An Autonomous Human Body Parts Detector Using A Laser Range-Finder

Lara Kornienko and Lindsay Kleeman
Monash University, Clayton
lara.kornienko@eng.monash.edu.edu

Abstract

This paper describes a human body-part detection system using a vertical laser scan. From a stationary Hokuyo laser range-finder, a scan is taken of the front-view of a human subject, where the scan falls on one side of their body to include the leg. As the subject walks towards the laser, the scan data is autonomously broken up into five body segments and the information compacted into feature points. Plotting the keypoints over scans shows the potential for them to be used as aids in a gait analysis and tracking system. Applications towards which these preliminary findings may contribute include tracking for surveillance operations and motion/behaviour analysis.

1 Introduction

The desire and need for sophisticated methods of observation and surveillance is increasing in modern societies. Companies are steering away from the man-powered systems of the past, which typically consist of human operators observing the information gathered by cameras at fixed positions. Often, the output is not of terribly high quality and this coupled with the reality of human fatigue and emotive responses when dealing with unclear data suggest that alternative means of observation must be explored for the future.

As computing power has increased and become more affordable, various ways of improving surveillance methods have been investigated. One such area is in the identification of individuals using methods from Artificial Intelligence.

Identification based on facial features is one of the more popular areas due to the uniqueness of the human face. However, this has its limitations since the face must be, at the very least, partially exposed to the system's data gathering component. Being such a small area, occlusions and the individual's orientation are problematic. Sunglasses, beards and hats can also confuse the system.

In order to increase the amount of usable information, and to build on the idea of face-recognition, methods that look at the whole human body have been explored. Of particular interest are methods of body-part detection and analysis and the related field of gait recognition.

This paper investigates a means of markerless human body-part detection using a vertical laser scan and the potential of the part descriptors to be used in a gait analysis and tracking system.

To begin the analysis, vertical laser scans of a person walking towards a stationary range-finder are autonomously partitioned into torso, thigh, calf, hip and knee regions (see Figure 1) and described via a set of feature points. Over a series of scans, positions of these feature points over scans are plotted and examined.

Section 2 gives a brief overview of what has been achieved in the field of gait recognition/analysis as a whole, while Section 3 will detail the physical set-up of the gait analysis system. Section 4 gives a description of the algorithms used for body-part detection and the results of this are given in Section 5, followed by a brief discussion in Section 6. A concluding statement is given in Section 7.

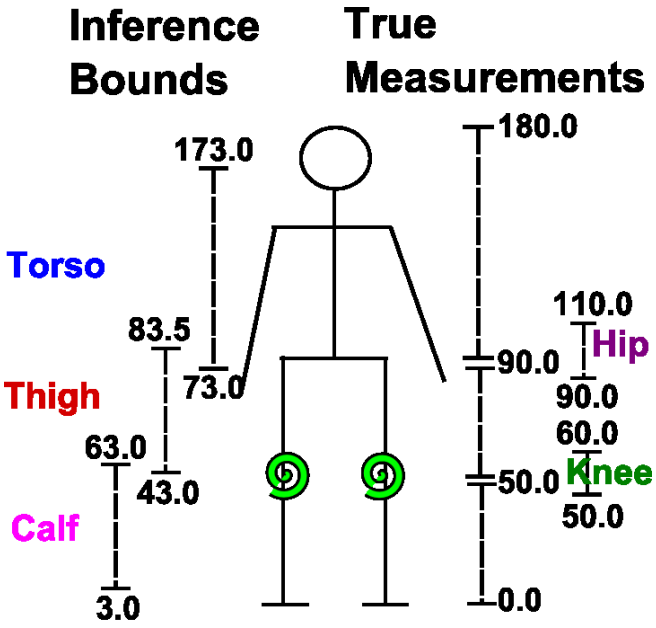

Figure 1: Keypoint Bounds and True Measurements of Subject

2 Part Detection and Gait Analysis Background

Gait has been defined as 'The coordinated, cyclic combination of movements that result in human locomotion' [Boyd and Little, 2003]. The analysis of human gait comes down to the observance of a subset of attributes characteristic of the way a human moves when they walk, but it is by no means restricted to a fixed number of these attributes. Those attributes used and the way in which they are used and measured depend very much on the application.

For instance, gait analysis for medical situations is used to detect abnormalities in a patients gait to aid in the rehabilitation of the patient or for the appropriate design of prosthetics for the patient. Typically, in such laboratory settings, markers are placed at strategic positions on the patient's body and they are then made to walk down a fixed path whilst several cameras capture their motion. The footage is then analysed in order to measure the amount of deviation in this gait in comparison to an 'ordinary' gait by looking at how the positions of the markers vary.

However, using markers on subjects has great disadvantages in surveillance operations or when subjects are to be viewed in a 'natural' environment. Thus, research into markerless systems has become quite popular as subjects can be viewed in any number of different environments and situations.

The majority of current markerless gait and motion-based methods use vision only. In these systems, a popular way to detect specific body parts is to use prior models, which may also be supplemented by the use of real biometric data such as joint angles. In [Dorthe Meyer and Niemann, 1998], body parts are abstracted into blocks that are successively identified and connected together to form a human hypothesis. Trajectories of these body parts are then classified via Hidden Markov models to obtain an estimate of the overall motion model.

An example of a technique which make use of both a parts model and joint and pose angles in the human form is in [Hedvig Sidenbladh and Fleet, 2000]. Here, a Bayesian framework is utilised to define a generative model of image appearance, then a detailed model using prior probabilities over pose and joint angles is devised to represent how people move. [Ramanan and Forsyth, 2003] models the human body as a 'puppet of coloured, textured rectangles' in order to track them across image frames and [Wagg and Nixon, 2002] uses anatomical data such as degree of hip, knee and ankle rotation and pelvic life to generate shape models which are consistent with normal body proportions.

There has also been work which uses feature points [Lowe, 2004; 1999; Lindeberg, 1998; 1994; Laptev and Lindeberg, 2003; Mikolajczyk and Schmid, 2004; Moreels and Perona, 2005] as a part detector in humans. For instance, in [K. Mikolajczyk and Zisserman, 2004], work has been done using image feature points/descriptors to describe people in static images as a parts model.

Less work in this area has been done for laser-based part-detection and gait analysis systems, though tracking problems are fairly common [Taylor and Kleeman, ; Ajo Fod and Mataric, 2002]. Combined laser and vision systems for tracking have also been developed [Brooks and Williams, 2003] and the advantages of using a laser in conjunction with a camera are being realised. For instance, a laser provides a good means of occlusion resolution due to the fact that they supply range information.

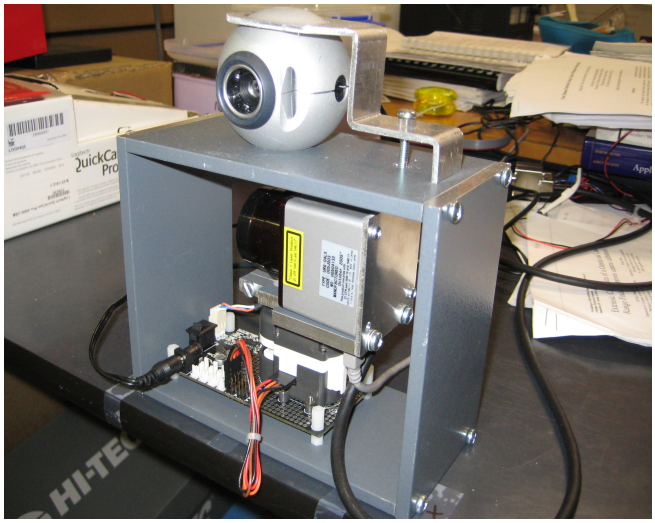

Figure 2: Laser Set-Up

The use of a laser alone or as an aid to another sensor in gait analysis has not been terribly popular perhaps due to the range limitations of the devices. However, as technologies improve, so will the range characteristics of lasers. Another deterrent may be in the somewhat limited information that they provide. However, this can be taken as being an advantage if one has a very specific type of information they want to extract and they want to do it quickly.

Furthermore, because the main information extracted by the laser is range, they are a perfect accompaniment to a single camera as an image does not convey depth. In a controlled setting, such as presented in this paper, a simple range-laser may easily provide information of the shape and size of a subject and their motion trajectories. Thus, these devices are certainly worth exploring further in the context of human shape analysis and tracking.

3 Physical Setup of the Gait Analysis System

The system described in this paper consists of a single Hokuyo range-finder laser on a table of height 720mm situated in a laboratory exposed to both natural and artificial light (see Figure 2). The Hokuyo laser is infrared, level 1 safety class (unobtrusive and non-detectable by the human eye), have a rotation speed of 10hz in a 240 degree angle range, 4m range accuracy and 1mm resolution. To obtain a vertical scan, it is positioned on its side

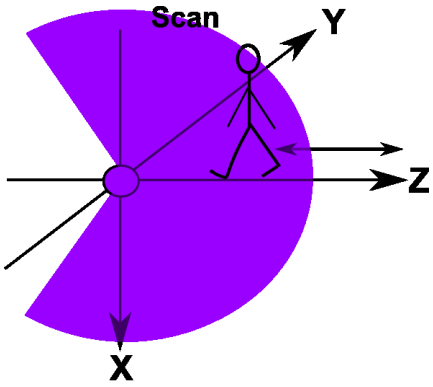

Figure 3: The Laser's Coordinate System

such that the x-axis of the laser becomes the vertical axis in the real-world space, the y-axis is the horizontal axis and the z-axis points outwards directly in front of the laser (see Figure 3).

The scan is aligned with a line on the ground, and the subjects are required to walk along it towards the laser at a regular pace. The alignment is such that the scan falls upon the left side of the subject's body, capturing the motion of the leg as it swings in the walking process. Since the laser has a maximum range of 4 metres only, the subjects must repeat the walk several times to provide enough information.

In order to minimise any 'junk' data, range data of the floor and the ceiling are filtered out by fitting lines to this data using RANSAC [Fischler and Rolles, 1981; Wang and Suter, 2005]. All those points that lie within 50mm of the floor estimate and 250mm of the ceiling estimate (to allow for light fittings) are ignored.

4 Human Body-Part Detection

The aim of the body-part detection system described in this section is to identify five regions of the human body from the data of a single vertical scan. All regions are described by a single keypoint, as will be discussed shortly. The first three of the five regions, the torso, thigh and calf, are called 'primary regions' or parts as they are determined from the laser data alone and independently of each other. The last two regions, the hip and the knee are 'secondary regions' or parts as they are estimated from the locations of the primary regions, thus if one of the primary regions do not exist, one of the secondary regions will not exist either.

The primary regions are found using a prior model on the position (that is, estimates of the height above the ground) of each part and orientation of the part. The height estimates are loose enough that they can work for people of average height (both men and women). The orientations only apply to the torso, as when a person is walking, one would not expect them to lean forward or backwards more than 45 degrees from the vertical. Whereas, the other body parts can have more variation in orientation, for instance, the thigh and calf can be positioned both horizontally and perpendicular to the ground, depending on the type of motion of the subject. Hence, it was not though necessary to pose orientation constraints on these parts. How the primary parts were found and represented will be discussed now in sections 4.1, 4.2, 4.3 and 4.4.

4.1 Finding Segments in the Data

In order to extract human body-parts in the laser range data, it was first roughly segmented into regions that may possibly contain the required parts. These estimated segments were then passed successively into RANSAC (Random Sample Consensus), which is a model-fitting algorithm highly robust to outliers [Fischler and Rolles, 1981], in order to fit straight-line models to the data. Those points which were considered 'inliers' to the models chosen by RANSAC were stored as separate clusters, to be processed further. Aside from clustering, this stage in processing also served to remove outliers in the data.

After all the clusters had been found, the next step was to go through and make sure there were no 'useless' clusters, that is, small groups of points that could easily be merged with other groups. Thus, if a 'small' cluster was found (less than 100 points), the positions of its endpoints were found and also its centroid. The small cluster was then merged with the cluster with the closest endpoint and centroid. Given that this was an iterative process, once a new cluster was formed, if it was also considered 'small', it may be merged with another cluster at a later stage.

4.2 Moments of Areas: Obtaining Keypoints and Orientations

In order to build the keypoint descriptors for the clusters, a more compact representation for each cluster was found. Firstly, the centroid of the cluster was calculated via its Moments of Area, where to find the Moment of Area, M_{pq} of a set of group of two-dimensional points,

$$m_{pq} = \sum_i \sum_j i^p j^q \tag{1}$$

This two-dimensional formulation could be used for the laser data as the y-axis (horizontal axis - see Figure 3) remained at zero.

Then, to find the Centre of Area, (i_0, j_0) (the centroid), for a set of n points one uses,

$$i_0 = \frac{m_{10}}{m_{00}} = \frac{1}{A} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} i \tag{2}$$

$$j_0 = \frac{m_{01}}{m_{00}} = \frac{1}{A} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} j \tag{3}$$

where $A = m_{00}$ represents the area of the region in question.

Secondly, the orientation, θ in radians with respect to the horizontal of each cluster was found using 1,

$$\theta = \frac{1}{2} \arctan\left[\frac{2(m_{00}m_{11} - m_{01}m_{10})}{(m_{00}m_{02} - m_{01}^2) - (m_{00}m_{20} - m_{10}^2)}\right] \tag{4}$$

Note: the indices of the moments in this expression may differ to other texts due to the axis arrangement of the laser system used in this paper.

Therefore, each cluster was replaced by a keypoint descriptor of the cluster's centroid, orientation and also the minimum and maximum range values for each dimension.

4.3 Identifying the Primary Regions/Parts

Now that the laser data was in a more compact form, the keypoints were analysed in order to assign them to a primary body part. As mentioned earlier, this was done via a prior model on the vertical positioning of body parts (using the positions of the centroids), but also on the size of the segments (via the minimum and maximum range values) and, in the torso's case, the orientation (via θ, see 4).

A keypoint, was assigned an i.d. of either 1 for torso, 2 for thigh, 3 for calf, 4 for 'something but

don't know what' or -1 if not assigned (for example, if there was no person in the scan or there was not enough data to represent a part). It was found that the parts were found whenever there was sufficient data to represent it. Thus, because the torso is such a large, stable area, it was found whenever there was a person present in the scan. However, the least stable part was the calf - scan data for this region was sometimes missing due to factors such as the creases in flared pants causing a fragmentation of the data, the fact that the calf has a narrow horizontal cross-section, the calf is furthest away from the centre of mass of the person and therefore more likely to have unstable motion, and that when people move their leg back, the calf tends to move to the centre of the body and then swings back to the front to its original position. This would lead to the calf being out of range of the vertical laser scan in some cases.

4.4 Identifying the Secondary Regions/Parts

Once the primary regions/parts have been assigned, the secondary regions/parts may now be assigned. These regions refer to the hips and knee of the subject. To find the hip, the torso and thigh parts must have been assigned because the possibly overlapping end-points of these regions define the estimate for the hip keypoint. The hip keypoint is assigned to the centroid of the difference between the lower part of the torso cluster and upper part of the thigh cluster. Similarly, the knee keypoint is assigned to the centroid of the difference between the lower part of the thigh cluster and the upper part of the calf cluster.

As one might expect, the presence and location of these secondary parts are not as reliable as for the primary parts. This is largely due to their reliance on the presence of two primary parts and the fact that they are calculated from endpoints in the clusters, which may not be as reliable as if one used central points. On noting this, more reliable means of finding the secondary parts will be investigated in the future, such as using the centroids of the primary parts.

Figure 4: View of Person from Laser's Point of View

5 Results

In order to test the body-part detection algorithm, a subject was asked to walk towards the laser at a regular pace such that their left foot was aligned with a line on the ground representing the line of the vertical laser scan. Figure 4 shows an image taken from a web-cam positioned on top of the laser of a person as they walk towards the laser. Figure 5 shows one scan for a particular walking sequence. The laser is represented by the small circle at the left of the image and the horizontal and vertical lines emanating from it are the z and x-axes of the laser respectively. Unfortunately, the range of the laser is only 4 metres, therefore, the subject could only take about four steps whilst within range. Furthermore, only about three steps could reliably be scanned because the last step was too close to the laser and did not result in a full body scan. The scan data itself was of limited reliability too most likely due to clothing variations and swaying in the gait motion, where significant gaps in the laser stripe were reasonably common. The subject was asked to repeat the walk thirteen times.

Figure 6 shows the inferred keypoints for the scan shown in Figure 5. Each number points to a keypoint corresponds to a Primary region, where keypoint 0 represents the torso, keypoint 1 the thigh and keypoint 2 the calf. Keypoints 3 and 4 point to junk data and were given the part ID of -1, meaning 'not assigned'. The Secondary regions, the hip and the knee, are found and labeled appropriately in the image. The red lines coming from the keypoints

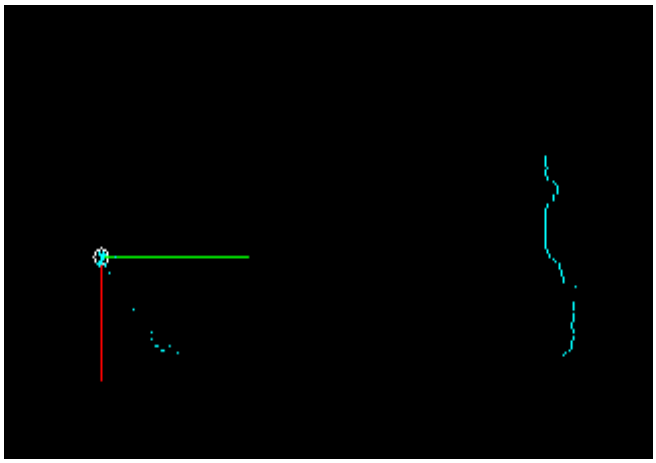

Figure 5: A Sample Scan

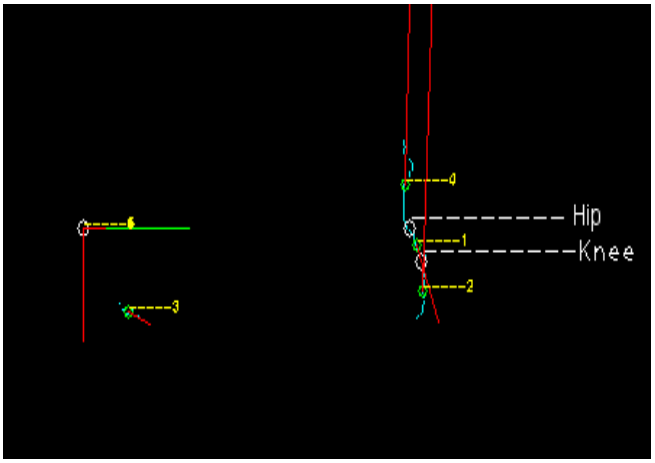

Figure 6: Keypoints Found in Scan

represent the orientation of the corresponding cluster.

To test the accuracy of the keypoint locations, the true body-part bounds of the subject were taken (see Figure 1). An error occurs if the inferred keypoint is outside the bounds of the true bounds. Note that the reason why stricter conditions for errors were not used (such as actual location of true centroid) is that for a part to be detected, the main condition is that a keypoint is *within* the relevant region. Hence, further restrictions on keypoint location are superfluous. The mathematical form of the error, ξ used is:

$$\xi = \frac{N_{out}}{N_{person}}$$

where N_{out} is the number of times a keypoint fell outside the true bounds and N_{person} is the total number of times a person was detected in the scan. This error is not a generous one as it also takes into account when a part was not found at all, where it can be argued that this is a problem with the laser and physical set-up of the system rather than with the algorithm itself. For instance, it does not take into account the times when the scan contained large areas of missing data, for example, if the scan beam did not fall on the calf of the subject. In response to this, future work will incorporate Bayesian inference of the locations of missing keypoints (see Section 6 for a further discussion of this).

Table 1 shows the error ξ for each keypoint type given the true bounds of the subject:

Torso	**Thigh**	**Calf**	**Hip**	**Knee**
0.0	0.0	0.0038	0.0097	0.1397

Table 1: Errors of Keypoint Location

To examine the overall displacement of the features over scans, it was found that the horizontal displacements of keypoints provided some information about the greater motion of the person - each peak corresponding to the 'start' of a scan (because the subject was asked to walk towards the laser) and the clusters of points around -1 indicate the scans when no person was in the scene or a keypoint was not found. Figure 7 shows the horizontal displacement of the torso. Furthermore, examining one of the walking trajectories (one walk to the laser) in more detail, Figure 8 indicates that it is possible to extract some information about the swing of the thigh and calf in comparison to the torso. It can be seen from this graph that the thigh and calf trajectories oscillate about the torso trajectory in a backwards and forwards motion. Examining this further, if we subtract the thigh and calf horizontal positions from the torso horizontal position, Figure 9 shows the oscillatory behaviour of these parts for a single walking trajectory.

The plots of the vertical displacements are slightly more difficult to analyse as differences in vertical keypoint component could be a mix of variations due to the data/algorithm and also due to the up/down motion of the subject as they walked.

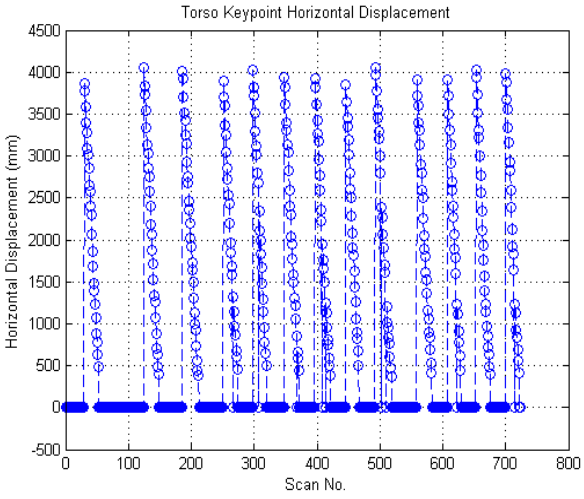

Figure 7: The Horizontal Displacement of the Torso Keypoint

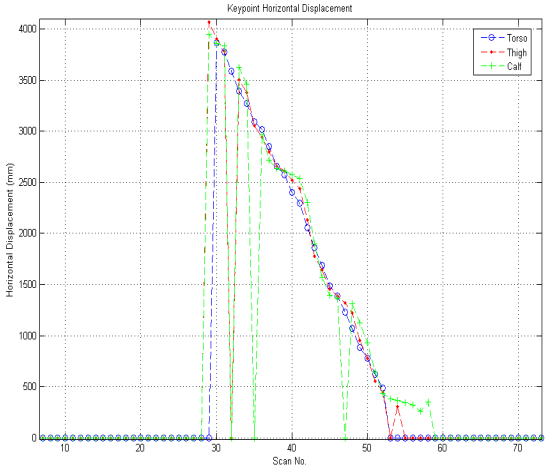

Figure 8: The Horizontal Displacement of the Thigh and Calf Against the Torso Keypoint

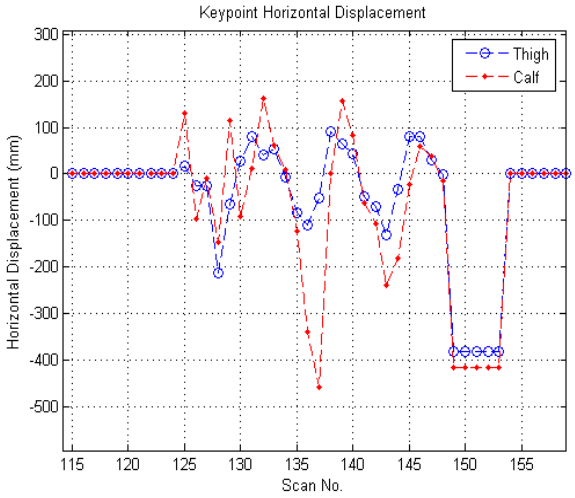

Figure 9: The Horizontal Displacement of the Thigh and Calf Relative to the Torso Keypoint

However, looking at Figure 10, it can be seen that the keypoint positions tend to cluster, where there is about a 200mm variation in the torso and calf vertical positions and a 100mm variation in the thigh.

6 Discussion

The plan for the method of body-part detection presented in this paper is to aid in the analysis of gait for a given subject, where the overall aim to to be able to build unique gait signatures for individuals. The reliability of the keypoints found by this method and the oscillatory motion present in the plots is encouragement enough to pursue this area further. Although this is a good starting point, the system has its limitations. For starters, the data sets are far too small. This is partially due to limitations in the Hokuyo laser, hence it may be replaced by one with larger range, higher scan-rate and better resolution to allow for more data of higher quality to be collected. An alternative would be to put the laser on a moving robot or sliding track to allow it to move with the person, eliminating the restriction on range. In the mean time, experiments will just have to be repeated several more times.

To deal with missing data, bayesian inference may be used to infer the position of missing datum and also to track persons over scans based on their motion. Camera data will also be introduced to the analysis. A set-up that is currently being worked on

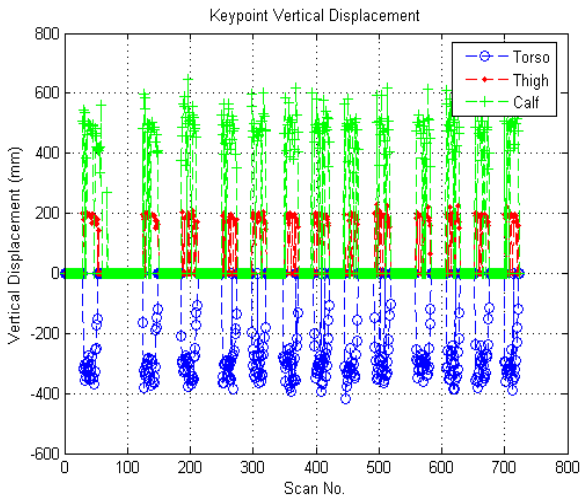

Figure 10: The Vertical Displacements of the Torso, Thigh and Calf Keypoints

combines a camera with the laser which is situated on a servo to move it to positions indicated by the camera data.

With reference to *how* the gait pattern of an individual may be analysed, an interesting study on the dynamics of human locomotion is given in [Perc, 2005]. This paper suggests that human gait may be analysed using simple nonlinear time series, and that the properties of the gaits tested are typical of deterministic chaotic systems. It would be of interest to pursue these findings in conjunction with the findings in this paper.

7 Conclusion

This paper presents a human body-part detection system using a vertical laser scan. The algorithm segments the range data into three primary regions: the torso, thigh and calf and two secondary regions, found using the primary regions: the hip and the knee. Each region is described by a single keypoint located at the centroid of the data segment/cluster, which also contains information on the orientation of the part and the extreme points of the cluster in three dimensions.

It was found that the keypoints were reliable and on further analysis, it was shown that their horizontal and vertical displacements over scans could be used to extract information about the subject's overall locations over time (i.e. the path they have taken). Interestingly, by examining the horizontal displacement trajectories further, it was found that the keypoints could also be used to analyse the motion of the thigh and calf relative to the torso. Future investigations will examine the potential for these keypoints to be used in a gait/motion analysis system and a gait-based tracking system for surveillance operations.

References

[Ajo Fod and Mataric, 2002] Andrew Howard Ajo Fod and Maja J. Mataric. Laser-based people tracking. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3024–3029, Washington DC, May 2002.

[Boyd and Little, 2003] Jeffrey E. Boyd and James J. Little. Biometric gait recognition. In *Advanced Studies in Biometrics: Summer School on Biometrics*, Alghero, Italy, June 2-6 2003.

[Brooks and Williams, 2003] Alex Brooks and Stefan Williams. Tracking People with Networks of Heterogeneous Sensors. In *Proceedings of the Australasian Conference on Robotics and Automation*, pages 1–7, 2003.

[Dorthe Meyer and Niemann, 1998] Josef Posl Dorthe Meyer and Heinrich Niemann. Gait classification with hmms for trajectories of body parts extracted by mixture densities. In *British Machine Vision Conference*, 1998.

[Fischler and Rolles, 1981] M.A. Fischler and R.C. Rolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In *Commun. ACM*, volume 24, pages 381–395, 1981.

[Hedvig Sidenbladh and Fleet, 2000] Michael J. Black Hedvig Sidenbladh and David J. Fleet. Stochastic tracking of 3d human figures using 2d image motion. In *European Conference on Computer Vision, ECCV'2000*, 2000.

[K. Mikolajczyk and Zisserman, 2004] C. Schmidt K. Mikolajczyk and A. Zisserman. Human detection based on a probabilistic assembly of robust part detectors. In *European Conference on Computer Vision*, volume 1, pages 69–81, 2004.

[Laptev and Lindeberg, 2003] Ivan Laptev and Tony Lindeberg. Space-time Interest Points.

In *Ninth IEEE International Conference on Computer Vision (ICCV 2003), Volume 1*, pages 432–439, 2003.

[Lindeberg, 1994] Tony Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994. ISBN 0-7923-9418-6.

[Lindeberg, 1998] Tony Lindeberg. Feature Detection with Automatic Scale Selection. Technical Report ISRN KTH/NA/P-96/18-SE, Computational Vision and Active Perception Laboratory (CVAP), Dep. of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), S-100 44 Stockholm, Sweden, August 1998.

[Lowe, 1999] David G. Lowe. Object Recognition from Local Scale-Invariant Features. In *International Conference on Computer Vision*, pages 1150–1157, Corfu, Greece, September 1999.

[Lowe, 2004] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[Mikolajczyk and Schmid, 2004] Krystian Mikolajczyk and Cordelia Schmid. A performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004.

[Moreels and Perona, 2005] Pierre Moreels and Pietro Perona. Evaluation of Features Detectors and Descriptors Based on 3D Objects. *International Conference on Computer Vision*, 1:800–807, 2005.

[Perc, 2005] Matjaz Perc. The dynamics of human gait. In *European Journal of Physics*, volume 26, pages 525–534, 2005.

[Ramanan and Forsyth, 2003] Deva Ramanan and D.A. Forsyth. Finding and Tracking People from the Bottom Up. In *Computer Vision and Pattern Recognition (CVPR)*, Madison, WI, June 2003.

[Taylor and Kleeman,] Geoffrey Taylor and Lindsay Kleeman. A Multiple Hypothesis Walking Person Tracker with Switched Dynamic Model. ARC Centre for Perceptive and Intelligent Machines in Complex Environments. Monash University, Australia.

[Wagg and Nixon, 2002] David K. Wagg and Mark S. Nixon. On automated model-based extraction and analysis of gait. In *Proceedings of 6th International Conference on Automatic Face and Gesture Recognition*, pages 11–16, Seoul, South Korea, 2002.

[Wang and Suter, 2005] Hanzi Wang and David Suter. Tracking and Segmenting People with Occlusions By a Sample Consensus Based Method. In *IEEE International Conference on Image Processing (ICIP)*, Genova, Italy, Sept 2005. Pages to appear.