Abstract

Computer Vision-based gait recognition has evolved into an active area of research and over the past two decades there have been rapid advancements in the field. Gait recognition is a long-term recognition task where the clothing conditions of a subject are expected to vary during the training and test data capturing phases. Due to this reason RGB information is generally not used to construct the gait features. Rather, gait signatures are traditionally derived from the binary silhouettes extracted from the RGB frames of a sequence since these can preserve proper information about the shape of the individual in the frames. Literature shows that there are two broad categories of approaches in gait recognition, namely model-based and appearance-based, out of which the appearance-based approaches have gained more popularity due to their simplicity, ease of implementation, and high accuracy. The appearance-based recognition approaches can be further categorized as either aggregation-based that combines information from the entire gait cycle or key pose-based (or, pose-based) that derives features at the granularity of fractional parts of a gait cycle. Among these, the key pose-based methods have been seen to provide more accurate performance due to effectively capturing the kinematics of a walking person. But the existing pose-based methods derive features from the complete binary silhouettes that contain a significant amount of redundant pixel information which is not very useful for gait recognition. Moreover, these require predetermination of a set of key poses prior to gait feature extraction. Although, several Deep Learning-based models have also been developed in the later years to perform gait recognition in presence of varying viewpoints, varying clothing conditions, etc., these are heavy-weight and require a significant amount of parameter tuning to achieve a desired level of performance. In this thesis, we aim to improve upon the limitations of the existing key pose-based gait recognition methods that include constructing features from the complete silhouette information and depend on a fixed set of key poses to perform

the classification. Specifically, we have developed plausible solutions to handle the above limitations as described in detail in Chapters 3, 4, and 5. We propose three new gait features, namely the Pose-based Boundary Energy Image (Pose-Based BEI), Generalized Active Energy Image (GAEI), and Dynamic Gait Energy Image (DGEI) and perform extensive evaluation of these features using two popular gait datasets, namely the CASIA-B and the TUMGAID. Among these, the Pose-based BEI feature discussed in Chapter 3 use averaged contour information corresponding to the different key poses in a half gait cycle. This feature has been seen to perform better than other existing key pose-based gait templates due to putting a higher emphasis on the contour pixels of the binary silhouettes, thereby preserving better kinematic gait information than those considering the entire silhouette shape information. However, similar to the previous key pose-based features, it also requires the key poses to be determined prior to extraction of the gait features and next maps an input sequence to the appropriate key poses through a spatio-temporal state transition model with the constraint that from a given state transition is possible only to the same state or to the immediately next state. The above mapping constraints have also been used in each of the previous pose-based gait recognition approaches. But these are very strict and are not effective if (i) the gait cycle of a test sequence has a fewer number of frames than the number of key poses, (ii) the test sequence is captured at a slightly higher frame rate or walking speed. The features proposed in the following two chapters attempt to reduce the strong dependency of the pose-based gait recognition approaches on the predetermined set of key poses. In Chapter 4, we consider a dictionary of key pose sets instead of a single key pose set. An input sequence of frames is mapped to each key pose set in the dictionary and contour-based gait features are constructed corresponding to each set and information fusion from the different key pose sets is carried out to perform gait recognition. Here, also we follow a similar state transition model and constraints as used in the previous chapter to map the input binary silhouette frames to each key pose set in the dictionary. This approach has been seen to provide promising results compared to that obtained using a single key pose set but at the cost of a high response time. It appears that relaxing the constraint in the state transition model by allowing additional transitions from a state to the next-to-next state can help in achieving a higher effectiveness in recognition even with a single key pose set, which we explore in the following chapter, i.e., Chapter 5. Here, a maximal set of unique poses in a half gait cycle is considered following which an input sequence of binary frames is mapped to the appropriate key poses using the above relaxed constraints of the spatio-temporal

state transition model. Experimental results show that this approach is more accurate than those proposed in the previous two chapters, and it also outperforms several other existing approaches that focus on gait recognition in the presence of co-variate conditions. Each of our proposed approaches in the thesis substantially improves upon the existing pose-based gait templates in terms of accuracy. Also, our approaches achieve a high accuracy mark using only non-deep hand-crafted features. In the future, our approaches need to be evaluated on more extensive real-life datasets. Further, efforts are needed to develop automated methods to handle occlusion, viewpoint variation, etc. Usually, end-to-end deep generative models have been seen to be quite effective in handling the above challenges, but these usually take as input the binary silhouette images only. Our work can serve as a basis for future researchers to study if the use of our proposed hand-crafted features in the thesis can boost up the performance of the existing deep learning models and result in an improved gait recognition accuracy.
Keywords: Gait Recognition, Binary Silhouettes, Pose-based Boundary Energy Image, Generalized Active Energy Image, Dynamic Gait Energy Image, Video Surveillance.