Developing Next-Generation Countermeasures for Homeland Security Threat Prevention

Maurice Dawson
University of Missouri-St. Louis, USA

Dakshina Ranjan Kisku
National Institute of Technology, India

Phalguni Gupta
National Institute of Technical Teachers' Training & Research, India

Jamuna Kanta Sing
Jadavpur University, India

Weifeng Li
Tsinghua University, China

A volume in the Advances in Information Security, Privacy, and Ethics (AISPE) Book Series

IGI GLOBAL
DISSEMINATOR OF KNOWLEDGE

www.igi-global.com

Published in the United States of America by
IGI Global
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Dawson, Maurice, 1982- editor.
Title: Developing next-generation countermeasures for homeland security threat prevention / Maurice Dawson, Dakshina Ranjan Kisku, Phalguni Gupta, Jamuna Kanta Sing and Weifeng Li, editors.
Description: Hershey, PA : Information Science Reference, [2017] | Includes bibliographical references and index.
Identifiers: LCCN 2016023422| ISBN 9781522507031 (hardcover) | ISBN 9781522507048 (ebook)
Subjects: LCSH: Terrorism--United States--Prevention. | Terrorism--Prevention. | National security--United States. | Internal security--United States.
Classification: LCC HV6432 .D48 2017 | DDC 363.325/170973--dc23 LC record available at https://lccn.loc.gov/2016023422

This book is published in the IGI Global book series Advances in Information Security, Privacy, and Ethics (AISPE) (ISSN: 1948-9730; eISSN: 1948-9749)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 8
On Using Gait Biometrics for Re-Identification in Automated Visual Surveillance

Imed Bouchrika
University of Souk Ahras, Algeria

ABSTRACT

As surveillance becomes ubiquitous in such modern society due to the immense increase of crimes and the rise of terrorism activities, various government and military funded projects are devoted to research institutions to work on improving surveillance technology for the safety of their citizens. Because of the rapid growth of security cameras and impossibility of manpower to supervise them, the integration of biometric technologies into surveillance systems would be a critical factor for the automation of identity tracking over distributed cameras with disjoint views i.e. Re-Identification. The interest of using gait biometrics to re-identify people over networks of cameras emerges from the fact that the gait pattern can be captured and perceived at a distance as well as its non-invasive and less-intrusive nature.

INTRODUCTION

Although personal privacy has emerged as a major concern for the deployment of large scale surveillance systems, research into automated visual surveillance has received remarkable interest within the computer vision community with potential integration of biometric technologies and human activity recognition systems. This is mainly due to the proliferating number of crimes and terror attacks as well as the vital need to provide safer environment. In fact, the inability of human operators to monitor the increasingly growing numbers of CCTVs installed in highly sensitive and populated areas such as government buildings, airports or shopping malls, has rendered the usability of such systems to be useless. According to the British Security Industry Association, the number of surveillance cameras deployed in the UK was estimated to be more than 5 million in 2015; this figure is expected to increase rapidly particularly after the terrorist attacks that London witnessed in July 2005. Despite the huge increase of monitoring systems, the question whether current surveillance systems work as a deterrent to crime is

DOI: 10.4018/978-1-5225-0703-1.ch008

still questionable. Security systems should not only be able to predict when a crime is about to happen but, more importantly, by early recognition of suspicious individuals who may pose security threats via the use of biometrics, the system would be able to deter future crimes as it is a significant requirement to identify the perpetrator of a crime as soon as possible in order to prevent further offences and to allow justice to be administered. The process of tracking people from one place to another place using surveillance networked cameras would be crucial for gathering valuable security intelligence. Moreover, queries can be made to search for possible locations of a given suspect that can indeed help security officers in their investigations and can lead to further evidence. Traditionally, it is impossible for human operators to work simultaneously on different video screens in order to track people of interest as well as analyze their behaviors across different places. Thus, it has become an essential requirement for research scientists from the computer vision community to investigate visual-based alternatives to automate the process for identity tracking over different views in addition to human activity analysis. Recently, various approaches were published in the literature to accomplish this task based on using basic features such as shape or color information. However, their practical deployment in real applications is very limited due to the complex nature of such problem. An alternative solution would be to employ biometric-based systems that can work at a distance and for low-resolution images such as gait and soft-based biometrics.

Biometrics is concerned with extracting and deriving descriptive measurements based on either the human behavioral or physiological characteristics which should distinguish a subject uniquely among other people. Such measurements are compared against computer records to either confirm or verify the identity of a person. The term *biometrics* is a composite Greek word of two parts: *bios* meaning life and *metrics* referring to the verb "to measure". Biometric measurements based on the physiological traits include face, ear, fingerprint and DNA whilst the behavioral features include gait, voice and signature. Apart from being unique, the biometric description should be universal and permanent. The universality condition implies that it can be taken from all the population meanwhile the permanentness signifies that the biometric signature should stay the same over time. As opposed to traditional identification or verification methods such as passports, passwords or pin numbers, biometrics cannot be transferred, forgotten or stolen and should be ideally obtained non-intrusively. Even though, people intuitively use some body characteristics such as face, gait or voice to recognize each other, it is proven a challenging task to extract and quantify such measurements into a biometric signature. Biometrics can work either in verification or identification mode. For verification, the system performs a one-to-one match for the newly acquired person signature against a pre-recorded signature in the database to verify the claimed identity. For identification mode, a one-to-many matching process is conducted against all people already enrolled in the database to infer the subject identity. Biometrics is now emerging in regular use being deployed in various applications such immigration border control, forensic systems and payment authentication. The choice of a specific biometric modality depends on the nature and requirements of the intended applications in addition to the availability of biometric features. Fingerprints, face and iris are among the most popular physiological traits used in commercial biometric systems with fingerprint capturing over 50% of the market share (Jain & Kumar, 2012). Biometric systems are sold mainly for the following purposes: physical access control, logging attendance and personal identification purposes.

It is no doubt that physiological-based biometric modalities enjoy higher recognition rates however their deployment in surveillance systems is not usually a favored option. In contrast, the suitability of gait recognition for surveillance applications emerges from the fact that the gait pattern can be perceived at distance from the camera even with poor resolution as opposed to other biometric traits where their performance deteriorates severely. Furthermore, the strength of gait recognition is its non-invasiveness nature and hence does not require the subject to cooperate with the acquisition system. This makes gait identification ideal for situations where direct contact with the perpetrator is not possible. Furthermore, as the purpose of any reliable biometric system is to be robust enough to reduce the possibility of signature forgery and spoofing attacks, the gait signature which is based on human motion is the only likely identification that can be taken in covert surveillance. Thus, gait analysis is more suited to forensic investigations as other biometric-based traits that could link to the presence in a crime scene can be obliterated and concealed as opposed to the gait pattern where the mobility of the perpetrator is a must as they have to walk or run to escape from the scene. In a recent empirical study conducted by Lucas (Lucas & Henneberg, 2015), they argued that a combination of eight body measurements is sufficient to achieving a probability of finding a duplicate to the order of 10^{-20} comparing such findings to fingerprints analysis. Interestingly, in one of the high profile murder cases in the UK where a child was abducted and killed, the identity of the murderer could not be revealed directly from the surveillance video footage. The only inspiring solution that could be employed to find out the suspect's identity in this situation was gait recognition as proposed by researchers from the University of Southampton (Nixon, Tan, & Chellappa, 2005a). The notion that people can be recognized by the way they walk has gained an increasing popularity and produced impacts on public policy and forensic practice by its take up by researchers at the Serious Organized Crime Agency.

HISTORY OF BIOMETRICS

For the history of biometrics, people have been identifying each other based on face, voice, appearance or gait for thousands of years. Though, the availability of automated biometric systems started to emerge since the last few decades due to the significant technological advances in the area of computer processing. Since the beginning of civilization, people have used faces to identify each other somehow unconsciously on a day-to-day basis as it was a rudimentary requirement of survival for people to know and differentiate between each other. There are different reports in the literature and archeological records tracing the first use of biometrics to different eras. Chinese records from the 221-206 BC of Qin dynasty revealed details on the use of handprints as evidence during burglary and theft investigations. There are also mentions of the use of fingerprints within the Babylonian civilization for business transactions recorded in clay tablets around 1900 BC. In the Persian book of "*Jaamehol-Tawarikh*" (Universal History) written by Rashid-al-Din Hamadani (1247-1318), he pointed to the practice of recognizing people from their fingerprints where most official government papers used to have fingerprint impressions on them writing "*Experience shows that no two individuals have fingers exactly alike*". In 1788, the German anatomist Johann Mayer argued that fingerprints are unique to each person.

However, the first systematic and scientific basis for people identification dates back to 1858 when William Herschel recorded the handprint of each employee on the back of a contract whilst working for the civil service of India (Berry & Stoney, 2001). This was used as a way to distinguish employees from each other on payday. It was considered the first systematic capture of hand and fingerprints that was used primarily for identification purposes. For the use of forensic biometrics, Alphonse Bertillon who was a French police officer was the pioneer to employ the first biometric evidence into the judicial system presented as anthropometric measurements of the human body to counter against repeat offenders who could easily fake or change their names. He introduced the *bertillonage system* such that each person is identified through detailed records taken from their body anthropometric measurements and physical descriptions as they are impossible to spoof or change them. The system was adopted by the police authorities throughout the world. In 1903, the Bertillon system was challenged on the basis that anthropometric measurements are not adequate to differentiate reliably between people. This was fuelled by the case of identical twins who have almost the same measurements using the Bertillon system. In 1890s, Sir Francis Galton and Edward Henry (Arbab-Zavar, Xingjie, Bustard, Nixon, & Li, 2015) described separately a classification system for people identification based on the person fingerprints taken from all ten fingers. The characteristics that Galton described to recognize people are still used today. The Galton system was adopted by New York state prison where fingerprints were employed for the identification of apprehended inmates. In 1936, ophthalmologist Frank Burch proposed the concept of using the iris pattern as a method to recognize an individual. The first attempt for semi-automated face recognition system was developed by Woodrow W. Bledsoe on a contract with the US Government during 1960s. The main setback of the biometric system is the requirement for an administrator to manually locate the different parts of the face such as eyes, ears, nose and mouth on a given image. This system relied solely on the ability to extract useable feature points. It calculated distances and ratios to a common reference point that was compared to the reference data. In 1969, the Federal Bureau of Investigation (FBI) started the initiative to develop a system for automating the process of people identification based on fingerprints.

GAIT BIOMETRICS

Gait is defined as the manner of locomotion characterized by consecutive periods of loading and unloading the limbs. Gait includes running, walking and hopping. However the term gait is most frequently used to describe the walking pattern. The rhythmic pattern of human gait is performed in a repeatable and characteristic way consisting of consecutive cycles (Winter, 1991). The gait cycle is defined as the time interval between successive instances of initial foot-to-floor contact for the same foot (Cunado, Nixon, & Carter, 2003), and the way a human walks is marked by the movement of each leg such that each leg possesses two distinct phases. When the foot is in contact with the floor the leg is at the *stance phase*. The time when the foot is off the floor to the next step is defined as the *swing phase*. Each phase is marked by a start and an end; the stance phase begins with the *heel strike* of one foot when the leg strikes the ground. The locomotion process involves the interaction of many body systems working together to yield the most efficient walking pattern. The locomotion system consists of four main subtasks that are fulfilled at the same time to produce the walking pattern (Woollacott & Tang, 1997). These four functions are: (i) initiation and termination of locomotor movements (ii) the generation of continuous movement to progress toward a destination (iii) adaptability to meet any changes in the environment or other concurrent tasks (iv) maintenance of the equilibrium during progression. Compared with quadrupeds, the maintenance of stability and balance for humans during walking is a particularly difficult task for the postural control system. This is mainly because for most of the gait cycle, the human body is supported by only a single leg with the center of mass passing outside the base of support provided by the foot in contact with the floor.

Gait analysis is the systematic study of human walking (Whittle, 2002) aimed at the quantification and understanding of the locomotion process. This study involves the observation of body movements, mechanics and muscle activities. Gait analysis is carried out for two main purposes (Whittle, 2002): firstly, the treatment of patients with gait abnormalities for rehabilitation purposes. Secondly, gait is studied to enrich the knowledge and understanding of human walking pattern to further explore other research avenues. The study of human gait dates back to the ancient times. Aristotle (384-322 BC) might be considered the first person to study and describe gait in his book "*De Motu Animalium*" (on the movement/motion of animals). Leonardo da Vinci (1452-1519) described the principles of walking and observed the complexity and symmetric nature of human gait in his famous anatomic paintings. During the 18th century, the Weber brothers in Germany conducted the first formal biomechanical experiment, giving clear description about the timing of the gait cycle. In 1892, Muybridge devised an apparatus with multiple trip wires attached to the camera shutters which was employed to record the locomotion process (Muybridge, 1979). Gait analysis has evolved greatly in the 20th century with the invention of many tools and instruments needed for the measurement and quantification of the gait pattern.

Interestingly, in one of the theatre plays (The Tempest: Act 4, Scene 1) written by Shakespeare (1564-1616) the following sentence was spotted: "*High'st Queen of state, Great Juno comes; I know her by her gait*". This gives a clear indication that gait recognition is not a novel concept and people intuitively identify each other at a distance via the observed gait pattern. Gait is an emergent biometric which is increasingly attracting the interests of researchers as well as the industry. Although gait recognition is still a new biometric and is not sufficiently mature to be deployed in real world applications such as visual surveillance, it has the potential to overcome most of the limitations that other biometrics suffer from such as face, fingerprints and iris recognition which can be obscured in most situations where serious crimes are involved. Face recognition in many cases has been proven to be unreliable for visual surveillance systems; this is due to the fact that people can disguise or hide their faces in addition to the fact that recorded video data can be inadequate for processing due to low resolution. In fact, gait recognition has been successfully deployed in numerous forensic investigations to confirm the identity of the perpetrators where traditional biometrics are unavailable (Bouchrika, Goffredo, Carter, & Nixon, 2011; Larsen, Simonsen, & Lynnerup, 2008; Yang, Larsen, Alkj\aer, Simonsen, & Lynnerup, 2013).

Medical Studies

Early medical investigations carried out by Murray *et al*. (1967) produced a standard gait pattern for normal walking subjects aimed at studying the gait pattern for pathologically abnormal patients. The experiments were performed on sixty people aged between 20 and 65 years old. Each candidate was instructed to walk for a repeated number of trials. For the collection of gait data, special markers were attached on every subject. Murray suggested that the human gait consists of 24 different components which render the gait pattern unique for every person if all gait movements are taken into account. It was reported that the motion patterns of the pelvic and thorax regions are highly variable from one subject to another. Furthermore, Murray observed that the ankle rotation, pelvic motion and spatial displacements of the trunk embed the subject individuality due to their consistency at different trials. Furthermore, psychological studies carried out by the Swedish psychologist Johansson (1973), revealed that people are able to perceive human motion from Moving Lights Display (MLD). An MLD is a two-dimensional video of a collection of bright dots attached to the human body taken against a dark background where only the bright dots are visible in the scene. Different observers are asked to see the actors performing various activities. Based on these experiments observers can recognize different types of human motion such as walking, jumping, dancing and so on. Moreover, the observer can make a judgment about the gender of the performer (Kozlowski & Cutting, 1978), and even further identify the person if they are already familiar with their gait (Goddard, 1992). Cutting argued that the recognition is purely based on dynamic gait features as opposed to previous studies which were confounded by familiarity cues, size, shape or other non-gait sources of information. Although the different parts of the human body are not seen in the points and no links exist between the bright dots to show the skeleton structure of the human body, the observer can recover the full structure of the moving object. Thereby, the motion of the joints contains sufficient information for the perception of human motion (Bingham, Schmidt, & Rosenblum, 1995; Dittrich, 1993).

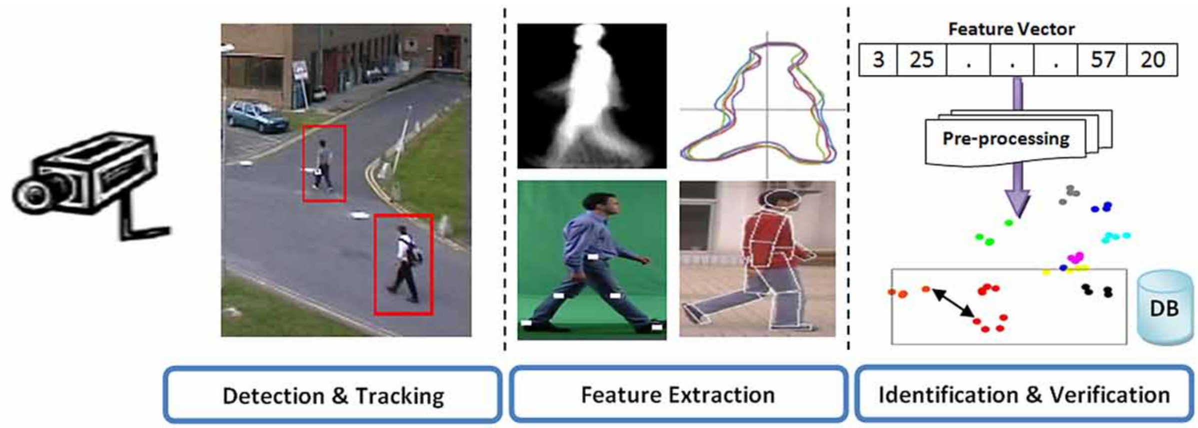
There is a wealth of research which strives to document the capability of the human visual system to perceive the human motion from a small number of moving points as argued by early medical studies by Johansson, Cutting and Murray. Nevertheless, the underlying perceptual process is poorly understood and there is still a lack of research which explains the underlying principles for representing and retrieving the biological motion (Troje, Westhoff, & Lavrov, 2005). Two main theories have been put forward for the perception of human motion from the MLD: *structure-based* and *motion-based* (Cedras & Shah, 1995). The former theory claims that the initial step is recovering the 3D structure from the motion information observed from the MLDs, and then uses the recovered structure for the purpose of recognition. In the motion-based approach, recognition is based directly on the motion information without recovering the skeleton structure of the human body from the MLD; instead the motion information is extracted from a sequence of frames. Although, there is a wealth of gait studies in the literature aimed for medical use with a few referring to the discriminatory nature of the gait pattern, none is concerned with the automated use of gait for biometrics and recognizing people. The gait measurements and results introduced by Murray are to be of benefit for the development of automated gait biometric systems. However, the extraction of the gait pattern is proven complex using computer vision methods as non-rigid human motion encompasses a wide range of possible motion transformations due to the highly flexible structure of the human body and to self-occlusion (Gavrila, 1999; Yoo, Nixon, & Harris, 2002). Furthermore, clothing type, segmentation errors and different viewpoints pose a substantial challenge for accurate localization of gait features using automated techniques.

Gait Recognition Methods

An automated vision-based system for people identification via the way they walk is designed to extract gait features without the need to use markers or special sensors to aid the extraction process. In fact, all that is required is an ordinary video camera linked to special vision-based software. Marker-less motion capture systems are suited for applications where mounting sensors or markers on the subject is not an option as the case of forensic analysis and visual surveillance. Typically, the gait biometric system consists of two main components: i) Hardware platform dedicated for data acquisition. This can be a single CCTV camera or distributed network of cameras. ii) Software platform for data processing and recognition. The architecture of the software side for gait biometric system is composed broadly of three main components: i) detection and tracking of the subject, ii) feature extraction and iii) classification stage. Figure 1 shows the flow diagram for gait identification outlining the different subsystems involved in the process of automated people recognition.

- **Subject Detection and Tracking:** a walking subject is initially detected within a sequence of frames using background subtraction techniques to detect moving objects or via the use of other methods such as the Histogram of Oriented Gradients (Bouchrika, Carter, Nixon, Morzinger, & Thallinger, 2010; Dalal & Triggs, 2005) which is capable of detecting people from still images at

Figure 1. Gait biometrics system overview



real-time. Subsequently, intra-camera tracking is performed to establish the correspondence of the same person across consecutive frames. Tracking methods are supported by simple low-level features such as blob size, aspect-ratio, speed and color in addition to the use of prediction algorithms to estimate the parameters of moving objects in the following frames. This is based on motion models which describe how parameters change over time. The most popular predictive methods used for tracking is the Kalman filter (Welch & Bishop, 2001), the Condensation algorithm (Isard & Blake, 1998), and the mean shift tracker (Comaniciu, Ramesh, & Meer, 2000).

- **Feature Extraction:** this is the most important stage for automated marker-less capture systems whether for gait recognition, activity classification or other imaging applications. This is because the crucial data required for the classification phase are derived at this stage. Feature extraction is the process of estimating a set of measurements either related to the configuration of the whole body or the configuration of the different body parts in a given scene and tracking them over a sequence of frames. The features should bear certain degree of the individuality of the subject for biometric applications. High-level features estimated at this level for gait recognition can be categorized into two types: *static* and *dynamic* features. Examples of static features include the subject height and other anthropometric measurements meanwhile dynamic or kinematic features are such joints angular measurements and displacement of the body trunk.
- **Identification or Verification Phase:** it is mainly a classification process which involves matching a test sequence with an unknown label against a group of labelled references considered as the gallery dataset. At this stage, a high-level description is produced from the features extracted during the previous phases to infer or confirm the subject identity. The classification process is normally preceded by pre-processing stages such as data normalization, feature selection and dimensionality reduction of the feature space through the use of statistical methods. A variety of pattern recognition methods are employed in vision-based systems for gait recognition including Neural Networks, Support Vector Machines (SVM) and K-Nearest Neighbour classifer (KNN). The latter is the most popular method for the classification due to its simplicity and fast computation.

As much of the interest in the area of gait analysis was limited to physical therapy, orthopedics and rehabilitation practitioners for the diagnosis and treatment of patients with walking abnormalities, gait analysis has become a challenging computer vision problem when it comes to deploying automated methods for the extraction of features. This is further motivated with the advent of gait as an attractive biometrics. Many research studies have aimed to develop a system capable of overcoming the difficulties imposed by the detection, tracking and extraction of biometric gait features. Yam *et al* (C.-Y. Yam & Nixon, 2009) grouped the challenges into two categories either related the environment or the person. The environment challenges include the recording viewpoint, nature of the walking ground and environmental conditions. For the subject related difficulties entail *physiological* conditions such as age and disability, *psychological* conditions or *external* factors subjected on the person such as carried load, clothing and footwear. Various methods were surveyed in (Nixon, Tan, & Chellappa, 2005b) and (C.-Y. Yam & Nixon, 2009). In the same way to the classification of approaches developed for the analysis of human motion, gait recognition methods can be divided broadly into two main categories which are model-based and appearance-based (model-free) approaches. This is based on the procedure for extracting and modeling gait features.

Model-Based Approaches

For the model-based approach, a prior model is established to match real images to this predefined model, and thereby extracting the corresponding gait features once the best match is obtained. Usually, each frame containing a walking subject is fitted to a prior temporal or spatial model to explicitly extract gait features such as stride distance, angular measurements, joints trajectories or anthropometric measurements. Although model-based approaches tend to be complex requiring high computational cost, these approaches are the most popular for human motion analysis due to their advantages (C.-Y. Yam & Nixon, 2009). The main strength of model-based techniques is the ability to extract detailed and accurate gait motion data with better handling of occlusion, self-occlusion and other appearance factors as scaling and rotation. The model can be either a 2 or 3-dimension structural model, motion model or a combined model. The structural model describes the topology of the human body parts as head, torso, hip, knee and ankle by measurements such as the length, width and positions. This model can be made up of primitive shapes based on matching against low-level features as edges. The stick and volumetric models are the most commonly used structural-based methods. Akita (1984) proposed a model consisting of six segments comprising of two arms, two legs, the torso and the head. Guo *et al* (Guo, Xu, & Tsuji, 1994) represented the human body structure by a stick figure model which had ten articulated sticks connected with six joints. Rohr (1994) proposed a volumetric model for the analysis of human motion using 14 elliptical cylinders to model the human body. Karaulova *et al*. (Karaulova, Hall, & Marshall, 2000a) used the stick figure to build a hierarchical model of human dynamics represented using Hidden Markov Models. For the deployment of structural model-based methods for gait recognition, Niyogi *et al*. (Niyogi & Adelson, 1994) was perhaps the pioneer in 1994 to use a model-based method for gait recognition. Gait signature is derived from the spatio-temporal pattern of a walking subject using a five stick model. Using a database of 26 sequences containing 5 different subjects, a promising classification rate of 80% was achieved.

The motion model describes the kinematics or dynamics of the body or its different parts thoughout time. Motion models employ a number of constraints that aid the extraction process as the maximum range of the swinging for the low limbs. Cunado (Cunado et al., 2003) was the first to introduce motion model using the Velocity Hough Transform to extract the hip angular motion via modelling human gait as a moving pendulum. The gait signature is derived as the phase-weighted magnitudes of the Fourier components. A recognition rate of 90% was achieved using the derived signature on a database containing 10 subjects. Yam *et al.* (C. Y. Yam, Nixon, & Carter, 2004) modeled the human gait as a dynamic coupled oscillator which was used to extract the hip and knee angular motion via evidence gathering. The method was evaluated on a database of 20 walking and running subjects, achieving a recognition rate of 91% based on gait signature derived from the Fourier analysis of the angular motion. Wagg *et al.* (Wagg & Nixon, 2004) proposed a new model-based method for gait recognition based on the biomechanical analysis of walking subjects. Mean model templates are adopted to fit individual subjects. Both the anatomical knowledge of human body and hierarchy of shapes are used to reduce the computational costs. The gait feature vector is weighted using statistical analysis methods to measure the discriminatory potency of each feature. On the evaluation of this method, a correct classification rate of 95% is achieved on a large database of 2,163 video sequences of 115 different subjects. Bouchrika (2008) proposed a motion-based model for the extraction of the joints via the use a parametric representation of the Elliptic Fourier Descriptors describing the spatial displacements. The following model is presented to describe the spatial and temporal motion of the joints through a gait cycle:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} s_x X(t) \\ s_y Y(t) \end{bmatrix} \tag{1}$$

Such that α is the rotation angle and S_x, S_y are the scaling factors. The joints positions for a walking subject are extracted using the described parametric equation using the Hough Transform where evidence gathering is employed. To reduce the computational load for the five parameters space, the heel strikes data is incorporated to deduce most of the parameters. A Correct classification rate of 95.7% is reported based on purely dynamic-based gait features experimented on a small dataset containing 20 subjects.

As most of the model-based are exploiting 2D images, there are recent work aimed for introducing 3-dimensional model for the extraction of gait features including the work of Urtasun (Urtasun & Fua, 2004) who described a 3D volumetric temporal model to track and estimate motion parameters that could be potentially used for people recognition. Ariyanto (Ariyanto & Nixon, 2012) described a 3D model-based method based on marionette and mass-spring models. The method combines the use of marionettes structure and motion principles with the physical mass-spring model in order to better fit the stick-figure model within the 3D voxel data obtained from the Southampton Biometric Tunnel. Based the estimated height and footprint features of walking people, a moderate recognition rate of 58.2% is reported using a dataset of 46 subjects with 4 sequences per person. Krzeszowski (Krzeszowski, Michalczuk, Kwolek, Switonski, & Josinski, 2013) proposed a view-invariant gait recognition system where the identification is achieved from data extracted using marker-less 3D motion tracking algorithm. The motion tracking was accomplished by a particle swarm optimization algorithm. Based on a small dataset of 88 sequences for 22 walking subjects, a promising CCR of 90% is reported. Figure 2 shows a number of examples for the model-based approaches used primarily for gait recognition.

Various model-based methods have been proposed in the literature to extract different types of features as well as to address particular problems encountered at the low-level extraction phase including self-occlusion, noise and appearance transformation factors. Incontestably, despite the fact that model-based methods are appealing and enjoy certain benefits as robustness and reliability, these approaches suffer from high computational cost due to the large number of parameters space and the issue of image quality sensitivity. Further, the robustness and success of such methods are totally dependent on the extraction

Figure 2. Model-Based Approaches for Gait Feature Extraction: (a) Karaulova (2000b). (b) Wagg (2004). (c) Wang (2004)

(a) (b) (c)

accuracy which is not guaranteed in real case scenarios. The main setback for model-based methods is the complexity involved for constructing a general model describing the structural or dynamical gait components which would span and translate easily to different configurations or settings. Although, state of the art research studies are claiming that model-based methods can be extended to cover a large number of cases for extraction of gait features, their performance is still questionable due to the low attained classification rates in addition to the evaluation protocol on limited cases applied on smaller datasets. For instance, the majority of existing *2D* models are constructed to work with the sagittal case where people walk normal to the camera as it is the most straight forward scenario to extract gait features (C.-Y. Yam & Nixon, 2009). There are recent attempts to devise a viewpoint-invariant approach capable of modeling and extracting the gait pattern regardless of the camera position (Bouchrika, Carter, & Nixon, 2014). Subsequently, the derived features are re-projected back into the sagittal plane for identification. However, the reported recognition rates are still modestly poor for critical applications as the system just fails for the case when people walk in the frontal or backward direction with respect to the camera (Goffredo, Bouchrika, Carter, & Nixon, 2010; Worapan Kusakunniran, Wu, Zhang, Ma, & Li, 2013).

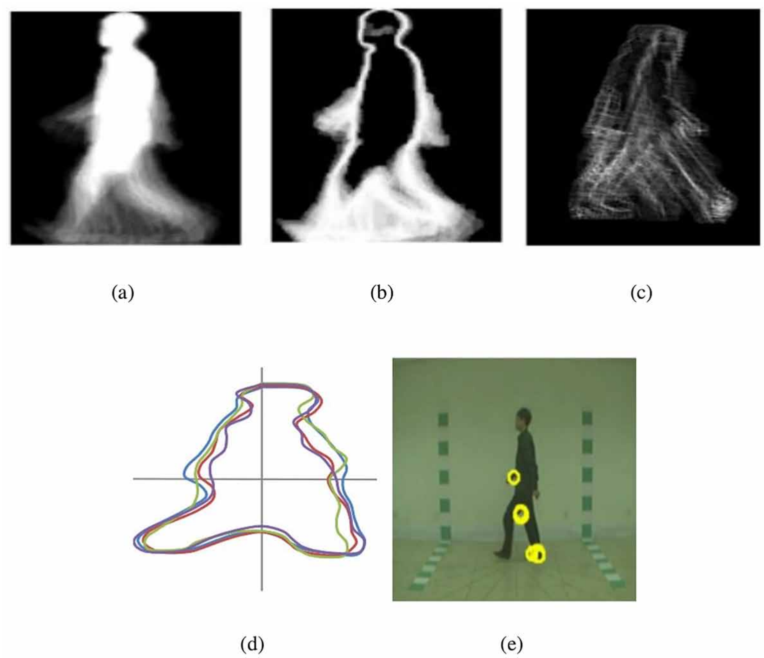
Appearance-Based Approaches

Appearance-based or model-free approaches for gait recognition do not need a prior knowledge about the gait pattern. Instead, features are extracted from the whole body without the need to explicitly extract the body parts. The majority of appearance approaches depends on data derived from silhouettes which are obtained via background subtraction. The simplest method is called the Gait Energy Image (GEI) introduced by Han and Bhanu (Han & Bhanu, 2006) in which gait signature is constructed through taking the average of silhouettes for one complete gait cycle. Experimental results confirmed that higher recognition rates can be attained to reach 94.24% for a dataset of 3,141 subjects (Iwama, Okumura, Makihara, & Yagi, 2012). However, such method performs poorly when changing the appearance as for instance recording from different viewpoint or changing the clothing. Gait Entropy Image (GenI) is a silhouette-based representation introduced by Bashir (Bashir, Xiang, & Gong, 2009) which is computed by calculating the Shannon entropy for each pixel achieving a correct classification rate of 99.1% on dataset of 116 subjects. The Shannon entropy estimates the uncertainty value associated with a random variable. Other similar representations include Motion Energy Image, Motion Silhouette Image, Gait History Image and Chrono-Gait Image. Hayfron-Acquah (Hayfron-Acquah, Nixon, & Carter, 2003) introduced a method for constructing a gait signature based on analyzing the symmetry of human motion. The symmetry map is produced via applying the Sobel operator on the gait silhouettes followed by the Generalized Symmetry Operator. On a dataset of 28 subjects, a higher recognition rate of 97.3% is attained. Further experiments confirmed that the symmetrical-based gait signature is relatively insensitive to noise and robust enough to handling partial occluded data as well missing frames. Choudhury (Choudhury & Tjahjadi, 2012) employed Procrustes shape analysis for silhouette contours using the elliptic Fourier Descriptor for encoding the gait biometric signature. Figure 3 shows the discussed examples for appearance-based gait recognition methods.

As the accuracy of silhouette-based methods depends on the background segmentation algorithm which is not reliable for the case of real surveillance footage in addition to the sensitivity issue to varied appearances, a number of appearance-based methods have emerged recently that use instead interest-point descriptors. Kusakunniran (2014) proposed a framework to construct gait signature without the need to extract silhouettes. Features are extracted in both spatial and temporal domains using Space-Time Interest Points (STIPs) by considering large variations along both spatial and temporal directions at a local level. Extensive experiments are carried out on the CASIA-B gait dataset reporting a moderate recognition rate of 52% for recognizing people walking normal to the camera plane (90°). For invariant viewpoint gait identification, an average recognition rate of 21.4% is achieved for a dataset containing 124 subjects recorded at 11 different views. Appearance-based Method relies pivotally on statistical methods to reduce or optimize the dimensionality of feature space using methods such as Principal Component Analysis. In addition, advanced machine learning methods are usually applied as multi-class support vector machine and neural networks.

In terms of the reported recognition rates, silhouette-based methods have generally superior performance as opposed to their counterpart. Furthermore, the biometric gait signature can be constructed through performing additional training to include different cases and settings such as various viewpoints or covariate factors without the need to create a model describing the gait pattern. The apparent drawback of such methods is that the recognition performance is highly susceptible to appearance changes that can utterly change the gait signature. Contentiously, recent investigations by Veres (Veres, Gordon, Carter, & Nixon, 2004) was published to answer and explore the question of what information is considered the most important for silhouette-based methods. Two approaches are investigated for deriving a gait signature using either averaging or differential silhouettes. Using Principal Component Analysis (PCA) and Analysis of Variance (ANOVA) for statistical feature significance of data derived from three major datasets, they reported that most of the discriminative features for appearance-based approaches are extracted from static components of the top part of the human body whilst the dynamic components generated from the swinging of the legs are ignored as the least important information. For the case of using the gait energy image i.e. average silhouette, the features around the contour of the head and upper part of the human body are confirmed to be adequate for human identification. For differential silhouettes, the inclusion of the leg motion improved slightly the recognition rate.

Figure 3. Appearance-based methods for gait recognition: (a) use gait energy image; (b) gait entropy image; (c) symmetry map; (d) procrustes shape analysis; (e) stip descriptors



PEOPLE RE-IDENTIFICATION

As surveillance becomes ubiquitous in such modern society due to the immense increase of crimes and the rise of terrorism activities, various government and military funded projects are devoted to research institutions to work on improving surveillance technology for the safety of their citizens. The deployment of surveillance camera networks becomes more and more prevalent in most public spaces including airports, shopping malls and underground stations. Because of the limitation that a single camera can only recover a limited field of view, multiple networked cameras are employed to monitor and record human activities over a large spatial area. Provided that intra-camera tracking is adequate and accurate, the tracking problem is reduced to associating two instances of the same identity seen in different cameras at the same time or at different times for the case of non-overlapping camera views. In fact, it is extremely desirable to identify a person of interest even anonymously from a large number of people within the entire camera networks, and afterwards track or re-identify this person when seen at different camera views. This task is called *People Re-Identification.* In visual surveillance, people re-identification is the process of recognizing whether an individual regardless of their true identity has already been observed over a network of cameras (Doretto, Sebastian, Tu, & Rittscher, 2011). The topic is recently attracting much interest from the computer vision research community because of the various applications primarily for security and forensic use. Tracing the whereabouts of people of interest for law enforcement officers and possibly inferring a semantic description of their activities based on visited locations would be a major milestone for visual automated surveillance technology. Re-identification can assist the indexing and retrieval of offline data using semantic queries applied against surveillance video footage from large camera networks. Examples of queries include searching for people who visited place *A* and then went to place *B* and so on. Furthermore, re-identification can be of benefit to the retail industry through providing viable information about customers' habits and activities within large complex shopping malls in order to improve customer service and shopping space management.

Whilst the tracking process of people within a single camera can be achieved robustly with good performance, it is proven to be an extremely challenging problem for a network of cameras to associate the same subjects observed at different places and times from different cameras with disjoint views. The main challenge for re-identification stems from the variation of people's appearance across different cameras (Bedagkar-Gala & Shah, 2014). There are other factors affecting the association process including low image quality, varying illumination, occlusion in crowded scenes. Low quality or obscured views can render re-identification techniques which are based on traditional biometric cues like fingerprints or face inaccessible. Furthermore, the articulated nature of the human is another prime cause for the person's appearance to change almost continuously. Networks of cameras span typically over large geospatial areas and have non-overlapping fields-of-view (FOVs) to provide enhanced coverage. Hence, the complexity of such systems increases in proportion to the scale of the camera network and the process of handling identification queries becomes cumbersome if not infeasible for camera networks of complicated temporal and spatial topology. There can be even *blind gaps* between the cameras where the surveillance system loses hold of the person of interest potentially changing severely their appearances. A number of datasets are being made publicly available for the research community to address the issue of people re-identification for surveillance scenarios. Table 1 summarizes the list of databases showing the number of different cameras used as well as number of subjects observed within the network of cameras. Most datasets provide still images for the exception of the original version of the iLids dataset which they provide raw videos taken from the surveillance system of Gatwick airport in the United Kingdom.

The majority of early research studies proposed for tracking people or moving objects over networks of cameras were limited in a way that they require information about the camera calibration parameters as well as overlapping fields to maintain correspondence or association between different views. Camera calibration is an expensive task and the availability of calibration parameters in real life surveillance applications is proven challenging to obtain or sometimes impossible. Other research studies relax the requirement for calibration information but still need the overlapping fields to maintain correspondences using different types of basic features such as color (Orwell, Remagnino, & Jones, 1999) or geometrical features such as people height. Cai *et al* (Cai, Aggarwal, Inc, & Seattle, 1999) proposed an approach to re-identify walking people from sequences of synchronized and calibrated cameras. The association of the same people across frames from different cameras is established via a set of feature points within a Bayesian probability framework. People are tracked using a single camera view until the system deduces that the active camera will no longer have a good view of the person. Basic features used for tracking include geometric properties such as the height of the person. Stein *et al* (Stein, 1999) proposed a novel method that does not require prior knowledge about the camera calibration. The camera calibration parameters are estimated by observing motion trajectories within the scene. Meanwhile, Javed (Khan, Javed, Rasheed, & Shah, 2001) described a simple system for tracking people across multiple un-calibrated cameras. Their approach is able to infer the spatial relationships between the camera fields of views and use such information to establish the correspondence between different perspective views.

The tracking of people across a network of cameras is of prime importance for smart automated surveillance systems as it provides vital intelligence to security agencies. There is a recent considerable research within the literature during the last few years (Bedagkar-Gala & Shah, 2014) dedicated to addressing this area of research for the case of non-calibrated cameras with non-overlapping views. Bedagkar-Gala

Table 1. Re-Identification datasets publicly available

	ViPER	iLIDS	Caviar4Reid	V-47	Grid
Number of Cameras	2	2	2	2	8
Number of People	632	119	72	47	250

(2014) surveyed recently different methods for people re-identification, datasets as well as performance metrics. Bedagkar-Gala categorizes the methods for non-contextual people re-identification into three main different areas: *Descriptor learning* focuses on learning the most discriminative features on the fly or instead a learning stage is put in place to generate a descriptive dictionary of features or words that better describe distinctively a person's appearance using bag-of-features techniques. *Distance metric* learning attempts to maximize the identification matching rate regardless of the choice for the appearance features or color calibration using various different types of features. *Color Calibration* methods rely upon modeling the color relationship between a given camera pair and update such model regularly. Color brightness transfer function (BTF) is used as the basis to establish the correspondence between the same people across a pair of cameras. Meanwhile, Das (Das, Chakraborty, & Roy-Chowdhury, 2014) grouped the approaches broadly into three main categories: *discriminative signature* methods, *metric learning-based* methods and *transformation learning* methods. In this study, three different types of features are being described for the task of people re-identification.

Low-Level Features

Low-level or basic features are purely based on color histograms, interest point or texture-based descriptors. Features can be either global where the whole image or body region is considered meanwhile local features refer to the characteristics which are extracted from smaller portions of the image. Prosser *et al* described a cumulative brightness transfer function (CBTF) in order to map colors between different cameras requiring only a sparse color training set (Prosser, Gong, & Xiang, 2008). The CBTF is preferred due to its merit for retaining the original brightness values which are not common. Therefore, uncommon brightness values from the initial training dataset can be accurately mapped across different cameras. The matching process for people re-identification is based on a newly bi-directional method for comparing people in order to reduce false matches. The method estimates the similarity between a pair of observations of the transfer functions via computing the Bhattacharya distance. Hamdoun argued that people re-identification can be performed through matching signatures of people taken using a locally developed variant of SURF interest-point descriptor collected from a sequence of frames (Hamdoun, Moutarde, Stanciulescu, & Steux, 2008). The proposed method is praised as computationally feasible processing a single match of two instances in less than half a second. Bak *et al* used a *haar*-like features and dominant color descriptors for the association of the same people (Bak, Corvee, Brémond, & Thonnat, 2010). The *haar*-based method relies on AdaBoost to find the most distinctive features whilst the dominant color representation is based on the upper and lower parts of the body. Experimental results performed on the CAVIAR dataset containing 10 subjects show that a promising recognition rate of 80% can be achieved using the *haar*-based features concluding its potency for handling viewpoint and pose changes. However, the dominant color descriptors achieve a poor identification rate of 40% on the same dataset. The main challenge of using low-level features is how to resolve the ambiguities resulting from similar visual features belonging to different objects. Further, appearance-based features are not guaranteed to stay the same across different cameras as people may change their clothing or carry extra objects that would severely alter the extracted low-level visual features.

Gait-Based Features

The interest of using gait biometrics for re-identification is partly due to fact that the gait pattern can be captured and perceived from a distance as well as its non-invasive and less-intrusive nature. For the automated extraction of gait features, a *Haar-like* feature template matching process is proposed by Bouchrika (2014) for the localization of the legs using motion information for a single walking subject. The approach does not depend on background subtraction for the derivation of gait features. This is because it is computationally an expensive and complex task to deploy background subtraction for real-time surveillance applications due the process of updating the background model which is influenced by various factors such as dynamic clutter, weather conditions and other environmental factors. Motion models are derived based on medical data describing the angular motion for both the knee and hip during a full gait cycle. As a first step, the proposed approach constructs the motion map image based on the change detection for the inter-frame difference. The only constraint of using frame differencing is that the camera must be in a stationary position. Moving pixels of a walking subject across consecutive frames are detected with the emphasis to produce better edge data. The motion map M_t at frame t is estimated as the absolute difference of two consecutive frames I_t and I_{t+1} as given in the following equation. Thresholding is thereafter applied on the resulting image M_t in order to reduce the artifacts. A sample motion image is illustrated in Figure for a walking subject taken from iLids dataset:

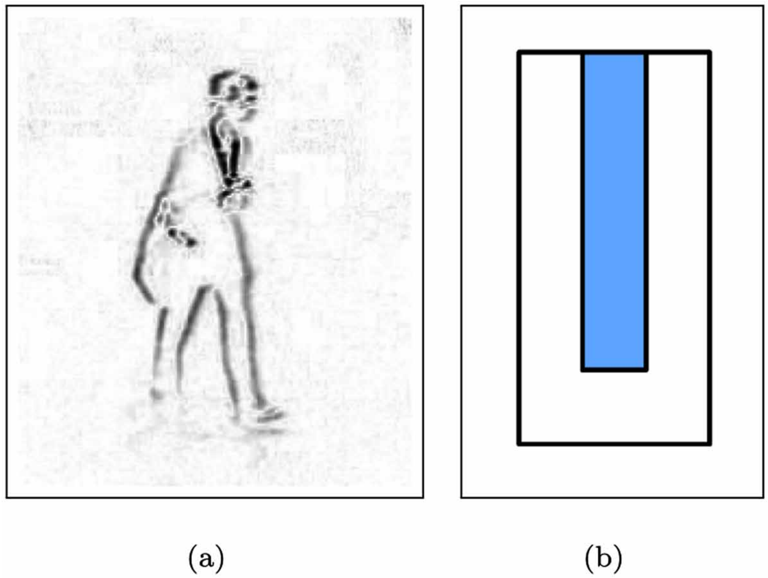
$$M_t = \left\| I_t - I_{t+1} \right\| \tag{2}$$

A Haar-like template is introduced for the recovery of the gait features due their fast and robust performance for real-time applications from object recognition to pedestrian detection. The template is shown in Figure 4 which is based on the outlier of the lower part of the leg. Let p^t_{ankle} is the candidate position of the ankle at t^{th} frame. To extract the ankle position, numerous templates are generated to account for the different possible appearance transformations such as translation and rotation defined by kinematical knowledge. The Haar-templates are superimposed against the motion map at the possible point p computing the match score S as given in Equation 2. The similarity score determines how well is the generated template is superimposed on top of the motion map. It is approximated as the sum of larger values inside the superimposed region divided by the accumulated lower values inside the area that are less than a predefined threshold which is experimentally set as τ=20.

$$S(x,y,\alpha) = \frac{\sum_{i \in P_{x,y,\alpha}} P_{x,y,\alpha}(i) \times Z(P_{x,y,\alpha}(i))}{\sum_{i \in P_{x,y,\alpha}} P_{x,y,\alpha}(i) \times \left| 1 - Z(P_{x,y,\alpha}(i)) \right|} \tag{3}$$

Such that α is the rotation angle and Z is given as follows:

Figure 4. Markerless gait feature extraction: (a) motion image. (b) haar-based matching template



$$Z(i) = \begin{cases} 1 & if\ i > \tau \\ 0 & otherwise \end{cases} \tag{4}$$

As opposed to using a per-frame method for pose estimation (Mori & Malik, 2006), a frame-to-frame approach is deployed instead for the recovery of gait features so that the results obtained from the previous frame are utilized to guide the matching process in the upcoming frames. In order to lower the search space for a candidate joint position and refine the extraction accuracy, kinematical and anthropometric constraints inferred from the spatial as well as angular data derived from the gait motion model described are enforced during the extraction phase. For instance, during the striking phase, one of the feet will be stabilized at the same place and therefore the ankle spatial movement is limited within a smaller part of the image whilst the rotation parameter α will be enclosed within a particular range depending on the phase of the gait cycle. The pose estimation method for the lower legs partially depends on the anatomical proportions of the body segments reported in the medical results of anatomical studies (Dempster & Gaughran, 1967) for a subject of height *H*:

$$\begin{aligned} y'_{hip} &= \min(y_{sil}) + 0.5 \cdot H \\ y'_{knee} &= \min(y_{sil}) + 0.75 \cdot H \\ y'_{ankle} &= \min(y_{sil}) + 0.90 \cdot H \end{aligned} \tag{5}$$

During the double-support stage of the walking cycle where the legs overlap, it is a challenging task to extract the lower limbs accurately partly due to the self-occlusion incurred during the overlap. Thus, the matching procedure is applied for the striking leg using kinematic gait constraints that can assist with the estimation of gait features. The swinging leg is not dealt with during the overlap due to self-occlusion. The overlapping begins when the Euclidean distance between the two ankles is less than a defined threshold which is related to the person height. The extraction of the swinging leg during the overlap is continued after a certain number of frames which is determined from the average gait cycle model. Experimentally, this number is set to six frames for a video recorded with a frame rate of 25 frames per second. To extract the joints trajectories as well the angular measurements when the legs do overlap, a 3rd order polynomial interpolation process is applied.

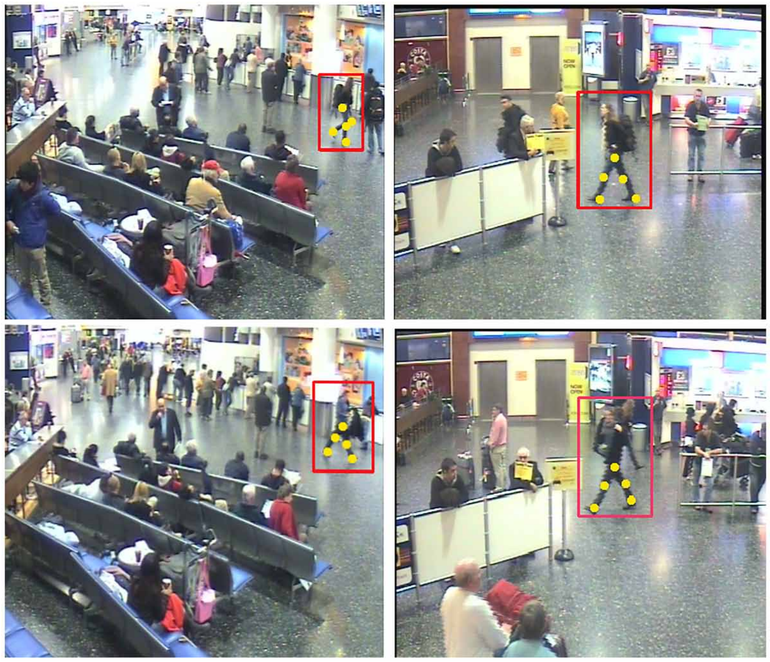
To determine the performance of the proposed method using real surveillance feed, the approach is evaluated using the iLids dataset recorded at Gatwick Airport in the United Kingdom. The dataset is taken from CCTV surveillance cameras with overlapping and non-overlapping views. A set of 20 different walking people who are seen on different cameras views (Camera 2 and 3) are manually annotated. The marker-less extraction algorithm is run to extract the angular data for the dataset of walking people in order to derive gait signatures from the gait and anthropometric features. Figure 5 shows an example of the extraction results of the joint trajectories for different camera views.

To investigate the efficiency of the proposed method in such surveillance cases, the leave-one-out cross validation with the *KNN* classifier is employed to determine the performance across all the 20 subjects across two different non-overlapping cameras. The attained classification rate (CCR) is 97% for the case of $k=1$. Furthermore, data from camera 3 are matched against data of camera 2 and vice versa in a probe to gallery mode. For the sake of imposing further challenges on the classification task, the size of the gallery dataset is increased for camera 3 to include more 10 subjects that are recorded and processed within camera 3 only. In the same way, the number of subjects for the dataset of camera 2 is increased to include 10 more subjects. The obtained average correct classification is 92.5% for the cross-camera matching process. This suggests that gait angular features can be used in surveillance systems for identity tracking and recognition across different cameras particularly for cases where it is impossible to derive robust features such as the face. However, the accuracy of such systems for now depends on the extraction reliability of gait features.

Soft Biometric Labels

The ultimate goal of a vision-based automated surveillance system is not only to track and re-identify people across different cameras, but preferably to provide a more effective interface to the operator. Central to this goal is the ability to search the surveilled scenes for a person of interest using simple search terms. Interestingly, soft biometrics is a new form of people identification which concerns the use of simple semantic search attributes or terms that people use to describe each other linguistically such as ‘tall’, ‘male’ and ‘bold’. The main purpose of soft biometrics is to bridge the semantic gap between biometrics and natural descriptions. Soft biometrics depends on describing a set of semantic attributes and assigning descriptive values for each attribute. A semantic attribute can be an observable property that can be described by a person. The attribute can be quantified by either binary values or comparative labels. (Reid, Nixon, & Stevenage, 2014). Despite the lack of discriminatory potency for single soft attributes, recent research studies have confirmed that combination of soft traits can be used as a biometric signature to identify people or at least to boost the recognition capabilities of traditional biometrics. Samangooei (Samangooei & Nixon, 2010) introduced a soft biometric approach which identifies people from video footage based only on verbal description made by human annotations. This description was composed of 23 absolute categorical attributes which are chosen to be universal, distinct and easily discernible from a distance and largely permanent labels. The selected soft biometric characteristics featured both

Figure 5. Feature extraction of gait features applied on the iLids dataset



categorical attributes, like hair color in addition to characteristics which are associated with metric-based values such as height. Both attributes were annotated using absolute labels. In another study by Reid, they described a method for human identification via the use of comparative descriptors derived from visual comparisons between subjects. Instead of using absolute labels for soft attributes, comparative categorical labels allows better objective annotations because of their benefit to infer more descriptive and continuous measurements (Reid et al., 2014). Reid *et al* showed that with relative measurements, a promising correct classification rate of 95% can be attained for people identification.

The use of soft biometrics for people re-identification has been recently reported in the literature. Le An (An, Chen, Kafai, Yang, & Bhanu, 2013) described a re-ranking approach for soft biometrics attributes to improve the recognition performance. The attributes are detected and then scores are calculated between pairs of images using a regression model. Meanwhile Layne (Layne, Hospedales, & Gong, 2014; Layne, Hospedales, Gong, & others, 2012) proposed a method for re-identification that learns a selection and weighting of mid-level semantic attributes to describe people. Ontology of useful attributes is determined by human experts. These high-level features are employed to complement the discriminatory potency of low-level feature representation significantly. Li (Li, Liu, Wang, Liu, & Yan, 2014) argued about the importance of using clothing appearance as the main cue for re-identification. The body parts and their local features are extracted initially for pose-misalignment. A latent SVM approach is presented to explain the relationship between the low-level part features, middle-level clothing attributes and high level labels.

CONCLUSION

Although personal privacy has emerged as a major concern for the deployment of large scale surveillance systems, the demand to employ biometric technology has become increasingly prevalent in highly populated and sensitive places. This is mainly due to the proliferating number of crimes and terror attacks as well as the vital need to provide a safer environment. Because of the rapid growth of security cameras and impossibility of manpower to supervise them, the integration of biometric technologies into surveillance systems would be a critical factor for the automation of identity tracking over distributed cameras with disjoint views i.e. Re-Identification. The interest of using gait biometrics to re-identify people over networks of cameras emerges from the fact that the gait pattern can be captured and perceived at a distance as well as its non-invasive and less-intrusive nature. Recent studies confirmed that gait angular features can be used in surveillance systems for identity tracking and recognition across different cameras particularly for cases where it is impossible to derive robust features such as the face. This way, we can identify people being seen in one camera view from data already derived from a different camera view. This is an important step in translating gait biometrics into real scenarios where prior knowledge about the camera calibration cannot be recovered such as in surveillance videos. However, the accuracy of such systems for now depends on the extraction reliability of gait features.

REFERENCES

Akita, K. (1984). Image Sequence Analysis of Real World Human Motion. *Pattern Recognition*, *17*(1), 73–83. doi:10.1016/0031-3203(84)90036-0

An, L., Chen, X., Kafai, M., Yang, S., & Bhanu, B. (2013). Improving person re-identification by soft biometrics based reranking. *Distributed Smart Cameras (ICDSC), 2013 Seventh International Conference on* (pp. 1-6).

Arbab-Zavar, B., Xingjie, W., Bustard, J., Nixon, M. S., & Li, C.-T. (2015). On forensic use of biometrics. Handbook of Digital Forensics of Multimedia Data and Devices.

Ariyanto, G., & Nixon, M. S. (2012). Marionette mass-spring model for 3D gait biometrics*5th International Conference on Biometrics* (pp. 354-359). doi:10.1109/ICB.2012.6199832

Bak, S., Corvee, E., Brémond, F., & Thonnat, M. (2010). *Person re-identification using haar-based and dcd-based signature.* Paper presented at the Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on. doi:10.1109/AVSS.2010.68

Bashir, K., Xiang, T., & Gong, S. (2009). Gait recognition using Gait Entropy Image*3rd International Conference on Crime Detection and Prevention* (pp. 1-6).

Bedagkar-Gala, A., & Shah, S. K. (2014). A survey of approaches and trends in person re-identification. *Image and Vision Computing*, *32*(4), 270–286. doi:10.1016/j.imavis.2014.02.001

Berry, J., & Stoney, D. A. (2001). The history and development of fingerprinting. *Advances in fingerprint*. *Technology (Elmsford, N.Y.)*, *2*, 13–52.

Bingham, G. P., Schmidt, R. C., & Rosenblum, L. D. (1995). Dynamics and the Orientation of Kinematic Forms in Visual Event Recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(6), 1473–1493. doi:10.1037/0096-1523.21.6.1473 PMID:7490589

Bouchrika, I. (2008). *Gait Analysis and Recognition for Automated Visual Surveillance*. University of Southampton.

Bouchrika, I., Carter, J. N., & Nixon, M. S. (2014). Towards automated visual surveillance using gait for identity recognition and tracking across multiple non-intersecting cameras. *Multimedia Tools and Applications*, 1–21.

Bouchrika, I., Carter, J. N., Nixon, M. S., Morzinger, R., & Thallinger, G. (2010). Using gait features for improving walking people detection.*20th International Conference on Pattern Recognition (ICPR)* (pp. 3097-3100). doi:10.1109/ICPR.2010.758

Bouchrika, I., Goffredo, M., Carter, J., & Nixon, M. (2011). On using gait in forensic biometrics. *Journal of Forensic Sciences*, *56*(4), 882–889. doi:10.1111/j.1556-4029.2011.01793.x PMID:21554307

Cai, Q., Aggarwal, J., Inc, R., & Seattle, W. (1999). Tracking human motion in structured environments using adistributed-camera system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(11), 1241–1247. doi:10.1109/34.809119

Cedras, C., & Shah, M. (1995). Motion-based Recognition: A survey. *Image and Vision Computing*, *13*(2), 129–155. doi:10.1016/0262-8856(95)93154-K

Choudhury, S. D., & Tjahjadi, T. (2012). Silhouette-based gait recognition using Procrustes shape analysis and elliptic Fourier descriptors. *Pattern Recognition*, *45*(9), 3414–3426. doi:10.1016/j.patcog.2012.02.032

Comaniciu, D., Ramesh, V., & Meer, P. (2000). Real-time Tracking of Non-Rigid Objects using Mean Shift.*Proceedings. IEEE Conference on Computer Vision and Pattern Recognition*, 2. doi:10.1109/CVPR.2000.854761

Cunado, D., Nixon, M. S., & Carter, J. N. (2003). Automatic Extraction and Description of Human Gait Models for Recognition Purposes. *Computer Vision and Image Understanding*, *90*(1), 1–41. doi:10.1016/S1077-3142(03)00008-0

Dalal, N., & Triggs, B. (2005). *Histograms of oriented gradients for human detection.* Paper presented at the Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. doi:10.1109/CVPR.2005.177

Das, A., Chakraborty, A., & Roy-Chowdhury, A. K. (2014). *Consistent re-identification in a camera network. In Computer Vision–ECCV 2014* (pp. 330–345). Springer.

Dempster, W. T., & Gaughran, G. R. (1967). Properties of body segments based on size and weight. *The American Journal of Anatomy*, *120*(1), 33–54. doi:10.1002/aja.1001200104

Dittrich, W. H. (1993). Action Categories and the Perception of Biological Motion. *Perception*, *22*(1), 15–22. doi:10.1068/p220015 PMID:8474831

Doretto, G., Sebastian, T., Tu, P., & Rittscher, J. (2011). Appearance-based person reidentification in camera networks: Problem overview and current approaches. *Journal of Ambient Intelligence and Humanized Computing*, *2*(2), 127–151. doi:10.1007/s12652-010-0034-y

Gavrila, D. M. (1999). Visual Analysis of Human Movement: A survey. *Computer Vision and Image Understanding*, *73*(1), 82–98. doi:10.1006/cviu.1998.0716

Goddard, N. H. (1992). *The Perception of Articulated Motion: Recognizing Moving Light Displays*. University of Rochester.

Goffredo, M., Bouchrika, I., Carter, J. N., & Nixon, M. S. (2010). Performance analysis for automated gait extraction and recognition in multi-camera surveillance. *Multimedia Tools and Applications*, *50*(1), 75–94. doi:10.1007/s11042-009-0378-5

Guo, Y., Xu, G., & Tsuji, S. (1994). Understanding Human Motion Patterns. *Pattern Recognition, Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on, 2*.

Hamdoun, O., Moutarde, F., Stanciulescu, B., & Steux, B. (2008). *Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences.* Paper presented at the Distributed Smart Cameras, 2008. ICDSC 2008. Second ACM/IEEE International Conference on. doi:10.1109/ICDSC.2008.4635689

Han, J., & Bhanu, B. (2006). Individual recognition using gait energy image. *Pattern Analysis and Machine Intelligence*. *IEEE Transactions on*, *28*(2), 316–322.

Hayfron-Acquah, J. B., Nixon, M. S., & Carter, J. N. (2003). Automatic gait recognition by symmetry analysis. *Pattern Recognition Letters*, *24*(13), 2175–2183. doi:10.1016/S0167-8655(03)00086-2

Isard, M. C., & Blake, A. C. (1998). CONDENSATION: Conditional Density Propagation for Visual Tracking. *International Journal of Computer Vision*, *29*(1), 5–28. doi:10.1023/A:1008078328650

Iwama, H., Okumura, M., Makihara, Y., & Yagi, Y. (2012). The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *Information Forensics and Security*. *IEEE Transactions on*, *7*(5), 1511–1521.

Jain, A. K., & Kumar, A. (2012). *Biometric recognition: an overview. In Second Generation Biometrics: The Ethical* (pp. 49–79). Legal and Social Context. doi:10.1007/978-94-007-3892-8_3

Johansson, G. (1973). Visual Perception of Biological Motion and a Model for its Analysis. *Perception & Psychophysics*, *14*(2), 201–211. doi:10.3758/BF03212378

Karaulova, I. A., Hall, P. M., & Marshall, A. D. (2000a). A Hierarchical Model of Dynamics for Tracking People with a Single Video Camera. In *Proceedings of the 11th British Machine Vision Conference*. doi:10.5244/C.14.36

Karaulova, I. A., Hall, P. M., & Marshall, A. D. (2000b). A hierarchical model of dynamics for tracking people with a single video camera.*Proc of the 11th BMVC* (pp. 262-352). doi:10.5244/C.14.36

Khan, S., Javed, O., Rasheed, Z., & Shah, M. (2001). Human tracking in multiple cameras.*Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings* (Vol. 1). doi:10.1109/ICCV.2001.937537

Kozlowski, L. T., & Cutting, J. E. (1978). Recognizing the Gender of Walkers from Point-Lights Mounted on Ankles: Some Second Thoughts. *Perception & Psychophysics*, *23*(5), 459. doi:10.3758/BF03204150

Krzeszowski, T., Michalczuk, A., Kwolek, B., Switonski, A., & Josinski, H. (2013). *Gait recognition based on marker-less 3D motion capture.* Paper presented at the Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on. doi:10.1109/AVSS.2013.6636645

Kusakunniran, W. (2014). Recognizing Gaits on Spatio-Temporal Feature Domain. *Information Forensics and Security*. *IEEE Transactions on*, *9*(9), 1416–1423.

Kusakunniran, W., Wu, Q., Zhang, J., Ma, Y., & Li, H. (2013). A new view-invariant feature for cross-view gait recognition. *Information Forensics and Security*. *IEEE Transactions on*, *8*(10), 1642–1653.

Larsen, P. K., Simonsen, E. B., & Lynnerup, N. (2008). Gait Analysis in Forensic Medicine. *Journal of Forensic Sciences*, *53*(5), 1149–1153. doi:10.1111/j.1556-4029.2008.00807.x PMID:18636978

Layne, R., Hospedales, T. M., & Gong, S. et al. (2012). Person Re-identification by Attributes. *BMVC*, *2*, 3.

Layne, R., Hospedales, T. M., & Gong, S. (2014). *Attributes-Based Re-identification*. Person Re-Identification.

Li, A., Liu, L., Wang, K., Liu, S., & Yan, S. (2014). Clothing Attributes Assisted Person Re-identification. *Circuits and Systems for Video Technology, IEEE Transactions on,* (99), 1-1.

Lucas, T., & Henneberg, M. (2015). Comparing the face to the body, which is better for identification? *International Journal of Legal Medicine*, 1–8. PMID:25666979

Mori, G., & Malik, J. (2006). Recovering 3d human body configurations using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(7), 1052–1062. doi:10.1109/TPAMI.2006.149 PMID:16792095

Murray, M. P. (1967). Gait as a Total Pattern of Movement. *American Journal of Physical Medicine*, *46*(1), 290–333. PMID:5336886

Muybridge, E. (1979). *Muybridge's Complete Human and Animal Locomotion: All 781 Plates from the 1887 Animal Locomotion*. Courier Dover Publications.

Nixon, M. S., Tan, T. N., & Chellappa, R. (2005). *Human Identification Based on Gait*. Secaucus, NJ: Springer-Verlag New York, Inc.

Niyogi, S. A., & Adelson, E. H. (1994). Analyzing and Recognizing Walking Figures in XYT. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. doi:10.1109/CVPR.1994.323868

Orwell, J., Remagnino, P., & Jones, G. (1999). Multi-camera colour tracking.*Second IEEE Workshop on Visual Surveillance* (pp. 14-21). doi:10.1109/VS.1999.780264

Prosser, B., Gong, S., & Xiang, T. (2008). *Multi-camera Matching using Bi-Directional Cumulative Brightness Transfer Functions.* Paper presented at the British Machine Vision Conference. doi:10.5244/C.22.64

Reid, D. A., Nixon, M. S., & Stevenage, S. V. (2014). Soft Biometrics; Human Identification Using Comparative Descriptions. *Pattern Analysis and Machine Intelligence*. *IEEE Transactions on*, *36*(6), 1216–1228.

Rohr, K. (1994). Towards Model-Based Recognition of Human Movements in Image Sequences. *CVGIP. Image Understanding*, *59*(1), 94–115. doi:10.1006/ciun.1994.1006

Samangooei, S., & Nixon, M. S. (2010). Performing content-based retrieval of humans using gait biometrics. *Multimedia Tools and Applications*, *49*(1), 195–212. doi:10.1007/s11042-009-0391-8

Stein, G. P. (1999). Tracking from multiple view points: Self-calibration of space and time.*Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Vol. 1, pp. 521-527). doi:10.1109/CVPR.1999.786987

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person Identification from Biological Motion: Effects of Structural and Kinematic Cues. *Perception & Psychophysics*, *67*(4), 667–675. doi:10.3758/BF03193523 PMID:16134460

Urtasun, R., & Fua, P. (2004). *3D tracking for gait characterization and recognition.* Paper presented at the Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on.

Veres, G. V., Gordon, L., Carter, J. N., & Nixon, M. S. (2004). What image information is important in silhouette-based gait recognition? *Computer Vision and Pattern Recognition, 2004. CVPR 2004.Proceedings of the 2004 IEEE Computer Society Conference on* (Vol. 2, pp. II-776).

Wagg, D. K., & Nixon, M. S. (2004). On Automated Model-Based Extraction and Analysis of Gait. In *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*. doi:10.1109/AFGR.2004.1301502

Wang, L., Ning, H., Tan, T., & Hu, W. (2004). Fusion of static and dynamic body biometrics for gait recognition. *Circuits and Systems for Video Technology*. *IEEE Transactions on*, *14*(2), 149–158.

Welch, G., & Bishop, G. (2001). An Introduction to the Kalman Filter. *ACM SIGGRAPH 2001 Course Notes*.

Whittle, M. (2002). *Gait Analysis: An Introduction*. Butterworth-Heinemann.

Winter, D. A. (1991). *The Biomechanics and Motor Control of Human Gait: Normal Elderly and Pathological*. University of Waterloo Press.

Woollacott, M. H., & Tang, P. F. (1997). Balance Control during Walking in the Older Adult: Research and its Implications. *Physical Therapy*, *77*(6), 646. PMID:9184689

Yam, C.-Y., & Nixon, M. (2009). Gait Recognition, Model-Based. In Encyclopedia of Biometrics, (pp. 633-639).

Yam, C. Y., Nixon, M. S., & Carter, J. N. (2004). Automated Person Recognition by Walking and Running via Model-Based Approaches. *Pattern Recognition*, *37*(5), 1057–1072. doi:10.1016/j.patcog.2003.09.012

Yang, S. X., & Larsen, P. K., Alkjaer, T., Simonsen, E. B., & Lynnerup, N. (2013). Variability and Similarity of Gait as Evaluated by Joint Angles: Implications for Forensic Gait Analysis. *Journal of Forensic Sciences*. PMID:24745080

Yoo, J. H., Nixon, M. S., & Harris, C. J. (2002). Extraction and Description of Moving Human Body by Periodic Motion Analysis. In *Proceedings of ISCA 17th International Conference on Computers and Their Applications*.

KEY TERMS AND DEFINITIONS

Biometrics: The process concerned with extracting and deriving descriptive measurements based on either the human behavioral or physiological characteristics which should distinguish a subject uniquely among other people.

Camera Handover: The task of tracking and following people through identity matching between different surveillance cameras.

Gait: The manner of locomotion characterized by consecutive periods of loading and unloading the limbs. Gait includes running, walking and hopping.

Gait Analysis: The systematic study of the human walking pattern which aims at the quantification and understanding of the locomotion process.

Gait Recognition: The biometric process to infer the person's identity through the use of their gait style or pattern.

Identity Verification: The process is confirming the claimed identity of a person based on biometric matching against database records.

Re-Identification: The process of recognizing whether an individual regardless of their true identity has already been observed over a network of cameras.

Soft Biometrics: A new type of people identification which concerns the use of simple semantic search attributes or terms that people use to describe each other linguistically.