

Head-to-shoulder signature for person recognition

Nathan Kirchner, Alen Alempijevic, Alexander Virgona

Centre for Autonomous Systems

University of Technology Sydney, Sydney, Australia

{nathan.kirchner, alen.alempijevic, alexander.virgona}@uts.edu.au

Abstract—Ensuring that an interaction is initiated with a particular and unsuspecting member of a group is a complex task. As a first step the robot must effectively, expediently and reliably recognise the humans as they carry on with their typical behaviours (*in situ*). A method for constructing a scale and viewing angle robust feature vector (from analysing a 3D pointcloud) designed to encapsulate the inter-person variations in the size and shape of the people's head to shoulder region (Head-to-shoulder signature - HSS) is presented. Furthermore, a method for utilising said feature vector as the basis of person recognition via a Support-Vector Machine is detailed. An empirical study was performed in which person recognition was attempted on *in situ* data collected from 25 participants over 5 days in a office environment. The results report a mean accuracy over the 5 days of 78.15% and a peak accuracy 100% for 9 participants. Further, the results show a considerably better-than-random ($1/23 = 4.5\%$) result for when the participants were: in motion and unaware they were being scanned (52.11%), in motion and face directly away from the sensor (36.04%), and post variations in their general appearance. Finally, the results show the HSS has considerable ability to accommodate for a person's head, shoulder and body rotation relative to the sensor - even in cases where the person is faced directly away from the robot.

I. INTRODUCTION

The reach of robotics - fuelled by the desire to see robots ubiquitous in society and by the growing capabilities and affordability of sensing, computing and actuation technologies - is expanding. For instance, a robot's potential to increase the quality of life of a range of individuals and to address several social issues have been demonstrated [1] and [2]. However, whilst the fundamental capabilities of robots to perceive and interact with the world are maturing, many challenges remain to be overcome before robots can effectively join society [3]. One such challenge is the recognition of people *in situ*; i.e. recognition of people engaged in typical behaviours and under typical environmental constraints.

For instance, consider the exemplar scenario presented in Fig.1. In this scenario the robot is tasked with delivering an object to a specific person at a specific location in an office environment, however upon arrival the robot encounters more than one person. Whilst at this point the robot could move the onus of responsibility onto the humans (by, for example, uttering the intended recipient's name or assuming that he/she will identify themselves) this potentially opens the interaction to undesirable outcomes, thus reducing the likelihood of task success. For example, a salient object may cause the human/s to misrepresent their identity in order to gain possession of the object, or to simply take the object from the robot [3]. At a minimum, this approach would

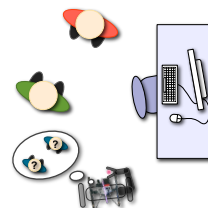


Fig. 1. The Exemplar Scenario: A robot tasked with ensuring a specific person is the recipient of an object encounters multiple persons engaged in an interaction upon arrival at the hand-over location.

require the humans to terminate their interaction in order to deal with the robot's intrusion. Clearly then, equipping the robot with person recognition technology capable of recognising the intended recipient *in situ* and whilst the robot approaches is desirable.

A number of potential solutions to the person recognition problem are available in the literature. Broadly speaking, these solutions can be dichotomised as near field or non-near field. An indicative example of a near field approach is the tactile approach presented in [4]. Whilst this technique was shown capable of person recognition it requires physical contact between the human and the sensor/robot. Another indicative example, one which loosens the near field constraint, is the audio based approach to person recognition presented and demonstrated viable in [5]. Whilst this approach accommodates for an increased human-sensor/robot separation the approach can still be considered near field. The detrimental effect of background noise and/or an audibly cluttered environment typically limits the operation range to around one meter or less [5]. The aforementioned is not a comprehensive review of near field person recognition research. However, it is sufficient to demonstrate that although these approaches can produce impressive results, they are inherently limited by their core technologies to be near field. Whilst this is suitable for a range of contexts, clearly a near field approach is not suitable for contexts such as those exemplified in Fig.1.

Several representative examples of non-near field approaches are [6], [7], [8] and [9]. In [6] the authors provide a treatment for several of the challenges that plague vision-based face recognition (partial occlusion and variations in illumination and facial expressions) through a training data selection process and demonstrate impressive results. However, this approach requires a prescribed set of training data (potentially problematic to generate in our context) and

relies on face-area detection (which requires the human to be somewhat pose). A considerably different approach to face recognition is presented in [7]. In this work, the authors exploit inter-person volumetric variations reconstructed from a dense 3D pointcloud of the face to recognise people. The use of 3D features from a pointcloud is shown to improve robustness against the aforementioned issues of variations in facial expressions and illumination. However this technique again relies on constrained observations (point of view and the human's pose) and the face being observable and detectable.

Another interesting approach to person recognition is through behavioural characteristics such as a person's gait. Indicative examples are [8] and [9] where the dynamic characteristics of a person's gait are derived from a series of silhouette images of the person walking. In both cases the silhouette images were constructed from a vision sensor and then utilised as the basis for classification. The coarseness of the underlying representation enabled recognition at greater distances and significantly improved robustness to variations in illumination, facial expressions and head-pose. However, prohibitive issues remain - these method cannot accommodate gross shape deformation (e.g. sitting), stationary (e.g. standing still) or reasonably foreseeable occlusions (e.g. a desk occluding a person's lower body).

Although it seems that non-near field approaches are more suited for our target context the aforementioned limitations prohibit direct application. Rather, there is a need for an approach to non-near field person recognition that is robust to the variable conditions reasonable to expect in the home/office type contexts exemplified, that is able to operate with typical *in situ* observations and that utilises *in situ* observable features for recognition. The method should employ a feature set that does not typically undergo gross shape deformation (either while standing/sitting or in motion) and that is generally not occluded. The typical visibility of the head-to-shoulder region of the human body and its relative stability of shape in our context (as can be intuitively appreciated from our exemplar scenario) makes it an attractive candidate for exploitation. Furthermore, [10] and [11] (although intended for person detection only) has shown the characterisation of people through geometric features computed on horizontal slices of range data to be robustly observable *in situ* and useful.

This paper presents a method of exploiting the inter-person variations in the size and shape of peoples' head, neck and shoulders to achieve robust *in situ* person recognition. A scale and viewing angle robust feature vector (constructed from analysing a 3D pointcloud) designed to encapsulate this information, here within referred to as the head-to-shoulder signature (HSS), was devised and utilised as the basis for Support-Vector Machine (SVM) based classification. Results from an empirical study designed to evaluate our HSS based person recognition's robustness to variations in the person's pose and attire/general-appearance are presented. These results demonstrate that our HSS based person recognition recognises people (N=23) with a single observation

at 78.15% mean accuracy while standing still (SS), with 52.11% mean accuracy while walking towards (WT) the sensor and with 36.04% mean accuracy while walking away (WA) from the sensor in the case that training data (naturally occurring and subsequently manually labelled) was collected on Day 0 and testing data was collected over the following 4 successive days. Further, the results show a considerably better-than-random (4.5%) result for when the participants were: in motion and unaware they were being scanned (52.11%), in motion and face directly away from the sensor (36.04%), and post variations in their general appearance. Finally, results are present that explore the limitations of our HSS to accommodate persons head and/or shoulder and/or body rotation relative to the sensor.

The breakdown of this paper is as follows: Section II discusses the potential of exploiting inter-person variations in the size and shape of peoples' head, neck and shoulders as a basis upon which to recognise people. Section III details our method for encapsulating this information within the HSS. An empirical evaluation of the HSS as a basis upon which to classify is presented in Section IV. Finally, conclusions are drawn and future work proposed in Section V.

II. BACKGROUND

Given the observability of the head-to-shoulder region, the first step is to identify a set of metrics of this region that will facilitate recognition. Obvious choices include facial and/or vocal features, however, as previously discussed these approaches have limitations that render them less than ideal for our target context. The remaining available features are geometric information contained within the size and shape of peoples' head, neck and shoulders. The question becomes: does such geometric information vary sufficiently within the population to be descriptive?

An anthropometric survey was conducted in 1988 by the US military to facilitate design of clothing, equipment and workspaces that accommodated the full range body sizes of its personnel [12]. This survey documents 180 anthropometric dimensions measured from almost 9000 soldiers. Amongst this set of measurements a number relate directly to the head-to-shoulder region of the body. For instance, Head Breadth, Head Length, Neck and Shoulder Circumference, to name a few. From [12] it seems that the standard deviation (σ) of these measurements tends to be 5% of the measurement or greater, indicating that significant variation in these anthropometric dimensions exists within this population.

Given the capabilities of the available sensors (relative to variations in anthropometric dimensions) a proof of concept study was undertaken in order to further investigate the possibility of person identification using physical dimensions of the human head-to-shoulder region. Firstly, 3D data from 50 pointcloud scans (via a SwissRanger SR4000) of 25 individuals standing in front of the sensor, at a distance of 2 metres, facing towards it were collected. The physical dimensions of the head and shoulders of each person were measured directly from the collected pointclouds. Results obtained for the head and shoulder spans noted in Fig. 2

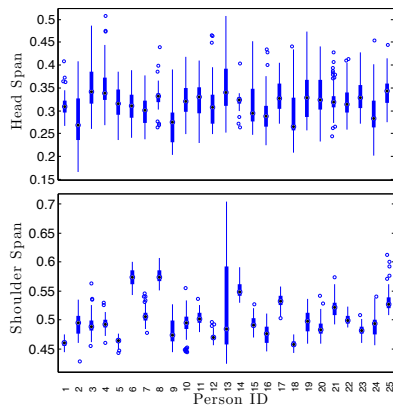


Fig. 2. Box plots of head and shoulder spans showing significant separation between individuals relative to the spread of data within each individual.

where the solid dots indicate the median measure for each individual and the box represents the 25th - 75th percentile range. Observing Fig. 2, there is a separation between the median result of each person’s head and shoulder span, suggesting potential usefulness for discriminating between people. Furthermore, the variation of results for each person was small relative to the separation between people. This proof of concept result suggests that not only is there significant variation between individuals, but that it appears to be measurable.

III. METHOD

With preliminary investigations demonstrating potential usefulness of the head-to-shoulder region for person recognition, a method for constructing HSS was devised and a person recognition system was developed. The system, illustrated in Fig. 3, consists of three stages: Person Detection, HSS Construction and Person Recognition. The Person Detection stage takes 3D depth data (e.g. from a Microsoft Kinect or SwissRanger SR4000) and analyses the scene to detect any people. Segments of the pointclouds determined to be people are then passed to the HSS Construction stage and HSS encapsulation of the head-to-shoulder region of the detected person are constructed. The HSS are then coarsely verified as belonging to a human, the positives are classified and the identity label outputted from the system. Each of these three stages are detailed in the follow sub-sections.

A. Person Detection

The first stage of our system robustly detects people in a 3D pointcloud of the scene and segments the pointcloud accordingly. As this stage is built upon the authors’ previous work [13] an overview will be presented here, for the sake of completeness, and the reader is referred to our original publication for full details. As can be seen in Fig. 3, the Person Detection process begins with the input pointcloud being projected onto a 2D horizontal plane via a bivariate histogram across the two horizontal axes (Scene Flattening). This results in a density image where the intensity of each pixel (the bins of the bivariate histogram) represent the

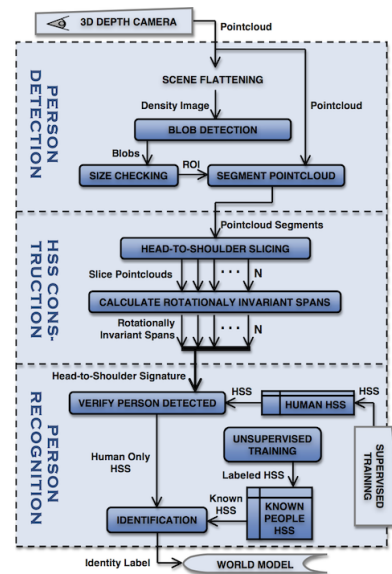


Fig. 3. Schematic of our HSS based person recognition system

concentration of the points that exist in that vertical column of the scene - vertical surfaces result in high intensity pixels. Blob Detection is then performed on the image to allow each apparent person (vertical surface of appropriate width - Size Checking) to be segmented from the scene (Segmented Pointcloud). This approach to scene analysis exploits the assumption that people appear in a scene as vertical surfaces with a relatively high point-density in the histogram, when compared with the many horizontal surfaces (floor, chairs, etc.). This assumption, as shown in [13], holds true.

Along with humans, common items identified as potentially human include: walls, doors, and tall items of furniture. Many of these false positives are eliminated by setting size constraints based on the expected minimum and maximum size of a person (Size Checking). The Size Checking is ad-hoc and serves only to limit the number of observations that are further processed in order to improve computational throughput. However, as seen in [13], false positives passing this stage of the detection process is preferable to false negatives as further discrimination is performed during the Person Recognition stage which generally removes the remaining false positives (details in Section III-C).

B. HSS Construction

The objective of the HSS construction stage is to encapsulate the characteristics likely to uniquely describe a person mentioned in Section II from the aforementioned segmented pointcloud of the person. Fig. 4 shows an indicative example of a segmented pointcloud; the characteristics of the person targeted for encapsulation are: width of head, width of neck, width of shoulders, height of head and height of neck. Although detecting the gross anatomy (head, neck, shoulders) and obtaining measurements directly is immediately appealing, such an approach would rely on accurate and robust identification/land-marking of physical features - a non trivial exercise in our intended operational context. For

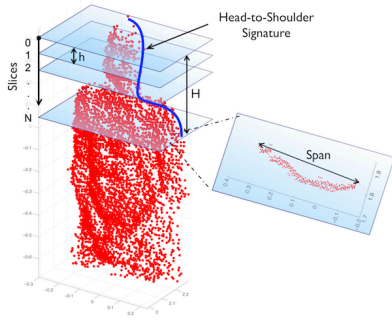


Fig. 4. Segmented pointcloud with the constructing elements of the HSS illustrated and a stylised HSS shown

this reason, a method for encapsulating the characteristics of the head and shoulders that does not rely on land-marking of physical features and that is robust to changes of viewing angle was devised.

As can be seen in Fig. 3, the first step of the HSS Construction stage is Head-to-Shoulder Slicing. Firstly, the point cloud was again segmented, the upper H is retained. The distance H is selected so that it covers from the top of the pointcloud (top of the head) down to the point where it includes the shoulders of most adults while minimising coverage of the arms, as shown in Fig. 4. $H = 0.4m$ was empirically determined to be suitable, although this may vary depending on the average size of the participants. In order to obtain lateral measurements the head-shoulders region of the pointcloud was divided into a number of horizontal slices of equal thickness h covering the region defined by H , as can be seen in Fig.4. Selection of the *slice thickness* is dependent upon the sensor used and the range of distances over which people are expected to be measured D_{max} . The equation, $h_{min} = S_{res}D_{max}\tan(\alpha)$, was used to determine the minimum slice thickness (h_{min}) given the vertical angular resolution of the sensor used (α), the number of points required for reliable measurement (S_{res}) and the maximum distance of people from the sensor (D_{max}). This ensures the possibility that data can exist in all slices.

At this point, each slice of the overall pointcloud (Slice Pointcloud) are analysed (Calculate Rotationally Invariant Spans) in turn to determine that slice's span. A slice of the pointcloud with its span labelled can be seen in the inset of Fig. 4, the span is defined as the largest distance that exists between any two points in the slice point cloud. As can be appreciated, in the case where the breadth of the person is observable from the point of view of the 3D depth sensor a representative span can be determined. In such, these span measurements are rotationally invariant. Clearly then, reiterating through each slice, lateral measurements of the head-to-shoulder region are explicitly measured.

The vertical characteristics are not explicitly measured however. Rather, the vertical characteristics are encapsulated in the HSS via the slice number in which the lateral measurements appear. For instance, consider Fig. 5 in which the pointcloud from two different people with 10 slice spans

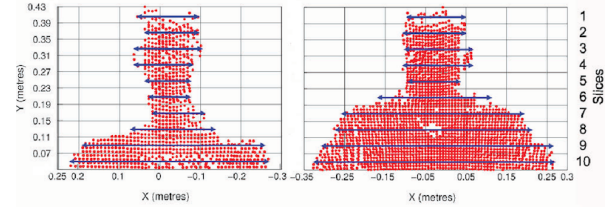


Fig. 5. Pointclouds of two different people with 10 slice spans illustrated

illustrated is shown (only 10 are noted for the sake of clarity). As can be seen from the figure, the slice numbers in which each individual's vertical characteristics varies. In this way, the vertical characteristics of the individual are implicitly encapsulated.

Finally, the HSS is defined as the ordered vector of the slice spans - which as previous demonstrated explicitly captures the lateral measurements and implicitly encapsulates the vertical characteristics (through the position of the lateral measurements within the ordered vector) of the head-to-shoulder region of an individual.

C. Person Recognition

The person recognition stage begins by verifying if the HSS supplied to it likely belongs to a person (Verify Person Detected). This is achieved via a single class support vector machine (SVM) trained (supervised) with HSS guaranteed to represent a range of people of diverse sizes and shapes. This broad classification is sufficient to remove the Person Detection false positives mentioned in Section III-A.

The remaining HSS are then passed to a multi-class SVM. The nature of our problem (number of features is less or equal to the number of samples) is suitable for a radial bias kernel function (RBF), gamma and cost parameters are optimised via cross-validation [14]. The multi-class SVM is trained (the robot autonomously selected observations to be used as training data) with HSS of detected people of unknown identity. The training data of each class of this SVM was guaranteed to have one novel person appear, but there was no guarantee of the person being observable in all observations, nor was there a guarantee of the person being in a particular or restricted pose. The Identification block (by means of the multi-class SVM) then produces a vector of probabilities that the observed HSS belongs to each of the known classes (people). Finally, the system outputs the Identify Label of the class corresponding to the largest high-confidence ($P > 0.8$) probability in the vector.

IV. RESULTS

This section presents results from an empirical evaluation of our HSS based person recognition system. There were 24 males and 1 female participants ageing between 22-45, the participants were PhD students and Academics from the Centre for Autonomous Systems (CAS) at the University of Technology, Sydney. No selection criteria was employed other than asking the participants if they believed that they



Fig. 6. The experiment setting

would be present at CAS on each of the 5 days of the experiment. Of the 25 participants, 2 (both male) have been excluded from the results as they only were present for 1 of the 5 days of the experiment.

The experiment was conducted as follows: Our HSS person recognition system was configured to operate using the Microsoft Kinect sensor on the RobotAssist platform [15]. The experiment was conducted in CAS, the robot was stationary throughout the entire experiment. The participants were then told to walk up to the ‘Stand Here’ line (which was 1.8m from the sensor), pause for 5 seconds and then walk away. The participants were told that the robot was going to attempt to recognise them, no further instructions were given. CAS does not require a uniform and the participants were not instructed/restricted on their attire. The experiment setting, the RobotAssist platform and a participant on the ‘Stand Here’ line can be seen in Fig. 6.

Data was collected for each participant from before they appeared in the scene, while they walked towards the robot, whilst at the line, while they walked away from the robot and for a period after they had left the scene. Participants were under the impression that data was only collected while they were standing at the line. Training data was collected for each participant at the line on Day 0. Testing data was collected on the following 4 successive days. The data was collected throughout the day depending on the time at which the participants were available; 18 of the 23 participants were present on all 4 testing days.

A. HSS Based Person Recognition

In order to gauge the general performance of our HSS based person recognition the data collected during the 5 days of the experiment were grouped into three categories: Days 1-4 standing on the line, Days 0-4 walking towards the line and Days 0-4 walking away from the line. The data from the participants standing on the line on Day 0 was used as the training data for the SVM. Cross validation of the SVM training data yielded 93.54% accuracy. The cross validation accuracy is mentioned as it is considered an indication of ‘best case’ performance. Fig. 7 shows the resulting confusion matrices. As can be seen by the strong diagonal (indicating that the label identified by our system matched the true label) the results for standing still on the line are reasonable. Specifically, the mean accuracy over the 5 days was 78.15%, the peak accuracy 100% (classes 1, 5, 7, 11, 12, 14, 16, 21 and 22) and the trough accuracy 0% (class 6). On inspection

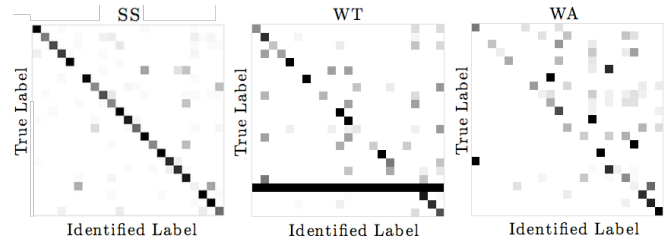


Fig. 7. The confusion matrices for the classifications of Walking Towards (WT), Standing Still (SS) and Walking Away (WA). The row number is the true label and the columns indicate the identified label; the pixels intensity represent the percentage of classifications (black = 100% and white = 0%). A solid black row indicates that no data was available for this participant.

it was found that participant 6 leant over, looked around and raised their arms considerably for the majority of their training data (so too, but to a lesser extent, participant 20). As for the WT and WA results, the mean accuracy was 52.11% and 36.04%, the peak accuracy 100% and 100% (for multiple classes) and the trough accuracy 0% and 0% (for multiple classes) respectively.

These results are of particular interest as they clearly show a considerably better-than-chance (1-in-23) result for when the participants were not aware that they were being scanned, while the participants were in motion and when the participants were face directly away from the sensor. This is significant as the participants were truly *in situ* and were not posing. At this point it must also be reiterated that these results are from single observations (at 30Hz) and no sensor-data processing or noise reduction has been performed. The end-results could be significantly improved by implementing a successive scan probabilistic approach. However, as the focus of this work is evaluating the HSS (our contribution) usefulness, such avenues have not been pursued.

B. Robustness to Variations in General Appearance

As previously mentioned, the experiment was run over a period of 5 successive days (with training occurring on Day 0). The experiment was constructed so as to capture data of the participants in which typical variations to their general appearance occurred. Fig. 8 shows the mean (over classes) accuracies for WT, SS and WA for each day of the experiment. As can be seen from Fig. 8 the mean accuracies were relatively stable for days 0 and 1, and for days 2, 3, and 4, for WT, SS and WA. However, there is a notable step between the two intervals. This is potentially due to a weather based gross change in attire; days 0 and 1 were on average 14.5°C and sunny where as days 2, 3, and 4 were 12.1°C and overcast. This will be explored further in future work. Regardless, this result shows that our HSS based person recognition system yields a considerably better-than-random (1-in-23) result for when the participants were measured post variations in their general appearance.

C. Further Exploration of HSS Ability to Accommodate Variations in Pose

In order to further explore the ability of the HSS to accommodate for variations in pose additional data was

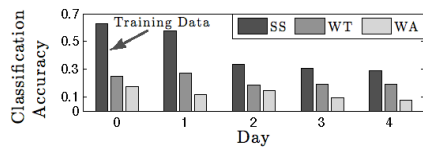


Fig. 8. The mean accuracies for SS, WT and WA for the experiment

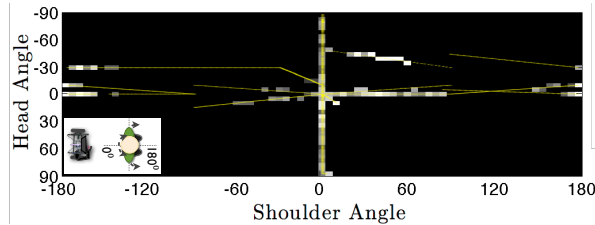


Fig. 9. HSS ability to accommodate head and shoulder pose variations. The line marks the space for which data exists. The underlying colour indicates the classification success (white = 100% and black = 0%)

collected of participant 22, this data was collected on Day 11. Participant 22 was selected as he was available at the time the experiment was ready to commence. The participant was instructed to stand in front of the RobotAssist platform and then slowly rotate their head and/or shoulders through a comfortable range, then repeat the process standing with their back to the robot. The data from this experiment was then run through our system, using the Day 0 training data.

The results of this are shown in Fig. 9, the participant's head and shoulder angles were estimated manually upon review of video taken from the point of view of the sensor. As can be seen from the figure classification tended to succeed for shoulder angles between -60° and $+60^\circ$ regardless of the head angle. Also of particular note is that classification tended also to succeed for shoulder angles between -140° and $+160^\circ$. This is of particular note as at this point the participant had their back to the sensor. The predominate limitation appears to be when the human being observed shoulder angle is between -140° and -60° , and $+60^\circ$ and $+160^\circ$. Although these results are limited (and should be considered an indication only) they clearly show the HSS has considerably ability to accommodate for a person's head and/or shoulder and/or body rotation relative to the sensor.

V. CONCLUSIONS AND FUTURE WORKS

This paper presented a method for exploiting the inter-person variations in the size and shape of peoples' head, neck and shoulders to achieve robust *in situ* person recognition. A method for constructing the HSS, a scale and viewing angle robust feature vector that encapsulates these inter-person variations, was presented. Furthermore, a method for utilising the HSS as the basis for Support-Vector Machine person recognition was detailed.

It was shown through an empirical study, in which HSS based person recognition was attempted on *in situ* data collected from 25 participants over 5 days in a office environment, that the mean accuracy over the 5 days of 78.15% and a peak accuracy 100% for 9 participants. Further, it

was shown that a considerably better-than-random result was achieved by our approach in the cases when the participants were: unaware they were being scanned, in motion, face directly away from the sensor, and post variations in their general appearance. Finally, the results showed that the HSS has considerably ability to accommodate for a person's head and/or shoulder and/or body rotation relative to the sensor - even in cases where the person is faced directly away from the robot. Specially, classification tended to succeed for shoulder angles between -60° and $+60^\circ$ (0° is facing the robot) regardless of the head angle and for shoulder angles between -140° and $+160^\circ$ (when the participant had their back to the robot).

Whilst the results presented here within are promising, a number of limitations remain to be addressed in future work. For instance, more thorough explorations of the effects of general appearance, variations in pose and the effect of a moving participant on the stability of the HSS are planned. Also, future work in which the systems performance compared to a skilled human is planned.

VI. ACKNOWLEDGMENTS

This work is supported by RobotAssist and the Centre for Autonomous Systems, University of Technology, Sydney.

REFERENCES

- [1] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: a survey," *Found. Trends Hum.-Comput. Interact.*, vol. 1, pp. 203–275, January 2007.
- [2] M. Woods, "Caring for older australians: Overview, final inquiry report," Productivity Commission, Australian Government, Tech. Rep. 53, 2011.
- [3] N. Kirchner, A. Alempijevic, and G. Dissanayake, "Nonverbal robot-group interaction using an imitated gaze cue," in *HRI '11: ACM/IEEE Int. Conf. on Human robot interaction*, Switzerland, 2011, pp. 1–8.
- [4] M. Kaitz, "Recognition of familiar individuals by touch," *Physiology and Behavior*, vol. 52, no. 3, pp. 565 – 567, 1992.
- [5] C. Sanderson, *Biometric Person Recognition Face, Speech and Fusion*. VDM Verlag, 2008.
- [6] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, vol. 31, no. 2, pp. 210 –227, Feb. 2009.
- [7] M. Irfanoglu, B. Gokberk, and L. Akarun, "3D shape-based face recognition using automatically registered facial surfaces," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4. IEEE, 2004, pp. 183–186.
- [8] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, vol. 25, no. 12, pp. 1505–1518, 2003.
- [9] D. Voth, "You can tell me by the way i walk," *Intelligent Systems, IEEE*, vol. 18, no. 1, pp. 4 – 5, Jan 2003.
- [10] L. Spinello and R. Siegwart, "A layered approach to people detection in 3d range data," *Learning*, vol. 55, no. 2007, pp. 1625–1630, 2010.
- [11] L. Spinello and K. O. Arras, "People detection in rgb-d data," in *Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS)*, 2011.
- [12] G. C. et al., "1988 anthropometric survey of U.S. Army personnel: Methods and summary statistics," United States Army Natick, Research, Development and Engineering Center, Tech. Rep., 1989.
- [13] D. Hordern and N. Kirchner, "Robust and Efficient People Detection with 3-D Range Data using Shape Matching," in *Proc. of the 2010 Aust. Conf. on Robotics and Automation*, 2010, pp. 1–8.
- [14] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.
- [15] N. Kirchner and et al., "Robotassist - a platform for human robot interaction research," in *Proc. of the 2010 Aust. Conf. on Robotics and Automation*, 2010, pp. 1–8.