Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

Research on Human Walking Feature Extraction and Identification Recognition System

Xiaodong Zhao, Hui Du*

School of Information Science and Engineering, HeBei University of Science and Technology, Shijiazhuang 050018, China

ARTICLE INFO

Article history:
Received 11 February 2019
Accepted 15 April 2019
Available online 30 June 2019

Keywords:
Feature extraction
BP neural network
Gait recognition

ABSTRACT

For the feature recognition of human gait in motion state, this paper uses a contour-based method to extract gait features, first separating the target from the background, and then extracting the overall contour of the person. The characteristic of a person during walking is the distance between the person's centroid point and the pixel points at the edge of the silhouette. The BP neural network algorithm is then used to perform gait recognition based on the gait database. The results show that the gait recognition algorithm in this paper is better than other algorithms.

1. Introduction

With the development of science and technology, people hope that computers can be more intelligent and complete some complex tasks that only human intelligence can do. Human walking feature recognition is a new field in biometrics. Its identification of identity is mainly based on the walking posture and leg characteristics when people walk. The walking feature recognition of the human body is non-invasive and difficult to disguise. It does not require physical contact, but fingerprint recognition, face recognition of the face, and iris recognition require close contact to extract the desired features. In recent years, the recognition of human walking characteristics has generated great interest among researchers.

Some traditional human walking feature recognition methods extract two-dimensional features and information, pictures of people walking by the camera, and extract some leg features in the picture when people walk, such as step spacing, walking cycle and so on. These are two-dimensional gait features that are ignored when people walk or exercise, so that they do not reflect the state process of the person. For these problems, this paper extracts 3D feature parameters from the contour map. The camera adopts a single-frame gait image sequence, and extracts the three-dimensional feature data of the human walking through the overall contour map of the human body according to the structural knowledge of the human body and the knowledge of the camera calibration, and then analyzes, thereby achieving the purpose of gait recognition.

* Corresponding author.
E-mail addresses: 362183949@qq.com (H. Du)

2. Gait feature extraction

Human gait feature information is a process of human's overall contour period change, including dynamic features and static features. The distance feature of the contour has characteristics such as rotation and translation invariance. The walking process of human beings is space-time movement, the spatial aspect is reflected in the motion contour in each frame of picture, and the time aspect is reflected in the process of static characteristics changing with time. The walking process of a person changes periodically, so the overall contour of a person is periodic with time, and the contours of different people are also different. The contour features can fully obtain the changing characteristics of a person during exercise.

To reduce the effect of variations, GAN is employed as a regressor to generate invariant gait images, that contain side view gait images with normal clothing and without carrying objects. The gait images at arbitrary views can be converted to those at the side view since the side view data contains more dynamic information. While this is intuitively appealing, a key challenge that must be address is to preserve the human identification information in the generated gait images.

The Gait GAN model is trained to generate gait images with normal clothing and without carrying objects at the side view by data from the training set. In the test phase, gait images are sent to the GAN model and invariant gait images contains human identification information are generated. The difference between the proposed method and most other GAN related methods is that the generated image here can help to improve the discriminant capability, not just generate some images which just look realistic. The most challenging thing in the proposed method is to preserve human identification when generating realistic gait images.

Domain transfer Generative adversarial networks (GAN) [4] are a branch of unsupervised machine learning, implemented by a system of two neural networks competing against each other in a zero-sum game framework. A generative model G that captures the data distribution. A discriminative model D then takes either a real data from the training set or a fake image generated from model G and estimates the probability of its input having come from the training data set rather than the generator. In the GAN for image data, the eventual goal of the generator is to map a small dimensional space z to a pixel-level image space with the objective that the generator can produce a realistic image given an input random vector z. Both G and D could be a non-linear mapping function. In the case where G and D are defined by multilayer perceptrons, the entire system can be trained with backpropagation.

2.1 Gait detection and segmentation

Moving target detection and segmentation is the extraction of moving targets from video images containing both foreground and background. It is generally detected by background difference method, interframe difference method, optical flow method and energy minimization method. The background difference method is used in this paper. Gait segmentation is an important step in gait recognition. It is also the first step after the camera acquires a series of gait images. The gait contour can be estimated by gait segmentation. The human walking detection method is as follows: the moving target is detected first and detected by video. The overall contour extraction of the moving image of the target is then performed. In this paper, the median method is used to restore the background. The background image is separated from the background image by the background difference method. The difference image histogram is equalized, the iterative threshold method is used for binarization, and finally the moving pixels in the current image are segmented. The result is shown in Figure 1.

Fig.1. Human body bottom edge contour extraction

The gait energy image [6] is a popular gait feature, which is produced by averaging the silhouettes in one gait cycle in a gait sequence as illustrated in Figure 1. GEI is well known for its robustness to noise and its efficient computation. The pixel values in a GEI can be interpretted as the probability of pixel positions in GEI being occupied by a human body over one gait cycle. According to the success of GEI in gait recognition, we take GEI as the input and target image of our method. The silhouettes and energy images used in the experiments are produced in the same way as those described in [22].

Fig.2. A gait energy image is produced by averaging all the silhouettes in one gait cycle

The input of the generative model can be an image instead of a noise vector. GAN can realize pixel-level domain transfer between input image and target image such as PixelDTGAN proposed by Yoo et al. PixelDTGAN can transfer a visual input into different forms which can then be visualized through the generated pixel-level image. In this way, it simulates the creation of mental images from visual scenes and objects that are perceived by the human eyes. In that work, the authors defined two domains, a source domain and a target domain. The two domains are connected by a semantic meaning. For instance, the source domain is an image of a dressed person with variations in pose and the target domain is an image of the person's shirt. So PixelDTGAN can transfer an image from the source domain which is a photo of a dressed person to the pixel-level target image of shirts. Meanwhile the transferred image should look realistic yet preserving the semantic meaning. The framework consists of three important parts as illustrated in Figure 3. While the real/fake discriminator ensures that the generated images are realistic, the domain discriminator, on the other hand, ensures that the generated images contain semantic information.

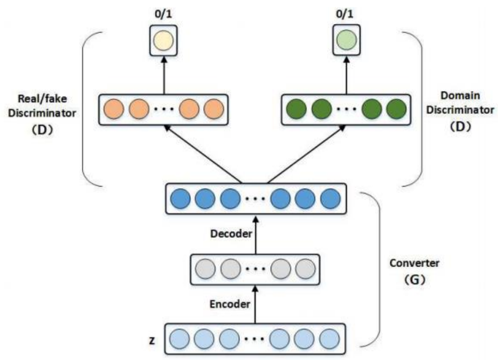

Fig.3. The framework of PixelDTGAN which consists of three important parts.

2.2 Contour-based feature extraction

The images of the obtained experimenters were detected and separated by walking, and then the images with clear and easy extraction of feature information were selected for dynamic walking feature extraction. The characteristics of different human walking processes are obviously different. This paper selects the characteristics of the leg of the experimenter as the key feature of gait recognition during the movement.

In general, human walking process recognition only requires walking characteristics of the walking sequence to be acquired within a walking cycle. Extract the contour map of the human contour based on the contour map of the lower edge of the human body and find the centroid. as shown in Figure 4:

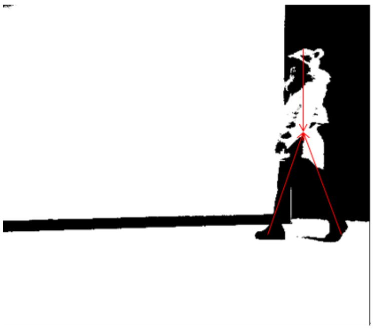

Fig.4. Centroid and contour edge map

The edge map of the human contour is two-dimensional feature information, and the edge contour can be expressed by the following formula:

$$Z_n = X_n + nY_n \tag{2-1}$$

The total number of pixels in the entire contour is n, n = 1, 2, …,N_s , and the pixel points on the edge of the contour are marked as(X_n, Y_n)。

Each contour of the human body can be expressed with information of N_s pixels. In order to facilitate the calculation, the two-dimensional gait silhouette feature can be represented by a one-dimensional feature vector. Specifically, according to the previous method, the contour map of the person is obtained, and the coordinate values of all the pixel points on the contour boundary are listed. The centroid coordinates (X_o, Y_o) of the human contour image can be calculated by the following formula:

$$X_o = \frac{1}{N_s}\sum_{n=1}^{N_s} X_n \tag{2-2}$$

$$Y_o = \frac{1}{N_s}\sum_{n=1}^{N_s} Y_n \tag{2-3}$$

Use the centroid as the origin of the coordinates, and then create an axis to calculate the distance D_o from the edge of each pixel(X_n, Y_n)to the centroid(X_o, Y_o)along a specific direction, ie:

$$D_o = \sqrt{(X_n - \boldsymbol{X_o})^2 + (\boldsymbol{Y_n} - \boldsymbol{Y_o})^2} \tag{2-4}$$

The distance signal vector of N_s pixels in the entire contour of a person is expressed by the following equation:

$$S = \{ D_1, D_2, D_3, \ldots\ldots, D_{N_s} \} \tag{2-5}$$

In Fig. 2, the point marked in the middle is the center of mass of the human body, and the arrow points to the distance signal from the centroid to the pixel edge of the human body. The size of the center distance of the contour is the distance from the centroid of the contour of the person to the edge of the edge pixel.

The distance feature vector is displayed by a histogram of **N** bins to form an N-dimensional distance vector, and the distribution of contour distance features of different people is calculated by using a histogram, thereby solving the trouble caused by the feature dimension and convenient for data processing. Since the walking process of humans is periodic, a gait sequence will contain multiple walking cycles, assuming that the eigenvectors obtained in the kth cycle are V_k, k=l, 2,..., n, then the walking characteristics of a sequence are as follows:

$$G = \frac{1}{n}\sum_{i=1}^{n} V_i \tag{2-6}$$

According to the human walking database provided by the Institute of Automation of the Chinese Academy of Sciences, the contour of the walking sequence during the walking process is extracted, and the distance signal of the contour is further studied by the histogram, as shown in Fig. 5, Fig. 6 and Fig. 7, from below. The histogram can be concluded that the histograms of different walking sequences of the same person are almost the same. The histograms of the walking sequence of different experimenters are obviously different, so according to this feature, we can identify the walking process of the human body.

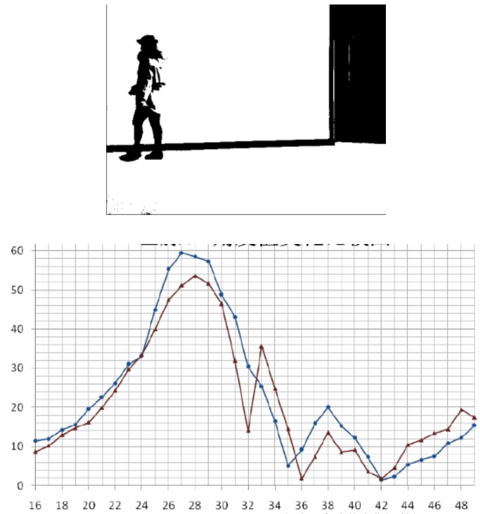

Fig.5. The gait sequence 1 of experimenter A contour map and contour center distance

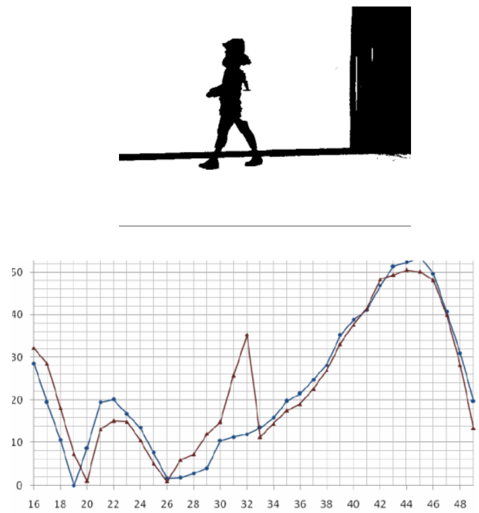

Fig.6. The gait sequence 2 of experimenter A contour map and contour center distance

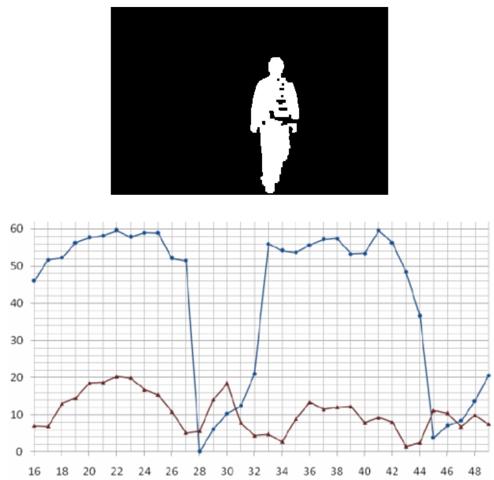

Fig.7. The gait sequence 1 of experimenter B contour map and contour center distance

2.3 Classification and identification

There are many methods for classifying the extracted features. The common ones include the following: support vector machine, K-nearest neighbor classifier and neural network classification. The BP neural network algorithm is used in this paper. The BP neural network gait recognition algorithm inputs the feature data extracted by the histogram into the neural network and then simulates it. Then, the walking sequence data of the test sample is imported into the model of the previously simulated neural network, and the similarity between the simulated network model and the input model is estimated, and the specific target with the greatest similarity between the two is found. The specific target is Test the ID number of the sample category. The gait recognition process based on BP neural network is shown in Figure 8:

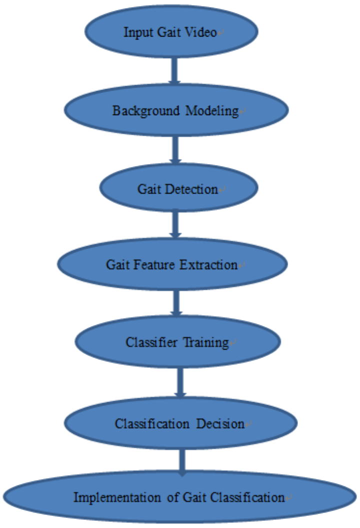

Fig.8. Gait recognition process based on BP neural network

The first important component is a pixel-level converter which are composed of an encoder for semantic embedding of a source image and a decoder to produce a target image. The encoder and decoder are implemented by convolution neural networks. However, training the converter is not straightforward because the target is not deterministic. Consequently, on the top of converter it needs some strategies like loss function to constrain the target image produced. Therefore, Yoo et al. connected a separate network named domain discriminator on top of the converter. The domain discriminator takes a pair of a source image and a target image as input, and is trained to produce a scalar probability of whether the input pair is associated or not. The loss function $L^A{}_D$ in [20] for the domain discriminator D_A is defined as:

$$L_A^D(I_S,I) = -t\log\left[D_A(I_S,I)\right] + (t-1)\log\left[1-D_A(I_S,I)\right]$$

$$s.t.\ \ t=\begin{cases}1 & if\quad I=I_T\\ 0 & if\quad I=\hat{I}_T\\ 0 & if\quad I=I_T^-\end{cases}$$

where I_S is the source image, $I_T{}^-$ is the ground truth target, the irrelevant target, and $\hat{I}_T$ is the generated image from converter。

Inspired by the pixel-level domain transfer in PixelDTGAN, we propose GaitGAN to transform the gait data from any view, clothing and carrying conditions to the invariant view that contains side view with normal clothing and without carrying objects. Additionally, identification information is preserved. We set the GEIs at all the viewpoints with clothing and carrying variations as the source and the GEIs of normal walking at 90◦ (side view) as the target, as shown in Figure 8. The converter contains an encoder and a decoder as shown in Figure9.

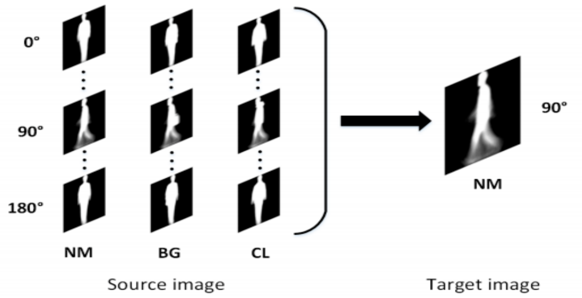

Fig.8.All the viewpoints with clothing and carrying variations as the source and the GEIs of normal walking at 90◦ (side view) as the target

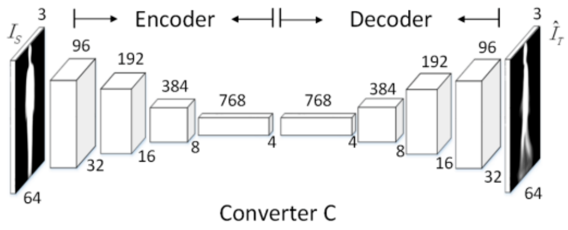

Fig.9. The converter contains an encoder and a decoder

With the real/fake discriminator, we can only generate side view GEIs which look well. But, the identification information of the subjects may be lost. To preserve the identification information, another discriminator, named as identification discriminator, which is similar to the domain discriminator in [20] is involved. The identification discriminator takes a source image and a target image as input, and is trained to produce a scalar probability of whether the input pair is the same person. If the two inputs source images are from the same subject, the output should be 1. If they are source images belonging to two different subjects, the output should be 0. Likewise, if the input is a source image and the target one is generated by the converter, the discriminator function should output 0.

The whole human body walking feature program is written and simulated by the software Matlab, and the gait recognition is performed according to the above method. The data in this experiment is based on data from the gait database of the Chinese Academy of Sciences. In order to ensure the accuracy of human body walking feature recognition, it is necessary to carry out repeated experiments. In this paper, the classification correct rate CCR is used as the evaluation standard for walking recognition:

$$CCR = \frac{N_r}{N} * 100\%$$

N_r is the correct number of samples, and N is the total number of tests samples.

Tab.1 gait recognition algorithm accuracy rate

Recognition algorithm	Gait perspective 0°	Gait perspective 45°	Gait perspective 90°
BP neural network	81.33%	87.33%	86.83%

Angle histogram algorithm	70.37%	73.15%	84.90%

From the recognition algorithm, the recognition algorithm of this paper is superior to the gait recognition algorithm based on angle histogram in the recognition rate, which proves that the neural network algorithm has great advantages compared with other gait recognition algorithms.

3. Experiments and analysis

3.1. Dataset

CASIA-B gait dataset [22] is one of the largest publicgait databases, which was created by the Institute of Automation, Chinese Academy of Sciences in January 2005. It consists of 124 subjects (31 females and 93 males) captured from 11 views.

The view range is from 0Æto 1 with 18 interval between two nearest views. There are 11 sequences for each subject. There are 6 sequences for normal walking (“nm”), 2 sequences for walking with a bag (“bg”) and 2 sequences for walking in a coat (“cl”).

3.2. Experimental design

In our experiments, all the three types of gait data including”nm”, “bg” and “cl” are all involved. We put the six normal walking sequences, two sequences with coat and two sequences containing walking with a bag of the first 62 subjects into the training set and the remaining 62 subjects into the test set. In the test set, the first 4 normal walking sequences of each subjects are put into the gallery set and the others into the probe set. There are four probe sets to evaluate different kind of variations.

3.3. Model parameters

In the experiments, we used a similar setup to that of[20], which is shown in Figure 8. The converter is a unified network that is end-to-end trainable but we can divide it into two parts, an encoder and a decoder. The encoder part is composed of four convolutional layers to abstract the source into another space which should capture the personal attributes of the source as well as possible. Then the resultant feature z is then fed into the decoder in order to construct a relevant target through the four decoding layers. Each decoding layer conducts fractional stride convolutions, where the convolution operates in the opposite direction. The details of the encoder and decoder structures. The structure of the real/fake discriminator and the identification discriminator are similar to the encoder’s first four convolution layers. The layers of the discriminators are all convolution layers.

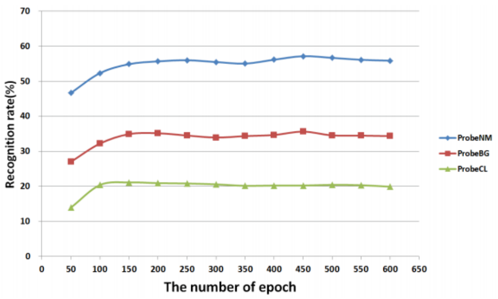

Fig.10. The recognition rate as a function of the number of iterations.

Normally to achieve a good performance using deep learning related methods, a large number of iterations in training are needed. From Figure 10, we can find that more iterations can indeed result in a higher recognition rate, but the rate peaks at around 450 epoches. So in our experiments, the training was stopped after 450 epoches.

4. Summary

This paper mainly studies the separation and detection of gait, the extraction of human walking characteristics and classification and identification. The feature extraction during walking is a key issue. In the gait separation detection, a background difference algorithm is used to detect the required part of the video image sequence, and then a relatively complete contour map of the person during walking is extracted. Then, the distance feature information of the gait sequence is extracted from the complete contour map of the human walking process, and the BP neural network algorithm is combined with the existing gait database to classify and recognize the walking data, thereby achieving the effect of human walking recognition. Accuracy is clearly superior to other gait recognition algorithms.

Acknowledgements

Xiaodong Zhao would like to thank the Hebei Provincial Education Department Youth Fund QN2018095.

References

Y. Liu, Na J., Yang J., and Gao G., Modified CRM-based Model Reference Adaptive Control with Reduced Peaking Phenomenon. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 70-78.

X. Luo, Ge S., Wang J., and Guan, Time Delay Estimation-based Adaptive Sliding-Mode Control for Nonholonomic Mobile Robots. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 1-8.

R. Meng, Adaptive Parameter Estimation for Multivariable Nonlinear CARMA Systems. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 96-102.

S. Shi, Liu M., and Ma J., Fractional-Order Sliding Mode Control for a Class of Nonlinear Systems via Nonlinear Disturbance Observer. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 103-110.

T. Zhao and Wang J., Multimodal Transport Path Planning based on Multilateration (MLAT) System Using a Pulse Neural Network Model. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 55-61.

Lily Lee Gait Analys is for Classification [R]. AI Technical Report Massachusetts Institute of Technology-Artificial Intelligence Laboratory, June 2003.

Deng Xiuchun, Wang Renfang. Gait Recognition Method Based on 3D Gait Characteristics[J].Computer Simulation, 2011,(4)

Zhang Heng. Profile-based gait recognition [D]. Nanjing University of Information Science and Technology, 2011.

Zhao Qinpei, et al. A new background-based infrared image segmentation method [J]. Computer Simulation, 2007, (5).

Wang L,Tan T, Ning H, et al. Silhouette Analysis−Based Gait Reco gnition for HumanIdentification[J]. Pattern Analysis & Machine Intelligence I-EEE Transactions on,20 03,25(12):1505-1518.

Ding Wei (Instructor: Qiu Qiqi). Human target detection and tracking based on gait recognition. Master's Thesis of Beijing Jiaotong University. 2017, (5)

Zhao Zijian; Wu Xiaojuan; Liu Yuncai's gait recognition algorithm based on angle histogram. Computer Engineering and Science.2006-06-30

Du Hui received her B.Sc. degree from Xuchang University in 2017，Now she is a master student in Hebei University of Science & Technology. Her main research interests include deep learning，computer vision， and etc.

Xiaodong Zhao male, born in 1973, Ph.D., associate professor, mainly engaged in research on signal detection, automatic control and computer technology. Now he is an associate professor and graduate supervisor in Hebei University of Science & Technology. His main research interests include augment reality， intelligent driving，and etc