Camera Network Based Person Re-identification by Leveraging Spatial-Temporal Constraint and Multiple Cameras Relations

Wenxin Huang[1], Ruimin Hu[1,2](✉), Chao Liang[1,2], Yi Yu[3], Zheng Wang[1], Xian Zhong[4], and Chunjie Zhang[5]

[1] National Engineering Research Center for Multimedia Software, School of Computer, Wuhan University, Wuhan, China
wenxin.huang@whu.edu.cn
[2] Collaborative Innovation Center of Geospatial Technology, Wuhan, China
hrm1964@163.com
[3] National Institute of Informatics, Tokyo, Japan
[4] School of Computer, Wuhan University of Technology, Wuhan, China
[5] University of Chinese Academy of Sciences, Beijing, China

Abstract. With the rapid development of multimedia technology and vast demand on video investigation, long-term cross-camera object tracking is increasingly important in the practical surveillance scene. Because the conventional Paired Cameras based Person Re-identification (PCPR) cannot fully satisfy the above requirement, a new framework named Camera Network based Person Re-identification (CNPR) is introduced. Two phenomena have been investigated and explored in this paper. First, the same person cannot simultaneously appear in two non-overlapping cameras. Second, the closer two cameras, the more relevant they are, in the sense that persons can transit between them with a high probability. Based on these two phenomena, a probabilistic method is proposed with reference to both visual difference and spatial-temporal constraint, to address the novel CNPR problem. (i) Spatial-temporal constraint is utilized as a filter to narrow the search space for the specific query object, and then the Weibull Distribution is exploited to formulate the spatial-temporal probability indicating the possibility of pedestrians walking to a certain camera at a certain time. (ii) Spatial-temporal probability and visual feature probability are collaborated to generate the ranking list. (iii) The multiple camera relations related to the transitions are explored to further optimize the obtained ranking list. Quantitative experiments conducted on TMin and CamNeT datasets have shown that the proposed method achieves a better performance to the novel CNPR problem.

Keywords: Person re-identification · Spatial-temporal constraint · Camera relation · Ranking optimization · Camera network

1 Introduction

Person re-identification is a task of visually matching person images, obtained from different cameras deployed in non-overlapping surveillance scenes [1–3] with

Q. Tian et al. (Eds.): MMM 2016, Part I, LNCS 9516, pp. 174–186, 2016.
DOI: 10.1007/978-3-319-27671-7_15

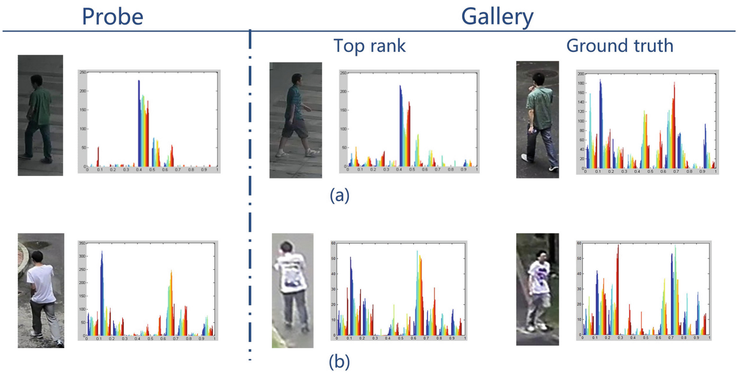

Fig. 1. Examples of the difficulties in exploiting visual features. The first column stands for the probe images, and the other two columns represent the candidates in the gallery. All the images have their color histograms in right boxes. The person in the second column is not the one in the first column but has the similar color histogram profile based on appearance information; the person in the third column is the ground truth but might be overlooked due to the large difference in the color histogram. (a) illustrates the difficulty caused by the different illuminations and (b) illustrates the difficulty caused by similar wearing.

place and time difference. The existing person re-identification problem is mainly regarded as a retrieval task on a pair of cameras [4,6,7,14,15,18,20]. With the rapid development of multimedia technology and vast demand for video investigation, camera networks have become increasingly deployed in public spaces such as airports, road intersections and campuses [14–18], which calls for the long-term cross-camera object tracking and human behavior analysis. However, the conventional Paired Camera based Person Re-identification (PCPR) methods cannot well solve this problem [11–13]. To achieve the retrieval task on multiple cameras, we introduce a new framework named Camera Network based Person Re-identification (CNPR).

As we know, some of the existing PCPR approaches focus on constructing visual features which are both distinctive and stable under various conditions [4–7], while others focus on learning an optimal metric [8–10], in which images of the same person are closer than those of different persons. All these methods mainly depend on the visual information, which faces intrinsic challenges caused by various changes in viewpoints, poses and illumination conditions (see Fig. 1(a)) in the practical surveillance environment [11,12]. To be worse, for special conditions where different persons wear highly similar clothes (see Fig. 1(b)), it is difficult to provide sufficient identity discrimination power when purely relying on visual information. According to the characteristic of CNPR, the spatial-temporal information exists among the images of persons, which motivates us to combine original visual features with it together to assist person re-identification. The distance of these two types of information cannot be simply fused and measured. Then, the proposed method introduces a spatial-temporal probability,

Table 1. Contrast between CNPR and PCPR.

	CNPR	PCPR
Characteristic	Multi-camera network	Pair of cameras
Method	1. Spatial-temporal information + Visual information 2. Probabilistic model	1. Appearance-based 2. Distance measurement
Evaluation	mAP [21]	CMC [19]

converts the distance of visual features to visual probability, and combines them together as a joint probability.

Meanwhile, we find two phenomena in the CNPR (the details will be illustrated in Sect. 2), that can be utilized to optimize the re-identification performance in the camera network. (i) **Spatial-Temporal Constraint.** A person of interest who appears in one camera cannot appear in another non-overlapping camera at the same time or in a certain time period. Based on this phenomenon, Hinge loss function is exploited to construct a filter model to reduce the query scope, and the Weibull Distribution is utilized to formulate the spatial-temporal probability model. Probabilities of probe-to-gallery images are all calculated, then the initial ranking lists are generated. (ii) **Multiple Cameras Relations.** Inspired by the idea of image retrieval re-ranking methods [18], the probe-to-gallery ranking list is influenced by the relationships of different gallery-to-gallery images from different cameras. However, the relationships are different originated from different relations of cameras. We consider that if two cameras are more relevant according to deployed locations and transition time, it reveals that the probability of the corresponding person should be elevated. Exploiting the gallery-to-gallery probabilities and camera relations, the initial ranking list is optimized.

We summarize the differences of CNPR and PCPR in Table 1, and further explain in the following: (i) PCPR only focuses on a pair of cameras in the network, neglecting the relationships among different cameras, while CNPR involves a camera network which includes many cameras deployed in different places. (ii) It is distinctly different to solve CNPR and PCPR problems. PCPR methods [8,10,22,23] regard PCPR as a ranking problem and solve it by measuring the differences among visual features, while the proposed CNPR exploits spatial-temporal information besides visual features. In order to solve the CNPR problem, a probabilistic model, instead of the distance metric, is used to compute the similarity of two images. (iii) The Cumulated Matching Characteristics (CMC) curve [20] is typically used in PCPR to evaluate the performance. This evaluation measurement is valid only if there is only one ground truth for a probe. For the CNPR problem, there may be more than one cross-camera ground truth for each query. Therefore, we adopt the mean average precision (mAP) [21] as the metric to evaluate the overall performance.

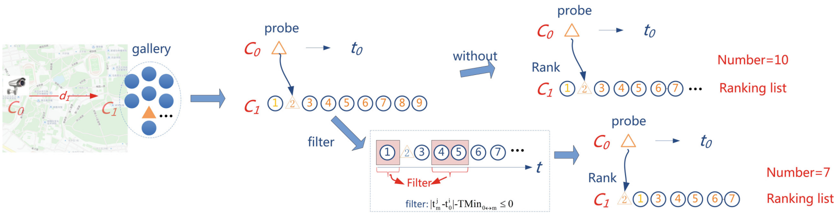

Fig. 2. Spatial-temporal constraint. There are always noises in the candidates in the gallery, especially in the case of ever increasing surveillance videos. Spatial-temporal constraint can effectively remove a number of noises in the non-overlapping camera network. That is to say, if the probe appears in camera C_0 at a certain time t_0, the images of candidates observed in camera C_1 at time t with the interval of $TMin_{0\leftrightarrow 1}$ can be filtered out. Here, $TMin_{0\leftrightarrow 1}$ is the minimized walking time which is related to the distance between C_0 and C_1.

Technical contributions of this paper are three-fold, as follows: (1) This paper puts forward CNPR as a new approach for the person re-identification problem. (2) Two phenomena, spatial-temporal constraint and camera relations, are exploited by a probabilistic model. (3) This paper adopts mAP as a new performance criterion particularly designed for CNPR. It considers both precision and recall, thus providing a more comprehensive evaluation than CMC widely used for PCPR.

2 Observations

In order to acquire additional information such as the time stamps and make full use of the captured videos, the association of the information of each camera can be explored since there are relations among a large number of videos. In this work, spatial-temporal constraint and relations of cameras in the network are investigated.

Spatial-Temporal Constraint. Spatial-temporal information represents more strict constraint to limit the query scope. It could help to improve the query efficiency and matching rate. There is a common sense that a person of interest appeared in a camera cannot be in another non-overlapping camera at the same time. In addition, significant spatial-temporal gaps exists since there is a distance interval between two cameras in a non-overlapping camera network. Based on the statement, it can describe when and where the target had stayed in the entire camera network. Therefore, we can reduce the confusing images which are in different places at the same time of interests. Figure 2 illustrates the phenomenon.

Multiple Cameras Relations. We assume that the walk of pedestrian is continuous, then a person will have a high appearance probability in two adjacent cameras. In CNPR problem, a person may appear in multiple cameras in the

Table 2. Proportion (%) of the ground truth appearing in the returned results with considering multiple cameras relations.

	Camera pairs with strong relation	Camera pairs with weak relation
top 1	17.64	10.53
top 3	64.71	26.32
top 5	88.24	36.84

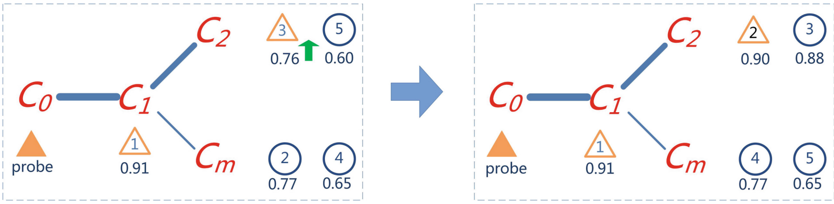

Fig. 3. Multiple cameras relations. There are relations of C_0, C_1, C_2 and C_m, where the degree of thickness of lines represents how the relations are. The numbers in the triangle and circle reveal the ranking number of the candidates. The numbers under the triangle and circle indicate the matching probabilities of the candidates. Left box indicates the situation of regular ranking and right box indicates the situation of the new ranking list after considering relations of cameras. Since the relation of C_1 and C_2 is closer than that of C_1 and C_m, the matching probabilities of C_2 rises.

network. In this assumption, the matching probability of the person is affected by the distance relations of cameras. A preliminary experiment conducted to explain this phenomenon. Two adjacent pairs of cameras and two furthest pairs of cameras in TMin [13] were selected respectively as the pairs with strong relation and the pairs with weak relation. For each pair, we choose one as the probe, and the other as the retrieval candidate. We performed statistical analysis on these data and the statistical result is illustrated in Table 2. We can find that a person will have a high appearance probability in cameras with strong relations. In the initial ranking list, the persons in the top results have high matching probabilities to be the probe person. For each top result, he can also obtain the corresponding pedestrians with high probabilities in other gallery cameras. These corresponding pedestrians may be highly possible to be the probe person. Therefore, the ranking number of these pedestrians in the initial ranking list should be elevated. The degree of elevation depends on the relation between the gallery camera of the top result and that of the selected pedestrian. Figure 3 reveals that if the relation of camera is closer, the evaluation will be higher. However, the selected pedestrians may have multiple relations with different cameras, then the elevation is related to multiple cameras relations.

In brief, the core idea of solving the CNPR problem is that we construct a probabilistic framework based on the whole camera network, further optimizing with spatial-temporal constraint and multiple cameras relations.

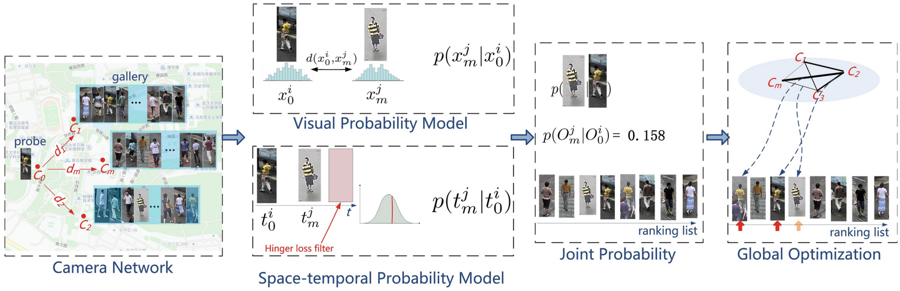

Fig. 4. Framework of the proposed method in CNPR

3 Our Approach

3.1 Problem Definition

This subsection gives a brief definition of the CNPR problem. We consider a camera network $C = \{C_0, C_1, C_2, ..., C_M\}$, which is composed of $M+1$ cameras with non-overlapping field of views. Here, C_0 denotes the probe camera, and M is the number of gallery cameras.

We denote the representative image of person j, captured by camera C_m as x_m^j, $x_m^j \in R^n$, and $0 < m \leq M$. Here, n represents the dimension of the visual feature of the image. The time of each observation is recorded as well. Then, an observation O_m^j can be described as a combination $O_m^j = (x_m^j, t_m^j)$, where t_m^j denotes the time person j walks into the view of camera C_m.

In PCPR, merely the visual feature is exploited. That is to say, the distance $d(x_0^i, x_m^j)$ is calculated between the probe image i and an image j in the gallery, where the gallery images are all from camera C_m. Then, a ranking list is obtained depending on the calculated distances. In comparison, for the CNPR problem, we exploit the probability theory instead of a distance metric to represent not only the similarity of visual features but also the similarity of spatial-temporal relationship. If the observation O_m^j gets a high conditional probability $p(O_m^j|O_0^i)$ based on the probe O_0^i, it will obtain a high ranking number. Besides, the probability is also related to the relations of cameras, as a result, we can exploit the relations of cameras to refine the observations under the entire camera network. The framework is shown in Fig. 4.

$$p(x_m^j|x_0^i) = e^{-\alpha \cdot d(x_0^i, x_m^j)} \tag{1}$$

Equation 1 converts the visual distance into a probability. Here, the distance can be obtained by any existing algorithm for PCPR.

3.2 Probabilistic Model with Spatial-Temporal Constraint

In a fixed camera network, the minimum walking time between each pair of cameras is given. It is assumed that the minimum walking time between C_0 and

C_m is $TMin_{0\leftrightarrow m}$. As discussed in Sect. 2, if $|t^j_m - t^i_0| < TMin_{0\leftrightarrow m}$, the spatial-temporal probability $p(t^j_m|t^i_0)$ will be zero for the person i appearing in C_m. In other words, if the person i appears in camera C_0, within the minimum walking time $TMin_{0\leftrightarrow m}$, he cannot appear in camera C_m. In this situation, we do not need to calculate the probability $p(O^j_m|O^i_0)$ any more. Here, we introduce hinge loss acting as a filter described in Eq. 2.

$$h(t^i_0, t^j_m) = \max(0, |t^j_m - t^i_0| - TMin_{0\leftrightarrow m}) \tag{2}$$

Considering the process of pedestrian's walk, when $|t^j_m - t^i_0| = TMin_{0\leftrightarrow m}$, the matching probability equals to zero; when $|t^j_m - t^i_0| > TMin_{0\leftrightarrow m}$, the matching probability increases at first and then reaches the peak value. As the time interval gets too long, the matching probability will get down and tend to zero with the assumption that the person is continuously walking. We assume the time for the transition between cameras follows a Weibull distribution [26], which has been successfully applied to nearly all scientific disciplines, such as biological, environmental, health, physical and social sciences. By fitting time data to Weibull distributions, Weibull analysis enables risk assessment and planning of corrective actions. Then, if we have two observations between the probe camera and another one in the fixed camera network, the conditional probability of the transition from C_0 to C_m is described as following:

$$p(t^j_m|t^i_0) = \frac{k}{\lambda}(\frac{h(t^i_0, t^j_m)}{\lambda})^{k-1} e^{-(h(t^i_0, t^j_m)/\lambda)^k}, \tag{3}$$

where $k > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter of the distribution.

Joint Probability Model. The visual distance between the image from a camera in the gallery and the probe image can be formulated as a conditional probability in Eq. 1. It is assumed spatial-temporal relations, independent of visual features, could assist to increase the ranking number. Then the conditional probability is obtained as shown in Eq. 4. The ranking list could be obtained after repeating the process on all the images in the gallery from the cameras in the network. Through this way, the ranking is related to both visual features and spatial-temporal constraint.

$$p(O^j_m|O^i_0) = p(x^j_m|x^i_0)p(t^j_m|t^i_0) \tag{4}$$

3.3 Optimization with Multiple Camera Relations

Optimization includes two aspects: (i) We exploit the relations between two cameras which are related to the spatial-temporal interval to construct spatial-temporal probabilities, and combine them with visual features probabilities to get the probabilistic model. (ii) We utilize the relations of entire camera network to improve the matching probability of candidates in the gallery. The optimization of the ranking can be concluded as follows: (i) By constructing the joint probability model, the initial ranking lists are obtained. (ii) With each image

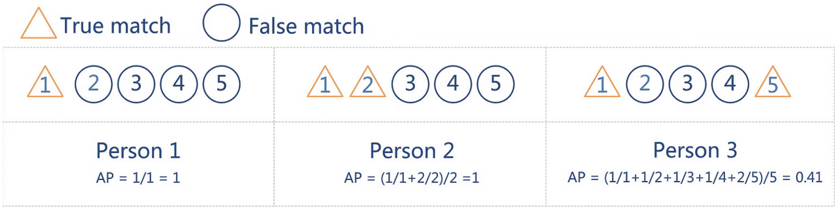

Fig. 5. Example of maP. This example indicates the value of AP. For all three persons, the CMC curve remains 1, which is not proper for CNPR but AP shows the differences. mAP is the mean average precision of the three persons which is $(1 + 1 + 0.41)/3 = 0.803$.

in top-k of the initial ranking list as a probe, we find its top-k similar images from the gallery. (iii) The relevance of two cameras which is related to spatial-temporal interval in the network can be obtained by the spatial-temporal difference as $r(m, m')$ $(r(m, m') > 1)$, and the closer the cameras are, the larger the value of $r(m, m')$ is. The matching probability of the candidates in the initial ranking list can be adjusted to $\sum_{m' \in KNN(m)} p(O_m^j | O_0^i) * r(m, m')$. Here, for an image in the initial ranking list, m is its camera ID, and m' is the camera ID of an image in its KNN. Finally, the new ranking list is achieved, which involves the relationship between cameras.

4 Experiments

4.1 Baselines

mAP. The CMC curve is exploited by most papers on the person re-identification problem [4,8–10,19]. The value of CMC which tells the rate of the correct match indicates the identification results for every pair of camera views [7]. As each probe image may correspond to multiple ground truths in the gallery, precision and recall over entire camera network should be considered as a metric while evaluating the CNPR problem. In this case, we adopt a metric named mean average precision (mAP) to indicate the percentage of the real matches over the camera network. There is an example of mAP shown in Fig. 5.

4.2 TMin Data Set

Our task requires a number of cameras, walking time between pair of cameras and the topology of cameras. As far as we know, there is no appropriate public database in non-overlapping multi-camera person re-identification field. Thus, we utilize two public databases, TMin [13] and CamNeT [24], often used in the multi-camera tracking field which can satisfy our requirements. The version 1 of the TMin database contains 1680 images from 30 subjects. All the images are extracted from 6 cameras and the video acquisition time starts at twenty

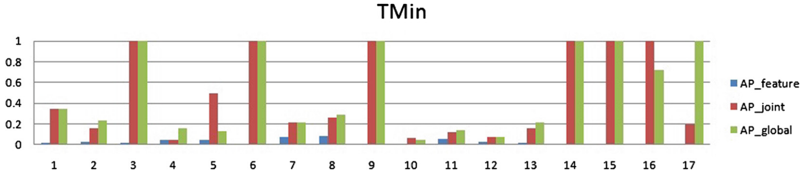

Fig. 6. Performance of different methods on the TMin dataset AP_ feature, AP_ joint and AP_ global represent the matching accuracy of each probe which are calculated by visual features, probabilistic model of combined spatial-temporal and visual features and optimization approach. Most of the results reveals that there is evident improvement after introducing spatial-temporal probabilistic model and relations of multiple cameras.

to twelve in the morning and ends at a quarter to two in the afternoon. Each pedestrian of TMin appears in at least two different cameras.

We employed camera 1# of TMin database which contains 17 of 30 pedestrians as our probe set in this paper. First, the noisy images were filtered out based on a filter introduced above. Then, we exploited Local Maximal Occurrence (LOMO) [25] as the feature representation which is effective and robust to illumination and viewpoint changes. We calculated the visual appearance probability of C_0 and C_m respectively, and unified them in a matrix. After that, spatial-temporal probability which was calculated according to hinge loss was joined with the appearance probability to construct the overall probability of an observation in the entire camera network compared with probe camera 1#. At last, we used the relations between cameras to refine the ranking. The results of visual features, joint probability and optimization are shown in Fig. 6. Moreover, in Table 3 we compare the performance of visual features, joint probability and global optimization in the range of the first 10, 20, 30, 40, 50, 75 ranks respectively. As can be seen, our approach greatly improves mAP compared with conventional methods, and the improvements are evident after introducing spatial-temporal probabilistic model and global optimization respectively[1].

4.3 CamNeT

CamNeT [24] is the database of non-overlapping camera network tracking data set for multi-target tracking, it consists of 5 to 8 cameras which cover both indoor and outdoor scenes at a university. The paths of around 10–25 people are predefined while several unknown persons move through the scene. There are 6 scenarios in which every scenario lasts at least 5 min with 5 to 8 cameras. We employed camera 1# of CamNeT database which contains 11 pedestrians as our probe set in this paper. To evaluate our method, the experimental process mainly followed the way for the TMin data set. The result is presented in Fig. 7.

[1] Here, we set $\alpha = 0.5$, $\lambda = 75$, $\beta = 2.5$, and $K = 5$, when evaluating on TMin data set.

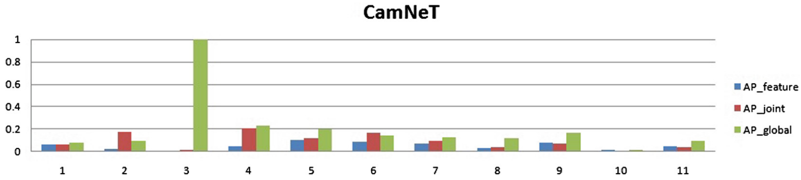

Fig. 7. Performance of different methods on the CamNeT dataset AP_ feature, AP_ joint and AP_ global represent the matching accuracy of each probe which are calculated by visual features, probabilistic model of combined spatial-temporal and visual features and optimization approach. Most of the results reveals that there is evident improvement after introducing spatial-temporal probabilistic model and relations of multiple cameras.

Table 3. Comparing mAP value (%) with different methods on top K (TMin).

K	10	20	30	40	50	75
mAP_ feature	2.39	2.28	2.39	2.36	2.46	2.57
mAP_ joint	50.08	49.10	48.44	47.97	47.83	47.73
mAP_ global	**53.96**	**51.92**	**51.06**	**50.50**	**50.33**	**50.20**

Our approach improves evidently. And we compared the performance of visual features, joint probability and optimization in the range of the first 10, 20, 30, 40, 50, 75 ranks respectively in Table 4[2].

However, we found that the joint probability of person 9 and person 11 which combined spatial-temporal and features had been decreased in Fig. 7. The reason is that there is excess layover time during the pedestrians' walk among cameras, which is beyond our assumption that the walk is continuous in this paper. For another unexpected situation, we also found that the optimization results of person 5 and person 16 had been decreased in Fig. 6. These two persons appeared only twice in other cameras, in other words, negative samples appears more which influenced the optimization effect. Although there are few disharmonious members, the overall results are still evident improved by the proposed method.

4.4 Running Time

In this subsection, we test the running time of each procedure. The additional running time, when compared with the running time of the baseline method using the distance of visual features, is no more than 1 ms even though we add the processes of calculating spatial-temporal probability and optimization. In fact, we use spatial-temporal constraint to filter 2.6 % and 27.35 % images of gallery of TMin and CamNeT, it somehow increases the efficiency. The experimental

[2] Here, we set $\alpha = 0.5$, $\lambda = 50$, $\beta = 1.5$, and $K = 10$, when evaluating on CamNeT data set.

Table 4. Comparing mAP value (%) with different methods on top K (CamNeT).

K	10	20	30	40	50	75
mAP_ feature	6.35	5.13	4.99	4.98	4.85	4.59
mAP_ joint	13.66	11.02	9.72	9.14	9.00	8.94
mAP_ global	**24.41**	**23.10**	**21.71**	**21.01**	**20.80**	**20.58**

environment is as follows [18]: our computer includes a dual core 2.80 GHz CPU and 2 GB RAM.

5 Conclusion

This paper puts forward a new framework CNPR as a fresh person re-identification solution. While solving the CNPR problem, two phenomena of spatial-temporal constraint and the relations among cameras in the network are investigated and leveraged in this work. On this basis, this paper proposes an approach of optimization, based on a probabilistic model taking into account spatial-temporal constraint and visual probability. The experiments, conducted on two public databases TMin and CamNeT, show significant improvement in efficiency and accuracy compared with conventional methods.

Acknowledgement. The research was supported by National Nature Science Foundation of China (61303114, 61231015, 61170023), the Specialized Research Fund for the Doctoral Program of Higher Education (20130141120024), the Technology Research Project of Ministry of Public Security (2014JSYJA016), the China Postdoctoral Science Foundation funded project (2013M530350), the major Science and Technology Innovation Plan of Hubei Province (2013AAA020), the Guangdong-Hongkong Key Domain Break-through Project of China (2012A090200007), and the Special Project on the Integration of Industry, Education and Research of Guangdong Province (2011B090400601). Nature Science Foundation of Hubei Province (2014CFB712). Jiangxi Youth Science Foundation of China(20151BAB217013).

References

1. Gong, S., Cristami, M., Yan, S., Loy, C.: Person Re-Identification. Advances in Computer Vision and Pattern Recognition. Springer, London (2014)
2. Wang, Z., Hu, R., Liang, C., Leng, Q., Sun, K.: Region-based interactive ranking optimization for person re-identification. In: Ooi, W.T., Snoek, C.G.M., Tan, H.K., Ho, C.-K., Huet, B., Ngo, C.-W. (eds.) PCM 2014. LNCS, vol. 8879, pp. 1–10. Springer, Heidelberg (2014)
3. Wang, Z., Hu, R., Liang, C., Jiang, J., Sun, K., Leng, Q., Huang, B.: Person re-identification using data-driven metric adaptation. In: He, X., Luo, S., Tao, D., Xu, C., Yang, J., Hasan, M.A. (eds.) MMM 2015, Part II. LNCS, vol. 8936, pp. 195–207. Springer, Heidelberg (2015)

4. Farenzena, M., Bazzani, L., Perina, A., Murino, V., Cristani, M.: Person re-identification by symmetry-driven accumulation of local features. In: Computer Vision and Pattern Recognition (CVPR) (2010)
5. Hu, Y., Liao, S., Lei, Z., Yi, D., Li, S.Z.: Exploring structural information and fusing multiple features for person re-identification. In: IEEE Workshop on Camera Networks and Wide Area Scene Analysis (in conjunction with CVPR 2013) (2013)
6. Gheissari, N., Sebastian, T.B., Hartley, R.: Person re-identification using spatiotemporal appearance. In: Computer Vision and Pattern Recognition (CVPR) (2006)
7. Wang, X., Doretto, G., Sebastian, T., Rittscher, J., Tu, P.H.: Shape and appearance context modeling. In: International Conference on Computer Vision (ICCV) (2007)
8. Prosser, B., Zheng, W.-S., Gong, S., Xiang, T.: Person re-identification by support vector ranking. In: British Machine Vision Conference (BMVC) (2010)
9. Zheng, W.S., Gong, S., Xiang, T.: Person re-identification by probabilistic relative distance comparison. In: Computer Vision and Pattern Recognition (CVPR) (2011)
10. Kostinger, M., Hirzer, M., Wohlhart, P., Roth, P.M., Bischof, H.: Large scale metric learning from equivalence constraints. In: Computer Vision and Pattern Recognition (CVPR) (2012)
11. Ma, L., Yang, X., Tao, D.: Person re-identification over camera networks using multi-task distance metric learning. IEEE Trans. Image Process. (TIP) **23**(8), 3656–3670 (2014)
12. Das, A., Chakraborty, A., Roy-Chowdhury, A.K.: Consistent re-identification in a camera network. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part II. LNCS, vol. 8690, pp. 330–345. Springer, Heidelberg (2014)
13. Hu, Y., Liao, S., Yi, D., et al.: Multi-camera trajectory mining: database and evaluation. In: International Conference on Pattern Recognition (ICPR) (2014)
14. Loy, C.C., Xiang, T., Gong, S.: Multi-camera activity correlation analysis. In: Computer Vision and Pattern Recognition (CVPR) (2009)
15. Javed, O., Shafique, K., Rasheed, Z., Shah, M.: Modeling intercamera spacetime and appearance relationships for tracking across non-overlapping views. In: Computer Vision and Image Understand (CVIU) (2008)
16. Wang, Y., Velipasalar, S., Casares, M.: Cooperative object tracking and composite event detection with wireless embedded smart cameras. IEEE Trans. Image Process. (TIP) **19**(10), 2614–2613 (2010)
17. Ding, C., Song, B., Morye, A., Farrell, J.A., Roy-Chowdhury, A.K.: Collaborative sensing in a distributed PTZ camera network. IEEE Trans. Image Process. (TIP) **21**(7), 3282–3295 (2012)
18. Leng, Q., Hu, R., Liang, C., et al.: Bidirectional ranking for person re-identification. In: International Conference on Multimedia and Expo (ICME) (2013)
19. Li, X., Tao, D., Jin, L., Wang, Y., Yuan, Y.: Person re-identification by regularized smoothing kiss metric learning. IEEE Trans. Circuits Syst. Video Technol. (TCSVT) **23**(10), 1675–1685 (2013)
20. Wang, Y., Hu, R., Liang, C., Zhang, C., Leng, Q.: Camera compensation using a feature projection matrix for person reidentification. IEEE Trans. Circ. Syst. Video Technol. **24**(8), 1350–1361 (2014)
21. Zheng, L., Shen, L., Tian, L., et al.: Person re-identification meets image search. arXiv (2015)
22. Zheng, W.-S., Gong, S., Xiang, T.: Re-identification by relative distance comparison. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **35**(3), 653–668 (2013)
23. Mignon, A., Jurie, F.: PCCA: a new approach for distance learning from sparse pairwise constraints. In: Computer Vision and Pattern Recognition (CVPR) (2012)

24. Zhang, S., Staudt, E., Faltemier, T., et al.: A camera network tracking (CamNeT) dataset and performance baseline. In: IEEE Winter Conference on Applications of Computer Vision (WACV) (2015)
25. Liao, S., Hu, Y., Zhu, X., et al.: Person re-identification by local maximal occurrence representation and metric learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
26. Liu, C., Ryen, W., Susan, D.: Understanding web browsing behaviors through Weibull analysis of dwell time. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (2010)