ISSN No:-2456 –2165

Post-Ranking Person Re- Identification using Discriminant Context Information Analysis

A.R.Atlin Beulah,
2n yr M.E (CSE), Student

Sworna Kokila.M.L
Associate Professor, Dpt of CSE

Abstract:-Person re-identification has become increasingly popular in the community due to its application and research significance. It aims at spotting a person of interest in other cameras.The discriminant information analysis which transforms the original feature vectors by removing the common information, thus defining the discriminant feature. A novel post-ranking framework for person re-identification (an unsupervised post-ranking framework) is proposed toimprove the first ranking results and outperforms the state-of-the-art approaches. The analysis of the similar appearances of the first ranks can be helpful in detecting, hence removing, and such visual ambiguities. Once the initial ranking is available, content and context sets are extracted. Then, these are exploited to remove the visual ambiguities and to obtain the discriminant feature space which is finally exploited to compute the new ranking. We demonstrate on two pedestrian benchmarks that by learning a more discriminative representation, our method significantly improves the first ranks results.

keywords:- Person Re-Identification, Discriminant, Extracted.

I. INTRODUCTION

Person re-identification is the problem of re-associating a same person moving between the disjoint Fields-of-View of a wide area camera network. Due to the inherent challenges present in a multi-camera setting, the person re-identification is still an open problem. In particular, when a person is sensed by the different viewpoints of disjoint cameras, his/her appearance undergoes significant illumination and color variations as well as poses changes. The non-rigid shape of the human body, as well as background clutter, introduces additional challenges. The algorithm that can reliably track multiple persons in a complex environment and provide metrically accurate position estimates. This is achieved by combining the POM, which provides a very robust estimation of the occupancy on the ground plane in individual time frame, with a global optimization of the trajectories of the detected individuals over 100-frame batches [1]. An algorithm that can reliably track multiple people in a complex environment .This is achieved by resolving occlusions and localizing people on multiple scene planes using a planar homographic occupancy constraint [2].The WASA framework for different types of contexts and critical scenarios. Traditional SA solutions to protect CIs against security breaches and malicious threats, their complex and critical nature make this adaptation difficult [3].The need for methods able to link the information acquired between the covered areas such that high-level semantics can be obtained. In particular, one of the most currently attractive issues that such blind areas have introduced is the problem of re-associating a same person that is moving in a wide environment and who might be detected at a different location and time [4].

II. RELATED WORKS

Post-ranking methods for person re-identification is a relatively unexplored area. Earliest works following the post-ranking approach exploited ranking SVMs[34], boosting techniques for feature selection[10] or additional cues coming from soft biometrics[2]. Ranked lists computed for multiple probe persons were exploited to refine a single probe ranking [30]. Therefore, the approach works only if additional rankings (minimum 3 or 4) besides the one obtained for the current probe are available. Bidirectional ranking [17] and a saliency-based matching scheme [4] were also introduced. In the former case, first direction is usual ranking of the probe with the gallery. Second direction is the ranking obtained by matching each gallery with the probe and the rest of the gallery. Hence, differently from our approach, the whole gallery for post-ranking is considered, and no focus is placed on the visual ambiguities shared between first ranks. In the latter, the saliency similarity is computed between the probe and the gallery only, not between galleries themselves. Such similarities are adopted to revise the initial ranking within a local gallery window. The post-ranking optimization was also studied by including human feedback in the loop. The end user had to identify both similar and dissimilar samples [1, 37], to provide relative feedback [31], or to select a single strong negative feedback to refine the ranking [23] in the deployment stage. In contrast to all such methods, we propose a single shot approach that does not require human intervention. A slightly different approach was recently introduced in [22], where an iterative extension to sparse discriminative classier was adopted to ensure that the best candidates are ranked at each iteration. However, such method did not directly consider the content and the context similarities of ranked individuals. It cast the problem by analyzing the reconstruction error and by partially ranking the gallery in terms of similarity to the probe. Two main differences between the proposed approach and all such existing works can be highlighted: (i) there is no human neither in the training nor in the deployment loops; (ii) most importantly, the proposed approach is the only one studying the visual ambiguities shared between first ranks to improve re-identification performance, thus reranking is performed on a subset of the gallery.

III. PROPOSED OPTIMIZED STABLE CLUSTERING ALGORITHM

Discriminative Context Information Analysis methods exploited multiple local and global features to compute the person representations. These representations were combined with reference sets patch matching strategies saliency learning and joint attributes. An interesting bag-of-words approach, together with a benchmark dataset, was proposed in. Both labeled and unlabeled data as well as collaborative representation were used to boost performance. Learning architectures and multiple frame analysis methods were also explored to extract the most relevant features.

Post-ranking methods for person re-identification is a relatively unexplored area. Earliest works following the post ranking approach exploited boosting techniques for feature selection or additional cues coming from soft biometrics. However, in such a case, there is the need to acquire reliable biometric features which is generally a challenge in surveillance scenarios. In, ranked lists computed for multiple probe persons were exploited to refine a single probe ranking. Therefore, the approach works only if additional rankings, besides the one obtained for the current probe, are available. Bidirectional ranking and a saliency-based matching scheme were also introduced.

In the former case, first direction is usual ranking of the probe with the gallery. Second direction is the ranking obtained by matching each gallery with the probe and the rest of the gallery. Hence, differently from our approach, the whole gallery for post-ranking is considered, and no focus is placed on the visual ambiguities shared between first ranks. In the latter, the saliency similarity is computed between the probe and the gallery only. Then, such similarities are adopted to revise the initial ranking within a local gallery window. The post-ranking optimization was also studied by including human feedback in the loop. The end user had to identify both similar and dissimilar samples to adapt the metric model, to provide relative feedback for the classifier training, or to select a single strong negative feedback to refine the ranking in the deployment stage.

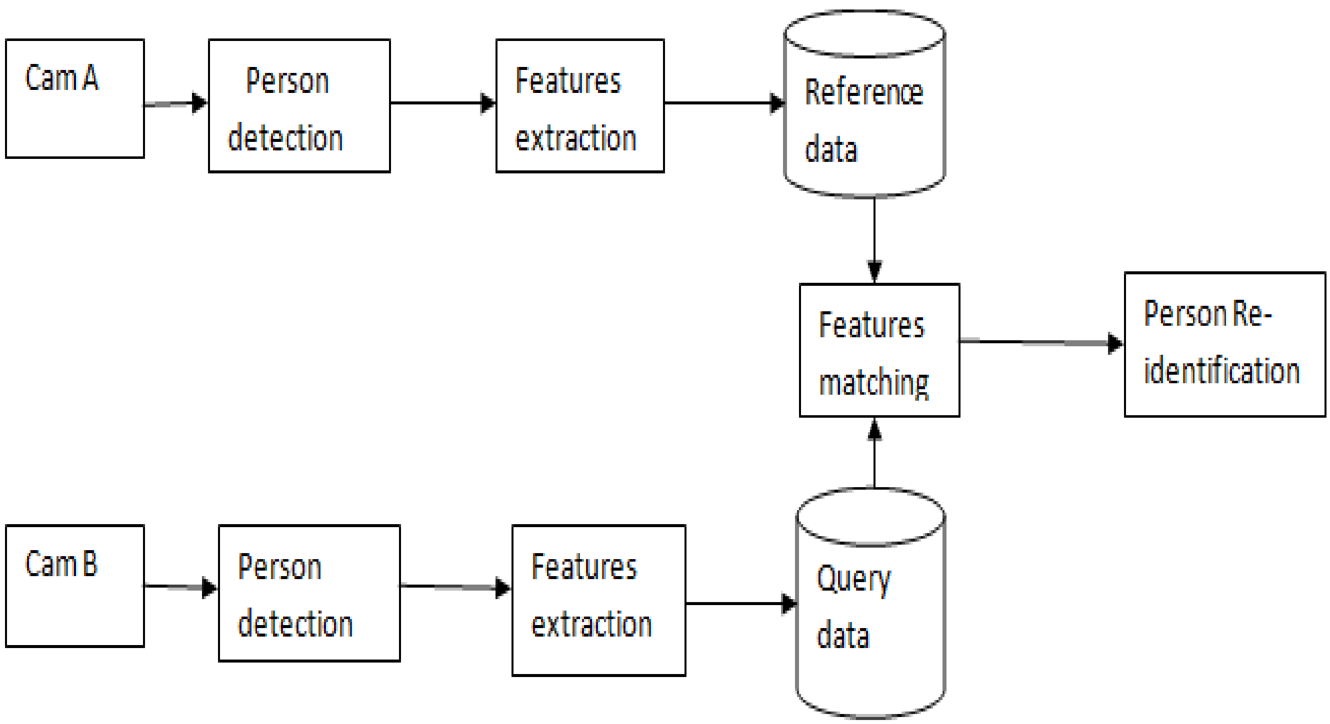

Figure 1: Overview of Proposed System

A. Person Detection

Human sensing (also called human detection or person detection) encompasses a range of technologies for detecting the presence of a human body in an area of space, typically without the intentional participation of the detected person.

B. Feature Extraction

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human.

Figure 2: Feature Extraction

C. Feature Detection

Feature detection (computer vision) In computer vision and image processing feature detection includes methods for computing abstractions of image information and making local decisions at every image point whether there is an image feature of a given type at that point or not.

D. Discriminant Context Information Analysis

Let A = {IAp}Np=1 be the set of N probe images and B = {IB g }Mg=1 be the set of M gallery images. Given a probe image IApits initial ranking is defined as Rp= {IB (r)}Mr=1 where m the gallery images IB (r) are sorted depending on the dissimilarityto the probe image IAp. Notice that, here, as well as in thefollowing, the subscript is used without parenthesis to denotea generic gallery person identity g, and it is enclosed withinparenthesis to indicate its position in a ranking list. In otherwords, d(IAp, IB(r)) < d(IAp, IB(r+1)), where r = 1, . . . , M − 1.Such an order is computed on the basis of a dissimilaritymeasure d(IAp, IBg) which is the result of the application of themodel LA,B on the feature vector pair (xAp, xBg).R = {Rp}Np=1denotes the set of such initial rankings computed for all the N probes.Our aim is to improve the rank of the true match in eachRp. Towards this objective we first select the content set, i.e.,a subset of gallery images Bcn⊆ B whose elements belong to the first ranks. Then, we compute the context set Bcx⊆ B,which contains gallery images that have small dissimilarity with respect to either the probe or an image in the content set.

To identify gallery images sharing visual ambiguities with a probe, *C*sa has to be computed. Towards this objective, we propose to use the k-means clustering algorithm to partition *B* into the three aforementioned sets, which minimize the following cost function,

In Equation(3.6) μjis the mean of the dissimilarities within the j -th partition. Before minimizing the cost function, we initialize the cluster means μsa= d(IAp, IB(1)), μda= d(IAp,IB(M/2)) and μoa= d(IAp, IB(M)). Thus, once the minimization isconcluded, the cluster Csa contains the top-m best matches,hence defines the content (cn) set Bcnp= {IB(1), . . . , IB(m)}.

E. Context Analysis

Context information can be defined as the object frequency appearance in a particular domain. In image retrieval, the context information is extracted from the set of images containing the target object. We provide a similar definition for the person re-identification problem: the context information is extracted from the K- common nearest neighbors of the probe and a gallery in the content set. For each image IB(r)∈ Bcnp,r = 1, . . . ,m, we compute Bcn(r) exploiting the content ranking R(r) which is obtained using the gallery set B∗ = B\IB(r). Given the content sets Bcn(p) and Bcn(r), we define _(r) =Bcnp∪ Bcn(r). The elements in _(r) are the images in B having high similarity with either the probe IApor the gallery IB(r).

Ranking results and corresponding content sets are shown in figure 4.2. Gallery images in the first ranks show visual ambiguities with the probe.

a). Initial Ranking

Euclidean distance to compute the matches for the initial ranking output. Test and trained features are given as input to the Euclidean distance.LA,B and ˜ LA,B define the set of parameters learned by such models trained with input feature vectors pairs in XTr and with discriminative feature vectors pairs in ˜ XTr.

b). Post Ranking Optimization

The post-ranking optimization is performed during the re-identification phase. Let us consider the ranking R computed using the original feature vectors pairs in XTe extracted from a test probe and every gallery image.

Then, the discriminant context information analysis is exploited to find the content and context sets, hence to obtain the discriminative feature vector set ˜ XTe. The final ranking R is obtained by sorting the dissimilarities d.

IV. RESULTS DISCUSSION

A. Performance Analysis

Discriminant context information analysis is a statistical analysis to predict a categorical dependent variable (called a grouping variable) by one or more continuous or binary independent variables (called predictor variables). The original dichotomous discriminant analysis was developed by Sir Ronald Fisher in 1936. It is different from an ANOVA or MANOVA, which is used to predict one (ANOVA) or multiple (MANOVA) continuous dependent variables by one or more independent categorical variables. Discriminant function analysis is useful in determining whether a set of variables is effective in predicting category membership.

Moreover, it is a useful follow-up procedure to a MANOVA instead of doing a series of one-way ANOVAs, for ascertaining how the groups differ on the composite of dependent variables. In this case, a significant F test allows classification based on a linear combination of predictor variables. Terminology can get confusing here, as in MANOVA, the dependent variables are the predictor variables, and the independent variables are the grouping variables

The achieved results are then provided in the form of a list of ranked matching persons. It often happens that the true match is not ranked first but it is in the first positions. This is mostly due to the visual ambiguities shared between the true match and other “similar” persons. At the current state, there is a lack of a study of such visual ambiguities which limit the re-identification performance within the first ranks.

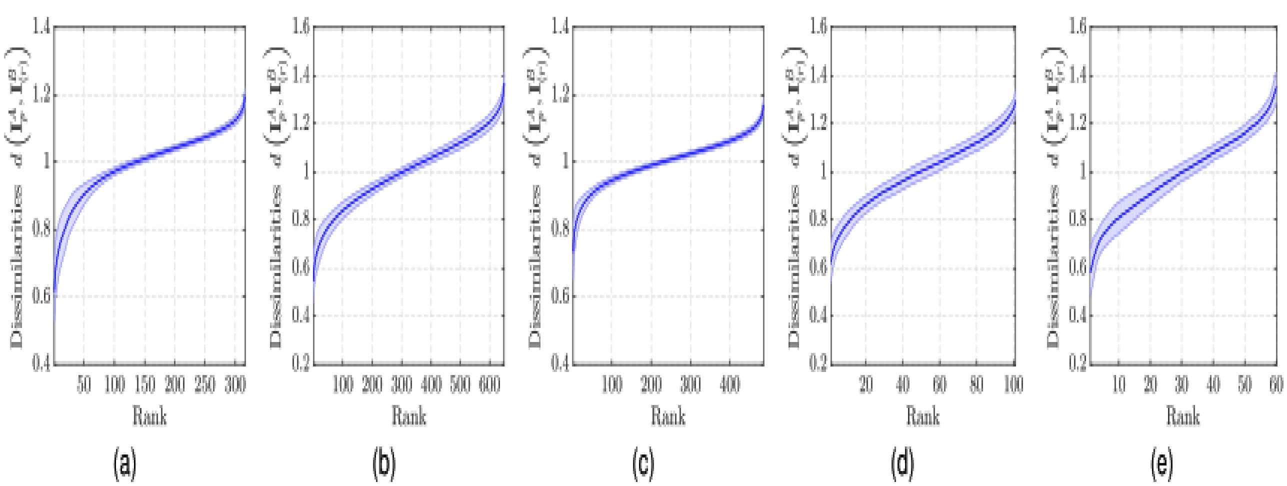

Figure 3: Ranks and Corresponding Dissimilarities

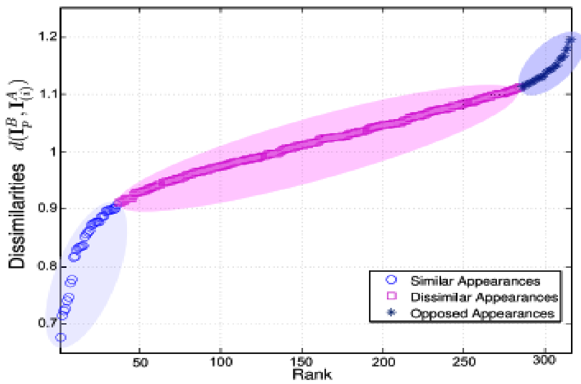

Figure 4: Result Applying K-Means

Figure 4: gives the example of the results obtained by applying k-means clustering algorithm to the dissimilarities computed between a probe and all the gallery images. Similar, dissimilar and opposed appearance clusters are represented by blue circles, magenta squares and dark blue stars, respectively.

a). Preliminaries and Dentitions

Our aim is to improve the rank of the true match in each p. Towards this objective we first select the content set, i.e.,R a subset of gallery images B cn⊆ B whose elements belong to the first ranks. Then, we compute the context set Bcx⊆ B, which contains gallery images that have small dissimilarity with respect to either the probe or an image in the content set.

b). Content Analysis

Existing methods try to locate the true match in the first ranking positions out from a large set of possible gallery matches. The visual ambiguities bring false matches in the first ranks, often before the true match. The selection process is inspired by the shape of the dissimilarities vs ranks plots. Fig. 4 shows that there exists a significant trend among all the datasets highlighting that: (i) at first ranks, dissimilarities with the probe image increases abruptly, then flatten (first elbow); (ii) from the first elbow, dissimilarities grow linearly till reaching high ranks, where they finally start increasing significantly (second elbow). According to such trend we have identified three classes of gallery images.(i) similar appearance class (Csa), which corresponds to gallery images with dissimilarities located before the first elbow; (ii) dissimilar appearance class (Cda) corresponding to gallery images having dissimilarities located in-between the two elbows and (iii) opposed appearance class (Coa), which corresponds to all the other gallery images. To identify gallery images sharing visual ambiguities with a probe, Csa has to be computed. Towards this objective, we propose to use the k-means clustering algorithm to partition B into the three aforementioned sets, which minimize the following cost function.

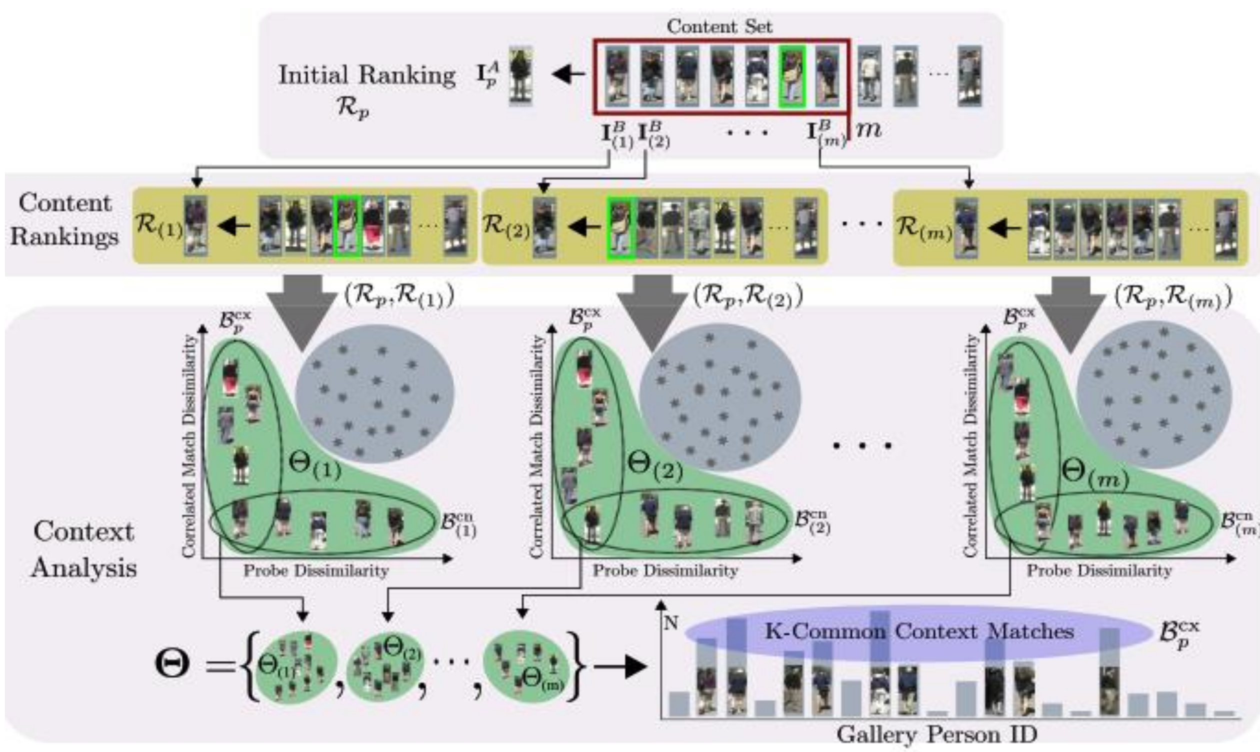

Figure 5: Representation of Content And Context Sets

Figure 5 represents that, Context information can be defined as the object frequency appearance in a particular domain. In image retrieval, the context information is extracted from the set of images containing the target object. We provide a similar definition for the person re-identification problem: the context information is extracted from the K -common nearest neighbors of the probe and a gallery in the content set.

c). Discriminant Information Analysis

The content and context sets include images with "similar" appearance. The goal of the following analysis is to detect, hence to remove, the visual ambiguities shared by these images. This allows us to focus on discriminant details that help to correctly locate the true match within the content set, thus improving its rank. We believe that visual ambiguities correspond to the global appearance which mainly contributes to the dissimilarity computation.

d). Post-Training

When an individual is certified as a peace officer, it means that they meet the requirements set up under POST, including successful completion of a training academy provided by a law enforcement, police or correctional POST-approved training center.

e). Post-Ranking Optimization

The post-ranking optimizations performed during the re-identification phase. Let us consider the ranking RRR computed using the original feature vectors pair's inXTe extracted from a test probe and every gallery image. Then, the discriminant context information analysis is exploited to find the content and context sets, hence to obtain the discriminative feature vector set ˜XTe. Thefinalranking˜R˜R˜Ris obtained by sorting the dissimilarities d computed by means of˜LA,Bfor each pair in˜XTe.

V. CONCLUSION

In this work, person reidentification have proposed a novel post-ranking framework. This introduced an unsupervised approach that exploits the visual ambiguities shared between first ranked persons. Two sets of gallery images are extracted from the initial ranking to model the content and context information. This is exploited by the discriminant information analysis which transforms the original feature vectors by removing the common information, thus defining the discriminant feature vector space. This is later exploited by a learning algorithm to discriminate between positive and negative image pairs. First rank matches are then re-ranked on the basis of the output of such a learning algorithm. Extensive evaluation performance have been conducted using three public benchmark datasets. Results demonstrated that, for every considered dataset, baseline models performance are always improved. This strongly supports person re-identification, i.e., that the initial ranking includes relevant information that can be used to improve first rank performance. Finally, comparisons with state-of-the-art methods have shown that rank 1 performance have been improved by more than 22% on two very challenging datasets.

VI. FUTURE WORK

Feature clustering can be added to reduce the feature space so that the speed of the recognition process gets reduced. The proposed DCIA framework represents one of the first efforts in improving re-identification performance by means of re-ranking. More specifically, our solution aims to identify and remove the global information shared between the first ranks such that the true match location can be improved. With an in-depth evaluation conducted on three datasets, person re-identification have demonstrated that, whenever the re-identification task relies on a metric computed in the feature space, DCIA applied on the top of any baseline method shows relevant performance improvement.

REFERENCES

[1]. F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multicamerapeopletracking with a probabilistic occupancy map," IEEE Trans. Pattern Anal.Mach. Intell., vol. 30, no. 2, pp. 267–282, Feb. 2008.
[2]. S. M. Khan and M. Shah, "Tracking multiple occluding people bylocalizing on multiple scene planes," IEEE Trans. Pattern Anal. Mach.Intell., vol. 31, no. 3, pp. 505–519, Mar. 2009.
[3]. C. Alcaraz and J. Lopez, "Wide-area situational awareness for criticalinfrastructure protection," Computer, vol. 46, no. 4, pp. 30–37,Apr. 2013.
[4]. A. K. Roy-Chowdhury, M. Kankanhalli, J. Konrad, C. Micheloni, andP. Varshney, "Introduction to the issue on signal processing for situationalawareness from networked sensors and social media," IEEE J.Sel. Topics Signal Process., vol. 9, no. 2, pp. 201–203, Mar. 2015.
[5]. E. Ricci, G. Zen, N. Sebe, and S. Messelodi, "A prototype learningframework using EMD: Application to complex scenes analysis,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 3, pp. 513–526,Mar. 2012.
[6]. B. Solmaz, B. E. Moore, and M. Shah, "Identifying behaviors in crowdscenes using stability analysis for dynamical systems," IEEE Trans.Pattern Anal. Mach. Intell., vol. 34, no. 10, pp. 2064–2070, Oct. 2012.
[7]. S. Kwak, B. Han, and J. H. Han, "On-line video event detection byconstraint flow," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 6,pp. 1174–1186, Jun. 2014. [8] W. Li, R. Zhao, T. Xiao, and X. Wang, "DeepReID: Deep filter pairingneural network for person re-identification," in Proc. CVPR, Jun. 2014,pp. 152–159.
[8]. Z. Wu, Y. Li, and R. J. Radke, "Viewpoint invariant human reidentificationin camera networks using pose priors and subjectdiscriminativefeatures," IEEE Trans.

Pattern Anal. Mach. Intell., vol. 37,no. 5, pp. 1095–1108, May 2015.

[9]. G. Lisanti, I. Masi, A. D. Bagdanov, and A. Del Bimbo, “Personre-identification by iterative re-weighted sparse ranking,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 37, no. 8, pp. 1629–1642, Aug. 2014.