



People Re-identification in Non-overlapping Field-of-views using Cumulative Brightness Transform Function and Body Segments in Different Color Spaces

A. G. Sorkhi^a, H. Hassanpour^{a*}, P. L. Mazzeo^b

^aFaculty of Computer and Information Technology, Shahrood University of Technology, Shahrood, Iran

^bInstitute of Applied Sciences and Intelligent System - CNR. Via della Libertà 3, 73010 Arnesano (Lecce), Italy

PAPER INFO

Paper history:

Received 23 May 2015

Received in revised form 10 October 2015

Accepted 24 December 2015

Keywords:

People Tracking

Camera Networks

Non-overlapping Field-of-view

Various Segments of Human Body

Cumulative Brightness Transform Function

ABSTRACT

Non-overlapping field-of-view (FOV) cameras are used in surveillance system to cover a wider area. Tracking in such systems is generally performed in two distinct steps. In the first step, people are identified and tracked in the FOV of a single camera. In the second step, re-identification of the people is carried out to track them in the whole area under surveillance. Various conventional features such as clothes and appearance of person have been used to identify peoples in the cameras. However, similarity between appearance and clothes of people causes unreliable results. Thus, much reliable features are still required to increase the ability of tracking system. The aim of this paper is to propose a novel algorithm to identify people in a network of cameras with disjoint views. In our proposed methodology, according to relative size of various parts, human body are divided into three portions of head, middle section and lower section. Considering histograms of these portions, matching process was done aiming to distinguish persons in FOV of cameras. The other main challenge in such multi-camera system is to deal with illumination changes and varying appearance of people at different camera's view. The cumulative brightness transform function (CBTF) was employed to alleviate the difficulty of illumination changes. Experiments were conducted in RGB, YCbCr and HSV color spaces and it was found that YCbCr color space gives a better performance result compared to other color spaces used in this work.

doi: 10.5829/idosi.ije.2015.28.12c.03

1. INTRODUCTION

The main purpose of a tracking system is to continuously detect exact location of an object as well as its trajectory once the object or camera moves. Tracking systems are used in various tasks including robotics, navigation, traffic surveillance and abnormal behavior detection. In this regard, different tracking systems have been developed in the past decades. Human tracking has dominated tracking systems for security or surveillance purpose. According to the studies conducted in this field, people tracking are a very complicated and difficult task as various features are required to be considered.

In video processing, due to 3D world projection into 2D images, information is often lost [1]. The loss of

information may be due to low camera resolution, complex nature of object motion and its geometry, complete or partial occlusion of an object with others. The other challenging issues for people tracking in multi-camera systems are scene illumination changes and various appearances of people at different camera views [2-4].

Object histogram matching techniques such as color histogram matching has been used as a signature in many published works [5]. These approaches have also been used for re-identifying the objects in a network of cameras without disjoint views [6, 7]. Porikli et al. [8] probability based Bayesian Belief Network was used to improve performance of the histogram-based approach for objects recognition and/or re-identification. Kettner et al. [9] Bayesian method was used for human tracking in cameras with disjoint view. In this paper, motion of people was assumed to be uniform, hence the system may fail once the object stops moving.

*Corresponding Author's Email: h.hassanpour@shahroodut.ac.ir (H. Hassanpour)

Histogram matching is a simple and a fast technique for tracking people but it may not be accurate in a network of cameras with disjoint views because of possible changes in people's appearance and scene illumination. In this paper, a methodology based on histograms of various parts of human body is proposed for a better people re-identification by increasing the accuracy of the histogram matching technique.

This paper also addresses the problem of dealing with illumination changes for appearance of people at different camera sites. Figure 1 shows the image of a person, along with its histogram observed in two views of different cameras under different poses in surveillance network. This figure shows that illumination differences of an object in different cameras in a camera network makes recognition or re-identification of people difficult. We use cumulative brightness transform function (CBTF) to alleviate this difficulty for people tracking in disjoint cameras.

2. RELATED WORKS

Various researches have been conducted on different aspects in human tracking. Moeslund et al. [10] almost 400 papers have been reviewed and discussed on objects tracking problems as well as rectification approaches such as human model initialization, tracking, position estimation and activity recognition. Yilmaz et al. [11] details of available methods for objects tracking have been introduced and summarized. In this study, they have categorized the tracking algorithms on the basis of the object shape, appearance, color and their motion representations. Enzweiler et al. [12] several object tracking systems has been reviewed and compared. Khakpour et al. [13] proposed a new method for kernel-based object tracking. In this paper, a definition for union image blob is defined and then mapped to a new representation which is called as potential pixels matrix.

Geronimo et al. [14] six steps have been considered for people tracking systems including pre-processing, foreground segmentation, object classification, verification, refinement and tracking. In this study, various tracking algorithms have been evaluated and discussed in more detail. These methods use several overlapping cameras for tracking purpose and hence, they could not be used to cover an extensive area. As a matter of fact, cameras with disjoint views (i.e. non-overlapping cameras) must be used to cover a large area. Human tracking with non-overlapping cameras have also been subject of many studies. Javed et al. [15] spatial parameters and object motion models were used for tracking tasks. In this paper, parameters were combined with people's appearance models to improve the object re-identification performance.



Figure 1. Same person observed in two views of different cameras under different poses in a surveillance network.

In addition, motion features like position, velocity and acceleration of people were used to re-identify objects in a network of cameras with disjoint views. For instance, the authors in literatures [16, 17] used the information gained from observing location and velocity of object moving across multiple non-overlapping cameras to determine spatial relationships between cameras. Rahimi et al.[18], it has been assumed that there is a possible relationship between objects in these cameras, and a Markov model was used for modeling people's dynamic. Javed et al. [19], spatial-temporal information as well as color and illumination transfer functions were used to establish relationships between people. These methods would be useful where distances between cameras are short and people have a consistent movement style.

Detecting and tracking active objects are two main issues in surveillance systems. Mahdavi et al. [20], a block-based similarity algorithm has been proposed to detect and track active objects in successive frames in different cameras. In this approach, the authors defined a similarity and cost functions based blocks' features in order to reduce the computational complexity of their algorithm. The results obtained through experimental investigations indicated that the developed approach is suitable for real time monitoring and gives a good performance for active objects in noisy environments.

A novel non-training approach was introduced by Xie et al. [21] where visual-spatial saliency is used to get query image. In this approach, images of segmented pedestrian are partitioned into small regions, so as hyper graphs which are employed to represent visual and spatial relationship among regions. The visual-spatial saliency is then formulated as a hyper graph via considering into human body and similarity of pedestrians. The visual-spatial saliency is ultimately

used in region-based matching to enhance the re-identification process.

Person re-identification is a process in which people across non-overlapping cameras are matched. Recent findings have revealed that metric learning is an effective approach for re-identification purpose. However, most of learning method metrics are not able to work properly with small sample size due to the limited number of training samples. Liong et al. [22], a new discriminative regularized metric learning (DRML) method was proposed for person re-identification. In this approach, discriminative information of training samples are extracted and used to regulate the eigenvalues of covariance matrices, which results to a better estimation of distance metric. Experimental results obtained from three widely used datasets indicated that this approach is so effective when is used for person re-identification.

Person re-identification methods are generally looking for robust person matching through combining different features. These features are often assigned with certain weights by assuming that all of them are equally good for all of the individuals included in re-identification process. Liu et al. [23] showed that only few of these features are important and the rest may just increase the computational time. To identify these certain features, they proposed a new unsupervised approach in which extracted features are weighted adaptively by their alien and appearance attributes.

Person re-identification across network of cameras with limited or non-overlapping fields of view is still a challenging subject. An investigation [24] evaluated performance of various re-identification techniques; including regularized Pairwise Constrained Component Analysis, kernel Local Fisher Discriminant Analysis, Marginal Fisher Analysis using both histogram-based and kernel-based features. Authors in this research found a significant improvements in identification, in terms of Cumulative Match Characteristic curves (CMC) and Proportion of Uncertainty Removed (PUR) scores.

3. THE PROPOSED APPROACH

Object re-identification is defined as observing an object in one camera and reidentifying it in the same or different cameras a little while after. An object may appear in a camera more than once. Re-identification of an object in a camera as well as obtaining information about the time and number of entrance would be very helpful in object attitude analysis and identification. Motion features (estimated positions and velocity [2]) would not be suitable to track objects in a network of cameras with disjoint view as there are too many entrance and exits point in such environments and motion feature depends mostly on the object position.

This means that objects may exit from a certain point and enter in the same view from a different point. Different entrances and exits points are shown in Figure 2. It is not rational to use spatial-temporal approaches in such environments as well since time and position of objects vary once they enter or exit. Similarly geometrical attributes would not be a good choice for re-identification due to low resolution images and lot of resizing processes performed on the images [2].

TABLE 1. Relative length of different human body parts

Segment	Relative length
Height	8
Forearm	2
Upper arm	1.5
Neck and head	1.25
Torso	2.5
Shoulder girdle	2
Pelvic girdle	2
Upper leg	2
Lower leg	2
Foot	1

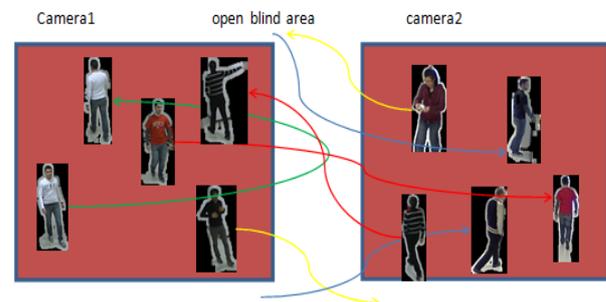


Figure 2. Depicting the inter-camera association between two non-overlapping cameras. Given tracked targets in each camera, our goal is to find the optimal correspondence between them, such that the associated pairs belong to the same object

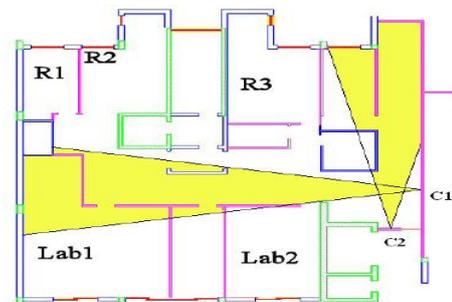


Figure 3. Generic topology of non-overlapping FOVs multi-camera surveillance system [25].

Figure 3 shows a general scheme of cameras with disjoint view. In this Figure two cameras are set in a building for surveillance.

The aim of this paper is to propose an approach which can improve the performance of re-identification to a significant level. In this approach, first the problem of changes in illumination in appearance of people at different camera views is resolved by using CBTF. Then, the histogram matching accuracy is improved via partitioning the image of human body into three parts. Histograms of each parts are compared with those from the previously observed people. The proposed method is described below in more details.

3. 1. Various Parts of Human Body Human body is a set of connected parts. To segment these parts we need to have their corresponding sizes which are measured relative to total height. In this study, relative distances, according to clinical and body proportions studies as well as HMA databases [26], are considered for each part of the body as they are frequently used in human motion and human body proportions analysis. Generally, each body can be considered to be eight times of the head and hence each part gives a size value relative to the total height of the body. Table 1 gives these values obtained for each part of the body. For the purpose of this study, body was partitioned into three portions including head, middle section and lower section. Figure 4 gives a representation showing how various parts of body can be portioned into different segments.

3. 2. Similarity Measure Various measures have been proposed to calculate the possible similarities of different bodies. It was shown that selecting a sophisticated measure can significantly improve the performance of a tracking system. In this study Bhattacharyya distance was used for tracking people in a network of cameras with disjoint views as it is one of best measure for matching purpose [25, 27]. Bhattacharyya measure is expressed as:

$$S(O_{i,k}, O_{j,l}) = 1 - D_B(O_{i,k}, O_{j,l}) \tag{1}$$

where $O_{i,k}$ gives persons observed in camera number i and frame number k while $O_{j,l}$ denotes persons observed in camera number j and frame number l. If H_i and H_j are considered to be the histograms of these people, parameter D_B is defined as [28]:

$$D_B(H_i, H_j) = \sqrt{1 - \sum_{v=1}^m \sqrt{H_i(v) \cdot H_j(v)}} \tag{2}$$

In the above equation, m denotes the number of bins in the histogram. Bhattacharyya distance gives a value between 0 and 1. The smaller the value, the lower the similarity.

3. 3. Re-identification Process

In re-identification process, an arrived person in a camera view is compared with those persons already identified by the system. The following equation is used to measure the similarity between their features for possible re-identification:

$$\arg \max_j \text{similarity}(\text{Similarity}(H_i, H_j)) \tag{3}$$

In the above equation, H_i corresponds to the histogram of the new person arrived in the frame of the camera (FOV) whereas H_j corresponds to the histograms of the persons already identified by the tracking system. In the following, we use the probability of the observations belonging to the same object as [25]:

$$P_{12}(O_{1,u}(a), O_{2,l}(a) | \lambda_{2,l}^{1,u}) = \prod_{ch \in \{R,G,B\}} \gamma^{D(O_{1,u}^{ch}, O_{2,l}^{ch})} \tag{4}$$

where γ is an arbitrary constant and D is a distance between an object appearance in C2 and C1. With $\lambda_{2,l}^{1,u}$ we consider a correspondence among the two consecutive observations, $O_{1,u}$ and $O_{2,l}$, i.e., exiting from one camera and entering into another one ($O_{1,u}$ shows object O in camera 1 and frame u). using the above equation:

$$\prod_{ch \in \{R,G,B\}} \gamma^{D(H_i^{ch}, H_k^{ch})} > \text{Threshold} \tag{5}$$

In this equation, D represents the discrepancy between the corresponding histograms of two persons; k is the person who has the maximum similarity with the new person. If the similarity measure from the color channels (with Equation (5)) exceeds the threshold, it means that the person has already been observed in the previous frames (re-identified); otherwise it is considered as a new person. In this paper, to compare the corresponding histogram of various parts of human body, Equation (6) was taken into consideration.

$$\left(\begin{array}{l} \prod_{ch \in \{R,G,B\}} \gamma^{D(H_i^{ch}, H_{k1}^{ch})} \\ \& \\ \prod_{ch \in \{R,G,B\}} \gamma^{D(H_i^{ch}, H_{k2}^{ch})} \end{array} \right) > \text{Threshold} \tag{6}$$

In the above equation, k1 and k2 represent the middle section and lower section of body, respectively.

3. 4. Cumulative Brightness Transform Function

In this paper, the cumulative brightness transform function (CBTF) introduced in [2] is used to reduce the associated effect of variation in severity of brightness while persons pass through the cameras.



Figure 4. Model and human body relative ratio.

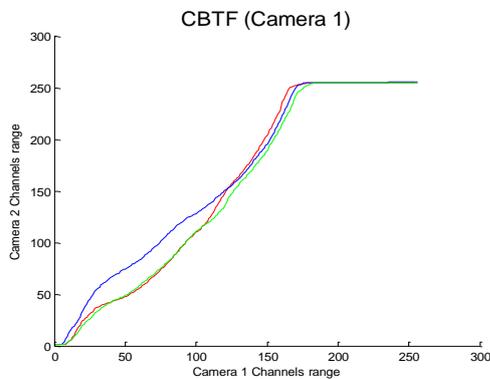


Figure 5. The transfer functions for the RGB color channels from camera C1 to camera C2 obtained using Cumulative Brightness Transfer Transform.

In the training step, we need to get information of the same person going through two cameras with disjoint views. Three RGB channels histograms are computed corresponding to each image in the two cameras, C1 and C2. All the 256 bins of each color channel are considered in computing the histograms. An accumulation of the brightness value training set is computed rather than calculation of a brightness transform function (BTF) for each training pair.

The cumulative histogram \hat{H}_1 of N training samples in camera C1 can be computed from the brightness values $B_1, \dots, B_m, \dots, B_M$ as:

$$\hat{H}(B_m) = \sum_{k=1}^m \sum_{i=1}^N I1(B_k) \tag{7}$$

where $I1(B_k)$ is the count of brightness value B_k in O_L . Note that this cumulative histogram must be normalized by the total number of pixels in the training set to alleviate the effect of size difference between views. The same is done for all the corresponding training images of camera C2 obtaining \hat{H}_2 . The CBTF equation is:

$$f_{i,j} = \hat{H}_j^{-1}(\hat{H}_i(B_m)) \tag{8}$$

Figure 5 shows the CBTF obtained based on the training set. It should be noted that mostly lower color values from camera C2 are being mapped to higher color values in camera C1, indicating that the same object is appearing much brighter in C1 as compared to C2. Figure 6 shows effect of CBTF on the same person in different cameras.

4. EXPERIMENTAL RESULT

In this section, performance of the proposed algorithm for object re-identification in non-overlapping camera networks will be briefly discussed. Algorithm 1 represents the sequences of tracking system. The first step of this algorithm is to calibrate the camera's colors sensitivity during the training time. As it is presented in Figure 7, in the testing step, objects are firstly detected by cameras and CBTF is applied on colors' histogram of the body. In the next step, person's body is partitioned into three portions including head, middle and lower sections. Tracking in this stage is divided into two steps of tracking/re-identification in cache and database. Cache here is the frames previously recorded for the same person, and the database consists of all the persons appeared in any camera at any past time. If the person cannot be found in the cache or databases, the person is considered to be a new person.

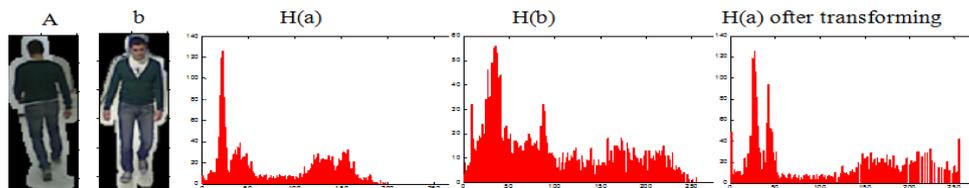


Figure 6. The observations are of the same object from camera C1 and camera C2, respectively (a and b). Their red channel histograms are also shown. The last histogram is obtained after transforming the colors (of first histogram) with a reconstructed CBTF from the subspace.

```

Input: Image sequences from camera C1 and C2
Output: Object re-identification and tracking
If time <= training time then
    Compute the CBTF for camera pairs C1 - C2 using
    algorithm in Section 4.
else
    □ Detect persons in cameras C1 and C2 ;
    □ Correct object's colors in cameras C2 using its
    corresponding CBTF ;
    □ Body's person partitioned into three portions;
    □ Match each objects in cache using equations (3)-(6);
if person is matched then
    Update its position, area, information of color's
    person and frame number in cache;
else
    Person is matched in database using equations (3)-(6);
if person is matched in database then
    Update its position, area, information of color and
    frame
    number in database;
else
    Create a new structure for the person in cache and
    database;
    add object's position, information of color, area
    and frame
    number in cache and database ;

```

Figure 7. Algorithm of Object tracking and re-identification in non-overlapping camera environment



Figure 8. Twelve men and three women extracted from the database.



Figure 9. Same person viewed in different sizes according to his distance to the cameras.

4. 1. Database The dataset used for current study was obtained from the interesting work published in [25]. They have used two non-overlapping cameras set at two different locations in an office as shown in Figure 3. Figure 8 shows twelve men and three women

extracted from the database using our proposed tracking system. The real size of the images were 640×480 which were resized to be 320×240 for reducing the processing time. We used a training set composed of 208 frames (for each view) of appearance of one person in both FOVs. As the test set, we have used 2434 frames coming from fourteen people. As it is shown in Figure 9, the same person might be viewed in different sizes according to his distance to the cameras. In this paper, we used normalized histograms to overlook the sizes of the persons.

TABLE 2. Obtained results using the proposed method in different color space

Methods	Accuracy	F1	RE	PR
Proposed method in RGB	58	81	76	87
Proposed method in HSV	59	80.9	80	82
Proposed method in YCbCr	65	87.6	80	97
Whole body in RGB	34.2	-	-	-
Without CBTF in RGB [25]	32.4	-	-	-
Without CBTF in HSV [25]	34.4	-	-	-
Applying CBTF in RGB[25]	54.16	-	-	-
Applying CBTF in HSV [25]	23.43	-	-	-

4. 2. Evaluation Measure Quantitative measures are rarely used to assess the efficiency of a tracking system in cameras with disjoint views. Accuracy is one of the quantitative measures introduced in several published works. This measure is used to determine the number of correctly detected objects per entire objects passed through a given scene [8, 19]. Several other criteria have also been introduced in literature. Precision (PR) and Recall (RE) are two of these criteria found to be reliable in image processing tasks. PR and RE of object re-identification for each algorithm is calculated by using equations

$$PR = \frac{TP}{TP + FP} \quad (9)$$

$$RE = \frac{TP}{TP + FN} \quad (10)$$

The three quantities TP, FP and FN being defined as follow:

True positive (TP): If an old label is correctly assigned to an object, or a new label is created for a new object.

False positive (FP): If an old label is assigned to a wrong object.

False negative (FN): If a new label is created for an already labeled object.

It is not straightforward to compare the classifiers using the two measures. The F1-measure introduced in literature [29] combines the recall and precision measures with an equal weight in the following form:

$$F\beta(r, p) = \frac{(\beta^2 + 1) \times PR \times RE}{\beta^2 \times (PR + RE)}, \text{ where } \beta = 1 \quad (11)$$

F1-Measure is an automated measurement that determines the precision and recall capabilities of a re-identification system. The F1 metric weights recall and precision equally, and a good re-identification algorithm will maximize both precision and recall simultaneously. Thus moderately good performance on both will be favored over extremely good performance on one and poor performance on the other.

4. 3. Tracking Results The data set was obtained by extracting people from whole FOVs of both cameras. Note that we did not consider any geometrical constraints on the exiting and entering areas of people moving in the observed scenario. We used a training set composed of 208 foreground patches (for each view) associated with the appearance of one person in both cameras. We have used 2434 foreground patches coming from fourteen people in the test set. Table 2 describes the results obtained on the test set detailed as correct matches for persons.

In this paper, experiments have been conducted using the color histogram to track and re-identify people. We separately compared the performance of the whole body color histogram and the performance of different parts' color histogram (proposed method) in some color spaces. Table 2 displays the results obtained from the experiments. Given in Table 3, the proposed method has a reliable efficiency in all of the color spaces used for this study. However, YCbCr color space was found to be the best one for tracking.

The experiments conducted by Mazzeo et al. [25], which has been done on this dataset, considered four persons for their study, but we used fifteen persons for our work.

Because of sufficient performance of CBTF in balanced brightness of different cameras views, people re-identification and tracking accuracy is more tangible than the way without using it. In other words, this method is able to reduce the unaccepted effects of environmental conditions and type of cameras. The other noticeable aspect is sufficient performance in occlusion of same part of person's bodies. In this case, we divide the body into three parts and separately compare each parts with those of previously identified people. Therefore, it can decrease the unacceptable effects of occlusion. Experimental results show the sufficiency and accuracy of the proposed method.

5. CONCLUSIONS

A network of cameras with disjoint view can significantly cover an extensive area as well as reduce the surveillance cost because of reducing the number of

cameras. People re-identification has a significant effect on tracking systems. Since severity in brightness of images can drastically reduce the efficiency of the systems, in this paper cumulative brightness transform function was used to compensate this effect so as to increase the ability of the tracking system in re-identification of persons. In this regard, we used the color histogram of different segments of body according to their relative distances in different color spaces such as RGB, YCbCr and HSV. In fact, we intended to improve the performance of the color histogram technique in this paper. Fortunately, the results indicate that our proposed method outperforms other existing methods. It is noteworthy that the proposed method has better results in YCbCr space than other color spaces.

6. REFERENCES

- Madden, C.S.T., "Tracking people across disjoint camera views", (2009). <https://books.google.com/books>
- Ilyas, A., "Object tracking and re-identification in multi-camera environments", (2011). <https://books.google.com/books>
- Yu, Q. and Medioni, G., "Multiple-target tracking by spatiotemporal monte carlo markov chain data association", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 31, No. 12, (2009), 2196-2210.
- Wu, B. and Nevatia, R., "Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors", *International Journal of Computer Vision*, Vol. 75, No. 2, (2007), 247-266.
- Pham, T.V., Worring, M. and Smeulders, A.W., "A multi-camera visual surveillance system for tracking of reoccurrences of people", in Distributed Smart Cameras., ICDSC'07. First ACM/IEEE International Conference on, IEEE, (2007), 164-169.
- Prosser, B., Gong, S. and Xiang, T., "Multi-camera matching using bi-directional cumulative brightness transfer functions", in BMVC. Vol. 8, (2008), 164.161-164.110.
- D'Orazio, T., Mazzeo, P. and Spagnolo, P., "Color brightness transfer function evaluation for non overlapping multi camera tracking", in Distributed Smart Cameras., ICDSC. Third ACM/IEEE International Conference on, IEEE., (2009), 1-6.
- Porikli, F. and Divakaran, A., "Multi-camera calibration, object tracking and query generation", in Multimedia and Expo, 2003. ICME'03. Proceedings. International Conference on, IEEE .Vol. 1, (2003), 653-656.
- Kettnaker, V. and Zabih, R., "Bayesian multi-camera surveillance", in Computer Vision and Pattern Recognition.. IEEE Computer Society Conference on., IEEE. Vol. 2, , (1999) 102-113.
- Moeslund, T.B., Hilton, A. and Krüger, V., "A survey of advances in vision-based human motion capture and analysis", *Computer Vision and Image Understanding*, Vol. 104, No. 2, (2006), 90-126.
- Yilmaz, A., Javed, O. and Shah, M., "Object tracking: A survey", *Acm Computing Surveys (CSUR)*, Vol. 38, No. 4, (2006), 1-45.
- Enzweiler, M. and Gavrilu, D.M., "Monocular pedestrian detection: Survey and experiments", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 31, No. 12, (2009), 2179-2195.

13. Khakpour, F. and Ardeshtir, G., "Using a novel concept of potential pixel energy for object tracking", *International Journal of Engineering-Transactions A: Basics*, Vol. 27, No. 7, (2013), 1023-1032.
14. Geronimo, D., Lopez, A.M., Sappa, A.D. and Graf, T., "Survey of pedestrian detection for advanced driver assistance systems", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 32, No. 7, (2010), 1239-1258.
15. Javed, O., Shafique, K., Rasheed, Z. and Shah, M., "Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views", *Computer Vision and Image Understanding*, Vol. 109, No. 2, (2008), 146-162.
16. Makris, D., Ellis, T. and Black, J., "Bridging the gaps between cameras", in *Computer Vision and Pattern Recognition, CVPR. Proceedings of the Computer Society Conference on, IEEE*. Vol. 2, , (2004), 205-210.
17. Zhu, L., Hwang, J.-N. and Cheng, H.-Y., "Tracking of multiple objects across multiple cameras with overlapping and non-overlapping views", in *Circuits and Systems, ISCAS. IEEE International Symposium on*, (2009), 1056-1060.
18. Rahimi, A., Dunagan, B. and Darrell, T., "Simultaneous calibration and tracking with a network of non-overlapping sensors", in *Computer Vision and Pattern Recognition, CVPR. Proceedings of the IEEE Computer Society Conference on, IEEE*. Vol. 1, (2004), 187 -194.
19. Javed, O., Shafique, K. and Shah, M., "Appearance modeling for tracking in multiple non-overlapping cameras", in *Computer Vision and Pattern Recognition, CVPR. IEEE Computer Society Conference on, IEEE*. Vol. 2, (2005), 26-33.
20. Mahdavi, M., Shahrouzi, S. and Hasanzadeh ,R., "A novel method for tracking moving objects using block-based similarity", *International Journal of Engineering-Transactions B: Applications*, Vol. 22, No. 1, (2008), 35-42.
21. Xie, Y., Yu, H., Gong, X., Dong, Z. and Gao, Y., "Learning visual-spatial saliency for multiple-shot person re-identification", *Signal Processing Letters, IEEE*, Vol. 22, No. 11, (2015), 1854-1858.
22. Liong, V.E., Ge, Y. and Lu, J., "Discriminative regularized metric learning for person re-identification", in *Biometrics (ICB), International Conference on, IEEE*, (2015), 52-57.
23. Liu, C., Gong, S. and Loy, C.C., "On-the-fly feature importance mining for person re-identification", *Pattern Recognition*, Vol. 47, No. 4, (2014), 1602-1615.
24. Xiong, F., Gou, M., Camps, O. and Sznajder, M., "Person re-identification using kernel-based metric learning methods", in *Computer vision-ECCV*, (2014), Springer, 1-16.
25. Mazzeo, P.L., Giove, L., Moramarco, G.M., Spagnolo, P. and Leo, M., "Hsv and rgb color histograms comparing for objects tracking among non overlapping FOVs, using CBTF", in *Advanced Video and Signal-Based Surveillance (AVSS), 8th IEEE International Conference on, IEEE*, (2011), 498-503.
26. Remondino, F. and Roditakis, A., "3d reconstruction of human skeleton from single images or monocular video sequences", in *Pattern recognition*, (2003), Springer, 100-107.
27. Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M. and Shafer, S., "Multi-camera multi-person tracking for easyliving", in *Visual Surveillance, Proceedings. Third IEEE International Workshop on, IEEE*, (2000), 3-10.
28. Comaniciu, D., Ramesh, V. and Meer, P., "Kernel-based object tracking", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 25, No. 5, (2003), 564-577.
29. van Rijsbergen, C.J., "Information retrieval", *London, Butterw*, (1979).

People Re-identification in Non-overlapping Field-of-views using Cumulative Brightness Transform Function and Body Segments in Different Color Spaces

A. G. Sorkhi^a, H. Hassanpour^a, P. L. Mazzeo^b

^aFaculty of Computer and Information technology, Shahrood University of Technology, Shahrood, Iran

^bInstitute of Applied Sciences and Intelligent System - CNR. Via della Libertà 3, 73010 Arnesano (Lecce), Italy

PAPER INFO

چکیده

Paper history:

Received 23 May 2015

Received in revised form 10 October 2015

Accepted 24 December 2015

Keywords:

People Tracking

Camera Networks

Non-overlapping Field-of-view

Various Segments of Human Body

Cumulative Brightness Transform Function

در سال‌های اخیر، از شبکه‌ای از دوربین‌ها با دید مجزا، به منظور نظارت بر ناحیه وسیع استفاده شده است. ردیابی در این نوع از سیستم‌ها، معمولاً بر پایه دو مرحله اساسی می‌باشد. در مرحله اول، افراد، در زاویه دید یک دوربین شناسایی و ردیابی می‌شوند. در مرحله دوم، بازشناسایی افراد به منظور ردیابی آنها در تمام ناحیه تحت نظارت انجام می‌گیرد. در این نوع از سیستم‌ها، از ویژگی‌های زیادی مانند لباس و ظاهر افراد به منظور شناسایی استفاده شده است. ولی شباهت بین لباس و ظاهر افراد نتایج قابل قبولی را از خود نشان نداده است. به همین دلیل استفاده از ویژگی‌های قابل اطمینان به منظور بالا بردن توانایی سیستم‌های ردیابی همچنان مورد نیاز می‌باشد. هدف از مقاله حاضر، ارائه الگوریتمی جدید به منظور شناسایی افراد در شبکه‌ای از دوربین‌ها با دید مجزا می‌باشد. در روش پیشنهادی، بر اساس اندازه نسبی قسمت‌های مختلف، بدن انسان به سه قسمت مجزا، سر، میان‌تنه و پایین‌تنه تقسیم شده است. عمل شناسایی افراد با توجه به هیستوگرام قسمت‌های مختلف بدن انسان در زاویه دید دوربین‌ها انجام می‌شود. یکی دیگر از چالش‌های موجود در این نوع از سیستم‌ها تغییر روشنایی و تغییر کردن ظاهر افراد در زاویه دید دوربین‌های متفاوت می‌باشد. تابع انتقال روشنایی جمعی به منظور کاهش مشکلات تغییرات شدت روشنایی در این نوع از سیستم‌ها به کار گرفته شده است. آزمایش‌ها در فضاها رنگی RGB، YCbCr و HSV به منظور ردیابی و شناسایی مجدد افراد انجام شده است که فضای رنگی YCbCr عملکرد بهتر در مقایسه با دیگر فضاها رنگی در این نوع از سیستم‌ها از خود نشان داده است.

doi: 10.5829/idosi.ije.2015.28.12c.03