

Adaptive image segmentation based on color clustering for person re-identification

Lixia Zhang¹ · Kangshun Li¹ · Yan Zhang¹ · Yu Qi¹ · Lei Yang¹

Published online: 6 May 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Person re-identification refers to identifying a particular person who has occurred in the monitoring network automatically by computer in the surveillance video, which is significantly important for the improvement of intelligence of video monitoring. But the research of person re-identification is not mature and facing many challenges. The following factors may lead to a certain difference for the same person in different monitoring video images; for example: the illumination changing in the monitoring environment, the shooting angle difference, and the posture difference. These may lead to low recognition accuracy. In this paper, a new appearance-based person re-identification method was proposed and an investigation was launched on the following topics for the improvement of recognition accuracy. First, a simple and feasible method for color invariants was proposed, so that the affection of color by change of illumination and shooting angle could be eliminated. Then, a highly adaptive image segmentation method based on color clustering and a color feature representation scheme for specific color characteristics were designed, which could help to extract color features in more reasonable points. Finally, an effective similarity measure criterion was obtained through QSF measure learning, which could ensure that the different pedestrians can be distinguished and was better able to capture the visual change of the same person. In addition, the traditional evolutionary algorithm was improved and applied to the process of iterative computation for QSF. The experimental results show that

our method is an effective way for the person re-identification problem.

Keywords Person re-identification · Adaptive image segmentation · Color invariants · Color clustering · Evolutionary algorithm

1 Introduction

Nowadays, surveillance systems have been widely used, which leads to the increasing need of determining if a particular person has been observed in the monitoring network. This problem is known as person re-identification and is an important research question with practical significance. It can be applied to criminals tracking and recognition, look for a lost child in public places, etc. It has received a lot of attention in the field of computer vision in recent years (Zheng et al. 2013).

Person re-identification technology can be divided into two categories according to the technical scheme. One is based on the biological characteristics, such as face, gait, etc., and the other is based on the appearance characteristics, such as the color of the clothes and so on. High requirements are needed for person re-identification based on biological characteristics, such as the resolution of the surveillance video and the shooting environment. Sometimes, these requirements cannot be met by the actual monitoring system, which will lead to an infeasible and unreliable recognition result. Person re-identification based on appearance characteristics has become the research hot spot. However, there are still many problems it must face. Influenced by many factors, such as illumination changes, shooting angle and people's posture, this may lead to the same pedestrians in different monitoring video images having a certain degree of difference as

Communicated by V. Loia.

✉ Kangshun Li
likangshun@sina.com

¹ College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, Guangdong, China



Fig. 1 The example of the same pedestrian's image in different monitoring videos with certain degree of difference

shown in Fig. 1, which will seriously affect the recognition result.

To address these issues, a new appearance-based person re-identification method is proposed in this paper. Color is used as the feature, which has good stability and high robustness, and an adaptive image segmentation approach based on color clustering is proposed in this method. To make the color feature more feasible and reliable, color invariants processing will be conducted first. Then, the color feature will be extracted in each area which has been segmented by the adaptive image segmentation approach. Finally, the QSF measure learning scheme will be adopted to measure the distance between those match objects.

Two obvious advantages can be found in this method. On one hand, a simple but effective color invariant algorithm is presented, which guarantees the reliability of the color feature. On the other hand, an adaptive image segmentation method based on color clustering is proposed, and the color feature is extracted according to the segmentation area, which ensures the accuracy of the identification.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature on person re-identification. Section 3 introduces the specific details of our approach. The experimental results are presented in Sect. 4. Section 5 discusses the implication of this research and concludes with future work.

2 Related work

Two key problems needed to be solved in person re-identification based on appearance characteristics are feature representation and feature matching. First, the right feature representation scheme needs to be found to characterize the person for recognizing; then, the distance between each pair of potential matches will be measured to determine whether it contains the same person. In this section, we will make

an introduction to the research status of these two issues. In addition, we use the clustering method to do image segmentation, and evolutionary algorithm to do iterative calculation, which means we will introduce some knowledge associated with clustering and evolutionary technology here also.

2.1 Feature representation

The goal of feature representation is finding a right scheme to represent the characteristics of a person's appearance, which would cope with the changes in illuminance, shooting angle, and body posture, and adopt an appropriate way to show the feature.

Features used in the person re-identification problem mainly include color, texture and shape. To obtain good recognition effect, a combination of variety features has been tried by many researchers (Farenzena et al. 2010; Bazani et al. 2013, 2012; Fan et al. 2013). For example, three kinds of symmetrical characteristic information have been used in references (Farenzena et al. 2010): weighted color histogram, maximum stability region and high-frequency complex structure. By using this scheme, an ideal effect could be obtained to some extent, but there are still some problems that cannot be ignored. First, to determine the weight for each feature is difficult. At present, we do not have a common method to judge them, but only rely on experience value. Using more features, greater amounts of calculation will occur in the feature matching phase, which will affect the practical application of person re-identification.

Color is the most basic visual characteristics for human perception. Compared with other features, color has good stability and high robustness; also, the method of feature extraction and calculation is relatively simple. That is why the color feature is widely used in person re-identification (Bak et al. 2010; Bedagkar and Shishir 2012). But it may be easily affected in the actual monitoring environment by factors such as illumination change and the shooting angle. How to keep color invariants is the problem needed to be resolved. The research of color invariants is not a new topic. Funt and Finlayson (1995) had proposed a color invariant descriptor known as "color index", which uses color logarithmic space derivatives to eliminate the influence of light. Many expanded and improved algorithms based on Funt's method have been proposed (Gevers and Smeulders 1999; Weijer and Sehnid 2006). In recent years, the research of color invariants has caused more and more attention, which has been applied to many fields, such as background modeling (Zhou et al. 2013), edge detection (Chu et al. 2013), and so on. The requirements of color invariants processing is different due to different applications, and many existing methods cannot be directly applied to person re-identification problem. Kviatkovsky et al. (2013) has proposed a specialized color invariant scheme for person re-

identification, but this algorithm is not robust enough for only considering the influence of illumination changes on color features.

In addition, some limitations such as the increase of false matching rate might occur using color feature only. For example, using color histogram to represent color feature, the extracted features are similar for one person who wears red coat and blue pants, and the other who wears red pants and blue coat. Therefore, image segmentation should be involved when using color feature for person re-identification. The pedestrian image will be divided into several regions in a certain way and color features extracted from each region, respectively. Some pedestrian image segmentation methods have been put forward (Farenzena et al. 2010; Bazzani et al. 2013; Bak et al. 2010; Yao et al. 2015; Garcia et al. 2011). But most of these methods did not consider the distribution of color information, and the number of the segmentation blocks was fixed, which might lead to lack of flexibility and adaptability. Some special image segmentation method is needed for segmenting the image adaptively, and the clustering method is a suitable one.

The clustering method can separate objects from others with similar characterizes. This function coincides with image segmentation. Therefore, the clustering method has been widely used in the image segmentation field. It has attracted many scholars to deeply research and improve these methods in recent years, and various successful image segmentation methods based on clustering algorithms have appeared. Many mature clustering methods have been used for image segmentation, such as k-means clustering method (Hartigan and Wong 1979), FCM algorithm (Bezdek et al. 1984), and hierarchical clustering algorithm (Johnson 1967). The traditional k-means clustering method has attracted much attention due to its simplicity and efficiency. But this method also has some limitations, for example, selection of the category number and initialing clustering center is difficult, and some initial center algorithm easily fall into the local optimum and make the clustering result unstable. The FCM algorithm belongs to the improved k-means method. It does not need human intervention for threshold setting, which might be more automatical. However, there are still some problems, such as the excessive calculation, sensitivity to noise, and over-segmentation phenomenon. Hierarchical clustering is also a classical clustering algorithm, which was widely used due to its simplicity and high quality of the clustering result. But it needs to compute the adjacent degree matrix, which may lead to high time and space complexity. So, it is not applicable for processing the image with large datasets.

Another problem is representation of color features. Color histogram (Baltieri et al. 2014) is the commonly used method at present. But, color histogram only describes the various color proportion in the whole image and leave the specific

location of each color in the image untouched. This may affect the accuracy of person re-identification to some extent.

Above all, there are still a lot of problems needed to be solved for the person re-identification problem when the color feature is used.

2.2 Feature matching

The key problem of feature matching is to find a similarity measure criterion, which can make the images of the same pedestrian have higher similarity than a different one.

Many mature and simple similarity measure functions have been proposed, such as K neighbor distance (Hahnel et al. 2004) and L1-norm-based distance (Wang et al. 2007). Most of them have been successfully used in the image retrieval field, but person re-identification problem is more difficult and good recognition effect is hard to achieve using these simple measurement functions directly. Therefore, many complex similarity measure criteria have been proposed to solve the person re-identification problem in specialty. For example, Prosser et al. (2010) proposed support vector sorting method, which used the kernel function of RankSVM (Joachims 2002) to project the feature space to make it distinguishable through separating the matched feature from the unmatched one. Zheng et al. (2013) proposed the relative distance model based on probability, which learned a Mahalanobis distance model (Maesschalck et al. 2000) and can make different contributions for different components of the features to distance. The key problem of these methods is measure learning, and some modified methods have been put forward such as references Du and Ai (2014) and Harandi et al. (2014). Du and Ai (2014) estimates the similarity of pedestrian images by learning a second power similarity function, and (Harandi et al. 2014) conducts measure learning through calculating and comparing the covariance descriptor. These research works have good reference value for measure learning, but there are still some limitations: some ignored the individual characteristics of apparent changes and some required the complex learning process.

Determination of the weight for each part when doing feature matching belongs to the objective optimization problem. Among the numerous objective optimization methods, evolutionary algorithm (Holland 1992) has attracted more and more attention of researchers. In recent years, it has been widely used in the area of image processing and machine vision, due to its high speed and reliability (Li et al. 2015).

Evolutionary algorithm is a bionic algorithm for solving complex optimization problems based on Darwin natural evolution and Mendel genetic variation theory. This method is developed from the evolutionary principle of "survival of the fittest". Basic genetics operation is used repeatedly to

the group which contains the possible solutions and generates new group constantly to make the population evolve. The characteristics of the evolutionary algorithm make it a suitable objective optimization method for image processing, such as image matching (Cuevas et al. 2013), image segmentation (Chang et al. 2012) and face recognition (Venkatesan and Madane 2012).

Although evolutionary algorithm has been widely used in the image processing area, some shortages also have been exposed in the process of application. First, the evolutionary algorithm has better robustness than other traditional search methods, but it is good at global searching and insufficient for local searching. Second, the convergence rate of the evolutionary algorithm will decline when there are many variables, large search scopes, or no given search range. Finally, the evolutionary algorithm is prone to premature phenomenon, namely, it may bind to the local optimal solution rather than the global one. Therefore, the above problems of evolutionary algorithm must be solved when used for image processing.

In conclusion, the research of person re-identification in the literature is at the exploration stage. A reliable and feasible solution scheme has not been found for feature representation and feature matching.

3 The approach

For the two key issues of appearance-based person re-identification, we pay more attention to the first one, that is, the feature representation problem. We have designed a simple, but feasible color invariants processing method to get color feature with high reliability and good computability. We also have proposed a highly adaptive image segmentation method and the corresponding color feature representation method based on color structure distribution, which could extract color feature in a more meticulous and reasonable way. We have not done much study for feature matching problem and basically refer to the QSF learning method (Du and Ai 2014), but we have made some improvements for the iterative calculation processing and threshold setting using the evolutionary algorithm. The whole system framework is shown in Fig. 2. It can be seen from Fig. 2 that the match results will be demonstrated from high to low, according to the similarity. The image with the highest similarity may be the potential target pedestrian image.

3.1 Color invariants processing

For the person re-identification problem, the main influencing factors for color invariants are shooting angle and luminance change. There is a simple and feasible method to solve any one of these problems separately, but it is very troublesome to address two issues at the same time.

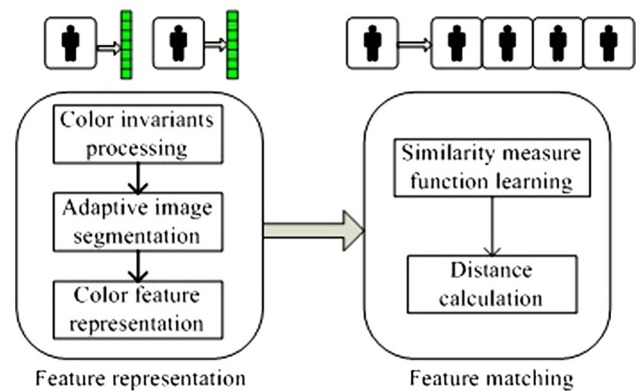


Fig. 2 The whole system framework of our person re-identification method

First, let us have a look at the shooting angle. In RGB color space, if the color descriptors of one pixel got through two different shooting angles (R_1, G_1, B_1) and (R_2, G_2, B_2) , respectively, then their relation can be shown as: $R_2 = sR_1$, $G_2 = sG_1$, $B_2 = sB_1$. Thus, to obtain color invariants, adjusting the RGB color space by performing the operation as in formula one to each pixel is needed:

$$\frac{R}{R+G+B}, \frac{G}{R+G+B}, \frac{B}{R+G+B}. \quad (1)$$

Then, another factor of luminance is to be considered. In RGB color space, if color descriptors of one pixel got in two different luminance conditions are (R_1, G_1, B_1) and (R_2, G_2, B_2) , respectively, then their relation can be shown as: $R_2 = rR_1$, $G_2 = sG_1$, $B_2 = tB_1$. Thus, to obtain color invariants, adjusting the RGB color space by performing the operation as in formula two to each pixel is needed.

$$\left(\frac{2R_1}{R_1+R_2}, \frac{2G_1}{G_1+G_2}, \frac{2B_1}{B_1+B_2} \right), \left(\frac{2R_2}{R_1+R_2}, \frac{2G_2}{G_1+G_2}, \frac{2B_2}{B_1+B_2} \right). \quad (2)$$

The method shown in formulas 1 and 2 can solve the shooting angle and luminance change problems, respectively, but neither of them can solve two problems at the same time. If we want low computational complexity and not involving complex color transform theory at the same time, using the above two formulas iteratively might be a feasible solution. That is, for the input image, formula 1 is to be processed first, and then formula 2. We perform this process repeatedly, until the difference between the last two processing results gets small enough. The detailed procedure can be defined as follows. TI and SI represent the target image and source image, respectively; P_1 and P_2 represent the process used in formula 1 and formula 2, respectively. Process 3 represents

the termination condition, that is, the process will terminate when the change is less than ε (a small positive number).

1. $TI_0 = SI$;
2. $TI_{i+1} = P_2(P_1(TI_i))$;
3. Repeat step 2 until $TI_{i+1} - TI_i < \varepsilon$.

The main question is whether the process is monotone convergence, that is, whether the result value can be obtained in finite iterations. It had been proved by mathematical method in references (Graham et al. 1998) and through an experimental way by us.

Another problem is the speed of iterative calculation. If the iterative calculation process is based on each pixel of the image, then more calculation time will be needed. We need to optimize the iterative calculation scheme to improve the speed. The optimization idea is as follows. The clustering method is used to divide the pixels with similar colors and will be put together for calculating, so that the processing time can be reduced. Due to the possible error caused by the clustering algorithm, the visual effect of the target image may be affected. So, this optimization method is suitable for the low-effect precision-demanded treatment. The pursuit of person re-identification is the accuracy of recognition for two persons' images, which means the optimization method is suitable.

3.2 Adaptive image segmentation

When doing the pedestrian image segmentation, we only aim at the extracted pedestrian foreground image and do not consider the prospects of background separation. The research of background separation is very mature and a lot of ready-made algorithms can be used for reference. In this paper, the method proposed by references (Zhu et al. 2014) will be adopted.

First of all, the SDALF method will be used to take out the head of the pedestrian image and separate the body into upper and lower. The reason for head taking is because it cannot provide distinguishable information for identification when using the color feature. The separating of the body is to avoid misrecognition of the situation such as one person wears red coat and blue pants, and the other wears red pants and blue coat. After these preprocessing works, adaptive segmentation will be made in the upper and lower body, respectively. "Adaptive" means that the number of division blocks is unfixed and it will be determined automatically according to the characteristic of the image needed to be segmented.

Using the clustering method for image segmentation is to perform image classification algorithm accurately and minimize the clustering index. In particular, first, some characteristics of the pixels are mapped to the feature spaces

which have been divided into several regions according to certain rules. Then, the regional category will be determined according to the nature of the pixel through searching the gathering status of points within the feature space. Finally, pixels with similar characterize will be marked to achieve the goal of image segmentation. In this paper, the k-means clustering method will be improved and used. The process is as follows:

1. Choose k points as the initial center.
2. Calculate the distances of each point to the k centers, find the minimum distance, and assign each point to the nearest center to form k clusters.
3. Reclassify to calculate each cluster center and modify the new center which is the mean value of this center.
4. Repeat steps 2 and 3 until the criterion function is convergent.

There are some problems that need to be further elaborated. First is how to determine the initial clustering center. The traditional k-means clustering algorithm selects the initial clustering center randomly, which may easily fall into the local optimum and make the clustering result unstable. The following method might be helpful for solving this problem. Formula 3 will be used to calculate the gradient of each pixel of the energy, where f represents the image and $e(x, y)$ is the energy value of pixel (x, y) in f . We first calculate the gradient norm for each channel (R, G, B) and then find the mean to obtain gradient norm, which is the image's energy. Then we select k representative pixels according to the distribution of energy as the initial clustering center. In particular, all pixels are divided into m intervals according to the pixel energy, such as: $[e_{\min}, e_1], (e_1, e_2], (e_2, e_3) \dots (e_{m-1}, e_{\max}]$. Here, e_{\min} represents the minimum energy value, and e_{\max} represents the maximum. Then, $e_i : e_{\min} < e_i < e_{\max}$. Assuming the number of pixels for the i th interval is n_i , and the number of all pixels is n , if $n_i/n < \varepsilon$, then the interval will be ignored and cannot be the initial clustering center. In other words, only those representative pixel intervals will be included in the initial clustering centers, and intervals with a few pixels will be abandoned. The exact value of m and ε depends on the actual demand; if more blocks are needed in the image with more detailed information, let m be larger and ε smaller as far as possible, and vice versa.

$$e(x, y) = \text{grad}(f) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}. \quad (3)$$

The second problem is how to calculate the distance between each point and the clustering center. The Euclidean distance will be chosen to do this work (show as formula 4), where D represents the distance between pixel p and q ,

and e_p and e_q represent the energy (calculated by formula three) of pixel p and q , respectively.

$$D(p, q) = |e_p - e_q|. \quad (4)$$

The final problem is how to select an appropriate criterion function. The classical k-means clustering method used SSE (sum of the squared error) criterion function (shown as formula 5). Where C_i represents the i th cluster, e_p is the energy of pixel p which belongs to the cluster C_i , and M_i is the mean energy value of C_i . This method is suitable for the situation that the density and size of the cluster are both similar. When differences of sizes are large, the value of SSE may become bigger for the cluster with more data. To minimize the value of SSE, a large cluster might be broken up, which may affect the quality of the clustering. So, we need to consider the impact number of object in one cluster on the value of SSE. In addition, to improve the effect of clustering, ensuring the maximization of difference between the clusters will be necessary. We improved SSE from the above two aspects. The improved calculation method is shown in formula 6, where w^t is the penalty factor which can adjust the effect to the value of SSE of different clusters. The key issue is the value of w and t . It is easy to determine the value of w , and we took it to be the reciprocals of the number of data object in cluster C_i in this paper. But there is no good way to determine the value of t , and the general experience value is 2. In fact, when the density and size of the dataset are similar, the value of t should be a bit smaller, such as close to 0. Otherwise, it should be a bit bigger, so that its impact on the weighted distance would become bigger. To achieve this goal, we adjust the value of t dynamically. Specifically, the program will determine how to select the value of t using formula 7. If the actual situation meets the requirement of formula 7, the experience value 2 will be selected; otherwise, t will be set to 0 which means the penalty factor has been cancelled.

The improved method can make the contribution for the value of SSE become smaller for the cluster with more data object by adding a penalty factor.

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} \|e_p - M_i\|^2 \quad (5)$$

$$SSE' = \sum_{i=1}^k \sum_{p \in C_i} w^t \|e_p - M_i\|^2 \quad (6)$$

$$\left| m_i - \sum_{i=1}^k m_i / k \right| > \mu. \quad (7)$$

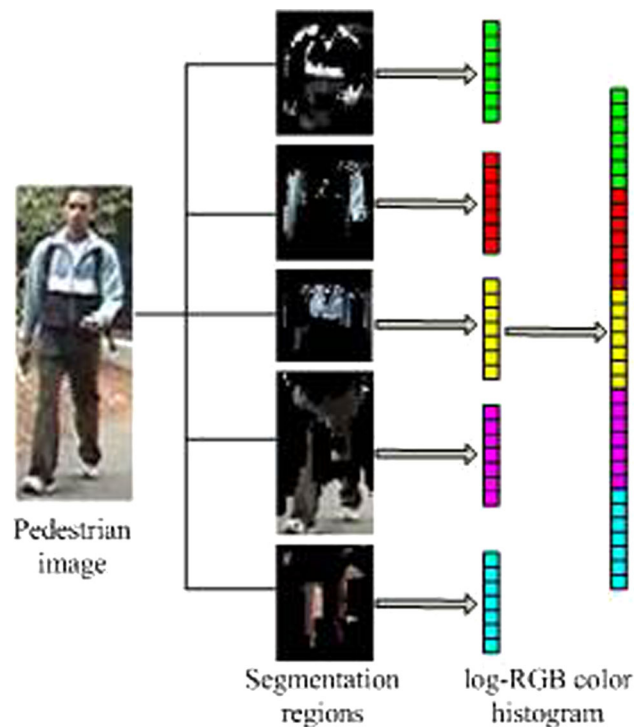


Fig. 3 The sketch for feature representation of our person re-identification method

3.3 Color feature representation

To solve the problem of commonly used color histogram method, a new effective scheme will be adopted for color feature representation. This method transformed the traditional RGB color space into a new one using formula 8, and it is called log-RGB color space. Through the above transformation, an x - y coordinate based on color could be obtained. Then, each color block segmented by the method proposed in Sect. 3.2 can be represented in the coordinates. Thus, the color features of each region and the shape rendered could be got in the coordinate system using this color feature representation scheme.

$$x = \log \frac{R}{G}, \quad y = \log \frac{B}{G} \quad (8)$$

To facilitate subsequent feature matching work, the log-RGB color space will be further converted into the representation format which is similar to the color histogram and can be named as LRCH (log-RGB color histogram). 16-Dimensional LRCH feature will be extracted from each color channel in the log-RGB color space, and all of them will be put together to form the feature of the pedestrian. The feature dimension being used in this paper is $32n$ ($16 * 2 * n$), where n is the number of the segmentation part. Figure 3 is the sketch for feature representation.

3.4 Feature matching

For feature matching, two key problems must be solved, the learning of a similarity measure function, and the determination of the final distance formula.

At present, learning a ‘‘Mahalanobis similarity function’’ (MSF) through measurement study is the dominant way for the first problem. This method fully considers the correlation between each component of the feature space, but it has inadequate ability to capture the change of individual apparent features. To strengthen the connection of the similarity function to each apparent feature, we learn a ‘‘quadratic similarity function’’ (QSF) as in (Cuevas et al. 2013). The general form of QSF is shown in formula 9, where X and Y represent the feature vector of the pedestrian image, A , B and W are the adjust matrix, and e is the offset value.

$$F(X, Y) = X^T AX + Y^T AY + X^T BX + Y^T BY + W^T X + W^T Y + e. \tag{9}$$

The key issue is to get four parameters’ values, and iterative optimization will be needed. The specific iterative process has been elaborated in (Cuevas et al. 2013), and unnecessary details will not be given in this paper. It is worth mentioning that the evolutionary algorithm is selected to replace the Newton iterative algorithm for iterative optimization in the iterative process. But, as described in Sect. 2.2, evolutionary algorithm must be improved when used for image processing due to some existing limitations. To obtain efficient and feasible objective optimization strategy, the evolutionary algorithm has been improved in the following several aspects.

First is the initial population set. The mapping function as in formula 10 will be used to initialize the population, which enables the uniform distribution of the initial population in the whole solution space. The value of θ can be identified according to the specific situation, and we use an experience value of 0.4 in this paper. Through the above process, not only the randomness of the population initialization can be remained, but also the diversity of the initial population. The individual searching periodicity can be improved as well.

$$\lambda_{n+1} = \begin{cases} \lambda_n/\theta & 0 < \lambda_n \leq \theta \\ (1 - \lambda_n)/(1 - \theta) & \theta < \lambda_n \leq 1 \end{cases} \tag{10}$$

Secondly, to reduce the computational complexity and improve the search speed, we will artificially control the search direction, which can make the search space of the optimal individuals to narrow down when the generation is increasing. In particular, for the optimal individuals x_j , its new value x_{j+1} can be calculated by formula 11, where P_c is the variable scale which is determined by the search scope and β_g is the compression factor which will be affected by

the search generation. The values of these parameters are all determined according to the specific situation.

$$x_j^{k+1} = (1 - \beta_g)x_j^k + \beta_g P_c. \tag{11}$$

Finally, to maintain the diversity of the population and prevent premature convergence, formula 12 will be used to measure the diversity of the population. When the population diversity is low (an appropriate threshold will be set according to the specific situation and relevant experience needed to determine whether the diversity is low), transfer operation (see as in formula 14) will be adopted to produce multiple new individuals, which will replace those individuals close to the optimal one. Thus, the diversity of the population can be maintained. In formula 12, n represents the feature dimension of the individuals, n_p is the population size, and s_i can be calculated as in formula 13, where $x_{j,max}$ and $x_{j,min}$ represent the upper and lower bound for the j th dimension data, respectively. In other words, they are the maximum and minimum values of the search space. For formula 14, $t\lambda$ is the adjustment coefficient, which could be calculated by the initialized mapping function (formula 10) according to the scope of interval mapping. P_{best} represents the optimal individuals, and P_{new} represents the new individuals obtained by the transfer operation.

$$div = \sum_{i=1}^n S_i / (n * n_p), \tag{12}$$

$$S_i = \sum_{j=1}^{n_p} (x_{j,i} - \sum_{j=1}^{n_p} x_{j,i} / n_p)^2 / (x_{j,max} - x_{j,min}), \tag{13}$$

$$P_{new}(j) = \begin{cases} P_{best}(j) + t\lambda(j) * (P_{best}(j) - x_{j,min}) & t\lambda(j) < 0 \\ P_{best}(j) + t\lambda(j) * (x_{j,max} - P_{best}(j)) & t\lambda(j) \geq 0. \end{cases} \tag{14}$$

When the similarity measure function has been obtained through the above learning process, this function can be used to calculate the distance in each segmentation region and get the final distance as shown in formula 15, where k is the number of block for the pedestrian image being segmented, and ε_i is the weight of each part. The key problem here is to determine the value of ε_i . The evolutionary algorithm is also adopted for objective optimization to get the optimal weight vector. The iterative process is the same as the general evolutionary algorithm, but the performance of the algorithm has been improved using the above scheme.

$$D(X, Y) = \sum_{i=1}^k \varepsilon_i F(X, Y) \tag{15}$$



Fig. 4 VIPeR dataset

4 Experiment and results

In this section, the experimental results will be presented to evaluate our approach. All of those experiments are based on two benchmark datasets which are VIPeR (Gray et al. 2007) and CAVIAR4REID (Cheng et al. 2011), respectively.

VIPeR dataset contains 632 pedestrian image pairs taken from different cameras under varying viewpoint and illumination conditions. All images are normalized to 128×48 pixels. Figure 4 shows a part of images of this dataset. It is considered the most challenging dataset available for person re-identification due to relatively low resolution, significant changes in shooting angle, illumination, etc.

CAVIAR4REID dataset is extracted from the well-known CAVIAR dataset. It is a relatively new dataset for person re-identification, which contains images of 72 pedestrians. About ten images are taken from different viewpoint and illumination conditions for each pedestrian. Figure 5 shows a part of images of this dataset. Almost 500 images for 50 pedestrians were captured by two actual surveillance cameras under various illumination conditions, which also make it a challenging dataset.

In the following part, the experimental results of our approach from several aspects will be presented. The first is the visual effect of color invariants processing, the second is the result showing adaptive image segmentation, and the last is the comparison of our completed person re-identification algorithm with some other state-of-the-art methods.

4.1 Experimental results for color invariants processing

Figure 6 is an example of the visual effect of our color invariants processing algorithm. Two images of different pedestrians are extracted from the experimental datasets randomly, and the images are obtained with the color changed through artificial processing, and then our color invariants processing algorithm is used for these images. Figure 6 shows the processing results of two sets of images. From left to right,



Fig. 5 CAVIAR4REID dataset

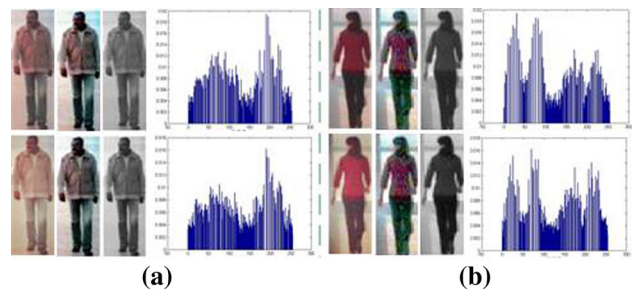


Fig. 6 Example of the experimental results for color invariants processing

they are: image being processed, color image after processing, gray image after processing, and histogram equalization after processing. We can see from Fig. 6 that our color invariants processing method has satisfactory performance, and the histogram equalization after processing is very similar.

4.2 Experimental results for adaptive image segmentation

Figure 7 is an example of the results showing our adaptive image segmentation method. From left to right, they are: image being processed, segmentation parts, and gray

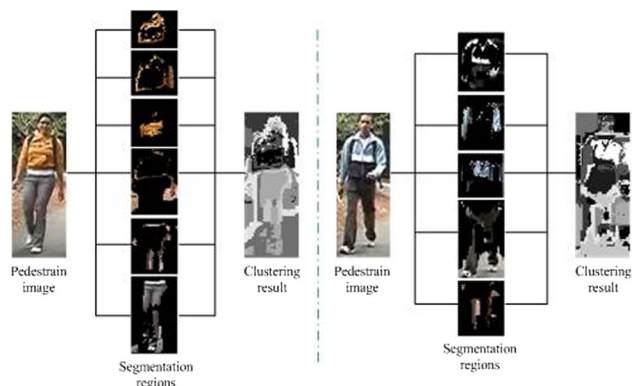


Fig. 7 Examples of the experimental results for adaptive image segmentation



Fig. 8 Examples of searching for the target pedestrian image using our method for the VIPeR and CAVIAR4REID datasets

image of the clustering result. We can see from Fig. 7 that our image segmentation method can segment the pedestrian image adaptively, that is, the number of segmentation blocks varies for the different color distribution of different pedestrian images.

4.3 Experimental results for person re-identification

The criterion used in this paper for measuring the performance of the person re-identification method is similar to most of the existing algorithms. That is, the popular cumulative match characteristic (CMC) curve is used in experimental results presented. In a given pedestrian image dataset, the CMC describes the probability of the correct match in the first n top-ranked search results. The first top-ranked match rate is important because it represents the real ability of recognition. When the value of n is very small, the n top-ranked match rate also has realistic meaning. In this situation, the first n searching results can be useful for the user's judgment.

In our experiment, some examples of searching for the target pedestrian image using our method (Fig. 8) will be given first, and then the comparison figure of our algorithm with some other state-of-the-art methods using the CMC curve (Fig. 9). These methods are SDALF (Farenzena et al. 2010),

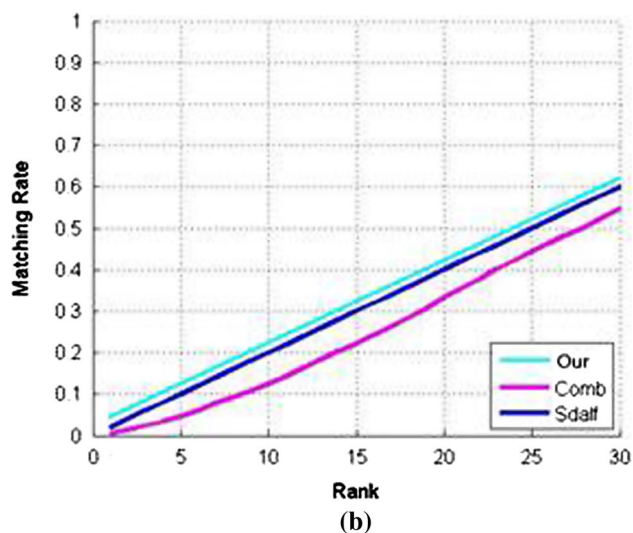
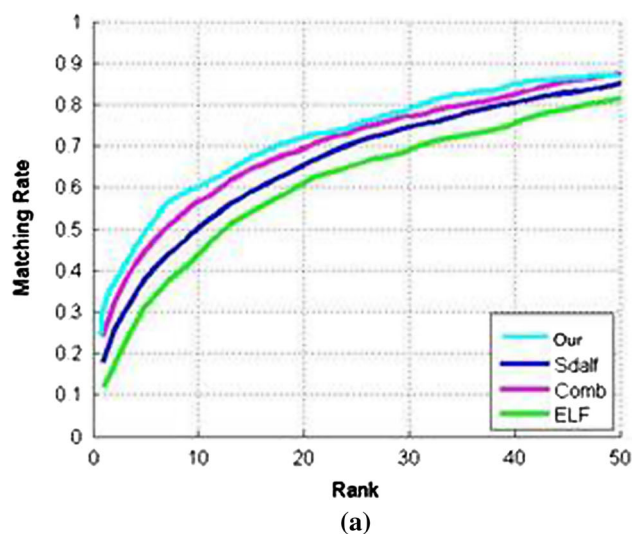


Fig. 9 The general comparison of our algorithm with some other state-of-the-art methods using the CMC curve

ELF (Gray and Tao 2008), and Comb (Kviatkovsky et al. 2013), and their CMC curve figures are obtained from (Kviatkovsky et al. 2013). For CAVIAR4REID dataset, there are a lot of images for the same person. To meet the requirements of the experiment, some images are extracted randomly from the dataset which ensures only two images for the same person.

Figure 8 shows some examples of matching results using our person re-identification method for VIPeR and CAVIAR4REID datasets. For each dataset, the leftmost column is the probe image, and the five top-ranked candidates are presented from the second to the rightmost column. Red highlighted images are the correct match. It can be seen from Fig. 8 that better matching results or valuable reference information could be given to the user by using our method.

Figure 9 shows a comparison of our method (shown as the blue curve, note for “Our”) with some other classic methods using the CMC curve. For VIPeR dataset, we compare our

method with SDALF, ELF, and Comb methods, and give the 50 top-ranked correct match rate. For CAVIAR4REID dataset, we only compare our method with SDALF and Comb methods, because of no available ELF CMC curve to use, and give the 30 top-ranked correct match rates. It can be seen from Fig. 9 that our method presents a better match rate than others. However, most images in the CAVIAR4REID dataset have low resolution, and the color information is not prominent, so the match rate of all algorithms used in our experiments are not perfect.

5 Conclusions

In this paper, a new appearance-based person re-identification method was proposed. This method used color as the feature and carried on the simple and effective color invariants processing. An adaptive image segmentation approach based on color clustering was proposed to divide pedestrian image adaptively into reasonable regions. The QSF learning method was adopted to do feature matching work, and some improvements had been made for the iterative calculation processing and threshold setting using the evolutionary algorithm.

Through the experimental results, our method has been proved to be an effective method for the person re-identification problem. The visual effect of color invariants is satisfied, the effect of image segmentation is very good, and the match rate is higher than some other similar methods.

While our work contributes to the research on person re-identification, this research has some limits. First, it is difficult to get good matching result for some datasets such as CAVIAR4REID using color feature simply. In addition, several iterative calculations are required in our algorithm. As a result, although we have controlled computational complexity in many aspects, relatively high complexity still shows up in the overall person re-identification process.

In the future work, we will further improve our algorithm to obtain a better result. At the same time, we will try to explore new calculation approaches to address the computational complexity problem.

Acknowledgements This work was jointly supported by the Natural Science Foundation of Guangdong Province of China (#2015A030313408), and Natural Science Foundation of China (#61573157).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Bak S, Corvee E, Bremond F, Thonnat M (2010) Person re-identification using spatial covariance regions of human body parts. In: IEEE Int'l conf. advanced video and signal based surveillance, pp 435–440
- Baltieri D, Vezzani R, Cucchiara R (2014) Mapping appearance descriptors on 3d body models for people re-identification. *Int J Comput Vis* 111(3):345–364
- Bazzani L, Cristani M, Perina A, Murino V (2012) Multiple-shot person re-identification by chromatic and epitomic analyses. *Pattern Recognit Lett* 33:898–903
- Bazzani L, Cristani M, Murino V (2013) Symmetry-driven accumulation of local features for human characterization and re-identification. *Comput Vis Image Underst* 117:130–144
- Bedagkar GA, Shishir SK (2012) Part-based spatio-temporal model for multi-person re-identification. *Pattern Recognit Lett* 33:1908–1915
- Bezdek JC, Ehrlich R, Full W (1984) Fcm: the fuzzy c-means clustering algorithm. In: *Intl. Conf. on image analysis and recognition*, vol 10, pp 191–203
- Chang DX, Zhao Y, Zheng CW, Zhang XD (2012) Genetic clustering algorithm using a message-based similarity measure. *Expert Syst Appl* 39:2194–2202
- Cheng DS, Cristani M, Stoppa M (2011) Custom pictorial structures for re-identification. In: *BMVC2011*, vol 2, pp 1–6
- Chu J, Miao J, Zhang GM, Wang L (2013) Edge and corner detection by color invariants. *Opt Laser Technol* 45:756–762
- Cuevas E, Echavarría A, Zaldivar D (2013) A novel evolutionary algorithm inspired by the states of matter for template matching. *Expert Syst Appl* 40:6359–6373
- Du YN, Ai HZ (2014) A statistical inference approach for person re-identification. *J Electron Inf Technol* 36(7):1612–1618
- Fan CX, Zhu H, Lin GF, Luo L (2013) Person re-identification based on multi-features. *J Image Graph* 18:711–717
- Farenzena M, Bazzani L, Perina A, Murino M, Cristani M (2010) Person re-identification by symmetry-driven accumulation of local features. In: *IEEE conference computer vision and pattern recognition*, pp 2360–2367
- Funt BV, Finlayson GD (1995) Color constancy color indexing. In: *IEEE trans on pattern analysis and machine intelligence*, vol 17, pp 522–529
- García J, Patricio MA, Berlanga A (2011) Fuzzy region assignment for visual tracking. *Soft Comput* 15:1845–1864
- Gevers T, Smeulders A (1999) Color based object recognition. *Pattern Recognit* 32:453–464
- Graham DF, Bert S, James LC (1998) Comprehensive colour image normalization. In: *Proceedings of the 5th European conference on computer vision*, vol 1, pp 475–490
- Gray D, Tao H (2008) Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: *ECCV2008*, pp 262–275
- Gray D, Brennan S, Tao H (2007) Evaluating appearance models for recognition, reacquisition, and tracking. In: *IEEE International workshop on performance evaluation for tracking and surveillance*, vol 3
- Hahnel M, Klunder D, Kraiss KF (2004) Color and texture features for person recognition. In: *Neural networks 2004 proceedings*, vol 1
- Harandi M, Salzmann M, Porikli F (2014) Regman divergences for infinite dimensional covariance matrices. In: *IEEE conference on computer vision and pattern recognition*
- Hartigan JA, Wong MA (1979) Algorithm as 136: a k-means clustering algorithm. *Appl Stat*, pp 100–108
- Holland JH (1992) Adaptation in natural and artificial systems: an introductory analysis with application to biology. In: *Control and arterial intelligence*

- Joachims T (2002) Optimizing search engines using click through data. In: Proceedings of the eighth ACM SIGKDD international conference, pp 133–142
- Johnson SC (1967) Hierarchical clustering schemes. *Psychometrika* 32:241–254
- Kviatkovsky I, Adam A, Rivlin E (2013) Color invariants for person reidentification. *IEEE Trans Pattern Anal Mach Intell* 35:1622–1634
- Li J, Li JW, Chen XF, Jia CF, Lou WJ (2015) Identity-based encryption with outsourced revocation in cloud computing. *IEEE Trans Comput* 64:425–437
- Maesschalck R, Jouan-Rimbaud D, Massart D (2000) The mahalanobis distance. *Chemom Intell Lab Syst* 50:1–18
- Prosser B, Zheng WS, Gong S, Xiang T (2010) Person re-identification by support vector ranking. In: Proc. British machine vision conf. vol 1, pp 1–5
- Venkatesan S, Madane SR (2012) Face recognition system with genetic algorithm and ant colony optimization. *Int J Innov Manag Technol* 1:469–471
- Wang X, Doretto G, Sebastian T, Rittscher J, Tu P (2007) Shape and appearance context modeling. In: Proc. IEEE Int'l Conf. computer vision, pp 1–8
- Weijer J, Sehnid C (2006) Blur robust and color constant image description. In: Proc. of Int. Conf. on image processing, pp 993–996
- Yao B, Hagrais H, Alhaddad MJ (2015) A fuzzy logic-based system for the automation of human behavior recognition using machine vision in intelligent environments. *Soft Comput* 19:499–506
- Zheng WS, Gong S, Xiang T (2013) Reidentification by relative distance comparison. In: IEEE transaction on pattern analysis and machine intelligence
- Zhou H, Chen Y, Feng R (2013) A novel background subtraction method based on color invariants. *Comput Vis Image Underst* 117:1589–1597
- Zhu W, Liang S, Wei Y (2014) Saliency optimization from robust background detection. In: IEEE conference on computer vision and pattern recognition, pp 2814–2821