

Person Re-identification Based on Fusing Appearance Features in Perceptual Color Space

Caixia Fan, Yajun Chen and Lei Cao

Abstract Person re-identification is to match pedestrian targets observed from different camera views of multi-camera surveillance systems. Aiming at the issues that the color similarity pedestrian affect the recognition result, a person re-identification method based on multi-feature fusion in perceptual uniform color space is proposed, which is according to the characteristics of human vision system. Firstly, the color space which is suitable for similar pedestrian recognition is selected from five color spaces CIELAB, S-CIELAB, IPT, LAB2000_{HL} and CAM02-SCD. Secondly, three kinds of pedestrian appearance features are extracted including spatial weighted histogram, local color and shape feature and global texture feature. Different distance measure methods are used to calculate the similarity of three different features. Finally, the linear fusion is performed by the adaptive weights. The experimental results based on VIPeR database and ETHZ database show that the proposed method is effective and can be used for the recognition of pedestrians with visual similarity.

Keywords Perceptual color space · Person re-identification · Multi-feature fusion

1 Introduction

Person re-identification is to determine whether an interest object in a camera view is in the other camera. But due to the low resolution, perspective changes, illumination changes, pedestrian posture changes and similar pedestrian's existence, result in a person in view of different monitoring looks very different. All of these

C. Fan (✉) · Y. Chen

School of Printing, Packaging Engineering and Digital Media Technology,
Xi'an University of Technology, Shaanxi, China
e-mail: fcx1981@xaut.edu.cn

L. Cao

College of Computer Science and Technology,
Xi'an University of Science and Technology, Shaanxi, China

© Springer Nature Singapore Pte Ltd. 2017

P. Zhao et al. (eds.), *Advanced Graphic Communications and Media Technologies*, Lecture Notes in Electrical Engineering 417,
DOI 10.1007/978-981-10-3530-2_34

273

make it a great challenge for person re-identification. At present, there are two main research ideas: (1) the method based on the appearance characteristic description; (2) the method based on metric learning [1].

The method based on the appearance characteristic description generally assumes that pedestrians don't change their appearance in a short time. The weighted HSV histogram and maximally stable color regions are extracted using as the color features, combined with the high frequency and complex structure blocks to identify pedestrians [2]. In the literature [3], the spatial color information and the structural information are fused. In order to establish the integrated model, the paper combines the different color features of pedestrians in the decision-making stage by using the method of measuring [4]. But at present, the person appearance characteristic description methods don't take into account for the perceptual process of human vision. This leads to recognition results, in particular, the recognition results of similar pedestrian targets are not consistent with human perception, and thus produce false recognition.

The distance measure of feature is concerned with how to make the characteristics of the same target have a higher degree of similarity when the appearance model is set up [5–7]. This kind of method needs to carry on the study of the choice of the sample and the distance measure, when the scene changes, the general need to retraining.

This paper proposes a person re-identification method based on multi-feature fusion in perceptual uniform color space. Firstly, the pedestrian images are transformed into the perceptual uniform color space, which is consistent with the human visual system, and the appearance characteristics of the target are extracted. Then the pedestrian recognition is performed by the method of adaptive weights. Experimental results based on VIPeR database and ETHZ database show that the proposed algorithm has a greater distinction.

2 Selection of Perceptual Color Space

In video monitoring system, the efficiency of person re-identification with color as the main feature can be severely affected by the pedestrians with similar clothing (i.e. small color difference images). The human visual system is able to identify the color of the pedestrians. Therefore, converting image to the perceptual color space which is consistent with human vision to extract the feature can be more effective distinguish between similar goals. So, the study of the performance of the different perceptual color spaces can provide the basis for the identification of similar pedestrians. At present, the color spaces and their corresponding color difference formulas for small color difference image are: CIEDE2000 color difference formula based on CIELAB color space [8], CAM02-SCD based on CIECAM02 color appearance model [9], S-CIELAB [10], color difference based on IPT color space [11] and color space based on LAB2000_{HL} [12]. In this paper, we choose the color

space which is suitable for person re-identification by visual experiments and pedestrian recognition results.

According to the characteristics of video surveillance scene, six standard images are selected as test images of visual experiments. These images cover still life, animals, figures and landscapes, contain the highlights and shadows, complicated texture changes, memory color and typical color. Detailed visual experimental procedure is seen in the literature [13]. By the comparison of the calculated color differences and the corresponding visual evaluation values and the STRESS values of statistical significance of color difference formulas, we can draw the following conclusions: for the small color difference image data, IPT color space is relatively stable. For all the attributes of the image transform its performance is better. The color difference formulas based on IPT color space is more close to the human eye’s visual evaluation results.

For person re-identification, the characteristics of different uniform color spaces are illustrated by experiments based on CAMPUS-Human database. There is a small change of illumination condition between the pedestrians in the database and there are many similar small color difference images [14]. Firstly, pedestrian images are transformed into different perceptual color spaces, then the recognition results are obtained by calculating Bhattacharyya distance of color histograms. The experimental results are evaluated by the ratio of finding the correct match in the top n matches. The experimental results in five different color spaces are shown in Table 1. The table includes the rank 1, 5, 10, 15, 20, 25 matching rates. It can be seen from the table, the poor performance of the CIEDE2000 color difference formula and S-CIELAB model when the rank is small, CAM02-SCD and LAB2000_{HL} have high recognition rates, the recognition rates of IPT color space of second. But with the increase of the rank, the pedestrian recognition rates based on IPT color space are rising fast, which show that the IPT color space is more stable and less affected by the environment.

Therefore, this paper will extract the person appearance characteristics to establish the person appearance model in the IPT color space.

Table 1 Rates of the target recognition in different color space

Color space	Top ranked matching rates%					
	r = 1	r = 5	r = 10	r = 15	r = 20	r = 25
CIEDE2000	9.37	15.23	17.78	27.18	36.23	40.88
S-CIELAB	9.37	15.42	18.96	25.55	35.40	39.51
CAM02-SCD	12.11	20.33	28.13	34.03	42.25	53.58
IPT	11.36	17.95	28.18	36.77	42.38	53.63
LAB2000 _{HL}	13.48	18.62	24.44	29.92	39.51	47.73

3 Person Appearance Model

3.1 Global Feature Based on Spatial Histogram

In order to describe the image spatial structure information, the second-order spatial histogram in IPT color space is used to represent the whole color characteristics of the target. Second-order spatial histogram of the image can be express as Eq. (1):

$$\mathbf{S}_I^{(2)}(b) = \langle \mathbf{n}_b, \boldsymbol{\mu}_b, \boldsymbol{\varepsilon}_b \rangle, \quad b = 1, \dots, B \quad (1)$$

In the formula, B is quantification series, \mathbf{n}_b is image quantification histogram, μ_b is the mean value of all pixels with the same color value, ε_b is the coordinate covariance matrix for all pixels with the same color value [3].

The similarity of the spatial histograms (\mathbf{S}, \mathbf{S}') can be calculated by the similarity weighted sum of the two histograms and the pixel positions.

$$\rho(\mathbf{S}, \mathbf{S}') = \sum_{b=1}^B \psi_b \rho_n(\mathbf{n}_b, \mathbf{n}'_b) \quad (2)$$

$\rho_n(\mathbf{n}_b, \mathbf{n}'_b)$ is the similarity between the histograms, which used Bhattacharyya distance to calculate, ψ_b is spatial similarity. It can be express as Eq. (3):

$$\psi_b = \eta \exp \left\{ -\frac{1}{2} (\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b)^T \hat{\boldsymbol{\varepsilon}}_b^{-1} (\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b) \right\} \quad (3)$$

η is Gaussian normalization constant, $\hat{\boldsymbol{\varepsilon}}_b^{-1} = (\boldsymbol{\varepsilon}_b^{-1} + (\boldsymbol{\varepsilon}'_b)^{-1})$.

3.2 Local Feature

In order to eliminate the influence of small areas, the person is divided into a number of regions with similar color by mean shift image segmentation algorithm; retain more than 30 of the total number of pixels as main color areas. The local features are described by the color and shape information [3].

$$F = (\hat{C}, \lambda \hat{H}) \quad (4)$$

F is the description of the combination of color and shape, \hat{C} is the normalized IPT color histogram, \hat{H} is 128 dimension normalized SIFT features, λ is the weight parameter, $\lambda = 0.6$. Principal component analysis (PCA) in feature space is performed in this paper, and the feature vectors V corresponding to the first 30 eigenvalues are obtained to describe the different regions, combined with the central

location information *Cent* and regional size information *Re size* of each region to describe the local characteristics of the image.

This paper adopts the improved EMD (Earth Mover’s short Distance) to calculate the similarity between the two local image characteristics [15]. Set $A = \{(a_1, w_{a1}), (a_2, w_{a2}), \dots, (a_m, w_{am})\}$ is the representation of image A with m cluster and a_i is the description of the cluster, w_{ai} is the weights of clustering. As well as $B = \{(b_1, w_{b1}), (b_2, w_{b2}), \dots, (b_n, w_{bn})\}$. $D = [d_{ij}]$ is distance matrix, d_{ij} is distance function between clustering a_i and b_j . The improved EMD in image A and image B can be defined as:

$$EMD_{\alpha}(A, B) = \left(\min_{f_{ij}} \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij} \right) + \left| \sum_{i=1}^m a_i - \sum_{j=1}^n b_j \right| \times \alpha \max_{i,j} \{d_{ij}\} \quad (5)$$

The solution method of the Formula (5) can be found in the literature [15].

In this paper, the representation method for the local characteristics of person is $O = \{(o_1, w_{o1}), (o_2, w_{o2}), \dots, (o_m, w_{om})\}$, o_i is the local characteristics of image regions, w_{oi} is the coordinates of the feature points, d_{ij} is the Euclidean distance between different areas. By experiment, this paper takes $\alpha = 0.3$.

3.3 Texture Feature Description

In this paper, we extract the LBP texture information to make up the deficiency of the overall color feature and the local feature [3]. The similarity of the texture features is obtained by Bhattacharyya distance.

4 Multi-feature Fusion

Person re-identification has two sets of image data, the candidate target P and the identify target Q . The similarity of the two targets is obtained by the linear fusion of feature similarities.

$$S(P, Q) = \alpha \cdot S_S + \beta \cdot S_L + \gamma \cdot S_{LBP} \quad (6)$$

S_S is the spatial histogram similarity, S_L is the similarity of local characteristics, S_{LBP} is the similarity of the texture features, α, β, γ are the weights.

An adaptive weight selection method is presented by comparing the color features and texture features of the target with the global features of all pedestrians. Firstly, The Bhattacharyya distances between the hue histogram of each candidate and global hue histograms are calculated, the greatest distance for DC_{max} . Then, the distance between the hue histogram of the person to be identified and global hue

histogram is obtained, referred to DC_p . Finally, the importance and distinguish ability of color characteristics are calculated through DC_p divided by DC_{\max} , recorded as S_{color} . If S_{color} is high, according to the color information can achieve higher recognition rate. Similarly the importance $S_{texture}$ of texture feature can be calculated. Based on the importance and the ability of distinguishing of different visual information, the weights of multi-feature fusion are obtained according to the Formula (7). Spatial histogram features and the local features are color based features, and select the same weight. The weight of the texture feature is selected according to its importance.

$$\begin{aligned}\alpha &= \frac{S_{color}}{2S_{color} + S_{texture}} \\ \beta &= \frac{S_{color}}{2S_{color} + S_{texture}} \\ \gamma &= \frac{S_{texture}}{2S_{color} + S_{texture}}\end{aligned}\quad (7)$$

α , β , γ are range between 0 and 1, and $\alpha + \beta + \gamma = 1$.

5 Experimental Results and Analysis

In this paper, the algorithm effectiveness is tested in VIPeR database and ETHZ databases. They include different issues that need to be addressed for person re-identification. The experimental results are evaluated by the cumulative matching characteristic cumulative (CMC) curve [16, 17].

5.1 VIPeR Database

There are 632 sets of 28 human targets in the VIPeR database. Images in the database are obtained from two different cameras, different views, pose, and illumination conditions [16]. The comparison of the algorithm and the literature [3] is shown in Fig. 1. It can be seen from the figure, in the first column based on the perceptual color space, person re-identification rate of 40.1%, compared with the literature [3] result 31.3% increased by 8.8%. Fifth column in the literature [3] the target recognition rate is 56%, and the result of this paper is 65.3%. In the first 15 columns, the recognition rates in the perceptual color space are growing fast. It shows that the extracted features in uniform color space are more distinguished, and they can distinguish visually similar pedestrian targets effectively.

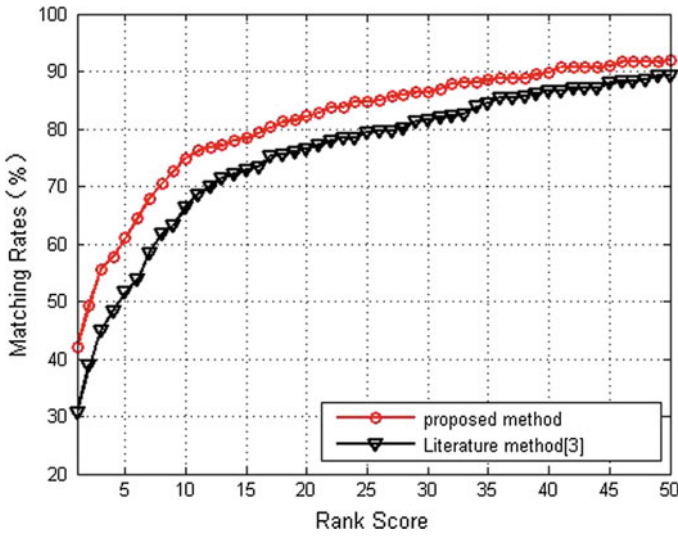


Fig. 1 Comparison of experimental results in VIPeR database

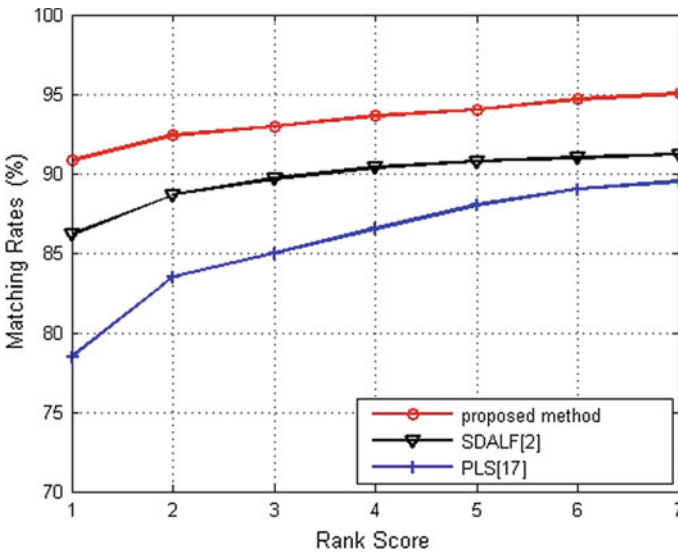


Fig. 2 Results in ETHZ1 database

5.2 ETHZ Database

The ETHZ database consists of three sequences: ETHZ1, ETHZ2 and ETHZ3 [17]. There are small change of attitude and light condition. The results of three data sets in the ETHZ database are shown in Figs. 2, 3 and 4. The appearance model of the candidate target is established by 5 frame key frames, and the target is matched with

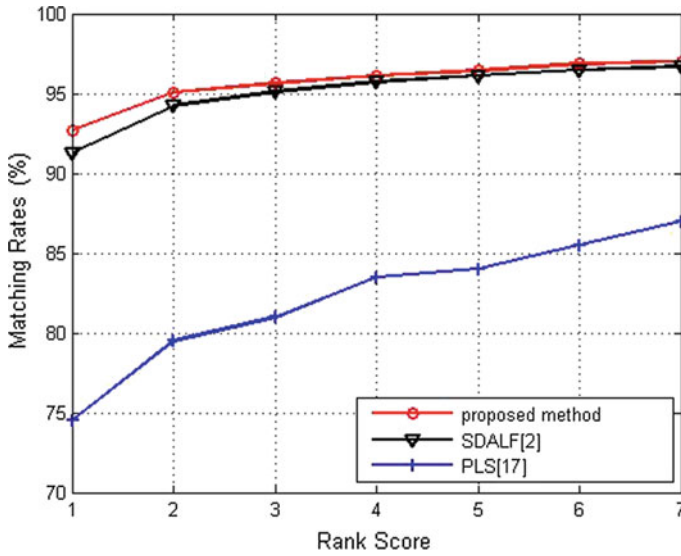


Fig. 3 Results in ETHZ2 database

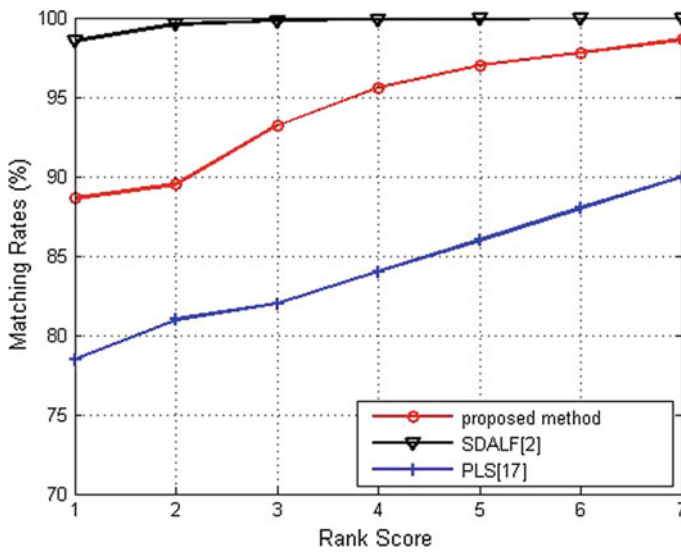


Fig. 4 Results in ETHZ3 database

1 frame. In order to get the stable experimental results, this paper selects 10 sets of targets for person re-identification, and the experimental results are the average of the results of several experiments. As can be seen from the figures, the algorithm can improve the recognition rate of the target in the top 5. In ETHZ1 image

sequences the algorithm is better than the SDALF algorithm and PLS algorithm. In the ETHZ2 data set, the algorithm and SDALF algorithm to identify the difference is not much, better than the PLS algorithm. In the ETHZ3 data set, this algorithm is better than PLS algorithm, but less than SDALF.

6 Conclusions

This paper proposes a new method of person re-identification based on multi-feature fusion in perceptual uniform color space. The method is based on the characteristics of human vision system, which mainly solves the problem of the influence of similar targets. The experimental results based on VIPeR database and ETHZ database show that the features extracted from the uniform color space have a stronger ability to distinguish the pedestrian recognition problem. The validity of the algorithm is illustrated. How to establish a more effective similarity evaluation criterion is the focus of follow-up study.

Acknowledgements This study is funded by the Doctoral scientific research start-up fund (108-451115001) and School science research program (2015CX021).

References

1. Bedagkar-Gala A, Shah S K (2014). A survey of approaches and trends in person re-identification. *Image and Vision Computing*, 32(4): 270–286.
2. Farenzena M, Bazzani L, Perina A (2010). Person re-identification by symmetry-driven accumulation of local features. *CVPR*: 2360–2367.
3. Fan Caixia, Zhu Hong (2013). Person re-identification based on multi-features. *Journal of Image and Graphics*, 18(6):711–717.
4. Liu X, Wang H, Wu Y (2015). An ensemble color model for human re-identification. *WACV*: 868–875.
5. Ahmed E, Jones M, Marks T K (2015). An improved deep learning architecture for person re-identification. *CVPR*.
6. Du Yu-Ning, Ai Hai-Zhou (2014). A statistical inference approach for person re-identification. *Journal of Electronics and Information Technology*, 36(7): 1612–1618.
7. Qi Meibin, Tan Shengshun (2016). Multi-feature subspace and kernel learning for person re-identification. *ACTA AUTOMATICA SINICA*, 42(2):299–308.
8. LUO M R, CUI G, RIGG B (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5): 340–350.
9. LUO M R, CUI G, LI C (2006). Uniform colour spaces based on CIECAM02 colour appearance model. *Color Research & Application*, 31(4): 320–330.
10. ZHANG X, WANDELL B A (1997). A spatial extension of CIELAB for digital color-image reproduction. *Journal of the Society for Information Display*, 5(1): 61–63.
11. EBNER F, FAIRCHILD M D(1998). Development and testing of a color space (IPT) with improved hue uniformity. In *proceedings of the 6th Color Imaging Conference*, 11. 17–20.
12. LISSNER I, URBAN P (2012). Toward a unified color space for perception-based image processing. *IEEE Transactions on Image Processing*, 21(3): 1153–1168.

13. Fan Caixia, Chen Yajun (2016). Evaluation of perceptual uniform color spaces and application in image processing. *Journal of Xi'an university of technology*, 32(2):191–198.
14. Xu Y, Lin L, Zheng W S, et al (2013). Human re-identification by matching compositional template with cluster sampling. In *proceedings of the IEEE International Conference on Computer Vision*: 3152–3159.
15. Pele O, Werman M(2008). A linear time histogram metric for improved sift matching. *ECCV*: 495–508.
16. Gray D, Tao H (2008). Viewpoint invariant pedestrian recognition with an ensemble of localized features. *ECCV*:262–275.
17. Schwartz W R, Davis L S(2009). Learning discriminative appearance-based models using partial least squares. In *Proceeding of Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing*: 322–329.