On a large sequence-based human gait database

J. D. Shutler, M. G. Grant, M. S. Nixon and J. N. Carter
University of Southampton
Department of Electronics and Computer Science
United Kingdom
e-mail: {jds1, mgg, msn, jnc}@ecs.soton.ac.uk

Abstract

Biometrics today include recognition by characteristic and by behaviour. Of these, face recognition is the most established with databases having evolved from small single shot single view databases, through multi-shot multi-view and on to current video-sequence databases. Results and potential of a new biometric are revealed primarily by the database on which new techniques are evaluated. Clearly, to ascertain the potential of gait as a biometric, a sequence-based database consisting of many subjects with multiple samples is needed. A large database enables the study of inter-subject variation. Further, issues concerning scene noise (or non-ideal conditions) need to be studied, ideally with a link between ground truth and application based analysis. Thus, we have designed and built a large human gait database, providing a large multi-purpose dataset enabling the investigation of gait as a biometric. In addition, it is also a useful database for many still and sequence based vision applications.

Introduction

Gait is an emergent biometric with increasing interest (e.g. [1, 2, 3] - primarily through DARPA's Human ID at a Distance program). Early gait databases contained very few subjects (primarily due to technological issues) and analyses resulting in high classification rates have prompted further research. For example, the UCSD (University of California San Diego) database consisted of six subjects with seven sequences of each. Several different publications have reported close to 100% classification on that data eg. [4, 5, 6]. Here we describe the design and capture of a new large sequence-based human gait database, detailing the ideas behind its structure through to areas of human psychology used to ensure good quality, consistent data.

Database design

A large gait database has been designed and built, providing both ideal laboratory data (indoor, controlled lighting) as an approximation to ground truth, and (non-ideal) outdoor data for application based analysis. Data was also derived using treadmills, to analyse the potential of capturing gait information in a confined space. High resolution still images were taken to enable face classification and to estimate a subject's body and size. Therefore, subjects extracted from indoor and outdoor scenes can be compared, both in terms of classification and extraction. The database consists of fronto-parallel and obliquely viewed imagery (allowing orientation independent analysis), extracted silhouette data, still images (side on and frontal) and ancillary data including subject specific information, camera setups and extraction parameters [7]. Further, a considerable amount of information redundancy was designed into the acquisition, storage and

final structure of the database. This allows for cross-referencing within the data itself, and eased the (manual and automatic) validation and consistency checking of the data.

Filming Issues

Human psychology plays a large part in collecting data of this type and magnitude. Firstly, to avoid affecting the subject's walking patterns, the treadmill training and filming took place after each subject had first walked outdoors, and then inside on the laboratory track. There is much debate in the use of treadmills for gait analysis concerning their suitability, speed and inclination. The speed and inclination were set at constant values derived by evaluation, however, it is worth noting that treadmills allow for capture of long high resolution continuous gait sequences. Further issues [7] included not informing subjects when the cameras were filming (reducing shyness issues by switching the cameras on prior to the subjects entering the laboratory), not talking to subjects as they walked (as invariably humans will turn their head to address the person), using a talk-only radio station for background noise (to reduce the impulse need of a human to break the talk-silence), removing the need for a person to control the cameras reduced the camera-shyness and talking issues further and placing a mirror in front of the treadmill aided balance and stopped the subject from looking at their feet and/or the treadmill control panel.

Imagery

A sequence-based database is rather a massive one. Our target was for over 100 subjects, with approximately 8 sequences of 1.5 steps per subject (as this is periodic data) together with background and supporting data. Given that digital video (DV) is now an established technology, we acquired the imagery via good quality off-the-shelf DV camcorders, having to experimentally examine the differences between cameras due to incomplete published specifications, as the units are designed for domestic-use. This analysis involved capturing data and looking specifically at colour and optical quality, frame rates and picture scan methods. To allow for 25 fps progressive-scan capture and good quality optical data, two different types of cameras were used, as at that time no single camera met the requirements (not all DV camcorders allow for 25 fps progressive scan data). Three Canon MV30i's were used for the progressive scan data while three Sony DCR TRV900E PAL's allowed for interlaced data with improved optical and colour quality (the Sony uses a three chip system, one for each colour, in comparison to the single-chip technology of the Canon). Two cameras filmed each scene (one progressive scan and one interlaced) allowing for analysis of viewpoint-independence. The progressive scan technology provides high resolution imagery (suitable for ground truth) whilst security video often uses interlaced data. The resulting database consists of over 100 subjects, filmed in 3 scenarios each from 2 different views, with accompanying high resolution still images. The subjects walked around a continuous track and were filmed walking in both directions, the track layout can be seen in Figure 1. The outdoor data presents a similar scenario, however with a greater distance between subject and camera, allowing scale independent analysis and an increased number of in-view gait cycles.

Data quality and post-processing

Blue screening (or chroma-keying) is the process of filming an object, or subject in front of an evenly lit, bright pure coloured backdrop. Object or subject extraction from the images is easily achieved through background removal, allowing an alternative background colour or scene to be used in its place. Here a bright green backdrop was used, mainly as it is an unlikely colour for the subjects to wear, although any pure colour could have been used. Thus, a purpose built chroma-key laboratory was constructed to allow controlled lighting conditions for the inside
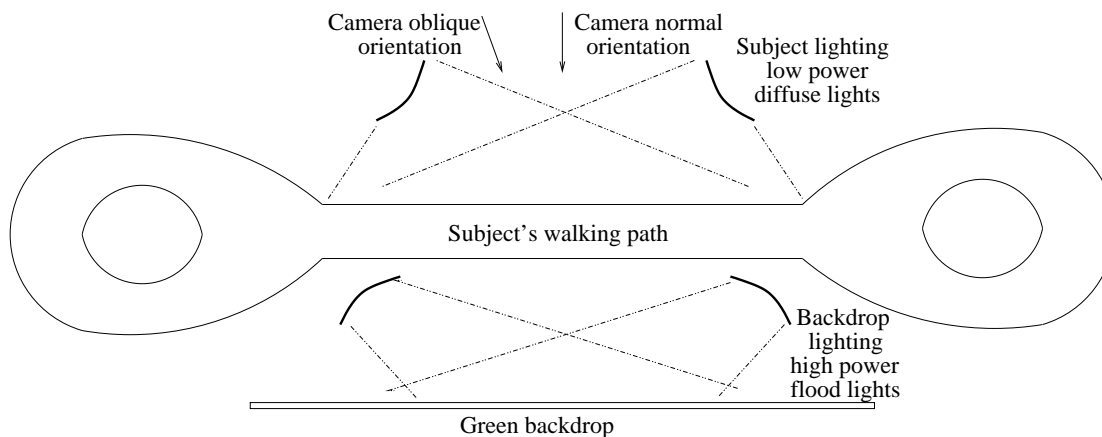
Figure 1: Laboratory lighting arrangement, enabling the separation of the two lighting schemes.

track and treadmill scenarios. This near perfect subject extraction provides an approximation to ground truth. Further ground truth includes the still photos and subject information. Iterative optimisation of the laboratory and equipment setups produced blur free data and significantly reduced shadows. This iterative process used Canny and Sobel edge detectors and a statistical subject extraction technique [8] enabling the assessment of the quality of the lighting and camera positions. A software implementation of chroma-key extraction allows for easy monitoring of the data quality between filming sessions, as invariably (over time) equipment can be knocked and the output from lights can change. Figure 2 shows example images and a chroma-keyed result (actual data is in colour). The complete database (captured in one month) includes details of all settings and session specifics, recorded on a per-filming session basis. To increase the available information for the ground truth (and to ease use), the data from the fronto-parallel track camera has been significantly described and labelled using XML, an example fragment of XML can be seen in Figure 3. Further information includes: less detailed labeling for the remaining viewpoints, camera sync information (between views) and parameters enabling radial distortion correction.

Analysis

To date, three different recognition approaches have been applied to the data, all with encouraging results. This analysis of the database suggests that it has indeed met its design objectives. First, high gait recognition performances have been achieved on the largest yet number of subjects for gait, an overview of these results can be seen in Table 1. The progression of these results reflects the gradual construction of the database and detailed explanations of these results can be found in [9, 10, 11]. The processing of the data used much of the available support material, enabling streamlined and in some cases automated analysis. These results used a selection of binary silhouettes and optical flow descriptions generated from the inside laboratory data. The use of the chroma-keyed binary silhouettes has also provided a test-bed for performance analyses of the various techniques eg. [11]. A preliminary analysis of the outdoor data confirms the increased variance of features extracted from application scenario imagery as opposed to ground truth [11], shown here for two subjects in Figure 4. In each plot the tight Gaussian represents the variance of the inside data, whereas the larger Gaussians (greater variance) are the outdoor data, the mean of each shows the mean feature point drift from the baseline (inside data). Finally, only data-quality checks have been applied to the inside treadmill data.
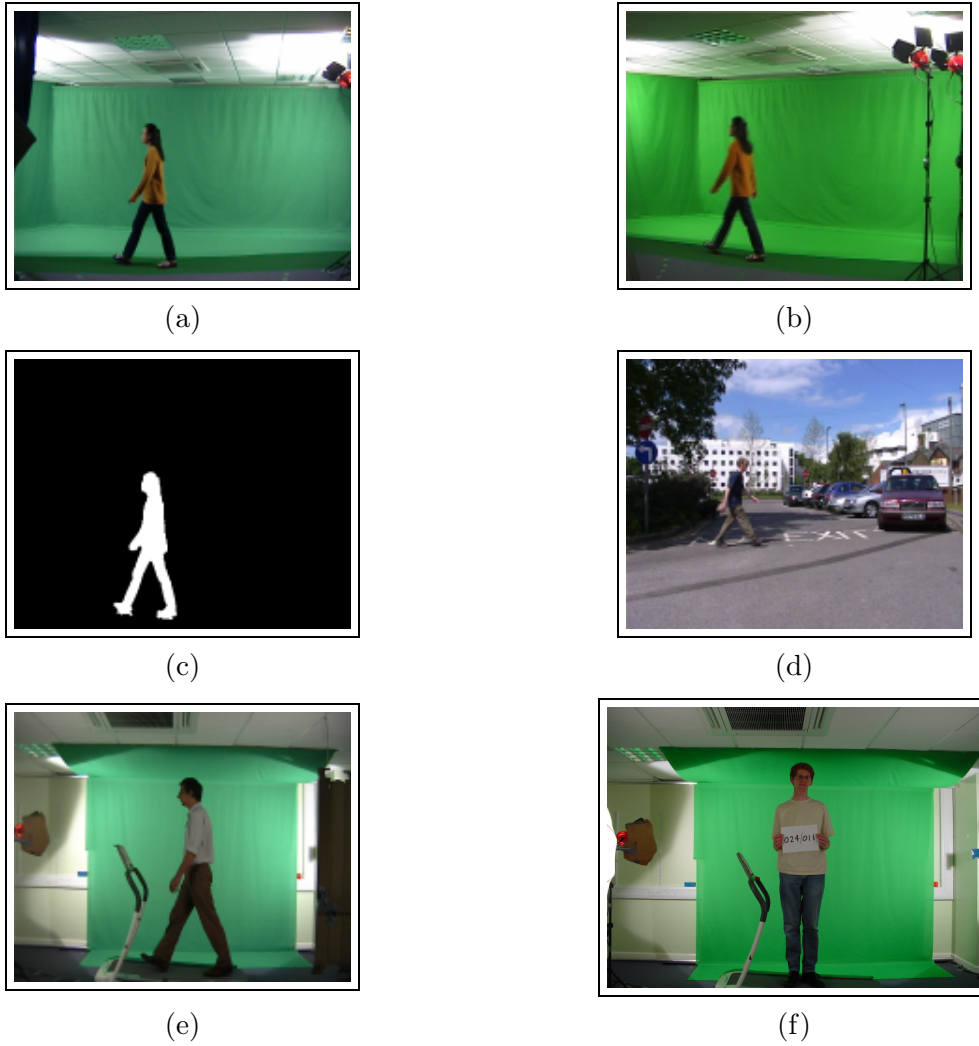
Figure 2: Example (a) inside data normal (b) inside data oblique, (c) processed silhouette from the inside data normal, (d) outdoor data, (e) treadmill and (f) high resolution still (actual data is in colour).

```
<?xml version="1.0" ?>
<data camera="a" date="23/05/01" direction="left"
filename="008a013s00L.dv" sequence="00L" session="008"
start_frame="2369" start_time=" 0 : 01 : 34.19" stop_frame="2546"
stop_time=" 0 : 01 : 41.21" subject_id="013" xmlcreatedby="Time Code
re-writer" xmlcreatedon="Mon Mar 25 11 : 51 : 16 2002">
  <comments>
    No comment
  </comments>
  <heelstrikes automatic="yes" checked="yes">
    <strike leg="left" time="62"/>
    <strike leg="right" time="77"/>
        .  .  .
```

Figure 3: Example fragment of XML for the normal-view camera.

Algorithm	No. of	No. of Sequences	Classifier Result	
Employed	Subjects	per Subject	$k = 1$	$k = 3$
Symmetry [9]	28	4	97.30 %	96.40 %
Velocity Moments [11]	50	4	97.00 %	96.00 %
Moment Based [10]	114	8	78.00 + %	-

Table 1: Early published results using the SOTON HiD database.

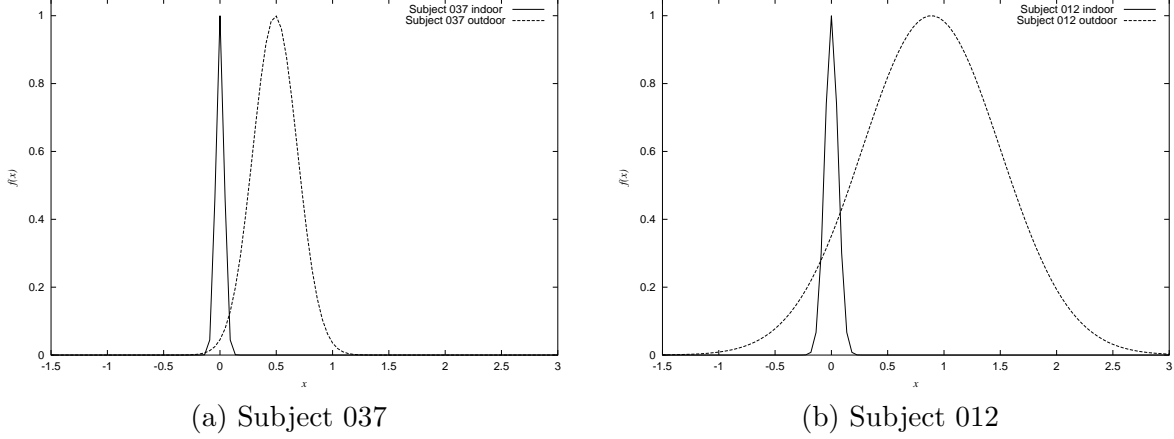(a) Subject 037

(b) Subject 012

Figure 4: Gaussian representation of the within-subject distributions for indoor and outdoor data for two different subjects.

Beyond gait research

The use of the database can be extended past that of gait-driven biometric research. In terms of technique development, the database allows for the evaluation of background removal techniques, high level moving-feature analysis techniques (both rigid and non-rigid) and face recognition analysis. As such, our new database allows not just for evaluation of the potential of gait as a biometric but also the evaluation of many other sequence based vision and image processing problems. The inside laboratory data enables the performance of an extraction technique to be compared with that of the near-ideal (baseline) chroma-keyed result. The outdoor data can be used to assess the performance of subject extraction techniques, providing data of single subjects, multiple subjects (both in the foreground and background) and rigid objects in the background eg. cars, buses, bicycles etc. The outdoor data allows for the analysis of these techniques under various weather environments as it includes data taken on bright cloudless days through to overcast, dull conditions. Further, the subjects themselves are non-rigid allowing the application of high-level feature analysis techniques. Finally, the high resolution stills are applicable to face recognition research, not only stand alone, but also allowing for research into the buttressing of gait and face biometrics, as face data exists for all gait subjects.

Conclusions

High recognition rates have already been observed on 100+ subjects using a variety of techniques, highlighting gait's potential as a biometric and justifying the detail behind the design and construction of the database. The chroma-key extraction does indeed allow for a basic ground truth that can be achieved with comparable ease, and the results can be compared with extraction by computer vision approaches. The new DV technology allows for ease in digitisation and has helped to mitigate the large storage requirements. This new database allows for estimation of

inter- and (limited) intra-subject variability on a large number of subjects, together with assessment of the efficacy of computer vision techniques that extract and describe moving objects, as the evaluation of gait covers many pragmatic factors. As such the data is applicable to many areas of interest within the vision community including moving object extraction algorithms, face recognition, and the buttressing of face and gait biometrics. Our future research will seek to capitalise the variability and extent offered by this new gait database and includes the construction of a new database designed specifically to enable the study of intra-subject variability, studying issues including, but not limited to, variations over time, effects of alternative footwear and the carrying of objects - this database is nearing completion. Early analysis of the outdoor data has confirmed the increased variance of features extracted from application scenario imagery as opposed to ground truth. Finally, further information and access to the database can be obtained at: *http://www.gait.ecs.soton.ac.uk/* and the database is available through contact with the authors.

Acknowledgements

References

[1] C. BenAbdelkader, R. Cutler, H. Nanda, and L. Davis. EigenGait: Motion-Based Recognition of People Using Image Self-Similarity. *Proc. AVBPA01*, :pp. 284–294, 2001.

[2] A. Y. Johnson and A. F. Bobick. A multi-view method for gait recognition using static body parameters. *Proc. AVBPA01*, :pp. 301–311, 2001.

[3] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. Baseline results for the challenge problem of human ID using gait analysis. *Proc. FGR02*, :pp. 137–142, 2002.

[4] P. S. Huang, C. J. Harris, and M. S. Nixon. Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, **13**:pp. 93–100, 1999.

[5] J. J. Little and J. E. Boyd. Recognising people by their gait: the shape of motion. *Videre*, **1**(2):pp. 2–32, 1998.

[6] J. D. Shutler and M. S. Nixon. Zernike velocity moments for the description and recognition of moving shapes. *Proc. BMVC01*, **2**:pp. 705–714, 2001.

[7] J. D. Shutler and M. G. Grant. The large Southampton HiD human gait database. Technical report, ISIS, University of Southampton, U.K., 2002.

[8] S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld. Detection and location of people in video images using adaptive fusion of color and edge information. *Proc. ICPR00*, **4**:pp. 627–630, 2000.

[9] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter. Automatic gait recognition by symmetry analysis. *Proc. AVBPA01*, :pp. 272–277, 2001.

[10] J. P. Foster, M. S. Nixon, and A. Prugel-Bennet. Gait recognition by moment based descriptors. *Proc. RASC02*, :accepted for publication, 2002.

[11] J. D. Shutler. Velocity Moments for Holistic Shape Description of Temporal Features. *PhD Thesis, University of Southampton, U.K.*, :, 2002.