Topological Modelling for Human Augmented Mapping

Elin A. Topp and Henrik I. Christensen
Centre for Autonomous Systems (CAS)
School of Computer Science and Communication (CSC)
Royal Institute of Technology (KTH)
10044 Stockholm, Sweden
Email: {topp,hic}@csc.kth.se

***Abstract*— Service robots designed for domestic settings need to navigate in an environment that they have to share with their users. Thus, they have to be able to report their current state and whereabouts in a way that is comprehensible for the user. Pure metric maps do not usually correspond to the understanding of the environment a user would provide. Thus, the robotic map needs to be integrated with the human representation. This paper describes our framework of Human Augmented Mapping that allows us to achieve this integration. We propose further a method to specify and represent *regions* that relate to a user's view on the environment. We assume an interactive setup for the specification of regions and show the applicability of our method in terms of distinctiveness for space segmentation and in terms of localisation purposes.**

I. Introduction

Service robots are designed to support people in their everyday life. This implies operating in close proximity to human users and sharing the environment with them. Even more, a mobile service robot needs to move within this environment from one location to another to provide its services. This requires navigation as well as localisation and mapping abilities. Robotic localisation and mapping is most often performed using geometric features being derived from sensory measurements. Simultaneous localisation and mapping (SLAM) methods allow a robot to navigate in an environment and acquire a map concurrently [21, for an overview]. This map can later be used for the navigation tasks enabling the robot to assist the user.

However, such feature based representations are typically different from the spatial models that humans use to define and reason about the same environment. This poses a challenge in particular if the system is to be operated by novice users without any robotics background.

Humans have a topological and mostly hierarchical representation of their environment [14]. In most domestic settings for service robots it can be assumed that users are familiar with the robot's operating area. Individual preferences and usage of this environment contribute to a *personalised* view on the surroundings, and in many cases the user's model of the environment is partial without a representation for all objects and places within the domain of operation. There is thus a need to reconcile the user's models of the environment with the robot's representation of the same space – a *shared* model that can be *personalised* is needed. We consider our framework of *Human Augmented Mapping* a possible way to approach this issue.

A central question is here how to partition the map derived from sensory data into *regions* that correspond to areas considered relevant by users. In other words, how can the underlying geometry be tied to a graphical model of the environment that represents the users understanding of the space.

The scenario considered here is one where a user as part of an initial installation of the system gives the robot a tour of the environment and provides guidance to rooms and locations within the environment – the relevant entities are labelled by the user, e.g., "this is the kitchen".

The objective of this article is the presentation of a strategy to partition space into *regions* that is the basis for user labelling and for tying such places to areas in the geometric map used by the robot for navigation.

We propose a statistics based definition of separate regions in the context of our system for Human Augmented Mapping. The method is evaluated in terms of the distinctiveness for the segmentation of a given environment and in terms of region classification or categorisation for localisation purposes.

A. Overview

The rest of this paper is organised as follows. In sections II and III we describe our framework for Human Augmented Mapping and refer to related work. Section IV explains our approach to the definition and representation of regions in an environment that is evaluated in section V. A conclusion and ideas for future work are presented in section VI.

II. System context

We propose a solution to the problem of representing regions in a topological map for a service robot. Since this is part of a conceptual approach to building a human comprehensible environment representation, the overall concept of Human Augmented Mapping is outlined to provide a context for the method described in this paper.

A. Human Augmented Mapping - concept

We assume a service robot designed to work in a domestic environment populated by humans - potential users. We also suppose that for appropriate communication about the robot's workspace a graph representation is needed, that can incorporate concepts into its nodes. On the other hand we assume that even an underlying metric representation is needed for the robot for exact localisation and navigation to perform its services. By taking the human knowledge and abilities into account and controlling the mapping process interactively, it is possible to integrate the human concepts and understanding of the environment into the resulting environment representation. This helps the human to communicate with the robot about its tasks and whereabouts according to the semantic and conceptual understanding the user has. For the robotic mapping process it is helpful to consider the user's information also for building a topological representation that forms the link between the conceptual graph representation and the metric map. It becomes possible to resolve ambiguities and answer questions related to different levels of the overall representation, e.g., "you mentioned one bedroom already, is this the same or a second one?" (conceptual/semantic level) or "was this a door we passed?" (topological level). This implies that a two way communication has to be made possible. Different types of events, i.e., external conceptual input from the user and internal detection of topologically significant structures, have to be considered. Figure 1 shows a possible system in a schematic way. The interaction part is crucial for the overall system to function but not relevant for the work presented here. Therefore the interaction functionalities (HRI) are not presented in detail. A form of topological

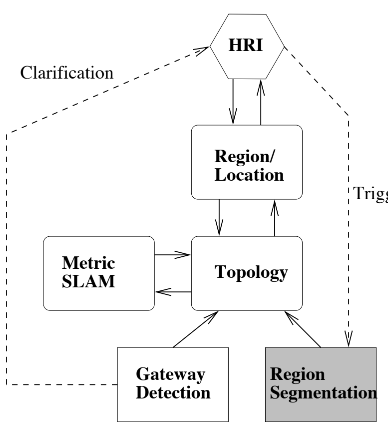

Fig. 1. Human Augmented Mapping (HAM) overview. The dashed arrows illustrate as a shortcut the information that is send from and to the user regarding the segmentation of the environment. Solid arrows represent information passed on internally between the respective system parts.

node distinction is needed, which is accomplished by the region segmentation module. This module can also be used for classification/categorisation to facilitate localisation on the basis of a previously acquired map and is described in detail in this article.

B. Representation

To classify spatial entities (e.g., rooms) it is necessary to separate them as spatial unit from the rest of the environment representation. We will term such spatial unit a *region*. Topologically speaking, a *region* can contain several distinct places (e.g., different in terms of the perceivable appearance of the environment at the respective location) that form the nodes of a topological graph representation. Another concept used in the place graph is formed by *locations*. Those represent specified poses (with respect to the current *region*) related to distinctive places or (mostly) static objects that can later be related to actions the service robot can perform to fulfil tasks.

C. Events

We consider two types of events that can trigger the system to segment the environment in the internal representation. One is to receive external input that annotates a certain spatial entity with a label (e.g, "... this is Elin's office..."). The other type of event is the data driven detection of a "new area". One possible solution is a gateway detection, but in this work we concentrate on investigating a feature descriptor based method for the segmentation of the environment and its possible use for the classification of specific *regions*.

III. Topological and hybrid mapping

Since our presented work concentrates on obtaining a topological partitioning of a given environment but is embedded in a context of hybrid (metric/topological) mapping we give an overview of related work in both areas.

A. Topological mapping

Approaches to topological environment representation, or map building, have been reported in the context of different presentation strategies or learning. One strategy is to predefine the topological structure of an environment and use this map for localisation and navigation purposes. Nourbakhsh *et al.* used this strategy for their implementation of a path planning system for "Dervish" [16]. The limitations of such an approach in the context of our Human Augmented Mapping framework and the arbitrary environment we assume are obvious: the complete possible working environment for the robot needs to be known in advance, including all possible transitions along the edges and measurements to describe doors and hallway intersections.

Other, more adaptive methods that assume the robot to acquire a topological representation of its environment are based on (sensory) data obtained while travelling. Those can be subdivided in unsupervised and supervised approaches.

An unsupervised/autonomous method for the detection of *places* is suggested by Beeson *et al.* [2]. The authors propose to use the extended Voronoi graph for the segmentation of the environment which was initially investigated by Choset *et al.* [4] and under a paradigm for an exploration strategy by Kuipers and Byun [10]. Their definition of a place suits the requirements and abilities of an autonomous system, but does not necessarily correspond to a personalised representation of a human user. This limitation can be observed also for other completely unsupervised methods of topology learning, as for instance proposed by Tapus *et al.* [19]. Here a method based on a combination of images and laser range data is used. The appearance of a certain area is captured as a "fingerprint", a string that represents different types of observed features (colour occurrences, vertical lines, corners, etc.) in the angular order they were perceived in. Such fingerprints are actually a rather concise description of a certain area that can be obtained on-line and can probably be triggered by both external or internal events. A disadvantage of the purely sequential representation though is that not the area (*region*) itself is captured as a spatial entity.

Another approach to segment different regions into a topological graph by autonomously detecting door passages is used to show the use of clarification dialogues in the context of Human Augmented Mapping by Kruijff *et al.* [9]. The graph obtained separates clearly only those regions from each other that have been labelled by the user, but so far the method does not incorporate the area information. The topological nodes are defined by associating travelled paths to the label.

The capturing of a complete area as one unit is suggested by Diosi *et al.* [5], who use a watershed implementation after interactively labelling positions that are then related to the areas that include them respectively. Compared to our approach, a clear difference lies in the assumption implicitly understood from Diosi *et al.* that all rooms and and other areas have to be specified in one complete tour to avoid merging of too many unlabelled regions into one "room". We consider this a strong limitation. In a pilot study [23] it was observed that potential users do not necessarily describe every room or area to a robot, but pick those that they personally consider important.

Mozos *et al.* show, how the *category* of a certain area (room, doorway, or corridor) can be determined with the help of supervised learning [15]. They generate a number of features from raw laser range data sets that were obtained at different locations corresponding to the named categories and use these features to form a training data base for the learning method. We adopt the idea of using a set of features to represent a laser range data set, that we obtain in *regions*, but use an even more concise set of features (see section IV for details).

For the representation of convex areas Kröse showed that it is possible to represent such regions reliably by obtaining only one sample range data set and transform it to its centre point and bearing with the help of a principal component analysis to anticipate future scans [8]. This method alone has the limitation of working robustly only in convex spaces but we believe that it is usable also for other areas as one method in a more complex framework. Our representation for *regions* is thus related to this proposed approach.

An approach to supervised learning of topology was reported by Althaus and Christensen [1]. They had a user guiding a robot through an office environment and assumed an explicit external trigger given by the user, when a new node in the topological graph had to be created. They assumed nodes as rooms that allowed for activities and doorways / gateways as edges. The triggering event had to occur exactly (metrically) where the link between nodes should be placed. The approach presented here offers a more lax strategy to supervised map annotation.

B. Hybrid mapping

A significant number of publications report on work in hybrid and hierarchical mapping. The motivation to build hierarchies of maps, e.g., to link metric or feature based local maps in global topological graph structures lies in two issues in robotic mapping. One is the computational complexity of metric maps, that can be reduced significantly by separating the environment representation into smaller maps that are linked by a topological graph [3], [13], [18]. The other issue is the incorporation of semantics that requires a topological or even conceptual level for the integration of the semantic information, but on the other hand needs in some cases to rely on a metric low-level representation for exact navigation purposes [7], [12]. One difference of these so far existing hybrid methods compared to ours is that the acquisition of the local maps is done either autonomously based on the particular specification on how the local map (e.g., a place) is defined [3], [12], [13], [18] or *a priori* as grid maps that are then linked into the hierarchy of maps and/or concepts by an anchoring technique [7]. Our concept requires a simple, concise description for the regions to be incorporated in our framework in an on-line fashion.

IV. STATISTICS BASED DEFINITION OF REGIONS

In this section we explain our approach to represent regions with data obtained from a laser range finder. We assume that the characteristics of an arbitrary region can be captured from a rather small data set (in our case a 360° range scan) obtained at one position. We assume also that for continuous updates (consistency checks) a data set covering 180° can be used in a similar way but leave that as subject to future investigations. For convex areas the validity of the initial assumption could be shown by Kröse [8], but we believe that characteristic properties of arbitrary areas can be captured with a similar approach. Mozos *et al.* [15] proposed the use of a feature set obtained from raw laser range for the supervised learning of "region categories". Our concept requires the acquisition of the environment representation without prior knowledge of the particular surroundings, i.e., it is not obvious that a respective system can previously be trained on some generic data set that suits the majority of possible working spaces. Thus, the description for regions we use to define the segmentation of the environment has to be even more concise so that it can be used in an on-line fashion. We are in fact able to show that even with a reduced set of statistic features a fairly precise categorisation of different regions can be obtained from only a very small set of interactively specified data samples.

We assume the axes of the largest ellipse fitting the range data as two characterising features and the mass (area) of the complete space covered by the scan as a third. The ellipse itself allows to decide which geometrically defined area belongs to the *region*. We are aware that this is only a rough estimate, since not all parts of a rectangular room can be covered by just one ellipse and in some cases areas outside the actual *region* might be assigned to it. We leave the solution to those issues to other methods like a gateway detection and clarification dialogues that can be invoked in ambiguous situations. We also assume that, although we present our approach as a "one-shot" method for the initial specification of a *region* in this paper, the representation can be continuously updated by the integration of features obtained from current poses in the overall context of our framework.

We investigate thus the three following features that characterise a laser range data set $\{X_i : 0 \leq i < N\}$, where N is the number of data points $X_i = (x_i, y_i)$. Those features are:

a) the area (or mass) m of the "visible" part of the represented region, and
b) the maximum range $l1$ and $l2$ along the two principle components of the data set (the axes of the "main" ellipse).

Locations according to section II can be integrated into the region with their relative position to the centroid $\bar{X} = (\bar{x}, \bar{y})$ of the data set.

Due to the angular sampling in laser range finders the spatial representation is non-uniform [1]. To compensate for this effect the centroid is computed as a range weighted average

$$\bar{X} = (\bar{x}, \bar{y}),$$

with

$$\bar{x} = \frac{1}{\sum_{i=0}^{N-1} r_i} \sum_{i=0}^{N-1} r_i x_i$$

and

$$\bar{y} = \frac{1}{\sum_{i=0}^{N-1} r_i} \sum_{i=0}^{N-1} r_i y_i$$

where $r_i = \sqrt{x_i^2 + y_i^2}$ is the distance of the data point from the origin of the data set, i.e., the position of the laser range finder. The data set is then transformed to the set $\{X_i' = (x_i - \bar{x}, y_i - \bar{y}) : 0 \leq i < N\}$ relative to the centroid. To compute the mass of the region an ordered data set is assumed, i.e., each data point X_i' is required to represent a smaller bearing angle α_i' as its neighbour X_{i+1}'. This allows estimation of the area m bordered by the data set to

$$m = \left(\sum_{i=0}^{N-2} m_i \right) + m_{N-1},$$

with

$$m_i = \frac{1}{2} tan(\alpha_{i+1}' - \alpha_i') {r_i'}^2$$

and

$$m_{N-1} = \frac{1}{2} tan(\alpha_{N-1}' - \alpha_0')(r_{N-1}')^2$$

where r_i' is the distance of the transformed point from the centroid. Since this estimated covered area is depending on objects that are placed in the region it represents an index of clutter, which is helpful to differentiate between regions of the same basic layout, but with different furnishing.

In order to obtain $l1$ and $l2$ a principal component analysis (PCA) has to be performed. The principal components correspond to the two eigenvectors E_1 and E_2 (to the corresponding eigenvalues λ_1 and λ_2) of the covariance matrix Q with

$$QE_i = \lambda_i E_i, \quad i = 1, 2$$

where

$$Q = \begin{bmatrix} C_{XX} & C_{XY} \\ C_{YX} & C_{YY} \end{bmatrix}$$

and

$$C_{XX} = \frac{N}{(N-1)\sum_{i=0}^{N-1} r_i} \sum_{i=0}^{N-1} r_i {x_i'}^2,$$

$$C_{YY} = \frac{N}{(N-1)\sum_{i=0}^{N-1} r_i} \sum_{i=0}^{N-1} r_i {y_i'}^2$$

and

$$C_{XY} = C_{YX} = \frac{N}{(N-1)\sum_{i=0}^{N-1} r_i} \sum_{i=0}^{N-1} r_i x_i' y_i'.$$

The covariances also have to be weighted due to the non-uniform sampling of the laser range data set[2]. We use the linear weights r_i, interpreting the original distances as the factor responsible for the *distribution* of the data samples around the laser range finder, which we have to compensate for. The two features $l1$ and $l2$ are now estimated as the maximum distances represented in the data set along the bearing angles of E_1 and E_2. To make sure that such a point is found, a tolerance threshold around the bearing angle is employed. The data set is now represented by the quadruple $reg = (name, m, l1, l2)$ and stored as a basis for comparisons. Also the obtained description is used to decide if already specified *locations* happen to be inside the covered area and thus can be assigned to the specified *region* in the place graph at the higher level of our framework.

To compare two region representations a simple nearest neighbour search over the Euclidean distances between obtained features m, $l1$ and $l2$ is performed. The distances are normalised with their average value to compensate for the significantly different orders of magnitude. In a number of tests with different distance measures and with different weight settings we found, that the overall results for the categorisation of rooms did not vary significantly which allowed us to choose this rather simple method.

V. Evaluation

The method we present to represent *regions* in a map can be evaluated in two different contexts. Firstly, we want to know about the distinctiveness or the segmentation power of our features. I.e., we need to know, how well the environment

[1] as a result of the equidistant angular resolution with which a laser range finder scans the environment objects in the direct vicinity of the laser range finder are represented with considerably more data points than objects that are further away

[2] we use the formula for weighted variances according to www.itl.nist.gov/div898/software/dataplot/refman2/ch2/weighvar.pdf.

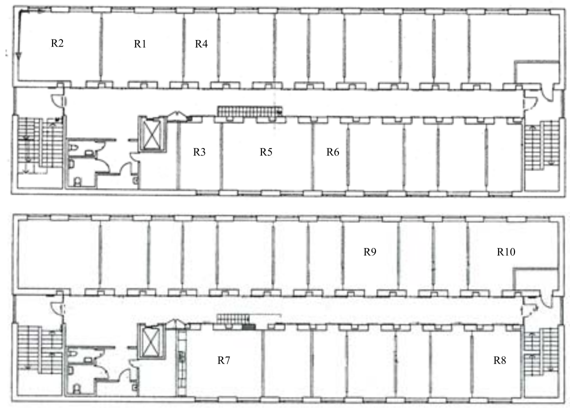

Fig. 2. The ten rooms of our office environment, that were used for the tests (R1-R10).

is described with regions that have been specified using the method described above. Secondly, we can use our feature sets for a classification/categorisation approach to facilitate localisation. Here different issues have to be considered. One is that we have a metric SLAM method integrated in the system that allows constant and exact localisation (in case that we are not dealing with "woken up" or "kidnapped" robot). What we thus are interested in is the ability for the system to report its current position in the context of the place graph or topological map it acquired, which means that also positions that are initially not included in the description of one region can be recognised as consistent with it. A second issue is in fact to facilitate localisation in a "waking up" or "kidnapped robot" scenario by reducing the search space for localisation on the basis of the metric map(s) or a topological exploration strategy. Assumed that the system knows to be either in room A or room B, but definitely not in any other room, can reduce the effort in exact localisation for large environments. We evaluate our proposed method in the two different contexts of distinctiveness and categorisation.

A. Categorisation

We tested our method in the context of classification and recognition of specific areas (rooms) for a number of rooms in our office environment. Figure 2 shows a schematic drawing. Looking at the picture it becomes obvious that certain groups of rooms can be identified considering the size. Within the groups the rooms are quite similar to each other as far as their size and shape are concerned. Since additionally the larger rooms correspond to robotic or vision laboratories and the kitchen, where the smaller rooms are offices and a workshop, they are also quite similarly furnished. Thus, it is not surprising that the results for the classification and recognition of particular rooms are not convincing. In a test setup we used stored region representations (feature sets) that were compared to all other feature sets available. The nearest neighbour according to the description in the previous section was picked as the recognised region. The overall recognition rate for this test was 40% which is clearly not sufficient for classification. For other test setups (using the average of the feature descriptions for each region or a one-shot presentation) we had similar low rates. We refrained thus from using the method for the classification of particular regions. Still, we were interested to see if it would be applicable for a *categorisation* that could be used to facilitate localisation.

Looking at the rooms in three groups or categories (i.e., "large open spaces" = R1, R2, R5, R7, "medium size cluttered/odd offices" = R8, R9, R10, and "small cluttered offices" = R3, R4, R6), we observe a recognition rate of 88%. Here it has to be noted that the categories were chosen only from the roughly estimated size of the complete room. We noticed that the errors mostly occurred for one of the medium size ones. Now, this office (R10) has a considerably different shape (L-shape) and is heavily cluttered with office furniture. This office can easily be perceived as several small, cluttered offices which were in fact the ones it got confused with. Since our framework assumes no prior knowledge of categories, a grouping of defined regions would have to be done according to a similarity measure. Such a measure could be the likelihood of confusing a particular feature set with another that belongs to a differently labelled region.

Grouping according to this measure (i.e., "large open spaces" = R1, R2, R5, R7, "medium size offices" = R8, R9, and "small/odd, very cluttered offices" = R3, R4, R6, R10) would result in a recognition rate of 94%. The remaining errors are mainly due to the fact that a previously correct "in group" recognition for R10 becomes an error by regrouping. These rates suggest, that it is in fact possible to give a rather strong estimate for the validity of a hypothesis for global localisation in terms of *categories* of rooms. We believe that this holds for most indoor environments in which at least two types of rooms can be found. The uncertainty for a global localisation in a "waking up" scenario could thus be reduced significantly before invoking either a metric localisation method or an exploration strategy to disambiguate the situation. Such strategies have been proposed already by Kuipers and Byun [10] and have been investigated later also by Seiz *et al.* [17].

B. Distinctiveness

The other issue to be evaluated is the distinctiveness of the method. Given that a particular region is presented to the system the question is, how dependent the acquired representation is on the current position of the robot. Intuitively and along the argumentation of Kröse [8] one would assume, that the data obtained in a simply structured (but not empty) convex room with only one door will be rather similar for different positions. Figure 3 shows such a room (R7) with the positions from where the 360° range data sets were taken (P1–P6). Additionally the positions where the system calculated the corresponding centre of the obtained laser range data are marked with grey dots and numbers 1–6. Not surprisingly they all fall into an area of about 35cm radius, but one (no.6). This particular data set was obtained very close to and in the line of the doorway, where a significant portion of the corridor could be perceived already. Table I shows the changes of different

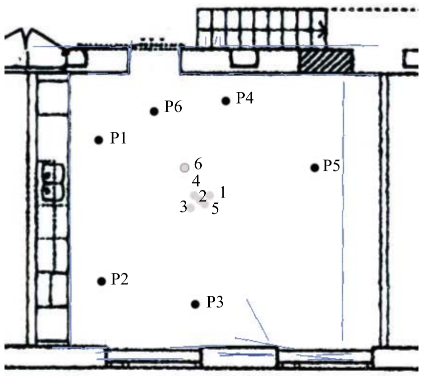

Fig. 3. One of the rooms (the kitchen – R7) with the positions (P1-P6) from where the data sets were obtained. The numbers 1-6 mark the corresponding centroids calculated for those sets. The thinner (blue) lines represent the line features extracted for the metric SLAM (partly caused by a sofa placed along one wall and thus not corresponding to the walls themselves.

features or measures from one position to another. From those measures it becomes obvious that for the major part of nearly convex regions the position to acquire the features for this region is arbitrary. In the immediate proximity of doorways though the representation becomes slightly unstable. This is still acceptable when interpreted in the sense of a human environment representation, where a door passage might be a transition not only in the spatial sense and thus is difficult to describe in a binary way as strictly “inside” or “outside”.

More interesting than the convex and nearly convex regions are actually those that are of particular shape or have a very distinct type of furnishing. This is in our set of rooms the case for R8 and R9 (furniture) and R10 (shape and furniture). Figure 4 shows similar to figure 3 the positions (P1–P5) from where data sets were taken together with the corresponding centroids (1–5) and the laser range data features. Additionally an illustration of the furnishing that makes R8 look like two cubicles connected by a corridor is shown. Still, since the room is only of medium size and thus the “cubicles” are not too deep, a large portion of the room can be perceived from at least positions P2, P3, and P4. Accordingly the feature sets are altered gradually along the path from P1 to P5. Table II shows the variation over a number of measures for rooms R8, R9, and R10. For those cluttered or heavily structured rooms the

TABLE I

STATISTICAL VALUES FOR R7

Feature	Mean	Variance
“Mass” / area (M) [m^2]	21.23	2.73
Length 1 (major axis) (L1) [m]	8.45	4.37
Length 2 (minor axis) (L2) [m]	5.10	0.13
Excentricity (E)	0.71	0.04
Distance between centroids (D) [m]	0.34	0.05
Angular difference between major axes (A) [rad]	1.11	1.04

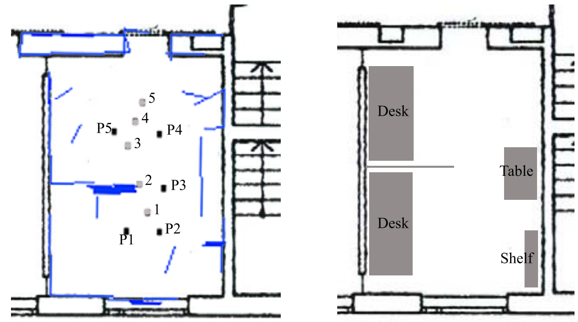

Fig. 4. Positions from where data sets were taken in R8 together with the laser range data features (thinner (blue) lines) (left) and a schematic drawing of the furnishing in R8 (right).

TABLE II

STATISTICAL VALUES FOR R8, R9, R10

	R8		R9		R10	
Feature	Mean	Var	Mean	Var	Mean	Var
m	18.04	10.71	13.63	9.93	23.45	364.62
l1	7.67	9.17	7.03	5.68	8.56	3.22
l2	2.76	0.53	3.42	0.44	2.39	0.38
E	0.82	0.05	0.84	0.01	0.95	0.00
D	1.15	0.33	1.02	0.26	1.47	0.47
A	1.33	1.57	1.02	0.63	1.14	0.73

higher variances (compared to table I) for our initial features (m, l1, l2) indicate an unstable geometric representation of the perceived area along the paths the robot took. Apart from those the distance of the centroids from each other can give quite good an indication for the change of the area perceived when used while travelling. The most significant change for R10 though can be observed in the area. The other features do not change as significantly as each part of the room represents an area quite similar to the others as far as the shape is concerned, but different in direction and size.

The angular distances can be interpreted as follows: In case of a generally low excentricity of the main ellipse an angular distance close to any multiple of $\pi/2$ does not represent a significant change since the ellipse is almost circular. In the case of a high excentricity an angular distance close to any odd multiple of π would indicate a significant change of the shape of the perceived environment. This means for our method, that measurements for “similarity” can be based on those features. The features displayed in the tables above can all be derived from the originally calculated features m, l1 and l2 together with the global position of the region represented.

C. Summary

From these results we conclude that our method to represent distinct regions works well as a categorisation approach for global localisation. More important, the distinctiveness for the segmentation of an environment is very good for simply structured regions as almost convex rooms. In strongly structured areas the representation is altered depending on the position the data set was obtained from. Still, since the observed changes occur gradually a similarity measure can be used here to identify ambiguities that can be resolved by the interaction with the user. The investigation, in how far our concise feature based representation can be used with continuously obtained data samples to identify transitions from one topologically consistent region into the next one, is subject to current work.

VI. Conclusion

With this paper we explained a method for a concise statistical feature based representation and segmentation of space in the context of Human Augmented Mapping. We presented the system context in which we see the applicability of this method and evaluated our approach in terms of region categorisation for localisation purposes and distinctiveness for segmentation.

Since we see the approach as part of a system context that can incorporate other, complementing methods, we can conclude that the method described in this paper is clearly applicable for our purposes. Given a respective measurement for similarity the method performed with a correct recognition rate of 94% for different categories of rooms. This is promising for the reduction of uncertainty in a global localisation approach. In terms of distinctiveness our method is a very concise approach to model geometrically stable regions as one area but also to detect transitions from one geometrically stable area into another one. This might result in an ambiguous situation as far as the direct relation to the human user's environment representation is concerned. Given the opportunity of interaction with a user that is provided by our framework, such possibly ambiguous situations can be resolved.

For our future work we consider the application of our statistical features in a continuously updated fashion to in fact define a topological map based on regions of geometrical stability with transitions into new regions. Such a map can then be compared and integrated with one derived from a pure gateway detection.

Another issue for future work is the application of an exploration strategy for global localisation. This can be based on an initial hypothesis calculated with the categorisation provided by our presented method.

Acknowledgement

The work described in this paper was conducted within the EU Integrated Project COGNIRON ('The Cognitive Robot Companion' - www.cogniron.org) and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

References

[1] P. Althaus and H.I. Christensen. Automatic map acquisition for navigation in domestic environments. *in Proc. of the International Conference on Robotics and Automation,* September 2003, Taipei, Taiwan.

[2] P. Beeson, N.K. Jong, and B. Kuipers. Towards Autonomous Topological Place Detection Using the Extended Voronoi Graph. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[3] M. Bosse, P. Newman, J. Leonard, S. Teller. SLAM in Large-scale Cyclic Environments using the *Atlas* Framework. *International Journal of Robotics Research,* 23(12):1113-1140, December 2004.

[4] H. Choset and K. Nagatani. Topological Simultaneous Localization and Mapping (SLAM): Toward Exact Localization Without Explicit Localization. *IEEE Transactions on Robotics and Automation,* 17(2), April 2001.

[5] A. Diosi, G. Taylor and L. Kleeman. Interactive SLAM using Laser and Advanced Sonar. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[6] J. Folkesson, P. Jensfelt, and H.I. Christensen. Vision SLAM in the Measurement Subspace *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[7] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J.A. Fernández-Madrigal, and J. González. Multi-Hierarchical Semantic Maps for Mobile Robotics. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* August 2005, Edmonton, AB, Canada.

[8] B. Kröse An Efficient Representation of the Robot's Environment. *in Proc. of Intelligent Autonomous Systems 6,* 2000, Venice, Italy.

[9] G.-J. Kruijff, H. Zender, P. Jensfelt, and H.I. Christensen. Clarification dialogues in human-augmented mapping. *in Proc. of the ACM Conference on Human-Robot Interaction,* March 2006, Salt Lake City, UT, USA.

[10] B. Kuipers and Y.-T. Byun. A robust qualitative method for spatial learning in unknown environments. *in Proc. of the National Conference on Artificial Intelligence (AAAI-88),* 1988, Los Altos, CA, USA.

[11] B. Kuipers. The Spatial Semantic Hierarchy. *Artificial Intelligence,* 119:191-233, 2000.

[12] B. Kuipers, J. Modayil, P. Beeson, M. MacMahon, and F. Savelli. Local Metrical and Global Topological Maps in the Hybrid Spatial Semantic Hierarchy. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2004, New Orleans, LA, USA.

[13] B. Lisien, D. Morales, D. Silver, G. Kantor, I.M. Rekleitis, and H. Choset. The Hierarchical Atlas. *IEEE Transactions on Robotics,* 21(3):473–481, 2005.

[14] T.P. McNamara. Mental Representations of Spatial Relations. *Cognitive Psychology,* 18:87–121, 1986.

[15] O.M. Mozos, C. Stachniss, and W. Burgard. Supervised Learning of Places from Range Data using AdaBoost. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[16] I. Nourbakhsh, R. Powers, and S. Birchfield. DERVISH An Office-Navigating Robot. *AI Magazine,* 16(2):53–60, AAAI, 1995.

[17] M. Seiz, P. Jensfelt, H.I. Christensen. Active Exploration for Feature Based Global Localization. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* November 2000, Takamatsu, Japan.

[18] N. Tomatis, I. Nourbakhsh, and R. Siegwart. Hybrid Simultaneous Localization and Map Building: a Natural Integrati on of Topological and Metric. *Robotics and Autonomous Systems,* 44:3–14, 2003

[19] A. Tapus, G. Ramel, L. Dobler, and R. Siegwart. Topology Learning and Place Recognition using Bayesian Programming for Mobile Robot Navigation. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 2004, Sendai, Japan.

[20] A. Tapus, S. Vasudevan, and R. Siegwart. Towards a Multilevel Cognitive Probabilistic Representation of Space. *in Proc. of the International Conference on Human Vision and Electronic Imaging X, part of the IST/SPIE Symposium on Electronic Imaging,* January 2005, San Jose, CA, USA.

[21] S. Thrun. Robotic Mapping: A Survey. *In G. Lakemeyer and B. Nebel, eds., "Exploring Artificial Intelligence in the New Millenium",* Morgan Kaufmann, 2002.

[22] E.A. Topp and H.I. Christensen. Tracking for following and passing persons. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* August 2005, Edmonton, AB, Canada.

[23] E.A. Topp, H. Huettenrauch, H.I. Christensen, and K. Severinson Eklundh. Bringing Together Human and Robotic Environment Representations - A Pilot Study. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* October 2006, Beijing, China.