simple mechanism, which was inspired by the human forearm. The optimization procedure presented here shows that the force-generating capacity of this mechanism is very sensitive to its kinematic parameters. However, following the optimization technique described here, the mechanism's RoM, inertia, and force-generating capacity match well with the human arm.

Since April 2009, we have been using BONES in a rehabilitation clinic to retrain arm-movement ability after stroke, with the use of the adaptive control algorithms that were described in [31]. BONES is allowing us to rigorously test whether functional transfer of robotic therapy is improved with the practice of naturalistic arm movements.

References

[1] M. Bergamasco, B. Allotta, L. Bosio, L. Ferretti, G. Parrini, G. M. Prisco, F. Salsedo, and G. Sartini, "An arm exoskeleton system for teleoperation and virtual environments applications," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1994, vol. 2, pp. 1449–1454.

[2] S. Coote, E. K. Stokes, B. T. Murphy, and W. S. Harwin, "The effect of GENTLE/s robot mediated therapy on upper extremity function post stroke," in *Proc. Int. Conf. Rehabil. Robot.*, 2003, pp. 59–61.

[3] A. Frisoli, F. Rocchi, S. Marcheschi, A. Dettori, F. Salsedo, and M. Bergamasco, "A new force-feedback arm exoskeleton for haptic interaction in virtual environments," in *Proc. First Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2005, pp. 195–201.

[4] S. Hesse, C. Werner, M. Pohl, S. Rueckriem, J. Mehrholz, and M. L. Lingnau, "Computerized arm training improves the motor control of the severely affected arm after stroke: A single-blinded randomized trial in two centers," *Stroke*, vol. 36, pp. 1960–1966, Sep. 2005.

[5] N. Hogan, H. Krebs, A. Sharon, and J. Charnnarong, *Interactive Robot Therapist*. Cambridge, MA: MIT Press, 1995.

[6] Y. Hurmuzlu, A. Ephanov, and D. Stoianovici, "Effect of a pneumatically driven haptic interface on the perceptional capabilities of human operators," *Presence: Teleoperators Virtual Environ.*, vol. 7, pp. 290–307, Jun. 15, 1998.

[7] L. Kahn, M. Zygman, W. Z. Rymer, and D. Reinkensmeyer, "Robot-assisted reaching exercise promotes arm movement recovery in chronic hemiparetic stroke: A randomized controlled pilot study," *J. NeuroEng. Rehabil.*, vol. 3, p. 12, 2006.

[8] T. Nef, M. Mihelj, G. Kiefer, C. Perndl, R. Muller, and R. Riener, "ARMin—Exoskeleton for arm therapy in stroke patients," in *Proc. IEEE 10th Int. Conf. Rehabil. Robot.*, 2007, pp. 68–74.

[9] D. J. Reinkensmeyer, L. E. Kahn, M. Averbuch, A. McKenna-Cole, B. D. Schmit, and W. Z. Rymer, "Understanding and treating arm movement impairment after chronic brain injury: Progress with the ARM guide," *J. Rehabil. Res. Dev.*, vol. 37, pp. 653–662, Nov. 2000.

[10] R. Riener, T. Nef, and G. Colombo, "Robot-aided neurorehabilitation of the upper extremities," *Med. Biol. Eng. Comput.*, vol. 43, pp. 2–10, Jan. 2005.

[11] R. J. Sanchez, E. T. Wolbrecht, R. Smith, J. Liu, S. Rao, S. Cramer, T. Rahman, J. E. Bobrow, and D. J. Reinkensmeyer, "A pneumatic robot for re-training arm movement after stroke: Rationale and mechanical design," in *Proc. 9th Int. Conf. Rehabil. Robot.*, 2005, pp. 500–504.

[12] T. G. Sugar, H. Jiping, E. J. Koeneman, J. B. Koeneman, R. Herman, H. Huang, R. S. Schultz, D. E. Herring, J. Wanberg, S. Balasubramanian, P. Swenson, and J. A. Ward, "Design and control of RUPERT: A device for robotic upper extremity repetitive therapy," *IEEE Trans Neural Syst Rehabil Eng.*, vol. 15, no. 3, pp. 336–346, Sep. 2007.

[13] M. Mihelj, T. Nef, and R. Riener, "ARMin II—7 DoF rehabilitation robot: Mechanics and kinematics," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007, pp. 4120–4125.

[14] C. Carignan, M. Liszka, and S. Roderick, "Design of an arm exoskeleton with scapula motion for shoulder rehabilitation," in *Proc. 12th Int. Conf. Adv. Robot.*, 2005, pp. 524–531.

[15] L. Q. Zhang, H. S. Park, and Y. Ren, "Developing an intelligent robotic arm for stroke rehabilitation," in *Proc. IEEE 10th Int. Conf. Rehabil. Robot.*, 2007, pp. 984–993.

[16] A. Kecskeméthy and A. Weinberg, "An improved elasto-kinematic model of the human forearm for biofidelic medical diagnosis," *Multibody Syst. Dyn.*, vol. 14, pp. 1–21, 2005.

[17] B. D. Adelstein, "Three degree of freedom parallel mechanical linkage," U.S. Patent 5 816 105, Oct. 6, 1998.

[18] J. Klein, S. J. Spencer, J. Allington, K. Minakata, W. Wolbrecht, R. F. Smith, J. Bobrow, and D. J. Reinkensmeyer, "Biomimetic orthosis for the neurorehabilitation of the elbow and shoulder (BONES)," in *Proc. BioRob*, Scottsdale, AZ, 2008, pp. 535–541.

[19] J. M. McCarthy, *Geometric Design of Linkages (Interdisciplinary Applied Mathematics)*, 1st ed. New York: Springer-Verlag, 2000.

[20] J. P. Merlet, "Jacobian, manipulability, condition number, and accuracy of parallel robots," *J. Mech. Des.*, vol. 128, pp. 199–206, 2006.

[21] R. Murray, Z. Li, and S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC, 1994.

[22] J. Klein, "Performance enhancing mechanisms for human manipulation," Ph.D. dissertation, Dept. Mech. Aerospace Eng., Univ. Calif., Irvine, 2009.

[23] E. T. Wolbrecht, D. J. Reinkensmeyer, and J. E. Bobrow, "Pneumatic control of robots for rehabilitation," *Int. J. Robot. Res.*, vol. 29, pp. 23–38, 2010.

[24] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. New York: Wiley, 2000.

[25] Mathworks, *MATLAB* R2009A ed., 2009.

[26] J. Rosen, J. C. Perry, N. Manning, S. Burns, and B. Hannaford, "The human arm kinematics and dynamics during daily activities—Toward a 7 DOF upper limb powered exoskeleton," in *Proc. 12th Int. Conf. Adv. Robot.*, 2005, pp. 532–539.

[27] Y. Saad, "Numerical solution of large nonsymmetric eigenvalue problems," *Comput. Phys. Commun.*, vol. 53, pp. 71–90, 1989.

[28] J. C. Perry, J. Rosen, and S. Burns, "Upper-limb powered exoskeleton design," *IEEE/ASME Trans. Mechatronics*, vol. 12, no. 4, pp. 408–417, Aug. 2007.

[29] J. P. A. Dewald and R. F. Beer, "Abnormal joint torque patterns in the paretic upper limb of subjects with hemiparesis," *Muscle Nerve*, vol. 24, pp. 273–283, 2001.

[30] C. A. Hautier, L. M. Arsac, K. Deghdegh, J. Souquet, A. Belli, and J.-R. Lacour, "Influence of fatigue on EMG/force ratio and cocontraction in cycling," *Med. Sci. Sports Exerc.*, vol. 32, pp. 839–843, 2000.

[31] E. Wolbrecht, "Adaptive, assist-as-needed control of a pneumatic orthosis for optimizing robotic movement therapy following stroke," Ph.D. dissertation, Dept. Mech. Aerospace Eng., Univ. Calif., Irvine, 2007.

Detecting Region Transitions for Human-Augmented Mapping

Elin A. Topp and Henrik I. Christensen

***Abstract*—In this paper, we describe a concise method for the feature-based representation of *regions* in an indoor environment and show how it can also be applied for door-passage-independent detection of transitions between *regions* to improve communication with a human user.**

***Index Terms*—Human–robot interface, semantic mapping, space segmentation.**

I. Introduction

In this paper, we aim to develop a case for a concise method for the segmentation of an indoor environment into a topological-graph

Manuscript received July 10, 2009; revised December 7, 2009; accepted April 19, 2010. Date of publication June 10, 2010; date of current version August 10, 2010. This paper was recommended for publication by Associate Editor D. Song and Editor L. Parker upon evaluation of the reviewers' comments. This work was supported in part by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

E. A. Topp was with the Centre for Autonomous Systems, Royal Institute of Technology (KTH), Stockholm 10044, Sweden. She is now with the Department of Computer Science, Lund University (LTH), Lund 22100, Sweden (e-mail: elin_anna.topp@cs.lth.se).

H. I. Christensen is with the College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0760 USA (e-mail: hic@cc.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2010.2049178

Fig. 1. (Left) "Hall" has been presented to the robot that now assumes the complete depicted area as "hall," since no "door passage" was passed while traveling. (Right) User wants the robot to understand that there is some part of the area that is NOT the hall, but, e.g., the "corridor." It seems natural to assume an "unspecified area" in the transition.

representation that is independent from particular "transition indicators," like door passages, and that allows the generation of a human-comprehensible environment representation for a mobile service robot. We assume such an environment representation as crucial to support meaningful interaction between a user and a supposed "general-purpose service robot" in and about the environment. We consider our framework of *Human-Augmented Mapping* (HAM) [1, Ch. 3] as a possible way to approach the issue of integration of robotic and human environment representation in general. The framework subsumes different aspects of robotic mapping, spatial representation, and human–robot interaction. Within the context of HAM, we assume an interactive scenario—a "home tour"—as the most natural way to provide the robot with the needed semantic information about the environment, as it is seen by the user. The human user guides the robot and gives names to things and places according to her personal preferences, while the robot builds a suitable (hybrid) map that is augmented with this information. In such a tour, it is not necessarily the case that the user will actively present all items [1, Ch. 6]; hence, the system-driven detection of transitions, e.g., from one room into another, is essential to make sure that the representation, which is generated by the robot, corresponds to the user's understanding of the environment. An obvious way to detect such changes is to find door passages. However, there are cases where the border between two structurally (and often also functionally) different rooms (or *regions*) is not described by an obvious separator, like a door passage. Fig. 1 visualizes such a "structural ambiguity." There are, of course, cases where one large room serves different functions, e.g., in very small studios with combined "living room" and "kitchen"; however, in this paper, we want to focus on a segmentation, which is based on structural features that can be observed in the environment. Such a feature-based representation should be suitable for the generation of *region* nodes in a topological-graph structure and support the detection of hypothesized transitions between *regions*. Ideally, the respective representation allows a robot to more or less immediately recognize a particular *region* as previously visited, even when it reaches it from a new "entry point." In this paper, we describe our concise method for the feature-based representation of *regions* as nodes of a topological-graph representation and show how it can also be applied for door-passage-independent detection of transitions between *regions*. We show the applicability of the method in different (interactive) contexts and give one "proof-of-concept" example for a successful "loop-closing" experiment.

II. Related Work

Since this study mainly deals with the issue of obtaining a topological partitioning of a given environment, we give an overview of related work in this area. We are aware of several works with the use of image-analysis techniques and object-recognition-based representations or categorizations for rooms or regions; however, due to the limited space, we mainly focus on structure-based methods in this brief overview.

One strategy is to predefine the topological structure of an environment and use this map for localization and navigation purposes [2]. The limitations of such an approach in the context of an interactive framework and the arbitrary environment that we assume are obvious: The complete possible working environment for the robot needs to be known in advance. Other, more adaptive methods that assume the robot to acquire a topological representation of its environment are based on (sensory) data obtained while traveling.

An unsupervised/autonomous method for the detection of *places* is suggested by Beeson *et al.* [3], which was based on earlier investigations in a related context [4], [5]. The definition of a "place" in these works suits the requirements and abilities of an autonomous system but does not necessarily correspond to a personalized representation of a human user. This limitation can also be observed for other completely unsupervised methods of topology learning like, for instance, the method proposed by Tapus *et al.* [6].

For the representation of simply structured convex areas, Kröse showed that it is possible to represent such *regions* reliably by obtaining only one sample range dataset and transform it to its center point and bearing with the help of a principal component analysis (PCA) to anticipate future scans [7]. Our representation for *regions* is closely related to this proposed approach, although we could move beyond the assumption of closed and simply structured areas. It is also well related to the approach that was presented by Buschka and Saffiotti [8], who detected "room-like" structures based on (sonar) range data, by the use of a very similar method. Because of the different nature of laser-range-finder data, our method appears less complex and easy to apply.

Mozos *et al.* show how the *category* of a certain area (room, doorway, or corridor) can be determined with the help of supervised learning [9], which was also used in another similar approach [10]. We adopt their idea to use a set of features to represent a (laser) range dataset, which we obtain in *regions*, but use an even more concise set of features [1, Ch. 4]. Further, Mozos *et al.* label places in the complete environment into a fixed number of categories, while we do not rely on any previously defined categories for the *regions* that can be specified by the user. This allows us to concentrate on the transition from one *region* into the other but not regarding to what category (in the sense of the mentioned work) the *regions* or the transition itself belong.

III. Representation of an Indoor Environment with Regions

In the HAM framework, we define a *region* as follows: *A region is a functionally and/or structurally delimited area of an indoor environment that can be a container for one or several particular locations and objects (other concepts in the framework, forming a hierarchy). A region offers enough space to navigate (typically regions correspond to, e.g., rooms, corridors, delimited areas in hallways).*

For this paper, we exclusively focus on *regions* and their structural properties. The *regions* that have been labeled in the assumed tour form the nodes of a topological-graph structure that (among other entries according to the HAM model) contain subgraphs, which represent known, viable paths (navigation graphs). We also introduce a generic node, i.e., the "generic *region*," as a starting point and to cope with situations in which entities of other conceptual levels (e.g., locations) are specified without the surrounding *region* being named earlier.

The general assumption is that a "*region* node" in the topological graph is generated when the user shows a particular *region* to the robot. This can also happen when the robot detects a significant change in the environment—a hypothesized transition from one *region* into another—and asks for clarification of the situation, while the user did not (yet) introduce any new region actively.

A new *region* node is linked into a topological-graph structure on two levels: A high-level edge describes the topological link between two nodes, i.e., the fact that it is possible to somehow travel from one *region* to a neighboring one. The so-far existing navigation graphs of these regions are rebuilt so that the high-level edge receives at least one concrete instantiation that describes *how* to travel. This concrete link is represented as a (metric) path vector relative to the *region* node's geometrical center point. In addition to the topological links between the graph nodes, each *region* node is described with its center point $\bar{X}$ and angle θ relative to the starting position of the tour. To derive these metric links, we make use of a (corrected) position estimate. Since the metric links between *region* nodes are described relative to the corresponding node "origin," we believe it would be possible to decouple small local (metric) maps from the global metric one if necessary. Hence, we assume an arbitrary, "classic" simultaneous localization and mapping (SLAM) method as suitable for the purpose of retrieving a sufficiently correct pose estimation.

A. Representation of Regions and Detection of Transitions

To actually compute the representation for the *region* nodes in our topological-graph structure, we rely on statistical features that are derived from laser-range datasets. This is a very concise, computationally rather inexpensive, and flexible method, and we propose to use it not only for the description of the *region* nodes that are specified by the user but also for continuous comparisons of hypothesized *region* representations for transition detection. The detection of transitions can then be handled independently from or as a complement to explicit cues, like, for instance, door detectors, as used by other approaches [11].

1) Region Representation: We represent specified *regions* with the help of a number of statistical features that are computed from a 360° laser-range dataset [1, Ch. 4], which are as follows.

1) mass m: the free space surrounding the robot ("clutter index");
2) length $l1$ and $l2$: the length along the two principal components of the dataset (overall "size");
3) excentricity e: the excentricity of the ellipse described by the two principal components (overall "shape");
4) center point $\bar{X}$: the centroid of the dataset;
5) angle θ: the angle of the first principal component relative to the origin of the map/the starting point of the tour.

The first three features describe the properties of the *region*, while the latter two are used to link the corresponding *region* node into the graph structure, as described previously. Although these features are related to each other, we found in some empirical tests that they all contribute to the distinctive power of the description [1, Ch. 5]. The features over a range dataset $\{X_i : 0 \leq i < N\}$, where N is the number of data points $X_i = (x_i, y_i)$, are computed as follows. To compensate for the distance-related distortion of the laser-range dataset, the centroid of the dataset 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 dataset, i.e., the position of the laser-range finder. The dataset is then transformed to the set $\{X'_i = (x_i - \bar{x}, y_i - \bar{y}) : 0 \leq i < N\}$ relative to the centroid. This allows the estimation of the area m bordered by the dataset 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 dependent on objects in the vicinity, it represents an index of clutter, which is helpful to differentiate between regions of the same basic layout but with different furnishing.

We perform a PCA to obtain the two eigenvectors E_1 and E_2 of the dataset. We then estimate the two features $l1$ and $l2$ as the maximum distances represented in the dataset 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 dataset is now represented by the quadruple $regDesc = (name, m, l1, l2, e)$ and stored as properties of the corresponding *region*.

2) Detection of Transitions While Traveling: While traveling through the environment, the available range datasets are continuously used to generate a "hypothesized *region*" representation of the surroundings, which is compared with a previously specified one to decide whether the environment has changed significantly so that it likely appears to have entered a new *region* [1, Ch. 4].

To compare two *region* representations, we compute a distance measure d from the relative differences in each of the descriptive features as follows:

$$d = \sqrt{\hat{m}^2 + \hat{l1}^2 + \hat{l2}^2 + \hat{e}^2}$$

with

$$\hat{f} = \left(1 - \frac{f_{\rm hyp}}{f_{\rm cur}}\right) \quad \text{for} \quad f \in \{m, l1, l2, e\}$$

where $f_{\rm cur}$ and $f_{\rm hyp}$ indicate the respective feature of the current and the hypothesized representation, respectively. We evaluated several distance measures in initial empirical tests and found the presented one most suitable to capture the changes that are being implied by the structure of the environment. If the distance measure d exceeds a threshold, a significant change in the environment representation is hypothesized.

To improve stability, we assume that the change has to be stable over a number of data cycles. Additionally, it is obvious that the robot cannot have entered a new *region* when it has not moved; hence, we apply a minimum-distance threshold between transition detections. These two conditions allow lowering of the computational effort and make the system more stable.

The hypothesized *region* representation is compared with the previously accepted current one. In the case that a significant change is detected, the hypothesized representation is checked against all other available representations (nodes in the graph) as to whether any of them matches sufficiently well and is not completely unlikely to have been entered, given its (metric) position. If none of the previously specified *region* representations match, the system is assumed to be back in the "generic *region*." Given appropriate interaction capabilities, a transition detection with the hypothesis for the actually entered *region* can lead to a confirmation dialogue with the user, which then can result in the specification of a new *region*. The corresponding representation

TABLE I
TRANSITIONS DETECTED WITH BIRON

counts & values	ap. 1	ap. 2	lab 1	lab 2	BIRON
min. movement	1m	1m	1m	1m	1m
data cycles	1	1	3	3	3
n	18	22	13	4	12
nCorr	18	24	13	4	5
nSens	18	20	13	4	9
% of nCorr	100	83	100	100	180
nSpurious	0	2	0	0	3
% of n	0	9	0	0	25
nMiss	0	4	0	0	0
% of nCorr	0	17	0	0	0

is then added to the graph and is used as the accepted current one for further comparisons.

IV. IMPLEMENTATION AND EVALUATION OF THE SYSTEM

We investigate our method for the representation of *regions* and detection of transitions in the context of three different implementations. Although the used implementations were slightly different regarding their integration into an interactive system (or lack thereof), the general system setup for both online runs and data collection consisted of a mobile nonholonomic platform with one laser-range-data finder mounted at about 30 cm (in one case 50 cm, but still below the top level of most furniture) above the ground. In all cases, time-stamped odometer readings and laser data were made available, and depending on the interaction capabilities, the user's labeling and the raw datasets used for the specification of *regions* together with the resulting feature-based representation are also available. As mentioned previously, we consider two types of events that are relevant to generate a new region representation in the topological graph. One is the—user-initiated—specification of a new region, the other is the—data-driven and robot-initiated—detection of a structural change in the environment. When the user through personal initiative or as a result of a clarification discourse specifies a *region*, the robot acquires a 360° range dataset by turning around once; for the continuous comparisons, we use "virtual scans" generated from a local map (part of the software tool package "CURE," courtesy of P. Jensfelt and J. Folkesson, Royal Institute of Technology (KTH), Stockholm, Sweden) to compensate for the fact that the used robots only have one laser-range finder available. Because of the different data sources, we refrained from an update of the representations in the presented experiments; respective possibilities were discussed in a different context [1, pp. 84–90].

A. Evaluation

We discuss our approach in the context of a number of datasets that were obtained in different indoor environments, which represent the range from "laboratory conditions" in an office building to the "real-world conditions" in a small, actually inhabited, apartment. In addition, the settings range from explicit test runs, where data collections from tours with a remotely controlled robot [12] were evaluated, to a fully interactively controlled run conducted by a test user. These different settings had, of course, influence on how the system could handle a detected transition *after* it stated its hypothesis; however, the main aspect was to use the *region* representation for the detection of transitions in the environment in all cases and evaluate the suitability of the method for the generation of a human-comprehensible representation

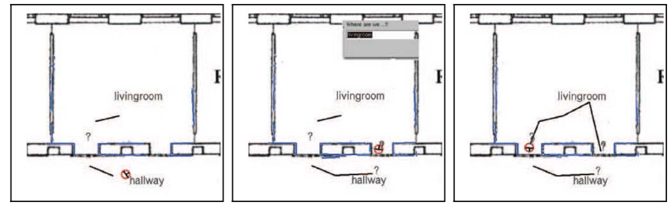
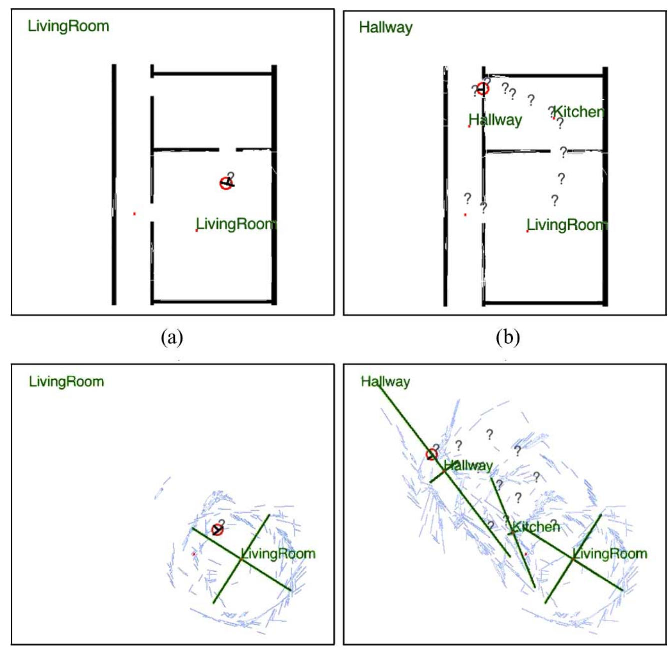


Fig. 2. (Left) Run starting in the "living room" (center) leaving it and coming back in (with the hypothesis "living room" in the dialog box) and (right) after merging the navigation graphs inside the room. Question marks indicate positions where the system asked for clarification, and solid black lines represent the navigation graphs. [Screen shot on architectural drawing.]



(a) (b)

(c) (d)

Fig. 3. Experiment with BIRON,which was visualized in post-hoc runs. Question marks indicate positions, where the robot asked the user for confirmation. In the upper left corner of each image, the system's hypothesis of the current *region* is shown. (a) and (b) Reconstruction with the help of a pose-estimation module with the room labels marked at the positions, where they were given to the robot by the user. (c) and (d) Visualization of the originally generated representation, which is based on raw odometer readings, where the crosses indicate the ellipses for the three specified *regions*. (a) or (c) Starting in the "living room." (b) or (d) Concluding the tour in the hallway after going through living room and kitchen twice.

of an indoor environment. We evaluated the runs (guided tours) in these different environments with respect to the following criteria:

1) consistency of the generated separation of *regions* in the environment with the "common understanding" of this separation;
2) detection of "obvious" transitions (doorways) and ambiguities, where, e.g., a hallway opens up into a larger area;
3) loop-closing ability on the conceptual/semantic level when coming back to a previously specified *region* through a new entry point;
4) overall number n of detected ambiguities/transitions (and requests for confirmation from the system for the fully implemented systems), with $nCorr$ being the number of expected transition detections between structurally different areas given the path of the robot;

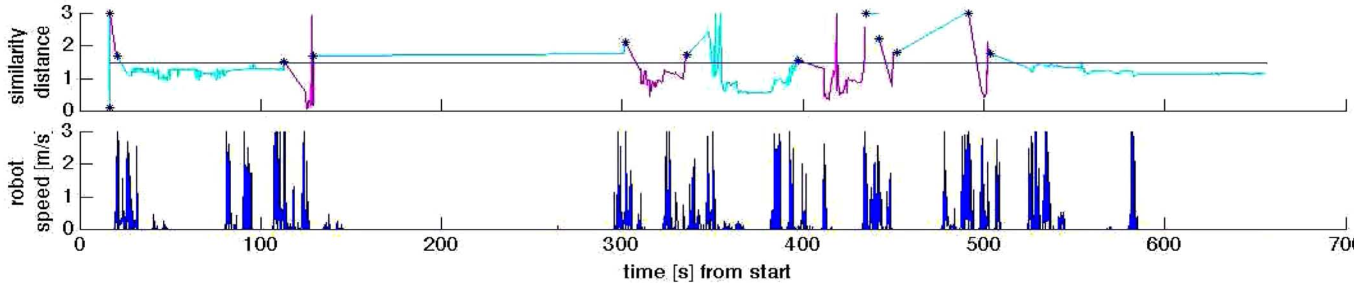


Fig. 4. Illustration of the similarity distance measure for the "BIRON" run as it is changing over time in relation to the robot's speed. A "star" marker and line color switch indicate a hypothesized transition. Extreme values (both for speed and distance measures) are cut for readability of the plot. "Dips" in the distance measure with low speed right before high values can be explained by the robot that moves slowly through doorways, where obviously, the surroundings change.

5) number $nSens$ of ambiguities that were detected in a sensible range (approximately 1–2 m in a standard indoor/domestic environment) from an obvious transition in the environment (e.g., a doorway);
6) number $nSpurious$ of obviously spurious (erroneous) detections of ambiguities (e.g., in the middle of an open area);
7) number $nMiss$ of obviously missed transitions into a structurally different area.

The generation of a new, explicitly specified, *region* was not considered to be a detected change; however, when this specified *region* was obviously left, a detection should have occurred; otherwise, a miss is counted.

In the following sections, we describe the test scenarios and the results according to our evaluation criteria being summarized in Table I as evaluation results for 1) two domestic settings (apartments 1 and 2); 2) test runs in the laboratory (laboratories 1 and 2); and 3) a test with a fully interactive system ("BIRON"—the "**Bi**elefeld **Ro**bot Companio**n**," Group for Applied Informatics, University of Bielefeld, Germany).

1) Domestic Environment, Mapping Subsystem Only: Two different domestic environments were considered: one being a rather small apartment with narrow passages and doorways and the second being a medium-sized flat with partially wide passages and after open spaces. In both the apartments, the living room, a bedroom, and the kitchen were presented to the robot. In the larger apartment as well, two runs were conducted, both being actual guided tours in a user study setup of the "home-tour" scenario.

2) Test Runs in the Laboratory/Office Environment: In the office environment, we evaluated two runs, one of which covered a large part of the corridor of one of our floors and two of the rooms. With the other run, the applicability for loop closing was tested by specifying the "living room" (one large laboratory room) and the connected hallway, where the robot was guided back into the "living room" through a different door than used when leaving it (see Fig. 2).

3) Test With a Fully Integrated Interactive System: To test the applicability of the transition-detection approach, the third implementation, which was integrated into the communication framework of the robot BIRON, an experiment in a laboratory environment that corresponds to a part of an apartment, which includes living room, kitchen, and hallway, was conducted. An interesting aspect to the integrated system was that for this experiment, no SLAM method was available to provide the corrected pose estimations that were usually assumed. We decided to use the experiment to investigate how far the purely topological mapping subsystem, including the transition detection, would be capable of representing the environment in a way that allowed meaningful interaction with a user, relying only on the feature-based representations.

Fig. 3 illustrates the guided tour with BIRON through the laboratory environment, which was conducted by a researcher that acts as "user." Since the pose-estimation error mostly depends on rotations of the robot platform [see the uncorrected illustration in Fig. 3(c) and (d)], the accumulated error was kept on a level that allowed to hypothesize the "hallway" correctly when it was reentered, since no significant turning movements "on the spot" had been made after its specification. The relatively high ratio of spurious detections for this experiment can be explained with the user that distorts the robot's "view" significantly—at this point, there was no feedback from the tracking module or interaction monitoring to the transition detection. Fig. 4 illustrates the similarity measures over time for the run by applying the same conditions for the detection of a transition as in the original run, i.e., a "new current *region* representation" is assumed (in this *post hoc* run, no confirmation question was actually posed) when a significant change ($d > 1.5$) is observed for more than three data cycles, and the robot is at least 1 m away from the point where the previous representation was accepted as the current one.

4) Summary: The results from the seven evaluated runs show that most of the obvious transitions in our test environments are detected rather reliably. As "obvious transitions," we consider door passages, junctions of hallways (available in the office settings), and hallways opening into a room (available in the two domestic datasets for the "medium-size apartment"). Most failures of the approach have to be counted regarding "false alarms." However, since we assume the user to assist the system, we consider this type of failure less critical than "false negatives." These occurred significantly less often and only in one "apartment" setting. Adaptive setting of the threshold values to the type of environment ("narrow apartment" versus "spacious laboratory") or the application of a more sophisticated change detection filter can be an option to cover such cases more appropriately. A number of spurious detections in one of the domestic settings can be explained with the user being very close to the robot (due to the interactive scenario) and, thus, covers larger parts of the laser-range finder's "field of view," Such spurious detections can obviously be avoided by increasing the number of data cycles that a change needs to last before a transition is hypothesized. This was done for the laboratory runs, where it seemed to have immediate impact in the sense that spurious detections did not occur frequently and is still being able to detect significant changes satisfyingly.

The second laboratory run (see Fig. 2) showed the advantage of using a feature-based representation both for the detection of transitions and the representation of *regions* so that it is not necessary to travel back to a previously observed path to hypothesize a loop closure, which would presumably be the case with a door detector in combination with pure metrical SLAM.

The aim of the integration of the mapping subsystem with the fully interactive framework on BIRON was to see if a meaningful interaction in and about the surroundings can be achieved with the proposed models and used representations. For this integrated system, it was decided to limit the functionality of the mapping subsystem to the rather basic situations, which were described earlier, i.e., the specification of *regions* and the detection of transitions together with the resulting requests for confirmation. Within this limited context, the question mentioned earlier can be positively answered, at least for the discussed environment. The robot detected all expected transitions and produced only a very limited amount of surprising questions.

As an overall result, we consider the approach for the separation of *regions* and detection of transitions between them as a useful tool to support the acquisition of a usable and understandable representation of an arbitrary indoor environment, which was suitable for a meaningful communication with the user, even without an underlying correction of the robot's pose estimation as it could be demonstrated with the last experiment presented.

V. Conclusion and Future Work

In this paper, we presented our approach to the separation of *regions* (one central spatial concept in our framework for HAM) in the environment and the detection of transitions between them. We assume an interactive guided tour in which a human user presents and explains a known environment to our robot.

We tested our implementation of the HAM framework, particularly, its subsystem for topological-graph building (region representation and transition detection), with ofline experiments as well as with a full interactive setup (graphical interface and tracking system included) and, in one case, integrated in a fully interactive framework, including dialogue abilities, in different online runs. With these experiments, the applicability of our method for *region* segmentation and transition detection could be confirmed. No prior knowledge of spatial categories is needed to generate a topological-graph representation of an arbitrary indoor environment that reflects the human user's conceptual understanding of the surroundings. This makes the approach very flexible. Our tests showed sufficiently good results in both office (or laboratory) and domestic environments. Other mentioned approaches aim to label an environment with spatial categories [9], [10], while our method can rather be considered to deal with transitions between any type of spatial categories. This makes it more flexible in situations where the spatial category is difficult to determine, even for a human user. Thus, we consider our approach as a fast and easy-to-apply complement to such categorizing methods.

So far, the corrections made by the user are not persistent in the system—the robot simply "forgets" that it has already asked the user about a particular detected transition. Thus, it has to be investigated how the topological-graph structure and respective representations of involved *regions* have to be changed persistently.

On a more detailed level, it would seem natural to investigate a more adaptive method to decide if, in fact, a transition has occurred. This should make the method better suitable to different types of environments (generally narrow or more open) without the need to manually adjust the parameters.

Acknowledgment

The authors would like to thank anonymous reviewers for their efforts, in particular, for finding a very old typo, and the Applied Informatics Group in Bielefeld for their cooperation.

References

[1] E. Topp, "Human-robot interaction and mapping with a service robot: Human augmented mapping," Doctoral dissertation, School Comput. Sci. Commun., Royal Inst. Technol., Stockholm, Sweden, 2008.

[2] I. Nourbakhsh, A. Powers, and S. Birchfield, "Dervish an office-navigating robot," *AI Mag.*, vol. 16, no. 2, pp. 53–60, 1995.

[3] P. Beeson, N. Jong, and B. Kuipers, "Towards autonomous topological place detection using the extended voronoi graph," presented at the IEEE Int. Conf. Robot. Autom., Barcelona, Spain, 2005.

[4] B. Kuipers and Y.-T. Byun, "A robust qualitative method for spatial learning in unknown environments," presented at the Nat. Conf. Artif. Intell., Los Altos, CA, 1988.

[5] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (slam): Toward exact localization without explicit localization," *IEEE Trans. Robot. Autom.*, vol. 17, no. 2, pp. 125–137, Apr. 2001.

[6] A. Tapus, G. Ramel, L. Dobler, and R. Siegwart, "Topology learning and place recognition using bayesian programming for mobile robot navigation," presented at the IEEE/RSJ Int. Conf. Intell. Robots Syst., Sendai, Japan, 2004.

[7] B. Kröse, "An efficient representation of the robot's environment," presented at the Int. Conf. Intell. Auton. Syst., Venice, Italy, 2000.

[8] P. Buschka and A. Saffiotti, "A virtual sensor for room detection," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2002, pp. 637–642.

[9] O. Martínez Mozos, C. Stachniss, and W. Burgard, "Supervised learning of places from range data using adaboost," presented at the IEEE Int. Conf. Robot. Autom., Barcelona, Spain, 2005.

[10] S. Friedman, H. Pasula, and D. Fox, "Voronoi random fields: Extracting the topological structure of indoor environments via place labeling," presented at the Int. Joint Conf. Artif. Intell., Hyderabad, India, 2007.

[11] H. Zender, P. Jensfelt, Ó. Mozos, G.-J. Kruijff, and W. Burgard, "An integrated robotic system for spatial understanding and situated interaction in indoor environments," presented at the Nat. Conf. Artif. Intell., Vancouver, BC, Canada, 2007.

[12] Z. Zivkovic, O. Booij, B. Kröse, E. Topp, and H. Christensen, "From sensors to human spatial concepts: An annotated data set," *IEEE Trans. Robot.*, vol. 24, no. 2, pp. 501–505, Apr. 2008.

Task Selection for Control of Active-Vision Systems

Yasushi Iwatani

***Abstract*—This paper discusses the problem of task selection in active-vision systems. It is shown that optimization of motion perceptibility does not work as the primary task. On the other hand, it is demonstrated that optimization of motion perceptibility under target tracking produces reasonable camera motion. The perceptibility measure is induced by a certain Jacobian matrix and not by the interaction matrix. The interaction matrix does not produce cooperative behavior of multiple active cameras.**

***Index Terms*—Active-vision systems, motion perceptibility, visual servoing.**

I. Introduction

An active-vision system is a robotic system with a camera (or cameras) mounted on the robot end-effector. The position and/or orientation

Manuscript received November 9, 2009; revised March 30, 2010; accepted May 9, 2010. Date of publication June 10, 2010; date of current version August 10, 2010. This paper was recommended for publication by Associate Editor E. Malis and Editor G. Oriolo upon evaluation of the reviewers' comments. This work was supported in part by KAKENHI under Grant 22760309. This paper was presented in part at the 2009 IEEE/Robotics Society of Japan International Conference on Intelligent Robots and Systems, St. Louis, MO, October 14, 2009.

The author is with the Graduate School of Science and Technology, Hirosaki University, Hirosaki 036-8561, Japan (e-mail: iwatani@cc.hirosaki-u.ac.jp).

Digital Object Identifier 10.1109/TRO.2010.2050517