Improving Person Re-Identification Rate in Security Cameras by Orthogonal Moments and a Distance-based Criterion

Ali Dadkhah[1], Saeed Nasri[2,3*]
1-Department of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: a.dadkhah@gmail.com
2- Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: s_nasri@iaun.ac.ir (corresponding author)
3- Member of Research Center for Digital Computing and Machine Vision, Najafabad Branch, Islamic Azad University, Najafabad, Iran.

Received: January 2020 Revised: May 2020 Accepted: August 2020

ABSTRACT:
Surveillance and security cameras help security forces in public places such as airports, railway stations, universities and office buildings to perform high-level surveillance tasks such as detecting suspicious activity or anticipating undesirable events. Re-Identification (Re-ID) is defined as the process of communicating between images of the person in different cameras in a surveillance environment. Changing the field of view of any camera presents challenges such as changing body posture, changing brightness, noise and blockage. This article focuses on extracting the most distinctive features to overcome these challenges. The features of Hu moment, Zernike moment in 9th order and Legendre moment in 9th order for each image are extracted and merged into a single feature vector to form a single feature vector for each image. Principal Component Analysis (PCA) was used to reduce the vector dimensionality and finally the Mahalanobis distance criterion was used for identification. The proposed method in the VIPeR database has achieved a re-ID rate of 96.5. Although the presented method is simple, the outcome has been superior compared to many of the state-of-the-art methods.

KEYWORDS: Person re-identification, Orthogonal Moments, Mahalanobis Distance, Security Cameras.

1. INTRODUCTION

Person Re-Identification (Re-ID) is defined as a process of communicating a person's images from different cameras in a surveillance environment. The main purpose of these systems is to identify one person in one camera and to re-identify him or her in all other cameras. In other words, it is determining whether the image recorded by different cameras belongs to one person? Fig. 1 shows an example of a surveillance area controlled by multiple cameras with non-overlapping field of view (FOV) [1]. Colored dots with numbers indicate different people. The dotted lines with the arrow indicate the directions that certain people move through the camera.

One person has been photographed from one camera FOV to another camera FOV, Re-ID is used to communicate between interrupted tracks to track multiple cameras. Thus, single-camera tracking with Re-ID on cameras enables one to reconstruct one's path across the big stage.

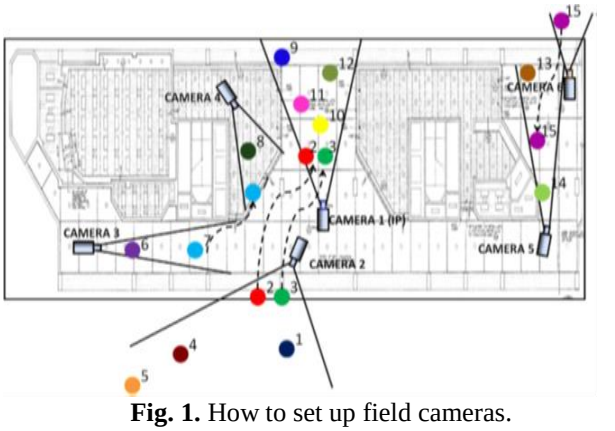

Fig. 1. How to set up field cameras.

Changes in brightness, changes in body position and changes in viewing angles due to changing cameras, is a challenging issue. Also, the images captured by surveillance system cameras are usually in

Paper type: Research paper
DOI: https://doi.org/10.29252/mjee.14.4.85
How to cite this paper: A. Dadkhah, S. Nasri, **"Improving Person Re-Identification Rate in Security Cameras by Orthogonal Moments and a Distance-based Criterion"**, *Majlesi Journal of Electrical Engineering*, Vol.14, No. 4, pp. 85-91, 2020.

low resolution, which makes many visual details such as facial components unrecognizable. Therefore, the appearance of different individuals may be similar, and these factors make the individual's visual appearance easily influenced and identifying a person becomes a difficult and complex issue. In Fig. 2, the first row shows the samples in one camera view and the second row in the other camera. Each part illustrates a challenge. 1) changes in brightness, 2) low resolution images, 3) changes in body posture, 4) identification among pedestrian, 5) changes in viewing angle. In this paper, due to the differentiating power of orthogonal moments, to overcome the above challenges, the combination of Zernike, Legendre, and Hu moments is used, which will be discussed below [4].

Fig. 2. Challenges in Re-ID Systems.

The rest of the article is organized as follows. Section 2 provides an overview of the previous works, Section 3 introduces and calculates the moments, Section 4 explains the proposed algorithm, and Section 5 presents the performance evaluation of the proposed method.

2. PREVIOUS WORKS

Recent efforts have focused on two aspects of the solution: 1) designing discriminative and robust visual descriptors to describe one's appearance; and 2) learning proper distance criteria that maximize the likelihood of a correct analogy [3].

Histogram is a good feature to represent an image. The reason for the widespread use of this tool is the speed of processes done on the histogram, such as computing the histogram, comparing two histograms. One of the main problems with the standard histogram is the lack of location information in the image. Authors in [11] have developed a new definition of color histogram, which is called probabilistic color histogram. The way it works is that for all the colors available, the usual colors that exist in different cultures are chosen as the primary colors. Then, using the tagged dataset of different images containing color in different environmental conditions, using the k clustering algorithm, the nearest fuzzy neighbor of the primary LAB color space is divided in a fuzzy way into 11 colors. As a result, by definition, a pixel can change more than one stack of histograms depending on its color. The main idea of the research [12] is to use the five properties of x, B, C, R and y in the histogram. So that a 5-dimensional histogram is created based on the five properties and then, for each pixel, according to the position and color, the degree of amplification is assigned to one of these stacks. This way, as in the previous method, a pixel will affect more than one stack of histograms.

In the research [5], a hybrid method was used to provide a description of multiple images simultaneously for humans. The main part of this algorithm consists of two parts: Histogram Plus Epitome (HTE). In the first step, the foreground areas are separated by the algorithm presented in [6]. The first part describes the histogram on the HSV (Hue-Saturation-Value) color model. The second part of the descriptor uses epitome analysis. Epitome analysis provides a collection of multiple images in the form of a reflective model, which incorporates texture, shape and appearance properties. This analysis is applied once generally and once locally, to clustered images.

Research [8] performs Re-identification in a multipurpose system. Therefore, we have a sequence of images for one person. For each of these images, the covariance descriptor is calculated. Then a graph is formed whose vertices, covariance matrices, and edges are the matrix distances mentioned in the previous section. The spectral clustering is then performed on the graph and, as a result, the mean k is calculated for each sequence and finally, at the re-identification stage, these k values are compared and the similarity is determined. In the study [9], the image is split into three HSV color channels and then is applied to each color channel with the use of Gabor filter in 8 bands and in each band in two different sizes. In this study, to speed up the computation, instead of calculating the covariance matrix for the whole image, by dividing the image into smaller areas and computing the matrix for each area, it has increased the speed of operation.

In research [10], authors use the Gabor filter and the LBP filter for the feature vector and also divide the person body area into four parts and calculates the covariance descriptor on the basis of these features on each area. In the research [7], special representations for humans have been used. In this representation, the body is divided into three parts: head, trunk and legs, using an innovative method. By slightly changing the same method, it is used to extract the vertical axis in each segment for which the symmetry assumption is true. Three features are extracted for each of these areas. The first feature is the weighted histogram in the HSV color model. Its way of weighing is that the more pixels in a person's image are distant from the symmetry vertical axis, the less weight it will get on the

histogram. The reason for this weighting lies in the fact that the longer we move away from the vertical axis, the greater the likelihood of pixels belonging to the background. The second feature used is the MSCR (Maximally Stable Colour Regions) descriptor. The MSCR descriptor actually performs clustering on the pixels of each region of the person image. These areas are selected during clustering, so that they are more distant from each other in color. Each cluster is represented by several dimensions, including area, center of the cluster, RGB color mean, and second impulse matrix. In multi-image mode, another clustering is performed on these areas for several images. The third used feature is the Recurrent High-Structured Patches (RHSP). The idea of extracting such areas is derived from the epitome analysis. In fact, in this method we are looking for areas with high entropy values and at the same time to become more visible under different transformations. In other words, areas that are more visible in the image under transformations such as rotation of the image are also resistant to human angular changes.

Studies [13], [15] try to extract areas of importance for a person. The method is that by using the concept of dense adaptation, with the help of SIFT (Scale-invariant feature transform) descriptor, by first comparing the image with a random set of human images, it identifies patches that are not seen in more than half of the persons as the important areas. Finally, in a hybrid and weighted relationship, these areas are more weighted and show that the detection rate has improved with this method. In another study [14], somehow the distance between samples in the learning process has been used. In this study, authors attempted to use untagged data for re-identification. Since sets of similar people are placed on a procedure, the closest person in the extracted existing data set can be obtained using the procedure ranking algorithms that propagate the label using the graph and neighborhood data. In fact, untagged data fills the procedural space and as a result, they provide us with the advantage of publishing labels through them.

3. THE PROPOSED ALGORITHM

The proposed method in this paper is based on a set of stable moment descriptors, principal component analysis, and Mahalanobis distance criterion classification. First, the Zernike, Legendre and Hu moment features of the training phase images (Camera A) are computed separately and their results are merged. Then a normalization and dimensionality reduction step is performed by PCA on the integrated feature vector and the data are trained by the Mahalanobis cluster. In the test phase, the test image (Camera B) goes through all the steps of the algorithm and in the last step, re-identification is performed by the trained cluster classifier.

3.1. Calculation of Feature Vectors by Moments

Moments are one of the most important and powerful regional feature extractors in the field of image processing. The moments are obtained by internal multiplication of the image function in a polynomial. Internal multiplication is a very important criterion for detecting similarity. The moments are divided into two categories of orthogonal and non-orthogonal. Non-orthogonal moments include central and geometrical moments that, if cleverly refined, will produce orthogonal moments that are resistant to many degradations such as torsion, blurring and scaling [16]. The most important orthogonal moments are Hu, Zernicke and Legendre.

3.1.1. Hu moment

The first orthogonal moment was introduced in 1962 by Hu using algebraic constants. This moment can be calculated up to a maximum of 3 order. Eq. (1) represents the geometrical moment in which the sum of p and q determine the order of the moment. f(x,y) is the image function, $x^p x^q$ defines the geometric moment function and N*M is the image size [16].

$$M_{pq=} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p\, x^q f(x,y) \quad (1)$$

(p + q) moment order

Zero order M_{00}
First order M_{10} and M_{01}
Second order M_{20} and M_{02} and M_{11}
Third order M_{30} and M_{03} and M_{12} and M_{21}
Increasing Order = Extracting Details

From Eq. (1) and (2) we can obtain a constant moment proportion to the transition (3).

The center of gravity of the image

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad , \quad \bar{y} = \frac{M_{01}}{M_{00}} \quad (2)$$

Constant moment proportion to transition

$$\mu_{pq=} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x-\bar{x})^p\, (y-\bar{y})^q f(x,y) \quad (3)$$

So the central moment μ_{pq} is constant over the transmission. Using normalization of μ_{pq} by equations (4-3), constant moment to scale change (Hu moment) can also be achieved:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}{}^{\gamma}} \qquad \gamma = \frac{p+q}{2} + 1 \quad (4)$$

Due to the normalization performed on the central moment; Hu moment, which is constant to transmission, scale change and rotation, can be achieved.

In the Hu moment, the length of the feature vector obtained for each image is 7. It should be noted that Hu moment is only calculated up to the 3rd order.

3.1.2. Zernike Moment

Zernike moments are orthogonal moments whose nucleus consists of Zernike polynomials, defined as polar coordinates and the circumference of a single circle. Due to the polarity of the workspace, these moments are independent of rotation. The Zernike moments of n order and the repetition of 1 for the image of $f(r,\varphi)$ are defined as the eq. (5). [16].

$$Z_{pq} = \frac{p+1}{\pi}\int_0^{2\pi}\int_0^1 V_{pq}^*\ f(r.\theta) r\ dr d\theta \qquad \begin{matrix} n \\ = 0,1,2,\dots \ell \\ = -n,-n+2,\dots . n. \end{matrix} \tag{5}$$

Where,

$$V_{pq}(r,\theta) = R_{pq}(r)e^{-jq\theta} \qquad ,j=\sqrt{-1} \tag{6}$$

And the value of radial polynomials, $R_{pq}(r)$, is calculated as (7):

$$R_{pq}(r) = \sum_{k=0}^{(n-|q|)/2} (-1)^k \frac{(n-s)!}{k!\left((n+|\ell|/2-k)!\right)\left((p-|q|/2-k)!\right)} r^{p-2k} \tag{7}$$

If the image size is N*N, Eq. (8) is approximated as follows:

$$Z_{pq} = \frac{4(p+1)}{(N+1)^2\pi}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} R_{pq} r' e^{-jq\theta'} \tag{8}$$

In order to map the image to the unit circle, the new values of x and y are calculated according to Eq. (9):

$$\begin{matrix} x' = \frac{\sqrt{2}}{n-1}X - \frac{1}{\sqrt{2}} \\ y' = \frac{\sqrt{2}}{n-1}y - \frac{1}{\sqrt{2}} \\ r' = \sqrt[2]{x'^k + y'^k} \\ \theta' = tan^{-1}\frac{y'}{x'} \end{matrix} \tag{9}$$

In the Zernike moment, the moment of the feature vector varies depending on the order of this moment. Given that the 9th order is the most optimal order for the feature vector length of each image.

3.1.3. Legendre moment

These moments are orthogonal moments that their core consists of Legendre polynomials. Legendre moment is defined by the (p+q) order in the interval of [-1 1] as Eq. (10), [16]:

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4}\int_{-1}^{1}\int_{-1}^{1} p_m(x)p_n(y)f(x.y)dxdy \quad m.n = 0,1,2,\dots . \tag{10}$$

Where, Legendre's polynomial function is of order P. Legendre moment in discrete mode is changed as Eq. (11):

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} p_m(x')p_n(y')f(x.y) \tag{11}$$

Given by mapping the image from N*N to 2*2 square with the coordinate center, the values of x' and y' from X and Y are calculated as follows:

$$x' = \frac{2}{N-1}X - 1 \qquad y' = \frac{2}{N-1}y - 1 \tag{12}$$

Legendre polynomials, $p_P(x)$, are a complete orthogonal set in finite of [-1 1] such that the nth legendary sentence is calculated by (13):

$$(x) = P_q = \frac{1}{2^q}\sum_{p=0}^{q/2}(-1)^q \frac{(2q-2p)!}{p!(q-p)!(q-2p)!}x^{q-2p} \tag{13}$$

Also in the Legendre moment, the moment of the feature vector varies depending on the order of this moment. Given that, the 9th order is the most optimal order for the feature vector length of each image.

3.1.4. Integration of Extracted Features

The Hu, Zernike and Legendary moment are applied to each image. The feature vector lengths are extracted from Hu 7 (order of 3), Zernike 190 (order of 9) and Legendre 100 (order of 9). Next, these 3 vectors are merged into a vector with the of length 297.

3.2. Normalization and Dimension Reduction

After extracting the feature, since some data are large and these numbers make classification difficult, we perform a normalization step on the extracted feature vector. After normalization, using the principal component analysis algorithm, a dimensional reduction step on the generated feature vectors is performed to

reduce the feature vectors length and reduce the computational and temporal complexity in the classification stage. Another important point to note about principal component analysis is that it eliminates the correlation between the values of the feature vectors, which improves the detection rate. Finally, the feature vector length is reduced from 297 to 100.

3.3. Classification

After reducing the dimensions of the feature vectors, these vectors are stored in a database called the feature database. The purpose of creating this feature database is to estimate the covariance matrix in relation to Mahalanobis or, in other words, the Mahalanobis Classification Training. The test image then goes through all the feature extraction and dimensionality reduction steps and finally, the feature vector distance generated from the test image is trained by the classifier and computed with all database images and the results are sorted, and the result is determined in different ranks according to the evaluation criteria.

Mahalanobis Distance is one of the most important learning methods, which is used to improve classification results and is shown as Eq. (14):

$$d_M(x_I, x_j)^2 = (x_i - x_j)^T M(x_i - x_j) \quad (14)$$

In this distance criterion, M is a semi-definite positive symmetric matrix. The problem that exists in practice in Eq. (14) is that the M matrix is unknown and must be estimated using the train dataset.

Given that the M matrix must be a positive semi-definite symmetric one, it can be subdivided into $M = LL^T$ which represents $T \in \mathbb{R}^{d*r}$ and r denotes the rank of the M matrix. So according to Eq. (15), it can be written:

$$\begin{aligned} d_M(x_i, x_j)^2 &= (x_i - x_j)^T M(x_i - x_j) \\ &= (x_i - x_j)^T L^T L(x_i - x_j) = \\ &\| L^T(x_i - x_j) \|^2 \end{aligned} \quad (15)$$

And through the Eq. (15), the M matrix or L matrix is directly estimated. It can be said that Mahalanobis distance learning finds a linear mapping. After mapping the data to the new space, Euclidean distance is used to compare the data.

3.4. Evaluation Criteria

The common performance criterion for evaluating a person re-identification is performed by examining the Cumulative Match Characteristic (CMC). The Cumulative Match Characteristic indicates the likelihood of identifying one person within the database, extracted from several persons, and is a measure to evaluate the performance of the biometric identification system. This curve reflects the results of the identification task, by plotting the probability of correct identifying (the y axis) versus the number of candidates received (the x axis) [17]. It is also called the Rank-k detection rate, which indicates how much the correct subject format is occurred in Rank-k and is defined Eq. (16).

$$\text{CMC} = \frac{\text{The number of correct matches that occur on matches above K}}{\text{Total number of test matches performed}} \times 100$$

K= Value of rank (16)

4. THE IMPLEMENTATION RESULTS

In person re-identification systems, images of existing databases are cropped and tagged. No pre-processing is done on images. In each database, 50% of the images are used as camera A (gallery) images for the training process and 50% of the images are used as camera B (probes) for the test process. In these experiments, the order of the Zernike and Legendre moments is considered 9 as the best. To overcome random effects, experiments were repeated 10 times and then average performance was reported. To make it easier to compare with other methods, the accuracy of Cumulative Match Characteristic (CMC) has been reported in the VIPeR dataset at Rankings 1, 5, 10 and 20 and the entire CMC curve has not been reported.

VIPeR dataset: The VIPeR dataset consists of two cameras and 632 pairs of images, each image is (in 128*48 pixels) for each camera display. Images are manually tagged. This dataset is taken outdoors, and therefore has complex lighting changes. With the exception of the light changes, each camera has a different perspective, which is usually more than 90 degrees between the two cameras field of view. In Fig. 3, the first row of images of each person is in the first camera and the second row of the same person image is in the second camera [2].

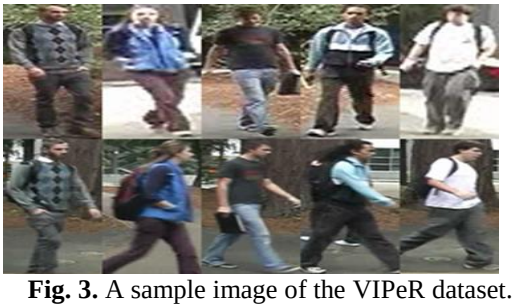

Fig. 3. A sample image of the VIPeR dataset.

4.1. Evaluation of the Proposed Method by Zernike and Legendre Moment Order Change

One of the parameters that has a significant effect on output is the order parameter in Zernike and

Legendre moments. In Table 1, the changes of this parameter and the accuracy of identification on the VIPeR database are examined.

Table 1. Changes of the moment order and its results on the VIPeR database in Rank 20.

Legendre-Zernike- Hu	Zernike	Legendre	Order
44.43	32.68	22.93	۲
81.50	60.75	55.60	۳
85.65	73.31	60.09	۴
92.54	76.64	64.92	۵
93.53	77.50	65.61	۶
94.41	80.64	66.24	۷
95.90	82.20	66.58	۸
96.50	82.20	67.00	۹

The results in Table 2 show that with the change of moment order, the results are improved but the computational cost increases dramatically. Also the result shows that the Zernike moments extract more distinctive features than other moments for image re-identification, which still holds the 9th order.

4.2. Evaluation of the Proposed Method (Moment Integration)

As can be seen from the results in Table 2, the merging of features derived from the three moments has yielded much better results in all ranks. It can also be deduced from these features that the Zernike moments extract more distinctive features than the other moments for image re-identification. Legendre moment and Hu moment are next, respectively.

Table 2. Re-identification results on the VIPeR Database.

Rank 20	Rank 10	Rank 5	Rank 1	Type of moment
54.70	46.10	32.20	8.50	**Hu moments**
67.00	60.50	40.70	20.70	**Legendre moments**
82.20	76.70	61.60	32.00	**Zernike moments**
81.50	74.60	60.00	31.50	**Hu-Legendre moments**
83.90	77.80	62.00	36.80	**Hu-Zernike moments**
92.10	85.30	69.80	42.10	**Legendre-Zernike moments**
96.50	89.10	76.90	46.60	**Integrating three moments**

4.3. Examining the Proposed Method and Comparing It with Other Methods

Table 3 summarizes the results of comparing person re-identification rates with those of other works.

Table 3. Comparison of the proposed method with previous works.

Method	Rank 1	Rank 5	Rank 10	Rank 20
PCCA	19.30	48.90	64.90	80.30
KISSME	19.60	48.00	62.20	77.00
LADF	29.30	61.00	76.00	86.20
SDALF	19.90	38.90	49.40	65.70
LOMO+XQDA	40.00	68.10	80.50	91.310
MetricEnsemble	45.90	77.50	88.90	95.80
Proposed method	46.60	76.90	89.10	96.50

5. CONCLUSION

Person Re-identification is an area of pattern recognition science that has received much attention in recent decades. It is important to note the challenges that have been identified for these studies. These challenges include re-identification in camera angles change, changing body postures, changing light, and more. Therefore, this paper aims to provide a way to respond to the challenges mentioned by integrating Zernike, Legendre, and Hu moments. Due to the differentiating power of these moments on the VIPeR database, acceptable results are obtained with the Cumulative Match Characteristic (CMC). In these results, the superiority of the Zernike moment over the other two moments is important.

REFERENCES

[1] D. Yi. Z. Lei. S. Liao. and S. Z. Li. "**Deep metric learning for person re-identification.**" in *2014 22nd International Conference on Pattern Recognition*. IEEE. pp. 34-39., 2014.

[2] C. Sun. D. Wang. and H. Lu. "**Person re-identification via distance metric learning with latent variables**." *IEEE Transactions on Image Processing.* Vol. 26, No. 1, pp. 23-34, 2017.

[3] A. Bedagkar-Gala and S. K. Shah. "**A survey of approaches and trends in person re-identification**." *Image and Vision Computing.* Vol. 32, No. 4, pp. 270-286. 2014.

[4] R. Vezzani. D. Baltieri. and R. Cucchiara. "**People reidentification in surveillance and forensics: A survey.**" *ACM Computing Surveys (CSUR).* Vol. 46, No. 2, pp. 29, 2013.

[5] L. Bazzani. M. Cristani. A. Perina. and V. Murino. "**Multiple-shot person re-identification by chromatic and epitomic analyses**." *Pattern Recognition Letters.* Vol. 33, No. 7, pp. 898-903, 2012.

[6] N. Jojic. A. Perina. M. Cristani. V. Murino. and B. Frey. "**Stel component analysis: Modeling spatial correlations in image class structure**." in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. pp. 2044-205, 2009.

[7] L. Bazzani. M. Cristani. and V. Murino. "**Symmetry-driven accumulation of local features for human characterization and re-identification**." *Computer Vision and Image Understanding.* Vol. 117, No. 2, pp. 130-144, 2013.

[8] J. Metzler. "**Appearance-based re-identification of humans in low-resolution videos using means of covariance descriptors**." in *2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*. IEEE. pp. 191-196, 2012.

[9] B. Ma. Y. Su. and F. Jurie. "**Bicov: a novel image representation for person re-identification and face verification**." in *British Machive Vision Conference*. pp. 11, 2012.

[10] Y. Zhang and S. Li." **Gabor-LBP based region covariance descriptor for person re-identification**." in *2011 Sixth International Conference on Image and Graphics*. IEEE. pp. 368-371, 2011.

[11] A. D'Angelo and J.-L. Dugelay. "**People re-identification in camera networks based on probabilistic color histograms**." in *Visual Information Processing and Communication II*. vol. 7882: International Society for Optics and Photonics. pp. 78820K, 2011.

[12] Z. J. Xiang. Q. Chen. and Y. Liu. "**Person re-identification by fuzzy space color histogram**." *Multimedia tools and applications.* Vol. 73, No. 1, pp. 91-107, 2014.

[13] R. Zhao. W. Ouyang. and X. Wang. "**Person re-identification by salience matching**." in *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2528-2535, 2013.

[14] C. C. Loy. C. Liu. and S. Gong. "**Person re-identification by manifold ranking**." in *2013 IEEE International Conference on Image Processing*. IEEE. pp. 3567-3571, 2013.

[15] R. Zhao. W. Ouyang. and X. Wang. "**Unsupervised salience learning for person re-identification**." in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3586-3593, 2013.

[16] J. Flusser. B. Zitova. and T. Suk. "**Moments and moment invariants in pattern recognition**." John Wiley & Sons. 2009.

[17] S. Paisitkriangkrai, Ch. Shen, and Anton van den Hengel. "**Learning to rank in person re-identi_cation** " with metric ensembles." CoRR, abs/1503.01543, 2015.