Markovian Decomposition of Range Images

Andreas Pichler[1] and *Robert B. Fisher*[2]

Abstract:
This paper describes a computational model for deriving a decomposition of objects from laser rangefinder data. The process aims to produce a set of parts defined by compactness and smoothness of surface connectivity. Relying on a general decomposition rule, any kind of objects made up of free-form surfaces are partitioned. A robust method to partition the object based on Markov Random Fields (MRF), which allows to incorporate prior knowledge, is presented. Shape index and curvedness descriptors along with discontinuity and concavity distributions are introduced to classify region labels correctly. In addition, a novel way to classify the shape of a surface is proposed resulting in a better distinction of concave, convex and saddle shapes. To achieve a reliable classification a multi-scale method provides a stable estimation of the shape index.

1 Introduction

Image segmentation is an important early vision task to group pixels with similar characteristic into homogeneous regions. Many high level processing tasks (object recognition, surface representation using volumetric models for example) are based on such a preprocessed image. Descriptions extracted by a system must reflect the characteristics of the environment. This paper is about computing such a description in 3D range data.

Segmentation of range images has been addressed by many researchers *e.g.* [2, 7]. Measurements are used to segment a depth map into surface patches which are either planar or curved and which are described formally by a function. These methods are restricted to process mainly CAD (Computer Aided Design) models. The recognition of complex objects in a cluttered scene without a feature extraction or segmentation step is presented by Johnson [9] who proposed a template matching procedure using spin images. Charvis et. al. [4] proposed point signatures for describing the local shape around a point. Both methods aim to extract local shape templates for every point on the model. Unfortunately, the methods described above cannot be used to extract a more generic description of an object. This is a significant step which needs be carried out focusing on object classification among other things.

[1] Profactor Produktionsforschung Gmbh, Im Stadtgut A2, A-4407 Steyr-Gleink, Austria,
e-mail: andreas.pichler@profactor.at
[2] School of Informatics, University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom,
e-mail: rbf@dai.ed.ac.uk

Psychological studies have shown an underlying regularity of part recognition [6]. A partitioning rule on the basis of the *Transversality Regularity* divides an object into its constituent parts along all contours of concave discontinuity of the tangent plane. The process described in our paper aims to produce a set of parts generated by a generic partitioning rule. The extraction of discontinuities does not always deliver a continuous or even closed boundary. Thus, boundary organization is needed to link spare discontinuity points. A problem is whether a putative point is labeled as boundary or dismissed due to noise in the image which often leads to misclassification of region labels. The approach proposed in this paper assigns a feature vector taking a shape index and curvedness estimation into account. The combination of both ensures compactness, smoothness and convexity of surface connectivity. A Markov Random Field (MRF) is used to incorporate spatial interaction as well as to to estimate slow spatial variations of the feature characteristic of the image. A flowchart of the algorithm is given in Figure 1. The paper is structured along the computational steps that define the procedure for decomposing an object from sensor data. In Section 2, a mathematical formulation is defined to decompose a scene. A novel computation method is presented for consistent shape index estimation. Furthermore, the spatial interaction of shape index and curvedness components of the feature vector is decribed.

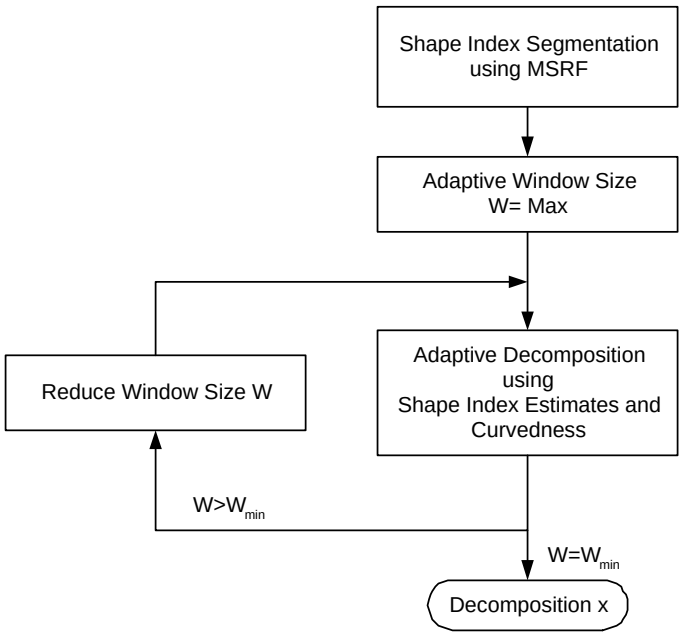

Figure 1: Adaptive decomposition algorithm.

2 PROBLEM FORMULATION

The decomposition scheme extracts convex constitutent parts of an object based on the local shape information. A quantitative measurement of the shape of a surface at a point s is the shape index $S_I(s)$ and the curvedness $C(s)$. According to Koenderink's definition of the shape index [10] and its adopted version [5] nine shape categories $\mathbf{S} = \{S_1, ..., S_9\}$ are distinguished.

The shape index provides a continous gradation between salient shapes such as convex, sattle and concave types. The curvedness is a measure of the scale or the amount of curvature of a region.

2.1 Multiscale Shape Index Segmentation

The computation of the shape index which is based on the curvatures of a surface, requires the estimation of a local quadric surface function at point s [13]. Shape index estimation on real range images is a difficult task because of the effect of noise and outliers. In order to overcome these issues a multiscale method [12] has been adopted to provide a robust method for a correct classification of the shape.

A multi-scale Markov random field (MSRF) is used, hierarchically ordered from the coarsest to finest level. The random variable fields are denoted by $x^{S(n)}$ at scale n, where $n = 0$ is the most detailed. They represent the estimated shape index labels for each layer. The observed image is given by y^S. In a stochastic image model two different random processes are defined. There is first an emission process, which takes the point's actual value into account. A stochastic emission model defines the probabilites for each point. A gaussian process is used to model the data related process. The initialization parameters are retrieved from the shape index characteristic. This data related term provides the evidence of the estimation.

The second process establishes homogenous regions. The region process is modeled by a multi-scale Markov random field. The random variables indicate the catagorized shape index class. The process starts with a coarse-scale random field. The field is successively refined until the final resolution of the image is reached. Labels at the finest scale give the categorization of the individual points into shape classes. The possible values of the values according to the shape index description are $x_s^{S(n)} \epsilon \{0, 1, ..., M-1\}$ with $M = 9$. The process expresses the assumptions about the shape classes for each layer, which are passed on from the coarser scale label to the finer scale label in a top-down fashion.

Thus, the likelihoods are calculated recursively. The emission probability is defined by

$$P(y^S) = \prod_s P(y_s^S | x_s^{S(0)}) \tag{1}$$

Using Bayes rule the probabilities are inververted to

$$P(x^{S(n)} | y^S, x^{S(n+1)}) = \frac{P(y^S | x^{S(n)}) P(x^{S(n)}) | x^{S(n+1)})}{P(y^S | x^{S(n+1)})} \tag{2}$$

The posterior probability is maximized over the set of possible values for $x_s^{S(n)}$. The transition probabilities $P(x_s^{S(n)}|x_{\mathcal{P}(s)}^{S(n+1)})$ introduce additional prior knowledge of the relationships between shape index classes over the overlapping pixel at the coarser scale.

$$P(x_s^{S(n)}|x_{\mathcal{P}(s)}^{S(n+1)}) = \begin{cases} \theta & : \quad x_s^{S(n)} = x_{\mathcal{P}(s)}^{S(n+1)} \\ \frac{1-\theta}{M-1} & : \quad otherwise \end{cases} \tag{3}$$

It assumes if there is certain shape class at coarse scale then with probability θ the same region type is present at the finer scale. The parameter θ was set to 0.8.

2.2 Decomposition based on Shape Index and Curvedness

A common approach to introduce spatial interaction is to model a scene x by a Markov random field (MRF). It is well known that an MRF is Gibbs distributed (Hammersley-Clifford theorem [1]). Hence its probability density $P(x_s)$ is given by a Gibbs distribution

$$P(x_s) = \frac{1}{Z} exp\left\{-U(x_s)\right\}. \tag{4}$$

with the normalization constant Z and the energy function $U(x_s)$. According to the Bayes rule, the posterior probability is computed by using the following formulation

$$P(x_s|y_s,\mu) = \frac{P(y_s|x_s,\mu)P(x_s)}{P(y_s)} \propto exp(-U(x_s|y_s)) \tag{5}$$

where $P(y_s|x_s)$ is the likelihood function of x for the observation y. The *maximum a posteriori* (MAP) estimate is equivalently found by minimizing the posterior energy $U(x_s|y_s,\mu)$.

The decomposition rule reflects the idea of splitting objects into their constituent parts along concave discontinuities which results in a set of surfaces of similar shape characteristic. Thus, the feature vector is made up of two components (shape index and curvedness) to model the local surface. The observation y_s^S is initialized with the model estimates of section 2.1.

$$U(y_s|x_s) = (y_s^S - \mu_s^{S_{x_s}})^2 + \eta(y_s^S)(y_s^C - \mu_s^{C_{x_s}})^2 \tag{6}$$

The effect of the curvedness component is weighted with a function $\eta(y_s^S)$ which depends on the local shape index at site s. A combined measurement makes sense if shape index values indicate concave or saddle points, whereas higher values represent convex shape classes, the curvedness term is omitted in that case. Since the shape models of constituent parts are not

constant, an adaptive estimation method [11] is adopted. Starting from global estimates, the segmentation results slowly adapts to local shape variations. To estimate the local mean $\mu_s^{S_{x_s}}$ and $\mu_s^{C_{x_s}}$, a window of size W is centered at each site s. The estimation process takes place within the window. By changing the window size from large to small, the final segmentation will be governed by local features.

The prior energy term consists of three terms: a convexity term, a spatial continuity and an edge discontuity term.

$$U(x_s) = V_1 + \sum_{\{s,q\}\epsilon C_2} V_2 + \sum_{\{s,q\}\epsilon C_2} V_3 \tag{7}$$

According to the decomposition rule, objects are represented by convex surfaces. In order to facilitate the extraction of convexity a single-site term is introduced to help minimize the number of concave regions.

$$V_1(x_s) = \begin{cases} -\alpha & : \quad f(\mu_s^{S_{x_s}}) = convex \\ +\alpha & : \quad otherwise \end{cases} \tag{8}$$

with the symbolic mapping funtion

$$f(\mu_s^{S_{x_s}}) = \{nonconvex, saddle, convex\} \tag{9}$$

which maps the shape index space into three shape descriptions.

The spatial component of the energy function is given by

$$V_2(x_s, x_q) = \begin{cases} +\beta_1 & : \quad x_s \neq x_q \\ -\beta_1 & : \quad x_s = x_q \end{cases} \tag{10}$$

and the edge component term is represented by

$$V_3(x_s, x_q) = \begin{cases} -\beta_2 & : \quad x_s \neq x_q, x_{s,q}^E = 1 \\ +\beta_2 & : \quad x_s = x_q, x_{s,q}^E = 0 \\ +\beta_2 & : \quad otherwise \end{cases} \tag{11}$$

where $x_{s,q}^E$ indicate an edge between sites s and q. Only jump edges [8], representing occlusions, are used in this approach. A dynamic edge linking procedure is adopted from [14].

Based on the energy terms above, a posteriori energy is defined as

$$U(x_s|y_s) = \frac{\chi^{x_s}}{T}\Big[U(y_s|x_s) + U(x_s)\Big] \tag{12}$$

The distortion χ^{x_s} describes the compactness of a region x_s and T is the tolerated distortion constraint. It is a global parameter which keeps region compact and reduces misclassification of region labels due to outliers.

The energy term $U(x_s)$ is minimized with a local enery minimization method based on highest confidence first (HCF) [3].

3 EXPERIMENTAL RESULTS

The performance of the algorithm is evaluated on a set of plastic models. All data are collected from a range scanner with a registered color image. Figure 2 shows the decomposition result of *Winnie the Pooh*. As can be seen the head is decomposed into several parts, which is correct since the cheeks, the nose and other parts have different shapes. The beak of the bird in Figure 3 is subdivided into two parts. It is not immediately obvious since the boundary along those parts is not a real concavity, but considering local distributions using an adaptive estimation, classifies saddle points correctly and along with the curvedness information a proper shape model is derived. Figure 4 and Figure 5 demonstrate similar results for this approach on other objects. The parameters are kept constant for all examples, which proves the robustness of the method.

4 DISCUSSION AND CONCLUSION

A new method is developed to decompose an object based on concave contours. Furthermore, a robust shape index estimation based on a multi-scale Markov random field allows to integrate stable classification of concavities along with range edge discontinuities into the decomposition framework. Using an adaptive model, shape models are more accurate near concave boundaries. The results are encouraging and will be further validated. Integrating a volumetric model estimation for decomposed regions will be the next step to get a high-level description.

References

[1] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, (36):192–236, 1974.

Figure 2: Part decomposition - range image (left), decomposed parts represented by different colors (right) with $\alpha = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 0.01$.

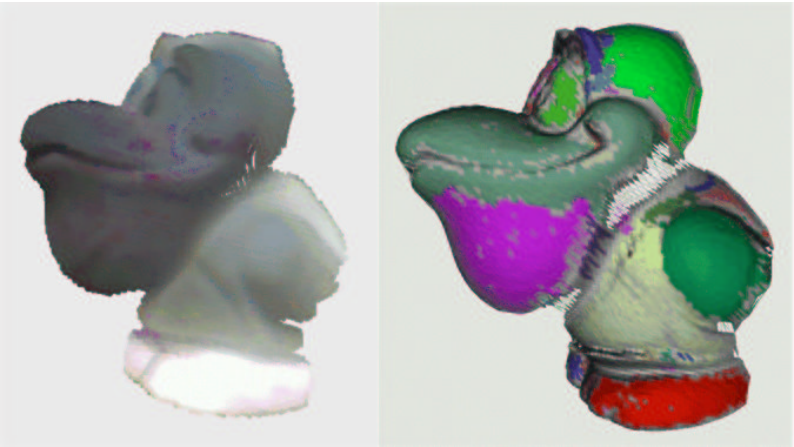

Figure 3: Part decomposition - range image (left), decomposed parts represented by different colors (right) with $\alpha = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 0.01$.

[2] Paul Besl and Ramesh Jain. Segmentation through variable-order surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2):167–192, 1988.

[3] P.B. Chou and C.M. Brown. The theory and practice of bayesian image labeling. *IJCV*, 4:185–210, 1990.

[4] S. Chua and R. Jarvis. Point signatures: a new representation for 3d object recognition. *IJCV*, 25(1):63–85, 1997.

[5] Chitra Dorai and Anil K. Jain. COSMOS - a representation scheme for 3d free-form objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(10):1115–1130, 1997.

[6] Richards W. A Hoffman D.D. Parts of recognition. *Cognition*, 18:65–96, 1984.

[7] A. Hoover, G. Jean-Baptiste, X.Y. Jiang, P.J. Flynn, H. Bunke, D.B. Goldgof, K.W. Bowyer, D.W. Eggert, A.W. Fitzgibbon, and R.B. Fisher. An experimental comparison of range image segmentation algorithms. *PAMI*, 18(7):673–689, July 1996.

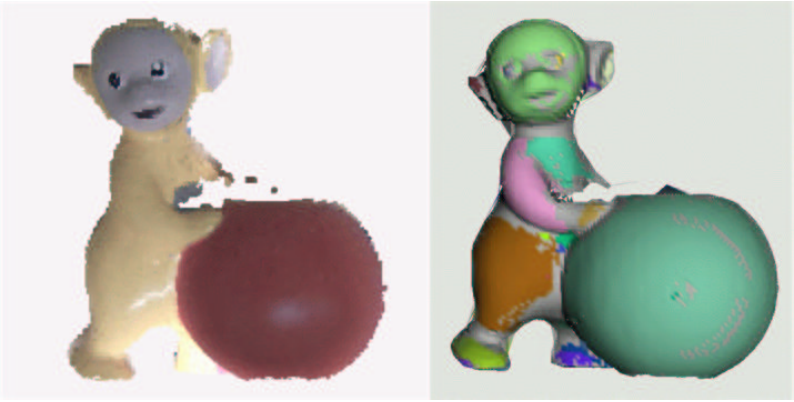

Figure 4: Part decomposition - range image (left), decomposed parts represented by different colors (right) with $\alpha = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 0.01$.

Figure 5: Part decomposition - range image (left), decomposed parts represented by different colors (right) with $\alpha = 0.1$, $\beta_1 = 0.01$ and $\beta_2 = 0.01$.

[8] X. Jiang and H. Bunke. Edge detection in range images based on scan line approximation. *CVIU*, 73(2):183–199, February 1999.

[9] Andrew Johnson. *Spin-Images: A Representation for 3-D Surface Matching.* PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.

[10] J.J. Koenderink and A.J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, pages 557–565, October 1992.

[11] T.N. Pappas. An adaptive clustering algorithm for image segmentation. *IEEE Transactions on Signal Processing*, pages 901–914, April 1992.

[12] H. Rehrauer, K. Seidel, and M. Datcu. Bayesian image segmentation using a dynamic pyramidal structure, 1998.

[13] Gabriel Taubin. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11):1115 – 1138, November 1991.

[14] S.W. Zucker, C. David, A. Dobbins, and L. Iverson. The organization of curve detection: Coarse tangent fields and fine spline coverings. *ICCV*, 88:568–577, 1988.