



PLACE CLASSIFICATION USING DEMPSTER-SHAFER THEORY

Barbara SIEMIĄTKOWSKA, Bogdan HARASYMOWICZ-BOGGIO *

Abstract. The paper presents a novel place labeling method. It is assumed that an indoor mobile robot equipped with a camera or RGB-D sensor ambulates an indoor environment. The places visited by the robot are classified based on objects which have been recognized. Each object (or set of objects) *votes* for a set of room classes. Data aggregation is performed using Dempster-Shafer theory (DST), which can be regarded as a generalization of the Bayesian theory. The possibility of taking into account the uncertainty of data is the main advantage of the described method. The classic Dempster's rule of data aggregation has been criticized because it can lead to non-intuitive results. Many alternative methods have been proposed and several were tested during our experiments. Most place classification methods assume a closed world model, i.e. a test sample is assigned to the most probable class even if its corresponding probability is very small. An advantage of our system is the intrinsic capability of giving *unknown* class as an answer in such situations, which can be used by the robot to take appropriate actions.

Keywords: Dempster-Shafer theory, place classification

1. Introduction

Mobile robots leave scientific laboratories and start to enter real life applications, so much attention is paid to social interactive robotics. However, communication between a robot and a human is not possible without semantic information about the environment. The ability to recognize and classify objects, places and events seems to be fundamental. Knowing its location, the robot can adapt to different situations. It can also perform tasks given in natural language, such as: *take this to the kitchen*. This is one of the main reasons why in this paper we address the problem of semantic localization.

*Warsaw University of Technology, Faculty of Mechatronics, Warsaw, Poland
{siema@mchtr.pw.edu.pl, mysticdrow@gmail.com}

A human indoor environment is divided into areas with different functionalities: living rooms, corridors, kitchens. Many authors considered the problem of adding semantic information to places on the basis of data obtained by the robot's sensors.

1.1. Related works

In [5, 7] methods of place labeling based on laser or ultrasonic scanners are presented but most of the algorithms require richer information given by cameras or RGB-D sensors. Usually the scene is described by a set of features, the authors compute various descriptors on laser range scans [11, 19] or images [21]. We can distinguish between methods which use global features [21, 18, 5, 7] or local descriptors [13] and algorithms which use deep features [31, 30]. Koenig and Simmons [14] apply a pre-programmed routine to detect doorways from range data. In [2] the line features are used in order to detect corridors and doorways. Malik et al. [21] propose that a simple texture analysis of the image can provide a useful cue towards rapid scene identification. Their model learns texture features across scene categories and then uses this knowledge to identify new scenes. In [25] an appearance-based algorithm is presented, it combines local features with support vector machines through an ad-hoc kernel. Shuai Yang et al. [27] present a place recognition scheme using the vocabulary tree [17] of multiple feature types. In the article [8] the Voronoi graph is extracted from an occupancy grid map generated with a laser range-finder, and then represent each point on the Voronoi graph as a node of a conditional random field. The labels are attached to each node. The labels provide a segmentation of an environment, with different segments corresponding to rooms, hallways, or doorways. Shrihari Vasudevan and Roland Siegwart [26] present a hierarchical probabilistic concept-oriented representation of space based on objects. This approach is based on learning from examples, clustering and the use of Bayesian network classifiers. In the paper [15] the authors present a novel top-down, multi- step visual place recognition system. The overall matching process is inspired by the increasingly selective and tolerant processing stream in the human brain.

Different learning techniques are used in order to identify different places: neural networks [32, 6], EM algorithm [3], AdaBoost [16]. In [24] an algorithm which uses Hidden Markov Models for learning places from image data is presented. In [28] a naive Bayes classification method is described.

We believe that most indoor places can be classified based on objects located within them. Object recognition is much easier than a few years ago, since many highly effective techniques have been introduced [32, 1]. The place classification method proposed in this paper, as opposed to purely geometrical models, holds promise of higher flexibility.

In order to aggregate the information we apply Dempster-Shafer theory (DST). Although our approach is supervised (the resulting labels correspond to user-defined classes), in the case of insufficient data (for example the robot sees a wall and a part of the floor) the answer *unknown* is given by the system. The proposed method has been tested using MIT dataset of real environments [32].

The paper is organized as follows: Section 2 reviews the state-of-the-art in Dempster Shafer theory. In section 3 we describe the idea behind the proposed method of place classification. In section 4 the teaching technique is described. Finally, section 5 presents the experimental results.

2. Dempster-Shafer theory

The information about the environment of a mobile robot is obtained by its on board sensors and is usually uncertain – some data can be missing, imprecise or inconsistent [23]. In [12] the following types of uncertainty have been distinguished: fuzziness (lack of strict definition, ambiguity), discord (disagreement in choosing among alternatives) and non specificity. In order to reduce uncertainty, probability theory is usually applied. But probabilities must be assigned even if no information is available, therefore that method is not capable of capturing epistemic uncertainty and it does not provide any means to distinguish ignorance and evidence conflicts. A proper measure of ignorance is very important in robotics because it allows us to verify whether the available knowledge is sufficient to make a decision. In practical applications we also need a quantitative measure of evidence conflict. A high degree of conflict usually shows that part of the evidence is unreliable or something unexpected occurs in the environment. Dempster-Shafer theory of evidence (DST) [4, 22] was designed in order to deal with uncertainty and ignorance. It also allows us to get information about the conflict level.

In DST knowledge is encoded by assigning masses m to subsets of the set T (power set) of all possible hypotheses.

$$m : 2^T \rightarrow [0, 1]. \quad (1)$$

If $T = \{A, B\}$ then $2^T = \{\emptyset, A, B, T\}$.

The mass function fulfills the following requirements:

$$\sum_{A_i \in 2^T} m(A_i) = 1, \quad m(\emptyset) = 0, \quad (2)$$

where \emptyset denotes the empty set. A belief measure is given by the function $bel : 2^T \rightarrow [0, 1]$,

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B). \quad (3)$$

A plausibility measure is given by the function pl :

$$pl : 2^T \rightarrow [0, 1] \quad pl(A) = \sum_{B \cap A \neq \emptyset} m(B). \quad (4)$$

If T is a set of all hypotheses then $m(T)$ represents the level of uncertainty. Belief and disbelief in a hypothesis need not to sum to 1 and both values can even be equal to 0, which would mean that there is no evidence for or against the hypothesis. The process of data aggregation according to DST consists of the following steps:

- Degrees of belief for particular hypotheses are obtained on the basis of facts, which are treated as information sources, for example a chair supports the set of hypotheses: $\{\text{classroom or living room}\}$. In comparison to Bayesian theory we are not forced to distribute the masses between all classes.
- Dempster's rule is applied in order to combine degrees of belief obtained based on different facts.

Dempster's rule of combination for sources $(1, 2 \dots N)$ is described as follows:

$$m_{1,\dots,N}(A) = \frac{\sum_{B_1 \cap \dots \cap B_N = A} m_1(B_1) \cdot \dots \cdot m_N(B_N)}{1 - K}, \quad (5)$$

where $A, B_1, \dots, B_N \subseteq T$, and

$$K = \sum_{B_1 \cap \dots \cap B_N = \emptyset} m_1(B_1) \cdot m_2(B_2) \cdot \dots \cdot m_N(B_N) \quad (6)$$

K represents the conflict between evidences, $1 - K$ is the normalization factor.

Shafer [4] introduced the discounting operation to handle the case when the source of some piece of evidence is lacking credibility. If $\alpha \leq 1$ is the credibility level of the source then it also becomes the credibility level of the piece of evidence, resulting in a new mass m^α defined by:

$$\begin{aligned} m^\alpha(A) &= \alpha m(A) \text{ for any } A \subset T, \\ m^\alpha(T) &= 1 - \alpha + \alpha m(T). \end{aligned} \quad (7)$$

The D-S method of aggregation gives counterintuitive results in the case of strong conflict between evidences. If $K \approx 1$ the conflicting beliefs management problem occurs. In order to solve the problem, different methods of aggregation are introduced [22]. Smets', Yager and the weighted average operator are examples of such methods – these methods were tested during our experiments. The Smets' rule of combination is the non-normalized version of Dempster rule.

For two sources of information it is given by:

$$m(\emptyset) = K = \sum_{A_1, A_2 \in 2^T, A_1 \cap A_2 = \emptyset} m_1(A_1) m_2(A_2), \quad (8)$$

and $\forall A \neq \emptyset, A \in 2^T$

$$m(A) = \sum_{A_1, A_2 \in 2^T, A_1 \cap A_2 = A} m_1(A_1) m_2(A_2) \quad (9)$$

In Yager's rule of aggregation [22] the measure of conflict (K) is added to uncertainty.

$$m(T) = m_1(T) \cdot m_2(T) + K. \quad (10)$$

$\forall A \neq T, A \in 2^T$ $m(A)$ is computed according to the formula 9.

Smets and Yager rules are cases of general weighted operator (WO). The combination rule using WO [22] consists of two steps:

- computation of total conflicting mass K (eq. 6),
- reallocation of the conflicting mass K according to:

$$m(\emptyset) = w(\emptyset) \cdot K, \quad (11)$$

and $\forall A \neq \emptyset$

$$m(A) = \sum_{\substack{A_1, A_2 \in 2^T \\ A_1 \cap A_2 = A}} m_1(A_1)m_2(A_2) + w(A) \cdot K. \quad (12)$$

In the weighted average operator (WAO) the $w(A)$ is computed as follows:

$$\begin{aligned} w(\emptyset) &= 0, \\ w(A) &= \frac{1}{N} \sum_{i=1}^N m_i(A), \end{aligned} \quad (13)$$

where N - is the number of independent sources to combine.

Different methods of aggregation have strong influence to classification result.

3. The system architecture

We believe that most indoor places can be recognized based on objects, so our place recognition algorithm is strictly connected with object identification. We have proposed a number of methods which allow object recognition [9, 10] on the basis on RGB-D sensors. Figure 1 presents sample scenes and objects which have been recognized by our algorithm.

The goal of our approach is to assign a place label(one of labels: A_1, \dots, A_M) from a set of M classes to an observation ($\{o_i, \dots, o_l\}$ a list of objects).

The algorithm consists of two stages:

- supervised learning - a model is built based on data set,
- classification - one of the rules of aggregation is used and winner take all method is applied.

3.1. Supervised learning

Training data set consists of N records:

$$\begin{aligned} &(\{o_i, \dots, o_l\}, A_1), \\ &\dots, \\ &(\{o_k, \dots, o_j\}, A_M), \end{aligned} \quad (14)$$

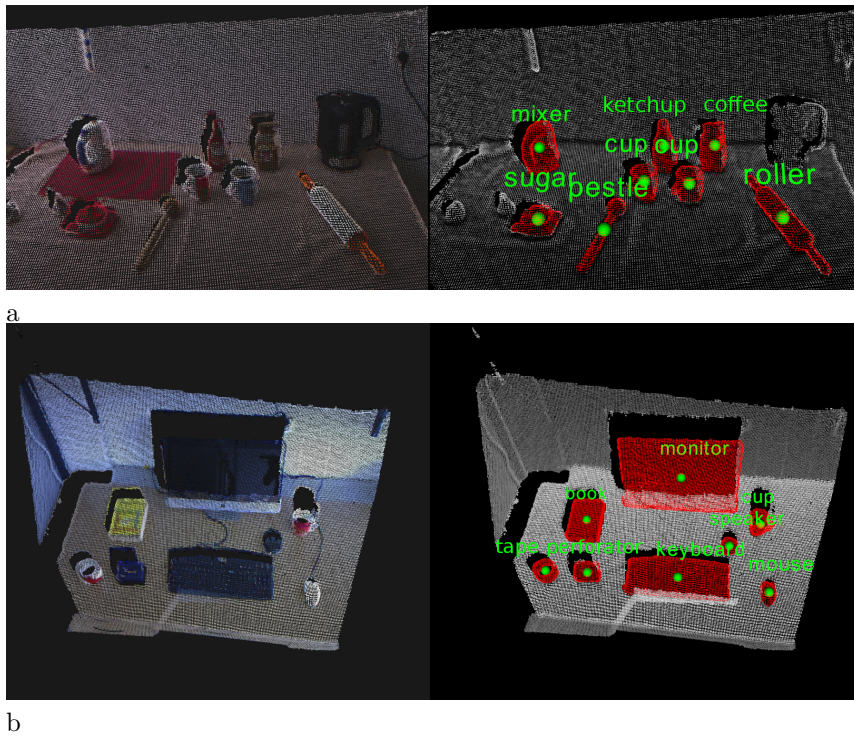


Figure 1: Sample objects a – kitchen, b – office

where: $\{o_k, \dots, o_j\}$ - the list of names of objects which have been found in place A_i . The training record for the room presented in fig. 1 can be described as follows:

($\{\text{mixer, cup, cup, sugar, roller}\}$, *kitchen*).

Usually M (number of classes) $\ll N$ (number of records).

When the training set is given $\forall i = 1, \dots, K$ and $\forall j = 1, \dots, M$ the values of the masses $m_i(A_j)$ are computed. Value $m_i(A_j)$ describes how much the object o_i supports the hypothesis (class) A_j .

In classic Dempster-Shafer method the masses are defined for all subsets of 2^T . In our approach the set of hypotheses has been reduced and consists of $\{A_1, A_2, \dots, A_M, T\}$ where A_1, \dots, A_M - classes of places, M - number of different classes, $T = \{A_1 \cup A_2 \dots \cup A_M\}$ represents uncertainty. In this paper T is described as A_{M+1} .

The masses are computed as follows:

- conditional probability $p(A_k | o_i)$ is defined,
- the masses are computed on the basis of $p(A_k | o_i)$,
- the masses are updated.

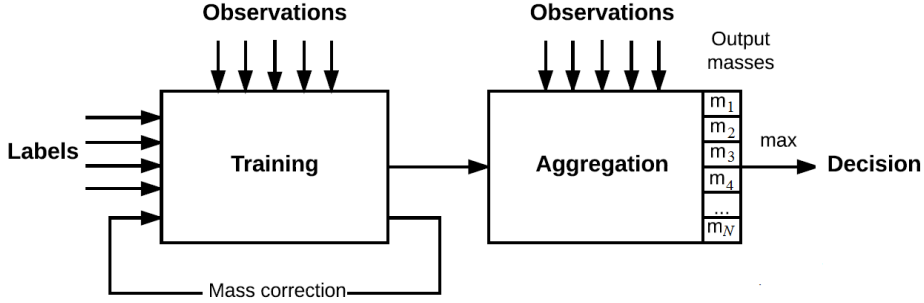


Figure 2: The place classification algorithm

Conditional probability $p(A_k | o_i)$ is defined by:

$$p(A_k | o_i) = \frac{n_k}{n_i}, \quad (15)$$

where: n_k - the number which describes how many times the object o_i has been seen in class A_k , n_i - the number which describes how many times the object o_i has been seen.

There are not any general rules on how to compute uncertainty in DS theory. Usually the mass of uncertainty is given by an expert. In computer science entropy is the common measure of uncertainty. If the probability of an event equals 1 that means that the event is certain, and the entropy equals 0.

The value of entropy computed for o_j is defined by:

$$e_j = - \sum_{i=1}^M p(A_i | o_j) \log_2(p(A_i | o_j)), \quad (16)$$

where M is the number of different places.

If the object o_j has been seen in one place then $e_j = 0$, if it exists in many classes then $e_j \gg 0$.

The credibility level α_j (eq. 7) has to be near 1, if entropy equals 0 and $\alpha \approx 0$ if $\frac{e_k}{M} \approx 1$. In our approach α_j is computed as follows:

$$\alpha_j = u \cdot (1.0 - \frac{e_k}{M}), \quad (17)$$

where u is a scaling factor. In our method $u = 0.9$, the parameter u is computed during learning process. This value minimizes the percentage of misclassified samples in training set.

The masses attached to the hypotheses are computed according the formulae:

$$\forall i = 1, \dots, K \quad \forall A_k \neq T \quad m_i(A_k) = \alpha_i \cdot p(A_k | o_i), \quad (18)$$

$$m_i(T) = 1.0 - \sum_{j=1}^M m_i(A_j), \quad (19)$$

where i – the index of the object.

3.2. The method of updating the masses

If the masses have been computed the process of classification is performed on the learning set. One of the methods described in section 2 can be used for data aggregation. The winner-take-all assignment is performed, the winning class A_k is defined as follows:

$$m(A_k) = \max_{A_j} m(A_j), \quad j = 1, \dots, M + 1, \quad (20)$$

where $A_{M+1} = T$. If the results of classification are correct then nothing is done. In case of failure, the masses are modified. If A_k is the winning class and A_t is the correct class, the list of objects which are used during classification is analysed (eq. 14). We look for an object o_i such that:

$$m_i(A_k) = \max_{o_j} m_j(A_k), \quad j = 1, \dots, L, \quad (21)$$

where L is the number of objects in the list.

The masses $m_i(A_k)$, and $m_i(A_t)$ are modified according to the formulae:

$$\begin{aligned} m_i(A_k) &= m_i(A_k) - w \cdot (m(A_k) - m(A_t)), \\ m_i(A_t) &= m_i(A_t) + w \cdot (m(A_k) - m(A_t)), \end{aligned} \quad (22)$$

where w is a learning factor. The described procedure is continued until the number of false classifications does not decrease.

4. Experiments

The proposed method has been tested using an MIT real environment dataset. The MIT dataset [20] consists of 67 indoor place categories, and a total of 15620 images – there are at least 100 images per category. The images were segmented and annotated with the objects that they contain. In our system the annotations are used. The dataset used in our experiments contained following classes: 1 – bathroom, 2 – bedroom, 3 – children room, 4 – closet, 5 – corridor, 6 – dining room, 7 – garage, 8 – greenhouse, 9 – living room. 10th class represents uncertainty. About 260 different objects were detected in the images.

The dataset was equally divided into two subset - learning set and testing set.

Figure 3 presents the masses computed for two different objects (before mass updating): a wall (fig. 3a) and shower (fig. 3b).

In the case of the wall the mass of uncertainty (class no. 10) is much bigger than the masses of other classes. In the case of the shower the uncertainty is small, while the *bathroom* mass is high.

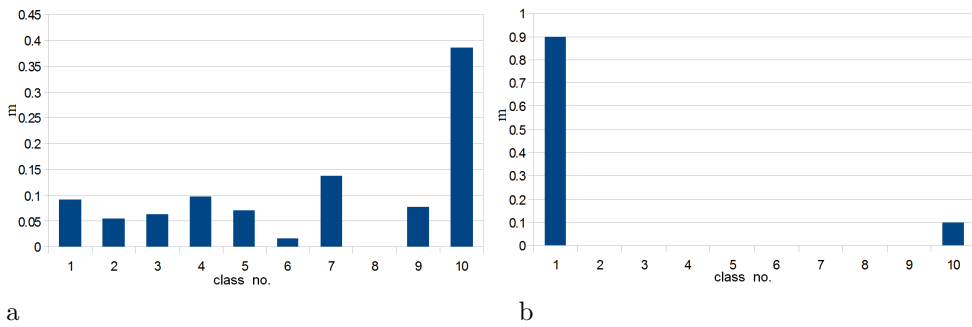


Figure 3: Masses computed for: a – a wall, b – a shower

Tab. 7–4 are confusion matrices. The results are given in percentage terms.

Tab. 1 presents the results of classification when Smets' method of aggregation is applied, in 2 and 3 the results of applying Yager and WOA rules are presented. The experiments have been done without mass updating.

Table 1: Confusion matrix for Smets' rule of aggregation

no.	1	2	3	4	5	6	7	8	9	10
1	96	2	0	0	0	0	0	0	2	0
2	2	82	2	4	2	0	0	0	8	0
3	0	0	94	0	0	0	0	0	0	6
4	0	0	0	90	6	0	0	0	0	4
5	4	2	0	2	48	2	2	0	10	30
6	0	0	0	4	0	68	0	0	26	2
7	0	0	0	0	0	0	98	0	0	2
8	0	0	0	0	0	0	0	97	0	3
9	0	0	0	0	0	4	0	0	94	2

Table 2: Confusion matrix for Yager rule of aggregation

no.	1	2	3	4	5	6	7	8	9	10
1	50	2	0	2	0	0	0	0	0	46
2	0	6	0	0	0	0	0	0	0	94
3	0	0	2	0	0	0	0	0	0	98
4	0	0	0	42	0	0	0	0	0	58
5	2	0	0	0	18	0	4	0	2	74
6	0	0	0	4	0	20	0	0	8	68
7	0	0	0	0	0	0	21	0	0	79
8	0	0	0	0	0	0	2	65	0	33
9	0	2	0	0	0	0	0	0	6	92

Smets' and WOA rules give similar results, but aggregation using WOA is more

Table 3: Confusion matrix for WAO rule of aggregation

no.	1	2	3	4	5	6	7	8	9	10
1	96	2	0	0	0	0	0	0	2	0
2	2	82	2	4	2	0	0	0	8	0
3	0	0	96	0	0	0	0	0	0	4
4	0	0	0	98	0	0	0	0	0	2
5	4	2	0	2	42	2	12	0	10	26
6	0	0	0	4	0	68	0	0	26	2
7	0	0	0	0	0	0	98	0	0	2
8	0	0	0	0	0	0	0	98	0	2
9	0	0	0	0	0	4	0	0	94	2

time consuming than Smets’ method. If the Yager method is applied the number of places that are misclassified is small but the answer *unknown* is given too often. We decided to use Smets’ rule of aggregation in our algorithm.

Tab. 4-5 present the results of classification when Smets’ method of aggregation is applied, and masses are updated. We consider two approaches. In the first one we do not distinguish between uncertainty (A_{M+1}) and labels of places ($A_i, i = 1, \dots, M$) during process of mass updating (tab. 4). In the experiment presented in Tab. 5 in the case of uncertainty the masses have not been updated. Fig. 4a and 4b represent the masses computed for the object *floor* before and after the process of mass updating.

Table 4: Place classification - updated masses

	1	2	3	4	5	6	7	8	9	10
1	99	0	0	0	0	0	0	0	0	1
2	0	93	2	0	0	0	0	0	4	1
3	0	0	99	0	0	0	0	0	0	1
4	0	0	0	98	0	0	0	0	0	2
5	0	1	0	0	78	5	2	0	2	12
6	0	0	1	0	0	93	0	0	2	4
7	0	0	0	0	0	0	99	0	0	1
8	0	0	0	0	0	0	0	98	0	2
9	0	1	0	0	1	2	0	0	94	2

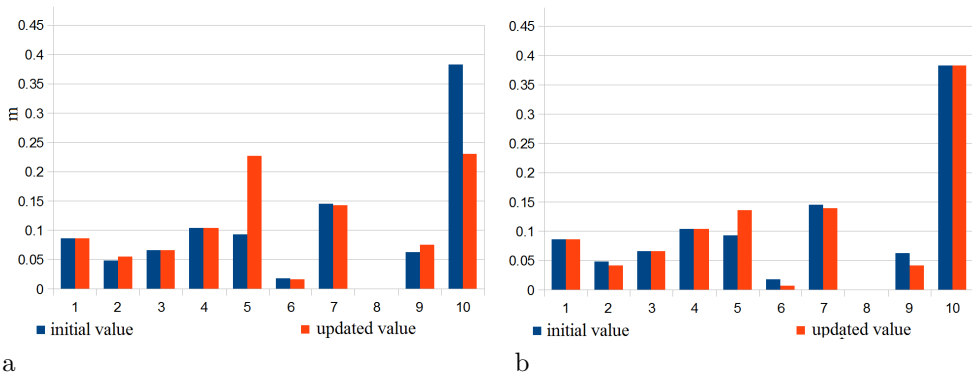
We can noticed that the number of misclassification in the first case is smaller than the in the second one.

In [31] the comparison of different methods of place classification is presented. Tab. 6 presents classification accuracy for MIT database (67 classes).

Our method of places classification was compared with two different method: naive Bayes and nearest neighbors. In all experiments the same data sets (training and testing) were used. Tab. 7 presents the result of places classification using a naive Bayes method and MIT database. We can notice that the naive Bayes classifier did not recognize *living room* correctly.

Table 5: Confusion matrix for updated masses, uncertainty has not been reduced

	1	2	3	4	5	6	7	8	9	10
1	99	0	0	0	0	0	0	0	0	1
2	0	93	2	0	0	0	0	0	4	1
3	0	0	99	0	0	0	0	0	0	1
4	0	0	0	98	0	0	0	0	0	2
5	0	1	0	0	67	4	2	0	2	24
6	0	0	1	0	0	93	0	0	2	4
7	0	0	0	0	0	0	99	0	0	1
8	0	0	0	0	0	0	0	99	0	1
9	0	1	0	0	0	2	0	0	95	2

**Figure 4:** The masses: a – uncertainty is not reduced during learning process, b – uncertainty is reduced during learning process**Table 6:** Classification accuracy

place classification method	accuracy
Place-CNN feature+SVM	67%
Image Net-CNN features+SVM	56%
Hybrid feature+SVM	76%

Table 7: Confusion matrix for naive Bayes

no.	1	2	3	4	5	6	7	8	9
1	94	4	0	0	0	2	0	0	0
2	0	90	0	2	0	8	0	0	0
3	3	0	97	0	0	0	0	0	0
4	2	0	0	98	0	0	0	0	0
5	12	0	0	2	52	12	14	0	8
6	10	0	4	0	0	86	0	0	0
7	0	0	0	0	0	0	100	0	0
8	0	0	0	0	7	0	0	93	0
9	0	30	0	0	2	68	0	0	0

Table 8: Confusion matrix for NN

no.	1	2	3	4	5	6	7	8	9
1	82	2	0	0	12	4	0	0	0
2	4	68	0	0	24	4	0	0	0
3	0	50	10	0	30	0	10	0	0
4	0	0	0	44	56	0	0	0	0
5	0	0	0	0	100	0	0	0	0
6	0	0	0	0	40	58	2	0	0
7	0	0	0	0	90	0	10	0	0
8	0	0	0	0	22	0	0	78	0
9	2	12	0	0	26	16	6	0	38

Tab. 8 presents the result of places classification using a nearest neighbor method (NN) and MIT database (9 classes). We can notice that for the class *corridor* false negative error equals 0% but false positive error is very large.

Table 9: Experimental results - classification accuracy

place classification method	accuracy
Naive Bayes	79%
NN	54%
Smets method	85%
Smets method, updated masses	94%

Table 9 summarizes the experimental results. It shows that our method gives much better results than Naive Bayes and NN classifiers. In order to compare our approach to SVM or neural networks, we would need a much larger dataset.

DST-based classifier also has the following advantages:

- the results are easy to interpret,
- fast training speed,
- performs well with small number of observation,
- we can combine different sources of information (sec. 5).

5. Future works - combining fuzzy rules and DST

In many situations the number of objects is a very useful source of information. Using the number of occurrences of various objects rather than just their individual occurrences is a more informative method of distinguishing between classes of places. For example if the robot is in a corridor it sees a small number of different objects. This sentence can be rewritten as follows: If *a small number* of different objects is seen then I might be in *a corridor*. A *small number* of different objects is the source of information and *might be in the corridor* is the consequence. A *small number* is a fuzzy set, its membership function depends on the exact number of objects (Fig. 5b). Fuzzy rules have the form *if A then B*, where A is called the premise and B is the consequence of the rule. In [29] is shown that fuzzy rules present tolerance to imprecision and uncertainty.

Based on an information source we can assign masses m to subsets of the set T (power set) of all possible hypotheses. Fig. 5a presents the masses attached to different places if *a small number* of objects has been detected.

The value of the membership function is the credibility level of the information source. Applying eq. 7 a new mass assignment is obtained. Fig. 6 presents the masses computed for different numbers of objects.

If during the learning process the described fuzzy rule is applied, the percentage of correctly recognized corridors increases to 82%.

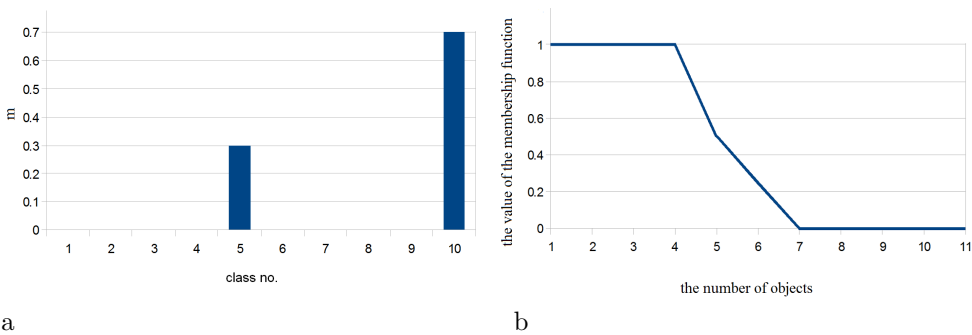


Figure 5: Masses and the membership function

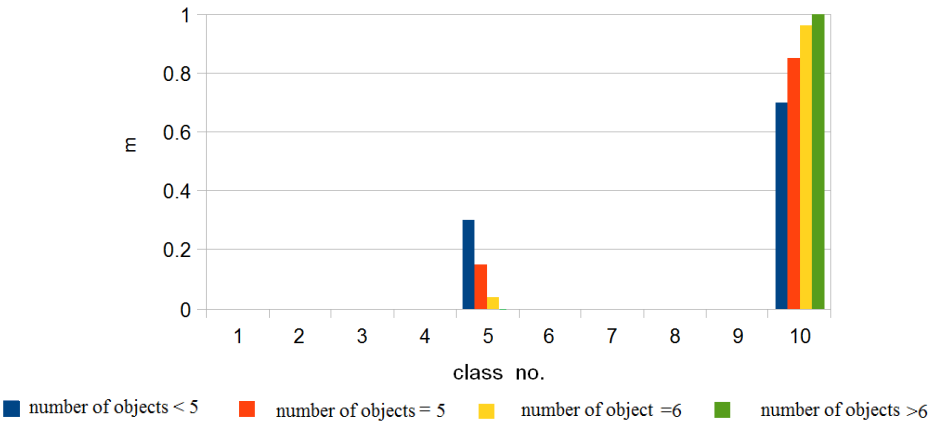


Figure 6: Masses computed for different numbers of objects

6. Conclusion

In this work we have proposed a new approach to place labeling by applying Dempster-Shafer theory. DST can be regarded as a generalization of the Bayesian theory and is able to deal with subjective uncertainty and ignorance. The experiment performed using the MIT public dataset proved the effectiveness of our approach - the learning process is very fast, the uncertainty is taken into account. We can combined data taken from many sources of information. Similarly to [16] the initial degrees of belief of hypotheses can be computed on the basis of history of past states of the robot and transitions between states. In future works we intend to use a wider range of features and objects and focus on real-time performance by exploiting GPU processing power.

References

- [1] <https://www.tensorflow.org/>.
- [2] Althaus P. and Christensen H. I. Behavior coordination in structured environments. *Advanced Robotics*, 17(7):657–674, 2003.
- [3] Anguelov D., Biswas R., Koller D., Limketkai B., Sanner S., and Thrun S. Learning hierarchical object maps of non-stationary environments with mobile robots. In *Proceedings of the 17th Annual Conference on Uncertainty in AI (UAI)*, Edmonton, Canada, 2002.
- [4] Bogler P. L. Shafer-Dempster reasoning with applications to multisensor target identification systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 17:968–977, 1987.
- [5] Buschka P. and Saffiotti A. A virtual sensor for room detection. In *Intelligent Robots and Systems (IROS)*, pages 637–642, 2002.
- [6] Chen Z., Lam O., Jacobson A., and Milford M. Convolutional neural network-based place recognition. *CoRR*, abs/1411.1509, 2014.
- [7] Dellaert F., Fox D., Burgard W., and Thrun S. Monte Carlo localization for mobile robots. In *IEEE Int. Conf. on Robotics & Automation (ICRA)*, 1998.
- [8] Friedman S., Pasula H., and Fox D. Voronoi random fields: Extracting the topological structure of indoor environments via place labeling. In *In Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2007.
- [9] Harasymowicz-Boggio B., Chechlinski L., and Siemiatkowska B. Nature-inspired, parallel object recognition. In Szewczyk R., Zieliski C., and Kaliczyska M., editors, *Progress in Automation, Robotics and Measuring Techniques. Control and Automation. Advances in Intelligent Systems and Computing vol. 350*, pages 53–62. Springer, 2015.

- [10] Harasymowicz-Boggio B., Chechliniski L., and Siemiatkowska B. Significance of features in object recognition using depth sensors. *Optica Applicata*, 45(4):559–571, 2015.
- [11] Himstedt M., Frost J., Hellbach S., Bhme H. J., and Maehle E. Large scale place recognition in 2d lidar scans using geometrical landmark relations. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5030–5035, Sept 2014.
- [12] Jousselme A.-L., Liu C., Grenier D., and Bosse E. Measuring ambiguity in the evidence theory. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 36(5):890–903, Sept 2006.
- [13] Jung H., Mozos O. M., Iwashita Y., and Kurazume R. Local n-ary patterns: a local multi-modal descriptor for place categorization. *Advanced Robotics*, 30(6):402–415, 2016.
- [14] Koenig S. and Simmons R. G. Xavier: A robot navigation architecture based on partially observable markov decision process models. In *Artificial Intelligence Based Mobile Robotics: Case Studies of Successful Robot Systems*, pages 91–122. MIT Press, 1998.
- [15] Milford M., Scheirer W. J., Vig E., Glover A., Baumann O., Mattingley J., and Cox D. D. Condition-invariant, top-down visual place recognition. In *The IEEE International Conference on Robotics and Automation (ICRA)*, June 2014.
- [16] Mozos O. M., Triebel R., Jensfelt P., Rottman A., and Burgard W. Supervised semantic labeling of places using information extracted from sensor data. *Robotics and Autonomous Systems*, 5:392–402, 2007.
- [17] Nistér D. and Stewénus H. Scalable recognition with a vocabulary tree. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 2161–2168, June 2006.
- [18] Oliva A. and Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vision*, 42(3):145–175, May 2001.
- [19] Premevida C. and Faria U., Diego R. and Nunes. Dynamic bayesian network for semantic place classification in mobile robotics. *Autonomous Robots*, 41(5), 2017.
- [20] Quattoni A. and Torralba A. Recognizing indoor scenes. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 413 – 420, 2009.
- [21] Renninger L. W. and Malik J. When is scene identification just texture recognition? *Vision Research*, 44(19):2301–2311, September 2004.
- [22] Smarandache F. and Dezert J. Information fusion based on new proportional conflict redistribution rules. In *Information Fusion, 2005 8th International Conference on*, volume 2, pages 8–pp. IEEE, 2005.

- [23] Teichman A. and Thrun S. Practical object recognition in autonomous driving and beyond. In *ARSO*, pages 35–38, 2011.
- [24] Torralba A. Contextual priming for object detection. *Int. J. Comput. Vision*, 53(2):169–191, July 2003.
- [25] Ullah M. M., Pronobis A., Caputo B., Luo J., Jensfelt P., and Christensen H. I. Towards robust place recognition for robot localization. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA08)*, 2008.
- [26] Vasudevan S. and Siegwart R. Bayesian space conceptualization and place classification for semantic maps in mobile robotics. *Robot. Auton. Syst.*, 56(6):522–537, June 2008.
- [27] Yang S., Mou W., Wang H., and Ge S. S. Place recognition by combining multiple feature types with a modified vocabulary tree. In *2015 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pages 1–6, Nov 2015.
- [28] Yi C., Suh I. H., Lim G. H., and Choi B. Bayesian robot localization using spatial object contexts. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 11-15, 2009, St. Louis, MO, USA*, pages 3467–3473, 2009.
- [29] Zadeh L. A. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [30] Zhou B., Khosla A., Lapedriza À., Oliva A., and Torralba A. Learning deep features for discriminative localization. *CoRR*, abs/1512.04150, 2015.
- [31] Zhou B., Lapedriza A., Xiao J., Torralba A., and Oliva A. Learning deep features for scene recognition using places database. In *Proceedings of the 27th International Conference on Neural Information Processing Systems, NIPS’14*, pages 487–495, Cambridge, MA, USA, 2014. MIT Press.
- [32] Zhou B., Lapedriza A., Xiao J., Torralba A., and Oliva A. Learning deep features for scene recognition using places database. In *NIPS*, 2014.

This is an extended version of the paper presented at the 14th National Conference on Robotics (KKR 2016), Polanica Zdrój, Poland, September 14-18, 2016