Chapter 11
Gait Recognition Using Deep Learning

Chaoran Liu
Auckland University of Technology, New Zealand

Wei Qi Yan
Auckland University of Technology, New Zealand

ABSTRACT

Gait recognition mainly uses different postures of each individual to perform identity authentication. In the existing methods, the full-cycle gait images are used for feature extraction, but there are problems such as occlusion and frame loss in the actual scene. It is not easy to obtain a full-cycle gait image. Therefore, how to construct a highly efficient gait recognition algorithm framework based on a small number of gait images to improve the efficiency and accuracy of recognition has become the focus of gait recognition research. In this chapter, deep neural network CRBM+FC is created. Based on the characteristics of Local Binary Pattern (LBP) and Histogram of Oriented Gradient (HOG) fusion, a method of learning gait recognition from GEI to output is proposed. A brand-new gait recognition algorithm based on layered fu-sion of LBP and HOG is proposed. This chapter also proposes a feature learning network, which uses an unsupervised convolutionally constrained Boltzmann machine to train the Gait Energy Images (GEI).

INTRODUCTION

Biometrics is based on unique physiological or behavioral characteristics of individuals. Biometrics is not as easily transferred or stolen as physical documents such as ID cards, which is secure, reliable, and convenient (Ahonen, Hadid, & Pietikainen, 2006). Gait recognition is based on walking posture of a human body (Little & Boyd, 1998). Compared with fingerprint recognition, speech recognition, face recognition and other technologies, gait recognition has the advantages of easy collection, long distance, less contact, and difficult forge, which is a hotspot research topic in the fields of biometrics, computer vision, intelligent surveillance, and etc. (Cunado, *et al*. 1997).

DOI: 10.4018/978-1-7998-2701-6.ch011

Computer vision is one of the research fields including gait recognition. Given an image sequence containing one or many walking gestures, gait recognition in a broad sense can be divided into four main phases: pedestrian detection, pedestrian region segmentation, pedestrian tracking, and pedestrian recognition (Boykov & Jolly, 2001) (Bulat & Tzimiropoulos, 2016) (Cao, *et al.* 2017) (Tan, *et al*. 2006).

At pedestrian detection phase, it locates the position of a pedestrian in a single video frame and determines the image size (Liu & Payandeh, 2018). At the pedestrian segmentation stage, pixel-level segmentation is performed based on the detection result, background information in the video will be removed. The pedestrian tracking phase determines motion trajectory of the target object and distinguishes individuals from the video sequence (Foresti, 1999).

At the classification and recognition phase, it assigns one of class labels to the test gait samples, which is to identify the person (Kusakunniran, 2014). The process is to calculate the similarity between the test sample and the registered sample so as to complete the classification according to machine learning rules. This stage is often combined with feature extraction to identify a suitable class for classification. This type of methods can be divided into two categories according to gait features in a time-series way.

RELATED WORK

In general, gait recognition refers to pedestrian recognition (Sarkar, et al. 2005) (Shiraga, *et al*. 2016), which utilizes the features extracted from the pedestrian silhouette map to identify a person (Lam, *et al*. 2011) (Murase & Sakai, 1996)(Wang, *et al*. 2004). In recent years, with the development of deep learning, such as Mask Region-Based Convolutional Neural Network (Mask-RCNN), it is possible to apply gait recognition to practical complex scenes (He, *et al.* 2017) (Lee & Grimson, 2002).

Different from directly using the gait silhouette as an input to the deep neural network, Shiraga applied Gait Energy Images (GEI) as the input feature (Tao & Maybank, 2007). GEI is a gait model of static and dynamic information in a sequence of mixed gait silhouettes. The energy of each pixel in the model is obtained by calculating the average intensity of the silhouette pixels in a gait cycle.

The LBP is based on the values of grayscale image pixels. The HOG mainly uses gradient size and direction of the pixels (Dalal & Triggs, 2005). Therefore, after the first round of operations, there is still an image having grayscale intensity changes. In order to obtain pretty rich and useful texture information and edge shape information from the grayscale images, a hierarchical LBP and HOG can be generated in Figure 2.

Figure 1. The flowchart of general steps of gait recognition