2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Incremental Object Part Detection toward Object Classification in a Sequence of Noisy Range Images

Stefan Gächter, Ahad Harati, and Roland Siegwart
Autonomous Systems Lab (ASL)
Swiss Institute of Technology, Zurich (ETHZ)
8092 Zurich, Switzerland
{gaechter, harati, siegwart}@mavt.ethz.ch

Abstract— **This paper presents an incremental object part detection algorithm using a particle filter. The method infers object parts from 3D data acquired with a range camera. The range information is quantized and enhanced by local structure to partially cope with considerable measurement noise and distortion. The augmented voxel representation allows the adaptation of known track-before-detect algorithms to infer multiple object parts in a range image sequence even when each single observation does not contain enough information to do the detection. The appropriateness of the method is successfully demonstrated by two experiments for chair legs.**

I. Introduction

In recent years, a novel type of range camera to capture 3D scenes emerged on the market. One such camera is depicted in figure 1. The measurement principle is based on time-of-flight using modulated radiation of an infrared source. Compared with other range sensors [1], range cameras have the advantage to be compact and at the same time to have a measurement range of several meters, which makes them suitable for indoor robotic applications. Further, range cameras provide an instant single image of a scene at a high frame rate though with a lower image quality in general [2]. The 3D information acquired with a range camera is strongly affected by noise, outliers and distortions, because of its particular measurement principle using a CMOS/CCD imager [3], [4]. This makes it difficult to apply range image algorithms developed in the past. Hence, the goal of this paper is to present an object part detection method adapted to range cameras.

Object parts – components with simple geometry – are quite proper features for object classification based on geometric models [5], [6], [7]. This approach can account for different views of the same object and for variations in structure, material, or texture of the objects of the same kind. The reason is that more or less the decomposition of the objects into its parts remains unchanged. The majority of the currently available approaches in the field of object classification are appearance based, which makes them very sensitive to the mentioned variations.

In general, range image algorithms depend on the robust estimation of the differential properties of object surfaces [8].

This work was partially supported by the EC under the FP6-IST-045350 robots@home project.

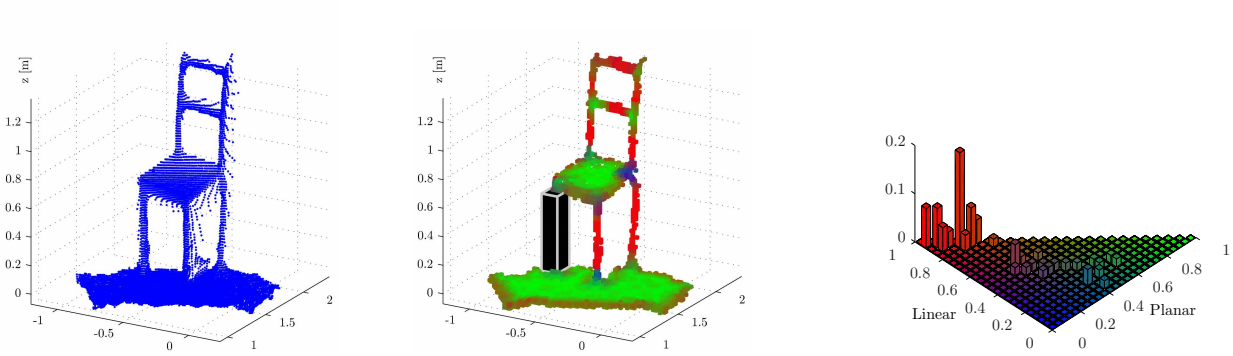
Fig. 1. SR-3000 Range Camera.

Given the noisy nature of images of range cameras, this can only be obtained with high computational cost. However, the detailed reconstruction of object surface geometry is not necessary for part based object classification as long as the parts are detected. On the other hand, object parts can be represented properly by bounding volumes [7], because the overall structure of an object part is more important and informative than the details of its shape or texture. For example, the concept of a chair leg is more related to its stick like structure than whether it is wooden or metallic, of light or dark color, round or square.

However, segmentation of range images into object parts remains the most challenging stage. Because of the low signal-to-noise ratio of the mentioned sensor, this is a particularly difficult problem. Using an incremental algorithm operating on several range images, can improve the performance. In fact, it is possible to skip segmentation and track hypothetical parts in the scene. This is a common approach in radar applications, where a target has to be jointly tracked and classified in highly noisy data [9], [10]. Hence, for each part category, a classifier is considered which incrementally collects the evidences from the sequence of range images and tracks the hypothetical parts. Therefore, the object part detection becomes the sequential state estimation process for multiple bounding-boxes at potential poses in the three-dimensional space. This is realized in the framework of a particle filter [10], which can cope with different sources of uncertainty, among them scene registration errors.

The contribution of this work lies in bringing well established algorithms from classification, tracking and state-estimation to the framework of object classification. In addition, to the best of our knowledge, this is a first work which addresses object part detection using a range camera. The presented work here paves the way toward incremental

978-1-4244-1647-9/08/$25.00 ©2008 IEEE.



(a) Point Cloud (b) Voxel Set (c) Shape Factor Histogram

Fig. 2. (a) Single point cloud and (b) a quantized version of a sequence of five registered range images at step $k = 25$ along with (c) the shape factor histogram of the right front leg. The bounding-box in (b) encloses all voxels that are considered to compute the histogram. The colors indicate the shape factors: red for *linear* like, green for *planar* like, and blue for *spherical* like local structures. Refer to the remaining part of the paper for the computational details.

part based object classification in the field of indoor mobile robotics. The approach presented here is quite general in handling different object parts with simple geometry. However, through out this paper, a chair leg is chosen as an example part to demonstrate the method.

II. Related Work

Part extraction from range images is a long standing issue in structure based object recognition and classification. Seminal work has been done by [11], where algorithms are presented that infer objects from surface information. Object parts are represented by surface patches. In the present work, bounding-boxes are adopted, which are more abstract volumetric representation than commonly used parametrical models based on surfaces [12], [13]. In addition, the quantization is achieved by the voxel representation which is related to occupancy grids, but less computationally intensive.

In [14], a method to capture local structure in range images is presented in order to classify natural terrain. In the present work, local structure is captured in the same way with shape factors. However, shape factors are calculated based on the voxel representation here.

The object part detection algorithm evolves from the work done in [15]. They developed a method for joint detection and tracking of multiple objects described by color histograms. Color-based tracking is a well researched topic in the vision community, see for example [16] and [17]. Here, these techniques are taken as inspiration to detect object parts in quantized point clouds using shape factor as color.

III. Range Image Quantization

One of the smallest range cameras in the market is the SR-3000 made by [18], see figure 1. For the work presented here, the SR-2 of the same manufacturer is used, which exhibits similar measurement performance for indoor applications. The SR-2 has a resolution of 124×160 pixels with maximum measurement range of 7.5 m. The intrinsic and extrinsic camera parameters are respectively calibrated based on the methods explained in [3] and [19].

Despite the calibration, the range image remains affected by noise, outliers and distortions. Main reasons include low emission power, scattering, and multiple reflections. A sample observation of a scene with a chair is shown in figure 2(a). Thus, a single observation does not contain enough information to detect object parts. On the other hand, registering different views over long runs accumulates alignment errors. Therefore, a sliding window containing the most recent five observations is considered. The corresponding point clouds are registered and quantized into a cubic voxel space with voxel size of 2 cm to reduce the computational burden. Voxels containing less than five points are neglected as outliers, see figure 2(b).

IV. Object Part Detection

The structural variability of objects is strongly related to the number and type of parts and their physical relationship with each other. Such relationships can be encoded within a probabilistic grammar in order to perform object classification [7]. Towards such an approach, object parts are modeled as probabilistic bounding-boxes to handle uncertain measurements of the range camera.

A bounding-box is a cuboid defined by the center point and the span length. The probabilistic extension assumes these parameters as random variables. Here, particle filter is used to estimate them. Each particle encodes hypothetical positions and extensions of some object parts, i.e. their bounding-boxes. The evolution of the particles over time enables the simultaneous detection and tracking of the object parts. Gradually, particles with realistic hypotheses survive, whereas the others die off. The fitness of each particle – the resampling weight – is obtained based on the shape factor histograms calculated in the image regions defined by the corresponding bounding-boxes.

A. Shape Factor

The shape factors characterize the local part structure by its linear, planar, or spherical likeliness. They are calculated for each voxel using its surrounding spatial voxel distribution

by the decomposition of the distribution into the principal components – a set of ordered eigenvalues and -vectors. Here, the standard principal component analysis is used.

In the literature, different methods are presented on how to compute the shape factors. In [20] and [21] a tensor representation is proposed to infer structure from sparse data. For the present work, the same scheme is used with a different normalization:

$$r_l = \frac{\lambda_1 - \lambda_2}{\lambda_3 + \lambda_2 + \lambda_1}, \quad r_p = \frac{2(\lambda_2 - \lambda_3)}{\lambda_3 + \lambda_2 + \lambda_1}, \quad r_s = \frac{3\lambda_3}{\lambda_3 + \lambda_2 + \lambda_1}. \tag{1}$$

where λ_i are the ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ obtained by the decomposition of the spatial voxel distribution. r_l, r_p, and r_s express local similarity to linear, planar, and spherical shapes respectively. Here, the shape factors are normalized by the sum of the eigenvalues [22] so that each lies in the range of $[0, 1]$ and their sum is one: $r_l + r_p + r_s = 1$. Another normalization scheme is to use the maximum eigenvalue [21]. The shape factors can also be defined by reasoning on the volume spanned by the eigenvalues [23]. Which of the shape factor computation methods is used, depends largely on their ability to characterize voxels distinctively according to the object structure at hand.

Figure 2(b) depicts a shape factor colored voxel set of a chair, where for each voxel the shape factor was computed according to (1). The computation was done within a neighborhood window of size $11\times11\times11$ voxels defining the scale of the local structure. As it is visible, this method correctly classifies the local structure; legs appear as linear, seat and back as planar, and joints as spherical structures.

B. Histogram as Feature Vector

The shape factor distribution in the region of interest defined by the bounding-box is approximated by a histogram to obtain a unique feature vector that models an object part. This approach is inspired by the work done in [16], where color histograms are used to track objects. In the present application, histograms have the advantage to be robust against the structural variability of object parts: rotation, partial occlusion, and scale have little effect on the model. In addition, the computational cost of histograms is modest.

Since the three positive elements of the shape factor sum up to one, they are constrained to a triangle in 3D space. Thus, it is sufficient to consider only two elements to populate a 2D histogram with $N_t = \frac{1}{2}(N_b^2 + N_b)$ bins, where the histogram shape is approximated by a triangular matrix of size N_b. Figure 2(c) depicts the shape factor histogram of the bounding-box volume enclosing the leg in figure 2(b). It is clearly visible that linear shape factors dominate indicating the general stick like structure of the object part. Because the number of bins N_t already becomes large for a small N_b, dimensionality reduction is applied on the 0 as feature vector. The dimensionality reduction is done by standard principal component analysis of the training set retaining the dimensions covering 95 % of the feature distribution mass.

Finally, six simple geometric features are added to the feature vector to account for the occupancy and eccentricity of the voxel distribution in the bounding-box.

C. Support Vector Classifier

In order to judge, if an object part in question is likely to belong to a certain class it is necessary to evaluate a quality measure. This can be done by computing a distance between a template and the generated feature vector. This is commonly done in color based tracking [16]. However, template matching might not be discriminative enough to cover an entire class of an object part. Using a classifier learned on a large amount of training data often results in a better detection performance. A suitable training method is the support vector machine (SVM), because it is less prone to overfitting, is applicable on high dimensional features, and the resulting classifier allows the estimation of meaningful posterior probabilities. In the present work, a support vector classifier with a polynomial kernel is trained using the framework provided by [24]. A training set of 2340 samples is generated. An equal number of positive and negative samples are used to avoid any bias in the learning. The 1170 positive samples of chair leg are manually extracted from voxel images from different views of twelve different chairs. The 1170 bad samples are randomly selected from a stream of voxel images containing background clutter or non-leg parts.

D. Incremental State Estimation

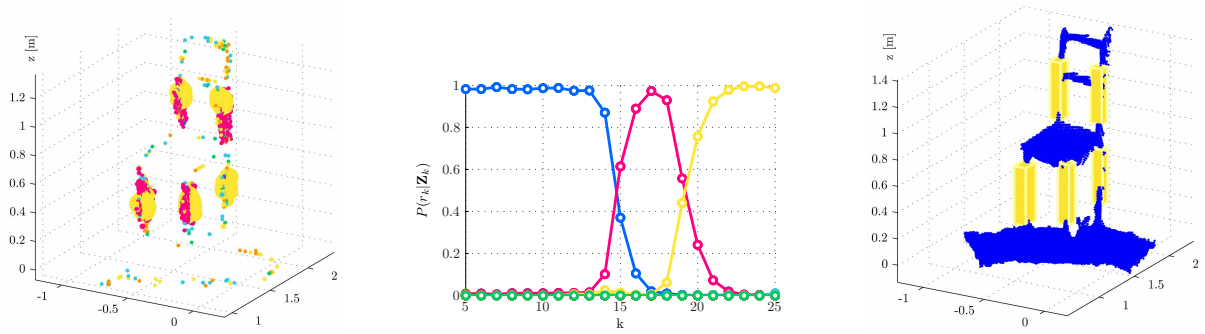
The aim is to incrementally detect object parts modeled by a bounding-box in a sequence of voxel images. The detection algorithm typically has to handle multiple object parts of the same type. Thus, the problem can be stated formally as follows:

$$p(\mathbf{y}_k|\mathbf{Z}_{k-1}) = \int p(\mathbf{y}_k|\mathbf{y}_{k-1}) p(\mathbf{y}_{k-1}|\mathbf{Z}_{k-1}) d\mathbf{y}_{k-1} \tag{2}$$

$$p(\mathbf{y}_k|\mathbf{Z}_k) \propto p(\mathbf{z}_k|\mathbf{y}_k) p(\mathbf{y}_k|\mathbf{Z}_{k-1}), \tag{3}$$

where $\mathbf{y}_k = [R_k, \mathbf{x}_{1,k}^\mathsf{T} \ldots \mathbf{x}_{r_k,k}^\mathsf{T}]^\mathsf{T}$. $\mathbf{y}_k$ is the augmented state, which contains the current estimate of number of object parts present in the view R_k and their bounding-box parameters $\mathbf{x}_{i,k}$ at step k. This incremental state estimation can be implemented by a particle filter. Here, the algorithm presented in [15] is used; an extension of the traditional particle filter [25], capable of tracking multiple targets. However, the transition and observation models have been adapted where necessary.

1) Transition Model: The object part number R_k is modeled by a Markov chain with a predefined transition matrix, where the state value at step k is a discrete number $r_k = \{0, \ldots, M\}$ with M being the maximum number of parts expected in each view, set to 8 here. The Markov chain defines three possible cases on how the number of parts can evolve over time: the number remains *unaltered*, *increases*, or *decreases* from step $k-1$ to k.



(a) Observation Density (b) Part Presence Probability (c) Estimated Parts

Fig. 3. Results for the first experiment. (a) Particle distribution at step k = 20 during the update. Particle size indicates its weight. (b) Evolution of the part presence probability over time. (c) Estimated object parts at step k = 25. The color indicates the number of hypothetical parts encoded by a particle: blue for 3, magenta for 4, and yellow for 5 states. Other colors indicate higher or lower number of states, where the maximum number of states is eight.

When the number of parts remains unaltered, their states are assumed to be affected by a process noise, which takes into account the measurement deficiencies and registration errors. Therefore, the proposal distribution for the bounding-box parameters is given by $p(\mathbf{x}_{i,k}|\mathbf{x}_{i,k-1}) = \mathcal{N}(\mathbf{x}_{i,k-1}, \mathbf{C}_u)$, where $\mathbf{C}_u$ is the covariance matrix assumed to be diagonal. In the experiments of this paper, for a chair leg, the diagonal entries for the bounding-box position are set to 1, 1, and 9 cm^2 and for the extension to 49, 49, and 64 mm^2, considering more uncertainty along the vertical direction.

When the number of parts decreases, r_k hypothetical parts are selected at random from the possible r_{k-1} with equal probability. The selected parts parameters are then affected by the process noise.

The crucial case is when the number of parts increases. Then, the current state of the particle has to be augmented by additional elements. For the r_{k-1} parts that continue to exist, again the process noise is considered. For the $r_k - r_{k-1}$ new hypothetical parts, the bounding-box position is uniformly sampled from occupied voxels, which have proper shape factors. In addition, to preserve consistency of different instances of a part, an intersection test is performed [26]. Hence, the initialized bounding-boxes for each particle keep a certain distance, here 10 cm.

2) Observation Model: The observation likelihood function generates the importance weights used to incorporate the measurement information $\mathbf{z}_k$ in the particle set. Since the parts have to be detected from various view angles out of sparse and noisy data, the observation model is a non-linear function of the part state and measurement noise. As in [27], instead of using a generative observation model, which is common in a Bayesian estimation framework, a discriminative one is selected, namely the learned support vector classifier presented previously.

In the detection framework, the observation likelihood is usually defined as a ratio of the probability that an object part is present to the probability of its absence. This is equivalent to the ratio of the classification probabilities computed with the learned classifier. Assuming that the classification can be done independently for each hypothetical object part, the observation likelihood for each particle is given by

$$L(\mathbf{y}_k) = \prod_{i=1}^{r_k} \frac{p(\mathbf{z}_k|\mathbf{x}_{i,k})}{1 - p(\mathbf{z}_k|\mathbf{x}_{i,k})}, \tag{4}$$

where $p(\mathbf{z}_k|\mathbf{x}_{i,k})$ is the classification probability for part i. Considering this probability as a distance $a_{i,k}$ in the range of $[0, 1]$ and an exponential function to compute a similarity measure, the unnormalized importance weight $\tilde{\pi}_k$ for each particle is computed as:

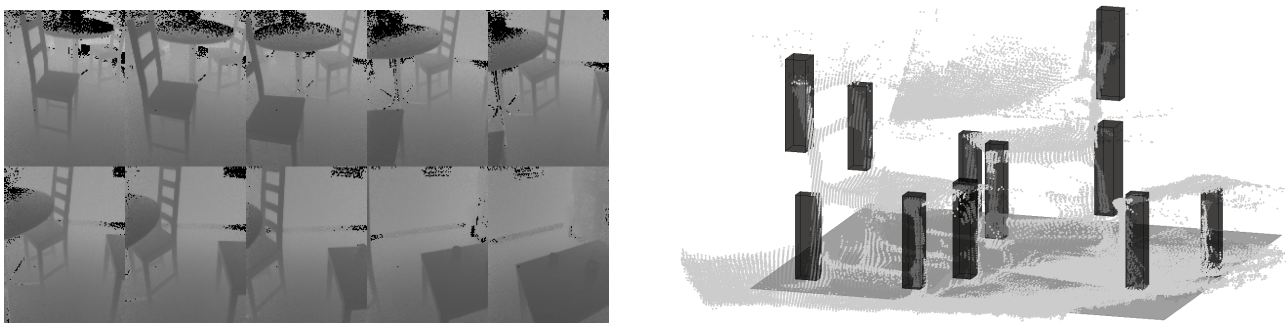
$$\tilde{\pi}_k = \begin{cases} 1, & R_k = 0 \\ \exp\left(-\frac{1}{b}\sum_{i=1}^{r_k}(1 - 2a_{i,k}) + r \cdot c\right), & R_k > 0 \end{cases} \tag{5}$$

where b is a parameter to adjust the observation sensitivity and c accounts for the a priori knowledge. Here, b = 0.21 and c = 0.21 are used, both determined experimentally.

When no a priori knowledge is considered, the weighting scheme defined above has a pivoting point for a classification probability equal to 0.5; meaning that particles with large number of hypothetical object parts and a probability only slightly greater than 0.5 are favored over particles with small number of parts but a high probability. Hence, the object part detection algorithm has an inherent tendency for exploration.

V. Experiments

The above discussed incremental object part detection method is exemplified by the detection of chair legs. Chair legs in reality are designed with various shapes and tilt angles. Here, they are modeled by a vertical bounding-box defined by its center point position $\mathbf{s} = [s_x, s_y, s_z]^\mathsf{T}$ and its extension $\mathbf{t} = [t_x, t_y, t_z]^\mathsf{T}$ to cover the overall shape for the class of chair legs. With the assumption of upright chairs, the rotations are neglected since they are not properly captured by the range camera. However, the effect of such rotations are generally small and pose minor variations with respect to the overall structure. For other parts of a chair, such as seat and back, the rotation around the z-axis has to be considered in the state.



(a) (b)

Fig. 4. Second experiment at cafeteria. (a) Samples of range image sequence at an interval of 50 steps starting at step 0 and ending at 450. (b) Detected stick like parts in the scene with a round dining table, two chairs and a coffee table. The black bounding-boxes indicate the estimated positions and extensions of stick like parts. Only two point clouds are depicted.

Two experiments are performed with the range camera mounted on a robot at height of about 1.1 m facing downward with a tilt angle of about 15°. In the first experiment, only one chair is in the scene while in the second experiment the robot is observing a round dining table, two chairs and a coffee table in the cafeteria of our lab. In both experiments, the robot slowly approaches the objects in the scene recording range images and odometry at about 2 Hz. Totally 200 and 450 range images are captured in the first and second experiment respectively. Because of occlusions and the narrow field of view of the camera, the number of hypothetical chair legs in the view varies considerably, see 4(a). Hence, the algorithm should dynamically adapt to what is present in the view.

Considering the complexity of the scenes, the number of particles is set to 750, which is rather low because of the intelligent initialization scheme. The outcome of the first experiment is summarized in figure 3. Figure 3(a) shows the observation density of the particle filter at step k = 20. The weights are represented by the size of the depicted particles. It can be seen that the particles at the chair legs and back columns are bigger than the ones at the seat. Therefore, the weighting based on the support vector classifier is successful. On the other hand in this figure, particles with different colors represent different number of hypothetical legs. Therefore, a competition between red and yellow particles corresponding to four and five legs is taking place at this step. This is a result of the explorative behavior of the transition matrix and necessary for discovering new parts, which may enter the scene. The same fact is depicted more properly in figure 3(b) versus time, where the probability of the number of object parts present in the view is approximated by the ratio of the number of particles sharing the same r value to the total number of particles. At step k = 25, three legs and two columns of the back support are successfully detected as can be seen in figure 3(c).

In the second experiment with a more realistic scenario, the robot is faced with the challenge of object part detection in the cafeteria. In figure 4(b), the estimated object parts are depicted overlaid with two original point clouds. Depicted are the hypothetical legs with the probability larger than 0.5. The observed deviation between the estimated bounding-boxes and the real parts are mainly because only two point clouds are depicted. If the whole 450 point clouds are considered together, the errors of range camera and odometry result in a messy accumulation of points where no geometry is detectable. As mentioned, this fact is one of the main motivations in using an incremental estimation method. In figure 5, upper graph, depicts the probability of the number of object parts present in the view. As can be seen, the

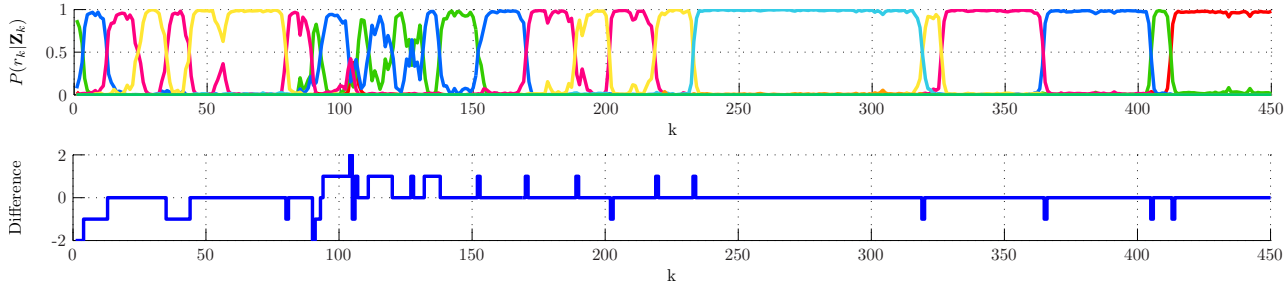


Fig. 5. Second experiment at cafeteria. In the upper figure is depicted the evolution of the part presence probability over time. The color indicates the number of hypothetical parts encoded by a particle: red for 1, green for 2, blue for 3, magenta for 4, yellow for 5, and cyan for 6 stick like parts. Other colors indicate higher or lower number of parts with the maximum of eight. In the lower figure is depicted the difference between the detected and the actual number of hypothetical legs in the view.

probabilities oscillate where the scene changes considerably – between step 80 and 140 – and the algorithm has to deal with many appearing and disappearing parts. This is also evident in figure 5, lower graph, where the difference between the detected and the actual number of hypothetical legs in the view is depicted. The actual number of parts is determined by visual inspection of the voxel images. In the end, all leg like parts are detected and no false positive remains.

VI. Conclusion

This paper presented an algorithm for object part detection using an extended particle filter as an estimation engine with a support vector classifier based observation function. The algorithm can handle multiple parts of the same class and deal with different sources of uncertainties. The provided experimental results show that using a limited number of particles it is possible to successfully estimate the position and extension of multiple chair legs – an exemplary object part – in an incremental process.This proves the accomplishment of the primary goal: accumulation of information in a sequence of noisy and sparse observations.

However, the method needs further testing and improvements for its robust application in robotics. First, the detection has to be extended to multiple classes of object structures by training the corresponding support vector classifiers. In addition, further investigation can be done on recently introduced support vector classifiers [27].

Finally, more informative constraints can be utilized in the particle filter by considering plausible object configurations. The presented algorithm is currently being integrated into a part based object classification system.

Acknowledgement

Thanks to Jacek Czyz for providing further insights into his particle filter implementation and to Luciano Spinello for the fruitful discussions on support vector classifiers.

References

[1] F. Blais, "Review of 20 years of range sensor development," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 231–243, January 2004.

[2] S. May, K. Pervoelz, and H. Surmann, *Vision Systems - Applications*. I-Tech, 2007, ch. 3D Cameras: 3D Computer Vision of wide Scope, pp. 181–202.

[3] T. Kahlmann, "Range imaging metrology: Investigation, calibration and development," Ph.D. dissertation, Eidgenössische Technische Hochschule Zürich, ETHZ, Diss ETH No 17392, 2007.

[4] S. A. Gudmundsson, H. Aanaes, and R. Larsen, "Environmental effects on measurement uncertainties of time-of-flight cameras," in *International Symposium on Signals, Circuits and Systems (ISSCS 2007)*, vol. 1, 2007, pp. 1–4.

[5] R. A. Brooks, "Symbolic reasoning among 3-D models and 2-D images," *Artificial Intelligence*, vol. 17, pp. 285–348, 1981.

[6] L. Stark and K. W. Bowyer, *Generic Object Recognition using Form and Function*, ser. Series in Machine Perception and Artificial Intelligence, H. W. P. Bunke, Ed. World Scientific, 1996.

[7] M. A. Aycinena, "Probabilistic geometric grammars for object recognition," Master's thesis, Massachusetts Institute of Technology - Department of Electrical Engineering and Computer Science, 2005.

[8] P. J. Besl, *Surfaces In Range Image Understanding*. Springer-Verlag Inc., New York, 1988.

[9] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques, and Software*. Artech House, 1993.

[10] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter - Particle Filters for Tracking Applications*. Artech House, 2004.

[11] R. B. Fisher, *From Surfaces to Objects - Computer Vision and Three Dimensional Scene Analysis*. John Wiley & Sons Ltd., Chichester, Great Britain, 1989, http://homepages.inf.ed.ac.uk/rbf/BOOKS/FSTO/ (14.9.2007).

[12] H. Rom and G. Medioni, "Part decomposition and description of 3D shapes," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, vol. 1, 1994, pp. 629–632.

[13] W. H. Field, D. L. Borges, and R. B. Fisher, "Class-based recognition of 3D objects represented by volumetric primitives," *Image and Vision Computing*, vol. 15, no. 8, pp. 655–664, August 1997.

[14] N. Vandapel, D. Huber, A. Kapuria, and M. Hebert, "Natural terrain classification using 3-d ladar data," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '04)*, vol. 5, 2004, pp. 5117–5122.

[15] J. Czyz, B. Ristic, and B. Macq, "A color-based particle filter for joint detection and tracking of multiple objects," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 2, 2005, pp. 217–220.

[16] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," in *Proceedings of the European Conference Computer Vision (ECCV)*, ser. LNCS 2350, A. H. et al., Ed. Springer-Verlag, 2002.

[17] H. Wang, D. Suter, K. Schindler, and C. Shen, "Adaptive object tracking based on an effective appearance filter," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, pp. 1661–1667, 2007.

[18] MESA Imaging AG, Switzerland, http://www.swissranger.ch/ (13.9.2007).

[19] S. May, B. Werner, H. Surmann, and K. Pervölz, "3D time-of-flight cameras for mobile robotics," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[20] C.-F. Westin, "A tensor framework for multidimensional signal processing," Ph.D. dissertation, Department of Electrical Engineering - Linköping University, Sweden, 1994.

[21] G. Medioni, M.-S. Lee, and C.-K. Tang, *A Computational Framework for Segmentation and Grouping*. Elsevier, 2000.

[22] C.-F. Westin, S. Peled, H. Gudbjartsson, R. Kikinis, and F. A. Jolesz, "Geometrical diffusion measures for MRI from tensor basis analysis," in *Proceedings of the 5th Annual Meeting of the International Society for Magnetic Resonance Medicine (ISMRM)*, 1997, p. 1742.

[23] S. Gächter, "Incremental object part detection with a range camera," Autonomous Systems Lab, Swiss Federal Institute of Technology, Zurich (ETHZ), Switzerland, Tech. Rep. ETHZ-ASL-2006-12, 2006.

[24] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," Department of Computer Science - National Taiwan University, Tech. Rep. July 18, 2007, 2007.

[25] M. Isard and A. Blake, "CONDENSATION - conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[26] T. Akenine-Möller and E. Haines, *Real-Time Rendering*, second edition ed. A K Peters, 2002.

[27] C. Shen, H. Li, and M. J. Brooks, "Classification-based likelihood functions for bayesian tracking," in *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS'06)*, 2006, p. 33.