenio Martinelli

Copyright © 2014 Dingbo Duan 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.

Person identification plays an important role in semantic analysis of video content. This paper presents a novel method to automatically label persons in video sequence captured from fixed camera. Instead of leveraging traditional face recognition approaches, we deal with the task of person identification by fusing information from motion sensor platforms, like smart phones, carried on human bodies and extracted from camera video. More specifically, a sequence of motion features extracted from camera video are compared with each of those collected from accelerometers of smart phones. When strong correlation is detected, identity information transmitted from the corresponding smart phone is used to identify the phone wearer. To test the feasibility and efficiency of the proposed method, extensive experiments are conducted which achieved impressive performance.

1. Introduction

With the rapid growth in storage devices, networks, and compression techniques, large-scale video data have become available to more and more ordinary users. Thus, it also becomes a challenging task to search and browse desirable data according to content in large video datasets. Generally, person information is one of the most important semantic clues when people are recalling video contents. Consequently, person identification is crucial for content based video summary and retrieval.

The main purpose of person identification is to associate each subject that appears in video clips with a real person. However, to manually label all subjects that appear in a large-scale video archive is labor intensive, time consuming, and prohibitively expensive. To deal with this, automatic face detection [1–3] and face recognition (FR) [4–7] were introduced. However, traditional FR methods are still far from supporting practical and reliable automatic person identification, even when just a limited number of people appear in the video. This is mainly due to the fact that only appearance information (e.g., color, shape, and texture) of a single face image is used to determine the identity of a subject. Specifically, variation in illumination, pose, and face

expression as well as partial or total face occlusion could all make recognition an extremely difficult task.

The main contributions of the proposed method are as follows. First, this method provides an alternative way towards automatic person identification by the integration of a new sensing model. This integration broadens the domain of semantic analysis of video content and will be catalyzed by the growing popularity of wearable devices and concurrent advance in personal sensing technology and ubiquitous computing. Second, the method is fully automatic without any need for establishing a predefined model or need for user interaction in the process of person identification. Moreover, the independence of any recognition technique makes the proposed method more robust with respect to issues mentioned above which degrade the efficiency and accuracy of FR techniques. Last but not least, the simplicity and computational efficiency of the method make it possible to plug into real-time systems.

2. Related Work

To improve the performance of person identification, contextual information was utilized in recent research. Authors

in [8] proposed a framework exploiting heterogeneous contextual information including clothing, activity, human attributes, gait, and people cooccurrence, together with facial features to recognize a person in low quality video data. Nevertheless, it suffers the difficulty in discerning multiple persons resembling each other in clothing color or action. View angle and subject-to-camera distance were integrated to identify person in video by fusion of gait and face in [9], only in situations when people walk along a straight path with five quantized angles. Temporal, spatial, and social context information was also employed in conjunction with low level feature analysis to annotate person in personal and family photo collections [10–14], in which only static images are dealt with. Moreover, in all these methods, a predefined model has to be trained to start the identification process and the performance is limited by the quality and scale of training sets.

In contrast to the above efforts we propose a novel method to automatically identify person in video using human motion pattern. We argue that, in the field of view (FOV) of a fixed camera, motion pattern of human body is unique. Under this assumption, except for visual analysis, we also analyze the motion pattern of human body measured by sensor modules in smart phone. In this paper, we use smart phones equipped with 3-axis accelerometers carried on human bodies to collect and transmit acceleration information and identity information. By analyzing the correlation between motion features extracted from two different types of sensing, the problem of person identification is properly handled simply and accurately.

The remainder of the paper is organized as follows. Section 3 details the proposed method. In Section 4, experiments are conducted and results are discussed. Concluding remarks are placed in Section 5.

3. General Framework

A flowchart of the proposed method is depicted in Figure 1. As can be seen, visual features of human body are first extracted to track people across different video frames. Then, optical flows of potential human body are estimated and segmented using the previously obtained body features. Meanwhile, accelerometer measurements from smart phones on human bodies are transmitted and collected, together with identity information. Motion features are calculated from both optical flow and acceleration measurements in a sliding window style which was depicted in Section 3.3. When people disappear from video sequences, correlation analysis starts the annotation process. Details of the method are illustrated in the following subsections.

3.1. Camera Data Acquisition. First of all, background subtraction (BGS) which is widely adopted in moving objects detection in video is utilized in our method. The main idea of BGS is to detect moving objects from the difference between current frame and a reference frame, often called “background image,” or “background model” [15]. In this subsection, we need to detect image patches corresponding to

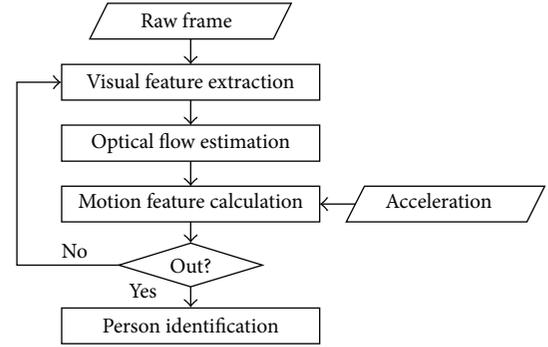


FIGURE 1: Flowchart of the proposed method.

potential human bodies moving around in the camera FOV. To this end, an algorithm of adaptive Gaussian mixture model [16, 17] is employed to segment foreground patches. This algorithm represents each pixel by a mixture of Gaussians to build a robust background model in run time.

When people enter into the camera FOV, image patches corresponding to potential human bodies are extracted and tracked by descriptors composed of patch ID, color histograms, and patch mass center in Algorithm 1. Moreover, we also include the frame index of first and last appearance of each patch in the descriptor in order to facilitate person annotation.

For patch p obtained from BGS, we try to associate p to previous patch descriptors. Histogram similarity between patches from consecutive frames is first analyzed. Normally image patches corresponding to the same subject are more similar to each other than those of different subjects. The comparison of color histogram of paths used in Algorithm 1 is defined in (1). The range of $s(H_a, H_b)$ is $[-1, 1]$. The larger $s(H_a, H_b)$, the more similar patches a and b . Then from the set of similar descriptors of p , the nearest one is selected to track p in terms of horizontal movement of patch center:

$$s(H_a, H_b) = \frac{\sum_{i=1}^N (H_a(i) - \bar{H}_a)(H_b(i) - \bar{H}_b)}{\sqrt{\sum_{i=1}^N (H_a(i) - \bar{H}_a)^2 \sum_{i=1}^N (H_b(i) - \bar{H}_b)^2}}, \quad (1)$$

$$\bar{H} = \frac{1}{N} \sum_{i=1}^N H(i), \quad (2)$$

where N is number of bins in histogram H .

For each patch p , we employ optical flow method to estimate motion pattern [18] and approximate patch acceleration as mean of vertical acceleration of keypoints within it, as defined in

$$y_{\text{-acc}_p} = \frac{1}{M} \sum_{i=1}^M y_{\text{-acc}_i}, \quad (3)$$

where $y_{\text{-acc}_i}$ is the second order derivative of y coordinate of keypoints with respect to time. M is the total number of keypoints within patch p .

Variables

$pDesc$: patch descriptor, $pDesc = \{id, frameStart, frameEnd, center, hist\}$.

$pDescs$: an array of patch descriptors.

$pCount$: patch descriptor counter, initialized to zero.

$frameIdx$: frame counter, initialized to zero.

id : the ID of a patch.

$frameStart, frameEnd$: frame index of first and last appearance of a patch.

$center, hist, accs$: center and color histogram of a patch.

$s_{thr}, d_{thr}, a_{thr}$: thresholds for histogram similarity, patch distance, and patch area, $a_{thr} \geq 0$, and $d_{thr} \geq 0, 0 \leq s_{thr} \leq 1$.

Procedure

- (1) Grab a video frame
 $frameIdx = frameIdx + 1$
- (2) Optical flow estimation
- (3) Background subtraction
- (4) **for** Each patch in current frame **do**
- (5) Calculate $pArea, pCenter, pHist, pWidth$
- (6) **if** $pArea < a_{thr}$ **then**
- (7) continue
- (8) **end if**
- (9) $pDescs^* = \emptyset$
- (10) **for all** $pDesc \in pDescs$ **do**
- (11) **if** $pDesc \cdot frameEnd + 1 == frameIdx$ **and** $s(pHist, pDesc \cdot hist) \geq s_{thr}$ **then**
- (12) $pDescs^* = \{pDesc\} \cup pDescs^*$
- (13) **end if**
- (14) **end for**
- (15) $d_{min} = d_{thr} * pWidth, pDesc_{min} = null$
- (16) **for all** $pDesc \in pDescs^*$ **do**
- (17) $d_p = |pCenter \cdot x - pDesc \cdot center \cdot x|$
- (18) **if** $d_p < d_{min}$ **then**
- (19) $d_{min} = d_p, pDesc_{min} = pDesc$
- (20) **end if**
- (21) **end for**
- (22) **if** $pDesc_{min}$ is null **then**
- (23) $pDesc_{min} = \{pCount, frameIdx, frameIdx, pCenter, pHist\}$
 $pDescs = pDescs \cup \{pDescs_{min}\}$
 $pCount = pCount + 1$
- (24) **else**
- (25) $pDesc_{min} \cdot frameEnd = frameIdx$
 $pDesc_{min} \cdot center = pCenter$
 $pDesc_{min} \cdot hist = pHist$
- (26) **end if**
- (27) Calculate and save vertical acceleration for $pDesc_{min}$
- (28) **end for**

ALGORITHM 1: Patch tracking and motion estimation.

Pseudocode of patch tracking and motion estimation is listed in Algorithm 1.

3.2. Accelerometer Measurements Collection. In this subsection we depict the procedure of acceleration measurements collection using wearable sensors. Android smart phones equipped with 3-axis accelerometers are utilized as sensing platforms. For the three component accelerometer readings, only the one with largest absolute mean value is analyzed in our experiments due to its best reflection of vertical motion pattern of human body. Three different placements are tested and compared in order to assess impacts of different phone placements on accuracy of motion collection. In each test, a

participant performs a set of activities randomly including standing, walking, and jumping while carrying three smart phones on body, with two phones placed in chest pocket and jacket side pocket, respectively, and one attached to waist belt, as shown in Figure 2. Results illustrated in Figure 3 qualitatively show that all three types of placement could correctly capture vertical motion feature of the participant with minor acceptable discrepancy. This test makes the choice of phone attachment more flexible and unobtrusive.

3.3. Feature Extraction and Person Identification. Noisy raw motion measurements of different sample frequency previously obtained from different sensor sources cannot be



FIGURE 2: Attachment of smart phones to human body. From left to right: jacket side pocket, chest pocket, and belt attachment.

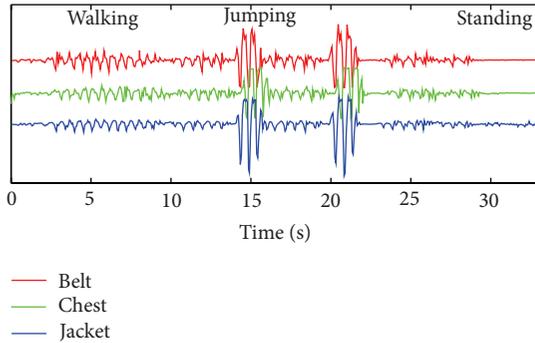


FIGURE 3: Acceleration measurements from three ways of phone attachment for the above mentioned activities.

compared directly. Instead, standard deviation and energy [19, 20] are employed as motion features for comparison after noise suppression and data cleansing. Energy is defined as sum of squared discrete FFT component magnitudes of data samples and divided by sample count for normalization. These features are computed in a sliding window of length t_w with $t_w/2$ overlapping between consecutive windows. Feature extraction on sliding windows with 50 percent overlapping has demonstrated its success in [21]:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (4)$$

To find out whether p represents a human body, correlation analysis is conducted. As a matter of fact, motion features extracted from video frames are supposed to be positively linear with those from accelerometer measurements of the same subject. We adopt correlation coefficient to reliably measure strength of linear relationship, as defined in (4), where X and Y are motion features to be compared, $\text{cov}(X, Y)$ the covariance, and σ_X and σ_Y the standard deviation of X and Y . ρ ranges from -1 to 1 inclusively, where 0 indicates no linear relationship, $+1$ indicates a perfect positive linear relationship, and -1 indicates a perfect negative linear relationship. The larger $\rho(X, Y)$, the more correlated X and Y . In our case, motion features of p are compared with each of those extracted from smart phones in the same period of time. Identity information of smart phone corresponding to the largest positive correlation coefficient is utilized to identify p .

4. Experiments and Discussions

In this section, we conduct detailed experiments in various situations to optimize Algorithm 1 and evaluate the proposed person identification algorithm. We use a digital camera and two Android smart phones for data collection. A simple GUI application is created to start and stop data collection on phones. Acceleration measurements are recorded and saved in text files on phone SD card and later accessed via USB. Video clips are recorded in the format of mp4 files at a resolution of 640×480 , 15 frames per second. The timestamps of video frames and accelerometer readings are well synchronized before the experiment. Algorithm 1 is implemented based on OpenCV library and tested on an Intel 3.4 GHz platform running Ubuntu 13.04. We recruit two participants, labeled as A and B, respectively, to take part in our experiments and place smart phones in jacket side pockets. We choose four different scenarios to perform our experiments, including *outdoor near field*, *outdoor far field*, *indoor near field*, *indoor far field*, as illustrated in Figure 8. In *near field* situations, the subjects moved around within a scope about five meters away from the camera. The silhouette height of human body is not less than half of the image height and human face could be clearly distinguished. In *far field* situations, the subjects moved around about twenty meters away where detailed visual features of human body are mostly lost and body height in image is not more than thirty pixels. In each scenario, we repeated the experiment four times and each lasts about five minutes. In all we collect sixteen video clips and thirty-two text files of acceleration measurements.

4.1. Tracking Optimization. Patch tracking is an essential step for motion estimation from camera video and directly affects accuracy and robustness of subsequent person identification. As listed in Algorithm 1, the aim of patch tracking is to estimate motion measurements for each patch that appeared in video frames. In the ideal case, a subject is continuously tracked in camera video by only one descriptor during the whole experiment and we could extract a sequence of acceleration measurements closest to that collected from the smart phone in terms of time duration, while in the worst case, we have to create new descriptors for all patches in each frame and the number of descriptors used for tracking a subject is as many as that of the frames of his appearance. We present a metric in (5) to measure the performance of Algorithm 1. The metric $L(i)$ is defined as a ratio between number of subjects in a video clip and number of descriptors used for tracking the subjects. The range of $L(i)$ is $(0, 1]$. The larger $L(i)$, the better the tracking performance. Moreover, we also provide a metric to evaluate tracking accuracy, as shown in (6). Accurate descriptor means that a descriptor tracks only one subject during its lifetime. The larger $K(i)$, the more accurate Algorithm 1:

$$L(i) = \frac{\text{\#subjects in video } i}{\text{\#descriptors in } i}, \quad (5)$$

$$K(i) = \frac{\text{\#accurate descriptors in video } i}{\text{\#descriptors in } i}. \quad (6)$$

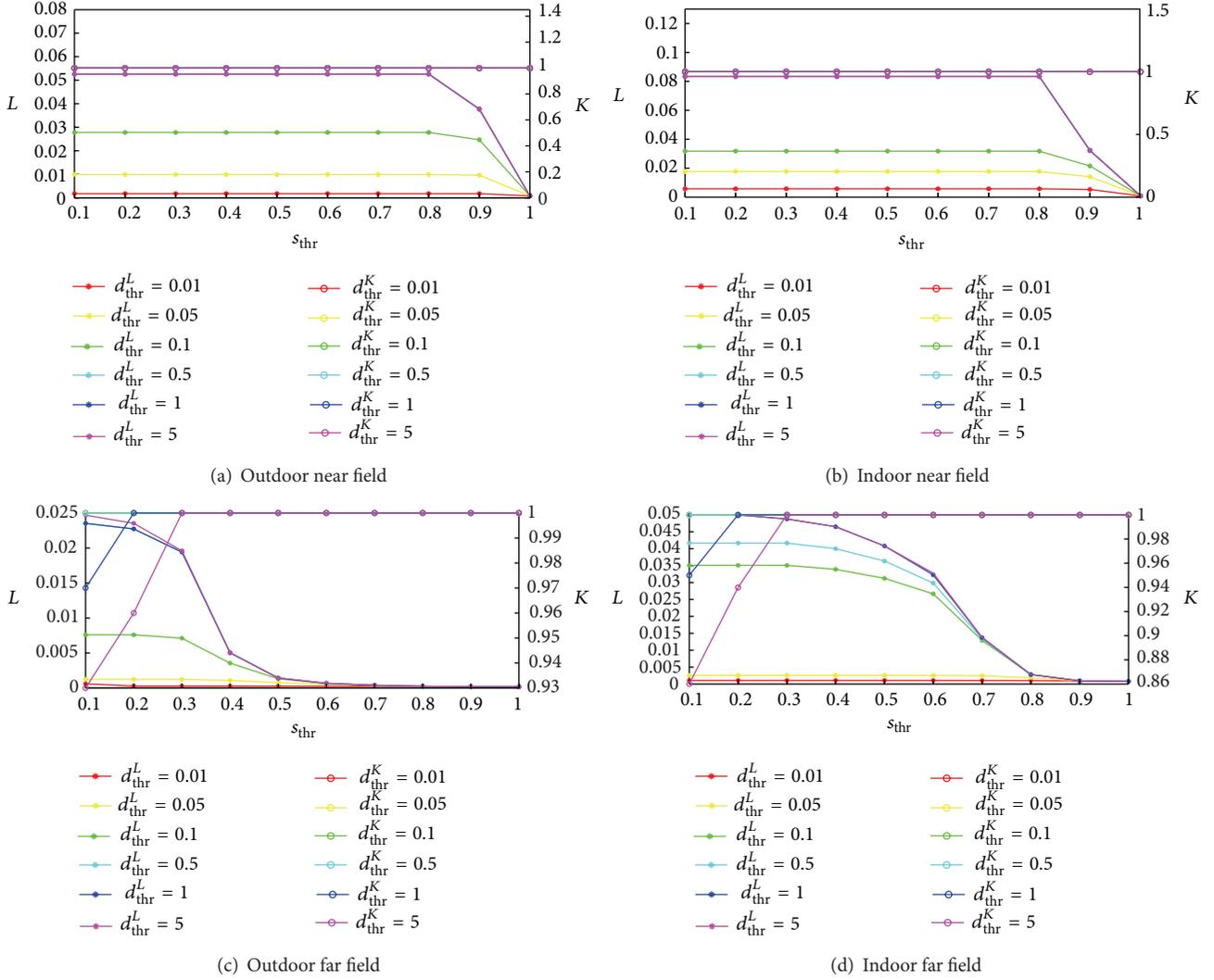


FIGURE 4: Tracking performance and performance in the four scenarios with different values of d_{thr} and s_{thr} .

As depicted in Algorithm 1, three parameters, a_{thr} , d_{thr} , and s_{thr} , affect L and K . a_{thr} indicates minimum area of a patch that potentially represents a subject. Patches with an area less than a_{thr} are filtered out. Generally, in a specific application scenario, the value of a_{thr} could be figured out empirically. In our experiments we set it to 150 which works fine. s_{thr} specifies a minimum histogram similarity between current patch p and potential descriptors of p . Each active descriptor that satisfies this requirement is tested in terms of horizontal distance to p . d_{thr} stipulates a distance threshold to rule out inappropriate alternative descriptors. A nearest descriptor satisfying this threshold is selected to track p if it exists. Otherwise we create a new descriptor for p . Moreover, many interference factors in the scenario including poor lighting condition, similar clothing color to the background, incidental shadow of human body, and unpredictable motion pattern of subjects like fast turning and crossing would also pose negative effects to patch tracking process. To rule out impacts of these factors and optimize patch tracking, from each of the four scenarios, we select a representative video clip

and run Algorithm 1 over the video with different s_{thr} and d_{thr} . Resulted L and K are illustrated in Figure 4. Extracted frames from video clips with labeled patches are listed in Figure 8.

Due to different motion patterns of the subjects, L may vary among video clips of different scenarios. However, from Figure 4 we can conclude that L drops dramatically when $s_{\text{thr}} > 0.8$ in near field scenario and $s_{\text{thr}} > 0.2$ in far field scenario with $d_{\text{thr}} \geq 0.1$. This is mainly caused by background subtraction noises. Histogram similarity of patches of the same subject from two consecutive frames is about 0.8 in near field in this situation. In far field scenarios with relatively smaller foreground patches, the negative impacts become more severe and threshold similarity degrades to 0.2. Patches of the same subject are associated with different descriptors when histogram similarity is beyond these thresholds. When $d_{\text{thr}} < 0.1$, the worst case occurs. We need to create new descriptors for patches in every frame as horizontal distance between patches of the same subject from two consecutive frames is mostly beyond this limit. As d_{thr} increases, L increases and converges at $d_{\text{thr}} = 5$.

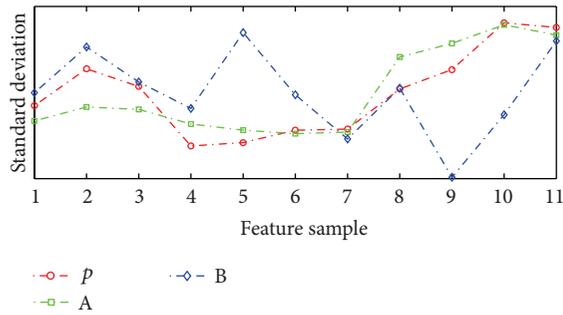


FIGURE 5: Standard deviation of accelerations of patch p and subjects A and B.

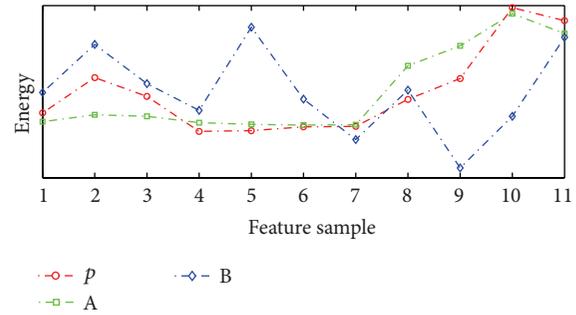


FIGURE 6: Energy of accelerations of patch p and subjects A and B.

In near field scenarios, Algorithm 1 achieves 100 percent accuracy with whatever s_{thr} and d_{thr} , while in far field scenarios, it does not perform so perfectly when $d_{\text{thr}} \geq 0.1$ and $s_{\text{thr}} \leq 0.2$. In the experiments, we found that this happened mostly in situations when subjects were close and the patch of one subject lost in the following frame.

To balance L and K , we set $d_{\text{thr}} = 5$, $s_{\text{thr}} = 0.2$, run Algorithm 1 over the sixteen video clips, and collect motion measurements for person identification in the following experiments. Statistics of the obtained descriptors are illustrated in Figure 7.

4.2. Person Identification. When motion measurements collection from video finished, we obtain a set of patch descriptors and each descriptor associates with a time series of acceleration data of a potential subject. Some descriptors within the set come with short series of motion data usually less than ten frames. This is possibly caused by subjects crossing each other, fake foreground from flashing lights, fast turning of human body, moving objects at the edge of camera FOV, and so forth. These insufficient and noisy data fail to reflect actual motion pattern of potential subjects and are filtered out in the first place. As shown in Figure 7, there are comparatively more noisy descriptors in far field scenarios, especially in outdoor far field scenarios where nearly 50 percent of descriptors are ruled out in each video.

Then we calculate a sequence of motion features for each descriptor and compare the feature sequence with each of those obtained from smart phones in the same period of time. Sliding window in motion feature calculation is closely related to subjects and application scenarios. It should be large enough to capture the distinctive pattern of subject movement but not too large to confuse different ones. In our experiments we set window size to 1 second empirically. Motion features from an example patch descriptor and those from the two smart phones in the same period are shown in Figures 5 and 6, where we could conclude that patch p represents subject B during its lifetime.

The total number of accurately identified patch descriptors in each video is listed in Figure 7. The proposed method achieves comparatively better performance in near field environment where we can capture more accurate and robust motion measurements of human body. The worst case

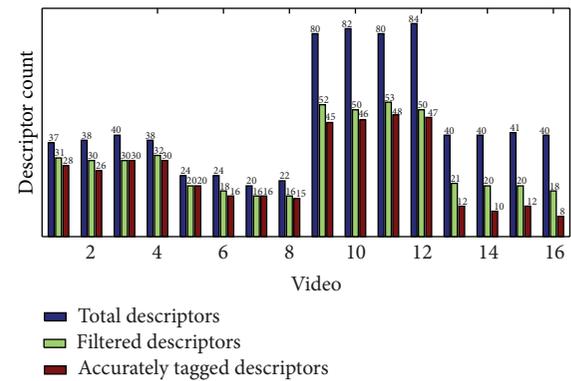


FIGURE 7: Obtained descriptors with optimized parameters, filtered descriptors, and accurately identified descriptors, where 1–4 represent the outdoor near field, 5–8 represent the indoor near field, 9–12 represent the outdoor far field, and 13–16 represent the indoor far field.

happens in outdoor far field scenario. In this case, there are less optical flows within each patch and less frames associated with each descriptor. We save the mapping between patch descriptors and their estimated identity and rerun Algorithm 1 with the same parameter configuration as before. The obtained patch identity is labeled in the video right after patch ID. As illustrated in Figure 8, the proposed method could maintain comparatively acceptable performance even under adverse situations.

5. Conclusions

In this paper, we propose a novel method for automatic person identification. The method innovatively leverages correlation of body motion features from two different sensing sources, that is, accelerometer and camera. Experiment results demonstrate the performance and accuracy of the proposed method. However, the proposed method is limited in the following aspects. First, users have to register and carry their smart phones in order to be discernable in camera FOVs. Second, we assume that phones stay relatively still with human body during the experiments, but in practice, people tend to take out and check their phones from time to time. Acceleration data collected during these occasions would damage the identification accuracy. Besides, the method



FIGURE 8: Screenshots of identification results, where (a)–(d) represent the outdoor near field, (e)–(h) represent the indoor near field, (i)–(l) represent the outdoor far field, and (m)–(p) represent the indoor far field.

relies heavily on background subtraction in the process of patch tracking. Thus a more practical and reliable strategy for motion data collection is needed. Third, subjects in archived video clips without available contextual motion information cannot be identified using the proposed method. Therefore, this method only works at the time of video capture. In the future, we plan to overcome the aforementioned constraints and extend the application of the proposed method into more complex environments.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Grant no. 61202436, Grant no. 61271041, and Grant no. 61300179).

References

- [1] P. Vadakkepat, P. Lim, L. C. de Silva, L. Jing, and L. L. Ling, "Multimodal approach to human-face detection and tracking," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 3, pp. 1385–1393, 2008.
- [2] C. Zhang and Z. Zhang, "A survey of recent advances in face detection," Tech. Rep., Microsoft Research, 2010.
- [3] C. Huang, H. Ai, Y. Li, and S. Lao, "High-performance rotation invariant multiview face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 671–686, 2007.

- [4] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [5] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [6] G. Shakhnarovich and B. Moghaddam, "Face recognition in subspaces," in *Handbook of Face Recognition*, pp. 19–49, Springer, 2011.
- [7] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2106–2112, 2010.
- [8] L. Zhang, D. V. Kalashnikov, S. Mehrotra, and R. Vaisenberg, "Context-based person identification framework for smart video surveillance," *Machine Vision and Applications*, 2013.
- [9] X. Geng, K. Smith-Miles, L. Wang, M. Li, and Q. Wu, "Context-aware fusion: a case study on fusion of gait and face for human identification in video," *Pattern Recognition*, vol. 43, no. 10, pp. 3660–3673, 2010.
- [10] N. O'Hare and A. F. Smeaton, "Context-aware person identification in personal photo collections," *IEEE Transactions on Multimedia*, vol. 11, no. 2, pp. 220–228, 2009.
- [11] Z. Stone, T. Zickler, and T. Darrell, "Autotagging Facebook: Social network context improves photo annotation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '08)*, pp. 1–8, June 2008.
- [12] D. Anguelov, K.-C. Lee, S. B. Göktürk, and B. Sumengen, "Contextual identity recognition in personal photo albums," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–7, June 2007.
- [13] M. Naaman, R. B. Yeh, H. Garcia-Molina, and A. Paepcke, "Leveraging context to resolve identity in photo albums," in *Proceedings of the 5th ACM/IEEE Joint Conference on Digital Libraries—Digital Libraries: Cyberinfrastructure for Research and Education*, pp. 178–187, June 2005.
- [14] M. Zhao, Y. W. Teo, S. Liu, T. S. Chua, and R. Jain, "Automatic person annotation of family photo album," in *Image and Video Retrieval*, pp. 163–172, Springer, 2006.
- [15] M. Piccardi, "Background subtraction techniques: a review," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, vol. 4, pp. 3099–3104, October 2004.
- [16] 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, IEEE, August 2004.
- [17] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," in *Video-Based Surveillance Systems*, pp. 135–144, Springer, 2002.
- [18] G. Farneböck, "Two-frame motion estimation based on polynomial expansion," in *Image Analysis*, pp. 363–370, Springer, 2003.
- [19] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
- [20] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones," in *Proceedings of the 8th International Conference on Intelligent Environments (IE '12)*, pp. 214–221, IEEE, June 2012.
- [21] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proceedings of the 20th National Conference on Artificial Intelligence and the 17th Innovative Applications of Artificial Intelligence Conference (AAAI/IAAI '05)*, pp. 1541–1546, July 2005.



Hindawi

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

