PEDESTRIAN RE-IDENTIFICATION USING COLOR FEATURE IN MULTI SURVEILLANCE VIDEO

[1]GUODONG ZHANG, [2]KAZUYUKI MATSUMOTO, [3]MINORU YOSHIDA, [4]KENJI KITA

[1,2,3,4]Faculy and School of Engineering, Tokushima University, Tokushima, 7708506, Japan
E-mail: [1]zhang-g@hotmail.co.jp, [2]mino@is.tokushima-u.ac.jp, [3]matumoto@is.tokushima-u.ac.jp, [4]kita@is.tokushima-u.ac.jp

Abstract— In this paper we present a system to solve the problem of moving pedestrian re-identification in surveillance video. Surveillance video has low-resolution, high video noise and limited monitoring scope. Our proposed framework must deal with several problems such as variations of illumination conditions, poses and occlusions. How to extract the robust feature that can adapt the problems have been the task. The people of global color approaches do not change in the process of monitoring. Our paper use the color histogram as feature descriptors and choose RGB HSV and UVW for color space. Traditional histogram method extract the global color approach as the feature. The object color structure information will be neglected. We use the SPM model supplement the structure information for the histogram. The results of a test from a real surveillance system show that our method can provide a probability of matching.

Keywords— Pedestrian Re-identification, UWV Color space, Color Histogram, SPM.

I. INTRODUCTION

With the development of society more and more surveillance cameras were set up in public places. Such as airport, supermarket and so on. Pedestrian re-identification is a hot area of research for several years. Many cameras were set in intersections for traffic surveillance. How to find the object in the mass data in the recorded data to meet the demands of users have been a problem.

The moving persons in the video due to the limited conditions, such as illumination, camera resolution, it is difficult to extract the facial information from the low resolution images. Face recognition also cannot be applied in Low Resolution Image.

Relative to texture and shape features have in low resolution[1-3], color feature has a better immunity to the noise jamming of images and has a good robustness for the image degradation and image size change .

We propose the color global approaches features. Extracting color information of person in a global feature makes the method clear and simple[9]. The color histogram has lost the information of color spatial distribution. We combine the SPM model to supplement the color histogram short coming, and test our method by the RGB, HSV and UVW color space in a real video images[4,5,8].

The outline of this paper is as follows. Section 2 gives an extensive literature overview, section 3 describes our method, section 4 shows the experiments and results, and finally, section 5 presents the conclusions.

II. RELATED WORK

Recently, the person re-identification problem became one of the most important tasks in video surveillance. J.yuk,k.wong first extracts moving objects from videos by an efficient motion segmentation method, and match the query with the features indexed in the database without re-processing the videos[6]. They used Kalman filter tracker to the object tracking. The system performs the retrieval by snapshot/photo images quite well especially when the objects have sharp dominant colors. However, if the dominant colors tend to be white/black/grey, the system may not be able to distinguish them due to the limitation of the hue component in differentiating gray levels. Thi-Lan Le base in the reach of Y.Ma[7], she propos the framework for surveillance video indexing using objects features and semantic events. And support an interactive module that allows users to formulate easily their queries by different ways (existing indexed objects, sub image example, feature generation); more specifically, interactive feature generation (currently color histogram and trajectory) gives a facility for users to make queries at different levels according to the a priori available information and the expected results from retrieval.

In this paper we base on the research status to try use different method to compare the methods of pedestrian re-identification in complex background

III. SYSTEM FRAME WORK

3.1. The View of System

- key-Frame extraction

We use the background subtraction and the frame difference for detecting the moving person in video. For the read frame in the video, we calculate the inter frame difference between frames. Save the key-Frame in video.

- Object detection

For the key-frames have saved we retrieve objects are the moving person, that is not the background image. We target by reducing background method to detect the prospect, and marked with the rectangular box.

For the image of Key-frame often contains multiple moving person. We save the corresponding object by marking tags as Ip1Image form that can differentiate the object in one key-frame .

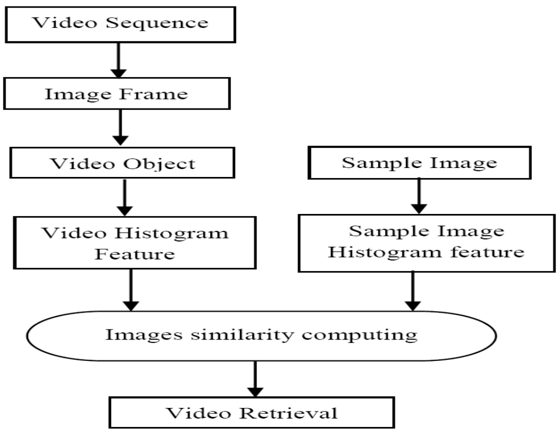


Fig.1. The Structure of Proposed System

- Feature extraction and similarity measure

The retrieval object would compute the image similarity with the object of key-frame one by one. We could sort the results in terms of similarity.

3.2. Perception-based Color Space histogram feature

Based RGB and HSV color space cannot solve the problem that is sensitive to the background illumination[10,11]. The color spaces always affect the computing accuracy of color histogram. We try to use perception-based color space in color histogram that have a good performance in image processing. As the name suggests Perception-based Color Space associated metric approximates perceived distances and color displacements capture relationships that are robust to spectral changes in illumination.

RGB color space can be transformed to Perception-based Color Space includes follows steps:

1. From RGB to XYZ color space transform formula as shown (3-1):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.177}\begin{pmatrix} 0.49 & 0.361 & 0.20 \\ 0.177 & 0.0812 & 0.011 \\ 0.00 & 0.01 & 0.99 \end{pmatrix}\begin{bmatrix} G(R) \\ G(G) \\ G(B) \end{bmatrix} \quad (3\text{-}1)$$

G( )is Gamma-correction function, and =2.0, Gamma correction function is put forward the color distortion, and rediscover the real environment to a certain extent.

2. Transform XYZ to UVW color space.

In UVW color space, the lighting condition influence is simulated by the multiplication of tristimulus values and scale factor. As shown in the formula:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \begin{bmatrix} U \\ V \\ W \end{bmatrix} = B^{-1}DB\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3\text{-}2)$$

D is a diagonal matrix, it only reck with illumination, and independent of the material. B is the transfer matrix from current color coordinates to base coordinates. The Nonlinear transfer can use(3-3):

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = A(\ln(B\begin{bmatrix} X \\ Y \\ Z \end{bmatrix})) \quad (3\text{-}3)$$

Where A and B are invertible 3 × 3 matrices and denotes the component-wise natural logarithm. The matrix B transforms color coordinates to the basis in which relighting (best) corresponds to multiplication by a diagonal matrix, while the matrix A provides degrees of freedom that can be used to match perceptual distances. Base on the similar color experiment in the database, A and B matrix-value formula shown as(3-4)and(3-5) :

$$A = \begin{pmatrix} 27.07439 & -22.80783 & -1.806681 \\ -5.646736 & -7.722125 & 12.86503 \\ -4.163133 & -4.579428 & -4.576049 \end{pmatrix} \quad (3\text{-}4)$$

$$B = \begin{pmatrix} 0.9465229 & 0.2946927 & -0.1313419 \\ -0.117917 & 0.9929960 & 0.007371554 \\ 0.0923046 & -0.046457 & 0.9946464 \end{pmatrix} \quad (3\text{-}5)$$

3.3. SPM Model

Lazebnik et al propose the partition-hierarchical model in 2006.SPM model coarse spatial information Based on spatial consider the information in color histogram in order to restrain the no order of space information.

The model divides the image into different levels. The divied levels also can be further refined.

SPM model space shown as (3-4)

As shown in fig 3, Level0 image P base on the original image feature information. But the image feature is based on the global unordered color information, Leve1 show image be separated as space geometry. P11 and P12 is expressed by spatial order that contain simple space information.

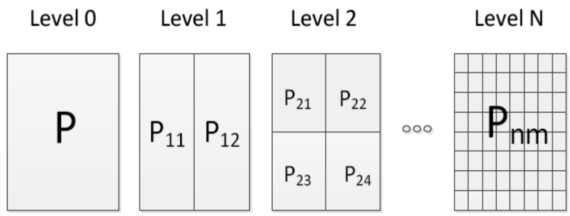


Fig.2. SPM Model Construction

P11 and p12 which also have not the space information internally part that in level 1.if we need internal space information in p11 and p12. We need to separate them as the same as before and so on. Level i+1 feature is divided by level i. The levels of the division is decided for the actual situation.

- The SPM histogram feature

Image similarity computing are computed by the levels corresponding parts in SPM model. As two images P and Q show the formula:

$$d(P,Q) = \sum k_{ij} d(p_{ij}, q_{ij}) \quad (3\text{-}6)$$

In the formula, $\mathbf{p}_{ij}$ means the image p histogram feature which of the part j in level i. $d\left(p_{ij}, q_{ij}\right)$ means the feature-similarity degree image P and Q. $\mathbf{k}_{ij}$ means the weight of similarity calculation. In case we focus the part j of level i. The weight of calculation should be set high.

3.4. GMM Model

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum A Posteriori (MAP) estimation from a well-trained prior model.

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (3\text{-}7)$$

The essential attribute of Gaussian distribution is used by mean values, and variance. Gaussian distribution is Gaussian probability density function. GMMs are often used in biometric systems, most notably in speaker recognition systems, due to their capability of representing a large class of sample distributions. One of the powerful attributes of the GMM is its ability to form smooth approximations to arbitrarily shaped densities. The classical uni-modal Gaussian model represents feature distributions by a position (mean vector) and an elliptic shape (covariance matrix) and a vector quantizer or nearest neighbor model represents a distribution by a discrete set of characteristic templates. A GMM acts as a hybrid between these two models by using a discrete set of Gaussian functions, each with their own mean and covariance matrix, to allow a better modeling capability. The GMM not only provides a smooth overall distribution fit, its components also clearly detail the multi-modal nature of the density.

IV. EXPERIMENT AND RESULT

- Evaluation Criterion

Effective of feature extraction and matching method are the crucial component of the video retrieval. It is important to find the way that can evaluate the multiple features and methods. Image correlativity and correlation are used to the video retrieval if there are two image retrieval schemes.

While there are K retrieval results that be sorted, we give different right-values by the sorted result w. The top result could get the higher weight. The top result has a better performance than the end-term. So on we can give K results the weight Wi to K, K-1,... Ki+1,…,2,1.

We define the relevancy of result Re is:

$$\mathrm{Re} = \sum_{n=1}^{k} W_n \bullet V_n \qquad (4\text{-}1)$$

And define the correlation Pr is:

$$\Pr = \sum_{n=1}^{k} W_n \bullet V_n \Big/ \sum_{n=1}^{k} W_n \qquad (4\text{-}2)$$

The relevancy and correlation can give an objective and comprehensive appraisal.

- Experimental setup

In this section, we evaluate the method in the data that we recorded in the crossroad of scene video.

Table 1: The Date size and object sample

Video Data		
Resolution	704 576	704 576
Object number	2486	3273
Object sample		
Retrieval object		

The resolution of the video data is 704 576, and it contains 6759 objects. We show the object sample in fig 3 as pictures in table 1. We list the histogram distance of retrieval projects in table 2 that based color space of RBG.

Table 2: The Result of Retrieval Object

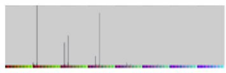
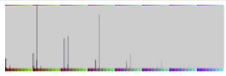
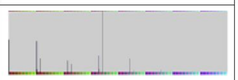
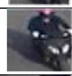
Retrieval Project	Sample	Histogram Distance	Histogram
Retrieval Object			
Top1		0.820	
Top2		1.035	
Top3		1.357	

Table 3: The Similarity Measure of Retrieval Object

Retrieval Object	Similarity Measure
	1.11102
	1.27508
	1.45655
	1.66188

CONCLUSIONS

In this paper we proposed an efficient method for re-identification. The re-identification is based on assigning a global appearance-based feature. We use the Perception-based Color Space to improve the weakness of the color histogram that color shift. The method is also higher than a method that uses only a color histogram as the signature. We also discuss the SPM model to improve the lack of spatial information. In some case it improved the accuracy of the image computing. We simply extract the main characteristics of color feature based GMM model, and have not analyze the information which be contained in the mean and covariance paramenters.in the future research we could consider the feature information. The GMM model also is based on the image of global information. We could challenge to combine the GMM and SPM model to improve the lack of spatial information.

ACKNOWLEDGMENTS

The research was partially supported by JSPS KAKENHI Grant Numbers 15K00425.

REFERENCES

[1] T. Madden, C., and Massimo Piccardi. "A framework for track matching across disjoint cameras using robust shape and appearance features." Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on. IEEE, 2007

[2] Chong, Hamilton Y., Steven J. Gortler, and Todd Zickler. "A perception-based color space for illumination-invariant image processing." ACM Transactions on Graphics (TOG). Vol. 27. No. 3. ACM, 2008.

[3] Lu, Yanyun, et al. "Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features." Neurocomputing 126 (2014): 132-140.

[4] Bąk, Sławomir, et al. "Boosted human re-identification using riemannian manifolds." Image and Vision Computing 30.6 (2012): 443-452.

[5] Metternich, Michael J., Marcel Worring, and Arnold WM Smeulders. "Color based tracing in real-life surveillance data." Transactions on data hiding and multimedia security V. Springer Berlin Heidelberg, 2010. 18-33.

[6] Mekonnen, Alhayat Ali, Frédéric Lerasle, and Ariane Herbulot. "Cooperative passers-by tracking with a mobile robot and external cameras." Computer Vision and Image Understanding 117.10 (2013): 1229-1244.

[7] Yuk, Jacky SC, et al. "Object-based surveillance video retrieval system with real-time indexing methodology." Image Analysis and Recognition. Springer Berlin Heidelberg, 2007. 626-637.

[8] Le, Thi-Lan, et al. "A framework for surveillance video indexing and retrieval."Content-Based Multimedia Indexing, 2008. CBMI 2008. International Workshop on. IEEE, 2008.

[9] Bazzani, Loris, et al. "Multiple-shot person re-identification by chromatic and epitomic analyses." Pattern Recognition Letters 33.7 (2012): 898-903.

[10] Berdugo, Guy, et al. "Object reidentification in real world scenarios across multiple non-overlapping cameras." Proc. Euro. Sig. Proc. Conf. 2010.

[11] Allouch, A. X. E. L. L. E. "People re-identification across a camera network." Master of Science Thesis, KTH Royal Institute of Technology, Stockholm, Sweden (2010).

★★★