PROCEEDINGS OF SCIENCE

Gait Feature Recognition Based on Multi-instance Learning

Chaoyu Yang[1,2]
School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing,100083 China
School of Economics and Management,Anhui University of Science and Technology,Anhui,232001, China
E-mail: `yangchy@aust.edu.cn`

Renjie Geng
School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing,100083 China

Qian Liu
School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing,100083 China

In this paper, we extend multi-instance learning, propose gait feature extraction and the recognition model based on multi-instance learning. We divide image into multiple regions. The image is regarded as multi-instance learning bag and the region is regarded as instance in the bag. We calculate the diverse density function value of each region in all images and extract possible positive region to constitute a set and use SVM to classify the gait feature. Experiments on image sets show that multi-instance learning is applied to the gait recognition and has better feature learning and classification effect.

CENet2015
12-13 September 2015
Shanghai, China

[1]Speaker
[2]The research work was supported by National Natural Science Foundation of China under Grant No.: 51404007, 51374114 and 51474007.

© *Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).*

1. Introduction

The gait recognition uses the walking style to identify different persons. The gait recognition is a new research field of computer vision. As the overall movement of the human body during walking, it is the basic activity in human life activity. Normal gait will not constrain walking at a normal pace, which features coordination and quasi periodicity[1]. The gait recognition system includes gait video sequence acquisition, gait detection, gait feature extraction and classification and recognition; and the gait study mainly depends on video to analyze the gait features. Shutler et al. proposed statistical method of time distance feature to recognize the gait features. HayfronAcquah et al. proposed a gait recognition algorithm based on human symmetry analysis to achieve automatic gait classification[2]. Ben-Abdlkader of University of Maryland extracted gait features by using gait sequence self-similarity graph. Carnegie-Mellon University's Colins relied on the collected information of human contour body to determine the human gait pattern. Foster from University of Southampton solved problem of gait recognition by using the method of region measurement strategy.

2. Multi-instance Learning Method

T. G. Ddietterich etc. proposed the concept of multi-instance learning in the research on drug activity prediction and gave three algorithms for carrying out experiments on Musk molecular data set in 1990's[3]. Afterwards, the research scholars proposed many practical multi-instance learning algorithms.

Multi-instance learning framework can be described as follows. The training set consists of a number of bags with tag. Each bag contains a number of no-tag instances. If a bag has a positive instance at least, the bag is tagged as positive. If all the instances in a bag are negative, the bag is tagged as negative. The multi-instance learning gets a learning system by learning from training bag and the learning system can predict tags of bags outside the training set[4]. The gait feature recognition is a multi classification problem. By introducing multi-instance learning ideas to the gait feature recognition, we use images containing gait features as a bag X_i and extract local area features in the image as the sample of the package by using the bag generation method. Bag X_i contains several instances $x_{i,1}, x_{i,2} \cdots x_{i,j}, \cdots$. Each instance $x_{i,j}$ is described by a number of attributes. $x_{i,j,k}$ represents the value of the first k attribute for the j instance of the i package. And then, according to the category of gait feature in the image, a concept tag Y_i is given to bag. Each gait feature classification has only one Y_i corresponding to it. (X_i, Y_i) is considered as a sample of multi-instance learning. The result of learning is to predict the new image and the classification of the bag so as to get the gait feature categories contained in the image.

3. Gait Feature Recognition based on Multi-instance Learning

In this paper, the image is segmented into a plurality of regions. Image is regarded as multi-instance learning bag, and region is regarded as instance in the bag. We calculate the diverse density function value of each region in positive image and extract possible positive region (region of interest) to constitute a set and use the SVM method for learning.

3.1 Image Segmentation and Feature Extraction

We use k-means clustering segmentation method to divide image into multiple regions. At first, the image is divided into 16*16 pixels blocks without overlapping. We extract block's average color feature of three dimension LUV. L encode brightness, U and V encode the color saturation information. In order to extract texture features of image, the small 16*16 pixel blocks carry on the first level Daubechies-4 wavelet transform, decompose into four sub bands. Each sub-band contains a 2*2 coefficient. The three sub-band coefficients of LH, HL and HH

PoS(CENet2015)084

reflect the variation information of small blocks in vertical, horizontal and diagonal directions respectively. We extract the root-mean-square values of their coefficients as one-dimensional features. Three dimension features have to be extracted. For example, the coefficient of the HL sub-band is assumed to $\left\{c_{k,j}, c_{k,j+1}, c_{k+1,j}, c_{k+1,j+1}\right\}$; thus a feature is. $f = (\frac{1}{4}\sum_{i=0..1}\sum_{j=0..1} c_{k+i,k+j}^{2})$ We use k-means algorithm to cluster blocks. The clustering blocks are considered as a region with similar color and texture features. Region feature is average value of multiple features of small blocks[5]. The k value of k-means algorithm is not specified, k value is set from the two values, then increase by 1 until the K value reaches a stop condition. k value is stopped and eventually k value is determined.

3.2 Select Positive Sample Collection

In the multi-instance learning, the tag information is labeled at the bag level, and the sample is not labeled. The positive bag has a positive sample at least and all samples in negative bag are negative examples.

DD algorithm of multi-instance learning regards each bag as a set constituted by instance. The algorithm finds a conceptual point in the feature space[6]. The concept point is close to an instance in positive bag, all instances in negative bag is far from that point. After we find the point, we can regard this point as a reference point to determine new bag tag. B_i^+ represents the i positive bag, B_{ij}^+ represents the j instance of i positive bag, and B_{ijk}^+ represents the first k attribute values for the j instance of i positive bag. Similarly, B_i^- represents the i negative bag, B_{ij}^- represents the j instance of i negative bag, and B_{ijk}^- represents the first k attribute values for the j instance of i negative bag. Concept point t indicates the most value point of DD function. The target concept point t is determined by the target concept point by maximizing objective function $\Pr(x = t \mid B_1^+, B_2^+, \ldots, B_l^+, B_1^-, B_2^-, \ldots, B_{l-}^-)$. We assume that each bag is independent of each other. According to Bias theory, the maximum point t of DD function can be determined by the following formula.

$$DD(t|L) = \arg\max_t \prod_i Pr(t|B_i^{+1}) \prod_j Pr(t|B_i^{-1}) \quad . \tag{3.1}$$

L is the training set, $\prod_i Pr(t|B_i^{+1}) \prod_j Pr(t|B_i^{-1})$ is called the diversity density function, that is, the DD function. In practical solution, the production term of the formula is quantified by using the Noisy-OR model that is converted into the following formula.

$$Pr(t|B_i^{+1}) = Pr(t|B_{i1}^{+1}, B_{i2}^{+1} \ldots B_{3}^{+1}) = 1 - \prod_j (1 - Pr(t|B_{ij}^{+1})) \tag{3.2}$$

$$Pr(t|B_i^{-1}) = Pr(t|B_{i1}^{-1}, B_{i2}^{-1} \ldots B_{3}^{-1}) = 1 - \prod_j (1 - Pr(t|B_{ij}^{-1})) \tag{3.3}$$

In the formula above, probability $\Pr(t \mid B_{ij})$ can be viewed as probability of potential target and distance between target concept point t and instance B_{ij} can be calculated by $\Pr(t \mid B_{ij})$. The formula is shown as follows.

$$\Pr(t \mid B_{ij}) = \exp(-\| B_{ij} - t \|^2) \tag{3.4}$$

DD algorithm uses the gradient ascent method to solve the maximum value. In order to avoid falling to the local optimal solution and make sure to find the maximum value, the initial point is set to each positive instance in all positive bags and each instance performs a search.

3.3 Multiple-sample and Multiple-valued SVM Classifier

Support vector machine(SVM) was firstly proposed by Vapnik and his collaborators. SVM shows many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition[7]. Generally speaking, SVM's goal is to find optimal classification surface. Its

PoS(CENet2015)084

means is quadratic programming. SVM is a machine learning method based on nonlinear mapping theory. SVM was firstly developed from the linear separable optimal classification. Taking two classifications of samples, we assume that existing sample $(x_1, y_1), (x_2, y_2), ...(x_i, y_i)......$, $x_i \in R^d$, d is sample dimension, i $y_i \in (+1, -1)$ s sample mark. Optimal hyper plane is divide into the positive and negative sample, and distance between positive and negative sample. Plane is furthest. The optimal hyper plane, that is, the classification function is shown as follows.

$$f(x) = sgn(\langle w \cdot x \rangle + b) = sgn(\sum_{i=1}^{N} y_i a_i \langle x_i \cdot x \rangle + b) \tag{3.5}$$

However, the problem of gait feature is often nonlinear. We map raw data to feature space of higher dimensions, design linear SVM in high dimensional space. The general form of nonlinear support vector machine is:

$$f(x) = sgn(\sum_{i=1}^{N} y_i a_i K \langle x_i + x \rangle + b) \tag{3.6}$$

$$b = -\frac{1}{2} \sum_{i=1}^{N} a_i y_i [K(x_i, x_r) + K(x_i, x_s)] \tag{3.7}$$

In this paper, we use SVM method of multi-instance learning to solve the problem of gait recognition. We transform the multi instance learning problem to the single instance learning problem so that each bag contains only one instance. SVM learning objective function is shown as follow.

$$\begin{aligned} &max\, imise\, W(a) = \sum_{i=1}^{N} a_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j a_i a_j K(\Phi(X_i), \Phi(X_j)) \\ &subjectto \sum_{i=1}^{N} y_i a_i = 0 \qquad a_i \geqslant 0,\ \forall i \end{aligned} \tag{3.8}$$

We consider that gait feature recognition is a multi classification problem, and SVM is two classification learning algorithm, so we construct SVM multi valued classifier[8]. We assume that M kinds of gait features need to be classified, X_i is the bag of multi-instance learning, and Y_i is the concept tag $Y_i \in \{1, 2, \ldots, m\}$. We construct multi-instance multi-value SVM classifier by using the decision tree method, which divides all classes into two sub classes at first, and then divides the two sub classes into two secondary sub-classes for further loop until all nodes contain only one class. Thus we constitute a decision tree structure identification system. We need to construct M-1 classifier for M kinds of classes of identification problem. The decrease of number of classifiers can reduce the risk of error classification in certain extent.; however, this structure determines upper classification error genetic, so it has lower fault tolerance. By defining the kernel function $K(x, x_i)$ which satisfies the Mercer condition, we convert the point production problem into the input space. In practical application, we use the Gauss radial basis function $K(x, x_i) = \exp(-\frac{|x - x_i|^2}{2\sigma^2})$.

4. Experiment Results and Analysis

In the public video data KTH, we compare the gait recognition rate of the method proposed in this paper with the gait recognition rate of common method, including BP artificial neural network and K-Nearest Neighbor. We set the value of K to 4, so we need to find four nearest training samples from test samples. Convolutional neural network uses four layers, the first layer S1 is convolution and virtual layer, the second layer C2 is sampling layer, and so on. The size of convolution kernel is 5*5. The comparison results are shown in Table 1. We can see the gait recognition rate of the method is higher than other common methods apparently.

PoS(CENet2015)084

No.	Action categories	K-Nearest Nerighbor	BP ANN	Multi-instance learning
1	Stand	87.3%	86.6%	93.8%
2	WalkSlowly	86.9%	85.8%	93.2%
3	WalkQuickly	85.2%	84.7%	92.6%
4	Upstairs	82.1%	83.2%	91.7%
5	Downstairs	82.6%	83.6%	91.4%
	Average	84.82%	84.78%	92.54%

Table 1: Comparison of Action Recognition Rate for Each Algorithm

Fig. 1 shows experimental result of algorithm proposed in this paper in KTH data set in different environments. Scene 1, Scene 2 and Scene 3 is the environment indoors, Scene4 is outdoor experiments. Because of external environment, Scene4's identification effect is the worst.

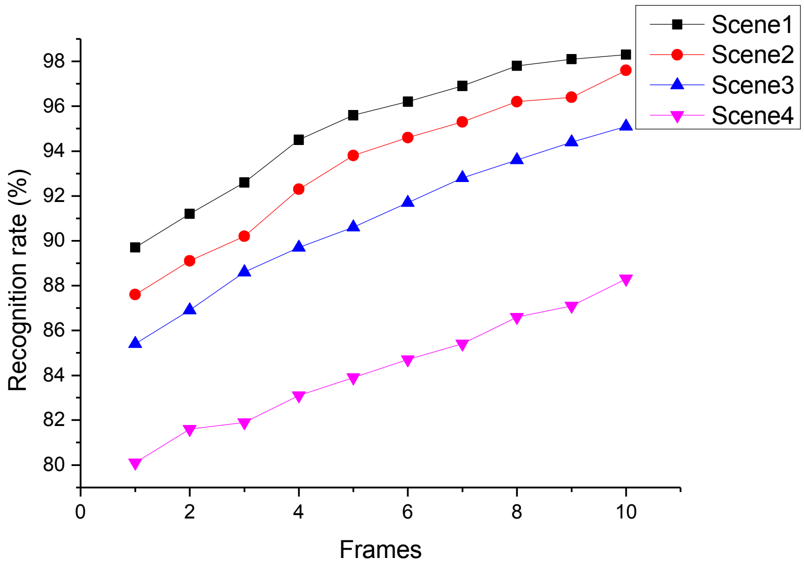

Figure 1: Recognition Rate of Different Scenes

Aiming at different walking motions, the experiment classified 100 images, judged the gait features, compared the consumption time of multi-instance learning algorithm and common algorithm. Experimental results are shown in Fig. 2.

PoS(CENet2015)084

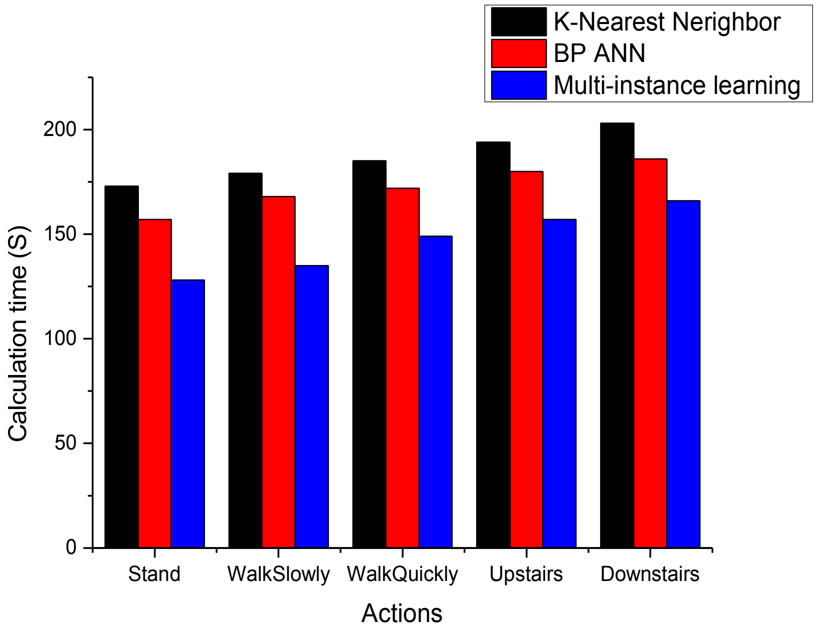

Figure2: Recognition Rate of Different Actions

The training under different quantities of training sets and different algorithms' identification accuracy are shown in Fig. 3.

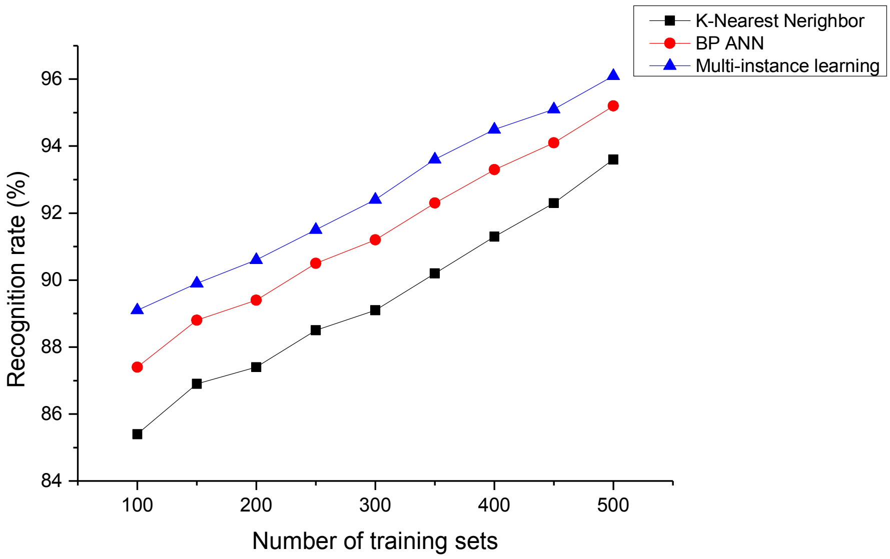

Figure 3: Recognition Rate of Different Quantity of Training Sets

5. Conclusion

Multi-instance learning is a new machine learning model. We propose the gait feature recognition method based on multi-instance learning. At first, we regard an image included gait feature as a bag, and then we use k-means method to cut apart image. We extract local features of the image as an instance of bag. After that, we mine instance of the concept tag of token bag by using the diversity density algorithm of multi-instance learning. Finally, we achieve the gait recognition by using the SVM algorithm to study simplified bag. The results show that the gait feature recognition based on multi-instance learning has better feature learning and classification ability.

PoS(CENet2015)084

References

[1] Srensen, L. Loog, M. Tax, Wan-Jui Lee, de Bruijne, M. Duin. *Dissimilarity-based multiple instance learning*[C]. Structural, Syntactic, and Statistical Pattern Recognition. Proceedings Joint IAPR International Workshop, SSPR & SPR. pp.129-138, (2010).

[2] J. J. Kavanagh, R. S. Barrett, S. Morrison. *Upper body accelerations during walking in healthy young and elderly men*[J]. Gait & posture. 20(3):291-298, (2004).

[3] M. L. Zhang, Z. H. Zhou. *Multi-instance clustering with applications to multi-instance prediction*[J]. Applied Intelligence. 31(1):47-68,(2009).

[4] Y. Yajima, T. F. Kuo, *Efficient formulations for 1-SVM and their application to recommendation tasks*[J]. Journal of Computers. 1(3):27-34,(2006).

[5] J. Foulds, E. Frank, *A review of multi-instance learning assumptions*[J], the knowledge engineering review. 25(1):1-24,(2010).

[6] A. K. Jain, A. Ross, S. Prabhaker. *An Introduction to Biometric Recognition*[J]. IEEE Transactions on Circuits and Systems for Video Technology. 14(1):4-20,(2004).

[7] G. Shakhnarovich, .L Lee, T. Darrell. Integrated face and gait recognition from multiple views[C]. IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA. pp:439-446,(2001).

[8] C. Frigo, R. Carabalona, M Dalla Mura, S Negrini. *The upper body segmentalmovements during walking by young females*[J]. Clinical biomechanics, 18(5):419-425(2003).

PoS(CENet2015)084