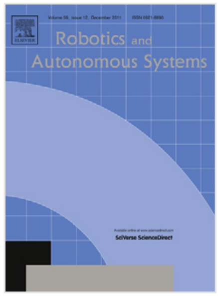
Accepted Manuscript



Novelty detection and segmentation based on Gaussian mixture models: A case study in 3D robotic laser mapping

Paulo Drews Jr., Pedro Núñez, Rui P. Rocha, Mario Campos, Jorge Dias

PII: S0921-8890(13)00111-5
DOI: http://dx.doi.org/10.1016/j.robot.2013.06.004
Reference: ROBOT 2147

To appear in: *Robotics and Autonomous Systems*

Received date: 28 February 2012
Revised date: 25 July 2012
Accepted date: 24 June 2013

Please cite this article as: P. Drews Jr., P. Núñez, R.P. Rocha, M. Campos, J. Dias, Novelty detection and segmentation based on Gaussian mixture models: A case study in 3D robotic laser mapping, *Robotics and Autonomous Systems* (2013), http://dx.doi.org/10.1016/j.robot.2013.06.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

***Research Highlights**

- Framework to detect and segment changes in large data sets, including robotic maps acquired by 3D sensors.

- New method based on Gaussian Mixture Models (GMM).

- Two alternative criteria for detecting changes in the GMM space are compared.

- The proposed method is compared with state-of-art methods using real datasets.

ACCEPTED MANUSCRIPT

*Manuscript
Click here to view linked References
ACCEPTED MANUSCRIPT

Novelty Detection and Segmentation based on Gaussian Mixture Models: a Case Study in 3D Robotic Laser Mapping

Paulo Drews Jr[a], Pedro Núñez[b], Rui P. Rocha[c], Mario Campos[d], Jorge Dias[c]

[a] *Center of Computational Sciences, Federal Univ. of Rio Grande (FURG), Rio Grande-RS, Brazil, Tel.: +55-53-3233-6654, paulodrews@furg.br*
[b] *Dpto. Tecnologa de los Computadores y las Comunicaciones, Univ. de Extremadura, Cacéres, Spain, Tel.: +34-927-25-7443, pnuntru@unex.es*
[c] *Instituto de Sistemas e Robótica, Dep. of Electrical and Computer Engineering, Univ. of Coimbra, Coimbra, Portugal, Tel.: +351-239-796-256, rprocha@isr.uc.pt, jorge@isr.uc.pt*
[d] *Dep. of Computer Science, Federal Univ. of Minas Gerais (UFMG), Belo Horizonte-MG, Brazil, Tel.: +55-31-3409-5860, mario@dcc.ufmg.br*

Abstract

This article proposes a framework to detect and segment changes in robotics datasets, using 3D robotic mapping as a case study. The problem is very relevant in several application domains, not necessarily related with mobile robotics, including security, health, industry and military applications. The aim is to identify significant changes by comparing current data with previous data provided by sensors. This feature is extremely challenging because large amounts of noisy data must be processed in a feasible way. The proposed framework deals with novelty detection and segmentation in robotic maps using clusters provided by Gaussian Mixture Models (GMMs). GMMs provides a feature space that enables data compression and effective processing. Two alternative criteria to detect changes in the GMM space are compared: a greedy technique based on the Earth Mover's Distance (EMD); and a structural matching algorithm that fulfills both absolute (global matching) and relative constraints (structural matching). The proposed framework is evaluated with real robotic datasets and compared with other methods known from literature. With this purpose, 3D mapping experiments are carried out with both simulated data and real data from a mobile robot equipped with a 3D range sensor.

Keywords: Novelty Detection, Gaussian Mixture Model, 3D Robotic Mapping

1. Introduction

The problem of detecting novelties comprises the comparison of current data with a priori data or expected behavior. This could be related with a training dataset to learn the normality or some information about the previous state of the data. Using this information, we intend to determine whether something changes, given that the data portion related to novelty is initially unknown. The expressions *novelty detection* and *change detection* are used interchangeably throughout the text to denote the identification of differences between pairs of datasets (*e.g.* robotic map of the same section of an environment at two different instant times). Segmentation of these novelties implies to determine which data is related to them. These problems are important in a large variety of areas, as shown in the work of Markou and Singh (2003) and, more recently and extensively, in the work of Chandola et al. (2009), where anomaly detection is focused. In that work, they made a distinction between anomalies and novelty detection, wherein the latter is related to a novel pattern, being typically incorporated into the model after being detected.

In robotic applications, it can be useful to take into account novelties in the scene (dynamic mapping), in order to accomplish other tasks relying on that knowledge about the environment. This is the topic of increasing interest in the robotic community and which is presented as a particular case study in this article. In robotic surveillance and security systems, for instance, changes in the environment affecting the robot's path can configure risky situations, which require the activation of some kind of alarms with which either the robot or a human operator should be aware of. In a similar way, robots exploring dangerous environments (*e.g.* abandoned mines, maintenance of nuclear reactors or oil pipes), should solve and warn about dangerous situations when a significant change is detected with respect to the known map. In this kind of applications, the robotic platform may be also equipped with a robotic arm and an artificial hand, in order to manipulate and grasp objects associated to changes in the environment. In all these situations, when the robot revisits some section of the environment, it is worth comparing current perceptual data with previously acquired one, in order to detect novelties in the scene (Drews Jr et al., 2010).

The problem addressed in this article is novelty detection and segmentation. It involves processing large datasets which is quite challenging and requires the development of specific techniques, aiming at achieving two interrelated goals (see Fig. 1): firstly, detecting whether there is some significant change; secondly, if some significant change exists, segmenting the data associated with it. In order to improve the accuracy and ensure the feasibility of the change detection process, sensor data must be transformed into a more compact form before comparing with previously acquired data. In this case, the chosen representation heavily determines the performance of the whole

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

task.

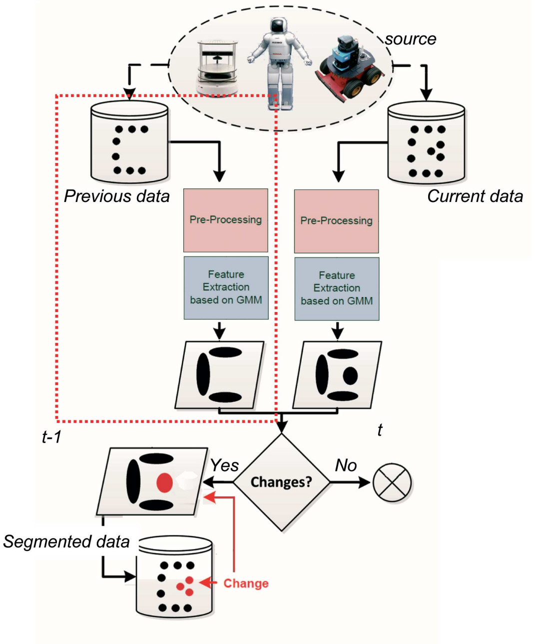


Figure 1: The system is aimed at detecting significant changes (novelties) and segmenting the data (set of points) related with novelties.The use of a high-level representation makes it easier to detect changes, represented in the flowcharts parallelograms.

This article proposes a framework to robustly detect and segment changes in the particular case of 3D robotic maps. Initially, the data to be compared is simplified and compressed without loosing essential information, by using a clustering method based on Gaussian Mixture Models (GMMs). Next, segmentation is accomplished by two alternative methods: a greedy algorithm (Drews Jr et al., 2010), which uses a metric based on Earth Mover's Distance (EMD); and a structural matching algorithm (Núñez et al., 2010). These two approaches are compared in order to detect changes and obtain a segmented point cloud representing those changes.

The main contribution of this article is the development of a practical method comprising techniques to detect and segment changes in large datasets. The method is validated with real data in the context of 3D robotic maps. Related to previous papers of the authors on novelty detection, a more detailed study and thorough experimental evaluation of our method is achieved in this article. The algorithm is evaluated through different experiments (*e.g.* different robot's point of views, different sizes of changes, *etc.*) and it is compared with other change detection algorithms. Besides, this work extends our previous contributions, pointing out new statistical results. A comparative study is made between a state-of-art method and the proposed method using real large datasets, which shows the advantages of the proposed method. Moreover, the use of GMMs as descriptive features of data allows describing and compressing datasets in an concise way. Furthermore, the use of geometrical features (surface variation) in 3D robotic maps is aimed at achieving even better data compression.

The techniques presented in this article open new opportunities to develop automatic processes for surveillance of infrastructures, by proposing a new and robust approach for searching and detecting changes within large amounts of data.

1.1. Related work

Extracting features from data is a challenging step. It involves removing irrelevant or redundant data in order to achieve some goal (*e.g.* detecting a novelty) with information compression (Yu and Liu, 2004). In spatial data, as in 3D robotic maps, the use of planar structures (Stamos et al., 2008; Mufti et al., 2012) and K-means clustering (Loménie, 2004) have been proposed. This problem can also be addressed by following different approaches, such as Principal Component Analysis (PCA) (Jolliffe, 2002) and Independent Component Analysis (ICA) (Comon, 1994), which enable a good dimensional reduction. Gaussian Mixture Models (GMMs) also exhibit interesting properties, namely good compression and description of data, as it was demonstrated in Núñez et al. (2009) and Drews Jr et al. (2010). GMMs are largely used in several applications, such as recognition, especially in audio (Reynolds and Rose, 1995), clustering (Fraley and Raftery, 1998) and classifiers for pedestrian recognition (Premebida et al., 2009). Due to their potential, GMMs were extensively used to detect novelties, but usually as classifiers (Tarassenko et al., 1995). See (Markou and Singh, 2003) for a thorough review of previous work on novelty detection using GMMs.

Different computational methods can be used to compare GMMs and to determine whether a change has occurred. Metrics to compute distance between GMMs, as KL-Divergence (Goldberger et al., 2003) and Euclidean Distance (Helén and Virtanen, 2007), are largely used in audio recognition and image matching. Tomasi et al. (1998) presented the Earth Mover's Distance (EMD) as a metric for comparing two different distributions; it is computationally efficient as an instance of the transportation problem. A large number of metrics are compared in the context of image retrieval in the work by Rubner et al. (2000). These metrics need to be associated to a method, such as the greedy algorithm used in the work of Drews Jr et al. (2010). The use of structural matching as an data association method can also be applied in this context, by using the maximum clique in a correspondence graph (Barrow and Burstall, 1976), as presented in the work of Núñez et al. (2010).

The behavior of an autonomous mobile robot working in dynamic environments has been extensively studied for the last decade. The common strategy has been to remove dynamic objects in order to improve the navigation and localization tasks (Thrun et al., 2001). However, these changes in the robot's surrounding may be actually relevant depending on the applications. In this sense, Andreasson et al. (2007) presented a system for autonomous change detection with a security patrol robot using 3D laser range data and images from a color camera.

Novelty detection in the context of surveillance robots was also addressed in the work of Marsland et al. (2002), where

ACCEPTED MANUSCRIPT

they use Grow When Required (GWR) nets together with habituation. They applied their method to sonar data. This work was extended by Vieira Neto and Nehmzow (2008) using visual colored data, where visual attention is applied through salience maps. Moreover, the use of incremental PCA is compared in this context providing similar results.

More recently, the work of Sofman et al. (2010, 2011) uses novelty detection to identify new classes in data. Features are selected using the Multiple Discriminant Analysis (MDA) approach and applied in NORMA algorithm with few adaptations. It was validated in an AGV acquiring features from LADARs and Cameras. Although this method is efficient, it requires extensive training data.

Novelty detection based on GMMs and EMD was addressed in Núñez et al. (2009), and later in Drews Jr et al. (2010). In a first stage, a GMM was computed to cluster the cloud of 3D points. Next, EMD was used to quantify changes in the data. A new greedy algorithm was proposed to segment changes. Despite the quality of the attained results, the computation times were still too large to make viable their practical use with large datasets. Moreover, the method is sensitive to the maximum number of Gaussians in the GMM. Bearing on the same essential ideas, this article extends previous work of the authors (Núñez et al., 2009; Drews Jr et al., 2010) with the aim of relieving the required computation burden and parameters specification. Furthermore, a more thorough experimental evaluation with real and large datasets is presented.

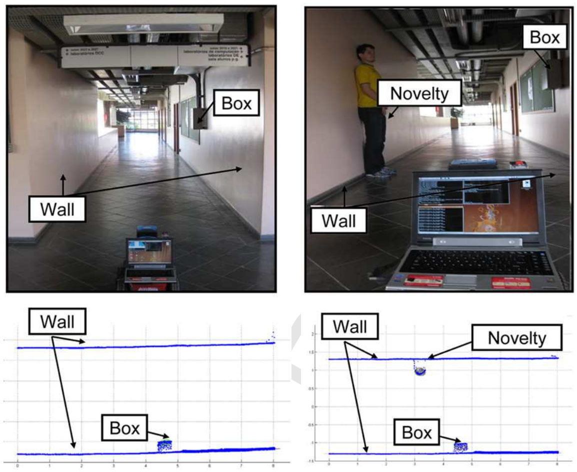
1.2. Organization of the article

The article is organized as follows. Section 2 presents in detail a solution based on Gaussian Mixture Models for the novelty detection and segmentation problem. Section 3 describes experiments to evaluate the proposed framework with large datasets acquired with real sensors. Finally, in Sec. 4, the main conclusions and future work are drawn.

2. Novelty detection and segmentation

This section describes the proposed method for detecting and segmenting changes in the environment. The crux of the novelty detection problem is *to identify any previously unknown feature* (Vieira Neto and Nehmzow, 2007). One of the main envisioned applications for our method is advanced perception for autonomous mobile robots equipped with 3D laser range finders, or an equivalent 3D range sensor providing dense 3D data of its working environment (*e.g.* depth camera). Fig. 2 illustrates a typical situation where an autonomous robot moves in its working environment. As shown in the figure, the environment may present sudden changes – in the example, a change occurred in the hallway. The robot must find and segment this novelty in order to cope with the environment's dynamics and it also needs to correctly update its model of the environment.

The several stages of the novelty detection and segmentation process that is proposed in this article were outlined in Fig 1. First, information from the environment is acquired by a 3D sensor (different laser scanners have been used in this paper).



(a) (b)

Figure 2: The aim of the proposed novelty detection method is to find and segment changes in the mobile robot's workspace (*e.g.* the appearance of a person in (b)).

This data is pre-processed in the next stage, by two consecutive methods with the aim of reducing the number of points in the 3D map: i) a simplification algorithm and ii) a sparse outliers and ground plane removal methods. The simplification method uses surface information and generates a multi-scale point cloud (Pauly et al., 2002). As most of the data contained in both datasets being compared are usually related with static structures, the sparse outliers and ground plane removal methods remove data which, although not being related to novelties, significantly increase the computation burden. Once the preprocessing stage is completed, the dataset is converted from the Euclidean space to the mathematical space of Gaussian Mixtures, *i.e.* features based on Gaussian Mixture Models (GMMs) are extracted from the simplified data. The use of these features allows the process to describe the environment information in a more convenient way, and easily detect and segment novelties. This GMM-based approach was seminally proposed by the authors in Núñez et al. (2009). Finally, novelties are computed and segmented in this mathematical space of GMMs, and two different algorithms will be described in Section 2.4. One algorithm is based on Earth Mover's Distance (EMD), as was described in Núñez et al. (2009) and Drews Jr et al. (2010); and second one is a structural matching algorithm based on graph theory recently proposed in Núñez et al. (2010). The latter one takes into account not only absolute constraints (similarity between Gaussians) but also relative constraints as regards local structural information. The following subsections describe in detail each one of these stages.

2.1. 3D laser data acquisition

Independently of the 3D laser sensor, range data provided by these sensors are typically represented in the form $\{P_l = (x, y, z)_{l|l=1..N_R}\}$, where $(x, y, z)_l$ are the cartesian coordinates of the l-th range reading and N_R is the total number of range readings in each 3D scan. The higher the laser scanner resolution,

ACCEPTED MANUSCRIPT

the larger is N_R, *i.e.* the cloud of points comprises a larger number of 3D points. Therefore, if the aim is to compare clouds of 3D points, first they have to be simplified and compressed without loosing essential information. Comparing directly these clouds of points is not feasible for typical values of N_R.

2.2. Pre-processing

The most important part of the preprocessing stage is the simplification algorithm that is used to decrease the density of points in the point cloud provided by the 3D laser scanner, without compromising its geometric properties. The proposed simplification method is based on the work by Pauly et al. (2002) in the domain of 3D surface modeling and reconstruction. This approach has the interesting property of reducing the amount of data with low computational cost and without compromising its geometric discriminative power. Results shown in Drews Jr et al. (2010) shows the importance of this part in the novelty detection system, because this method reduces significantly the computational burden without degrading the novelty detection ability.

The simplification method computes a multi-scale point cloud using binary space partition in a top down approach. The use of covariance analysis enables the method to compute the surface variation, σ, based on eigenvalues, λ, from each point cluster and defined by (1). The point cluster P is then split if the cardinality of P, $|P|$, is larger than a predefined value and the surface variation, σ, is above a maximum threshold σ_{max}. In this article, $\sigma_{max} = 0.1$ and the range of σ is $[0, \frac{1}{3}]$; these parameters were empirically selected for a typical commercial laser data density.

$$\sigma = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}, \quad \lambda_0 \leq \lambda_1 \leq \lambda_2. \tag{1}$$

This hierarchical clustering simplification process builds a binary tree by splitting each region. The split plane is defined by the centroid of P and the eigenvector associated to the greater eigenvalue (λ_2). The point cloud is always split along the direction of greatest variation. The multi-scale representation is based on the restriction level imposed to the tree. The tree grows until the cluster is just one point and the scale is then chosen by setting values for the size of $|P|$ and for σ_{max}. Fig. 3 illustrates this hierarchical clustering algorithm. Fig. 4 shows an illustrative result of the characteristics of the simplification method, wherein it reduces the number of points in flatter regions but not in regions with lower curvature ratios.

In a typical point cloud acquired by a laser scanner, the ground plane is almost always present. The subset of points representing the ground plane is one of the types of static information which is not useful for the novelty detection process, and unnecessarily increases the computational burden. It can be easily filtered out by using RANSAC to fit a ground plane (Lai and Fox, 2009). Moreover, sparse outliers in the 3D scan laser data are removed by using the technique of Rusu et al. (2008).

2.3. Feature extraction based on GMMs

Computing changes in the Euclidean space, *i.e.* by processing directly clouds of points, presents several pitfalls, including

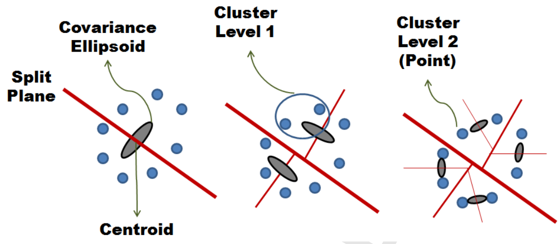


Figure 3: 2D sketch of the hierarchical clustering algorithm: on the left, split plane within the set of points (blue), based on covariance ellipsoid and centroid; on the middle, the result after an iterative step, generating the cluster in level 1; on the right, the last step which generates the lower level clustering, wherein each cluster represents a single point.

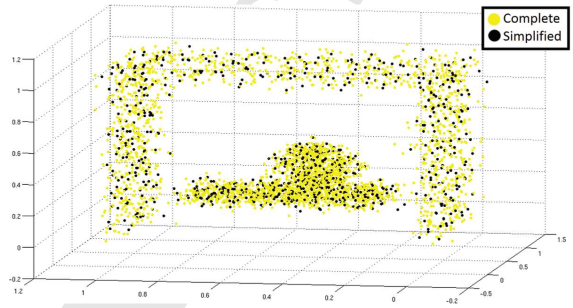


Figure 4: Simulated point cloud with white noise with zero mean, where the simplification method is applied. It illustrates, in black, the simplified point cloud and, in yellow, the original point cloud. Flatter regions suffer larger reduction in the number of points.

computation cost and scale variance. Therefore, it is required to represent data in a more convenient mathematical space for novelty detection and segmentation. The method proposed here uses the mathematical space of Gaussian Mixtures.

A *Gaussian Mixture Model* (GMM) is a probability density function described by a convex linear combination of Gaussian density functions (McLachlan and Peel, 2000), with the form:

$$f(x, \Theta) = \sum_{k=1}^{K} \pi_k \, g(x; \mu_k, \Sigma_k) \quad \left(x \in \mathbb{R}^M\right), \tag{2}$$

where the functions g are Gaussian densities with parameters $\mu_k \in \mathbb{R}^M$ and Σ_k, the mean matrix and the covariance matrix, respectively. The coefficients π_k are usually denoted by *mixing probabilities* and satisfy the condition:

$$\pi_k > 0 \quad \text{and} \quad \sum_{k=1}^{K} \pi_k = 1. \tag{3}$$

The GMM can be seen as a collection of Gaussian features providing a good model for clusters of points: each cluster corresponds to a Gaussian density whose mean is located about the centroid of the cluster and whose covariance matrix estimates the spread of that cluster.

Conversely, given a set of points in $\mathbb{R}^M$, one can try to find the mixture of Gaussian functions Θ that best fits those points,

ACCEPTED MANUSCRIPT

using an approach based in the *Expectation Maximization* algorithm, as shown in Algorithm 1. Parameter K denotes the number of Gaussians included in the GMM. It is selected by using K_{max} and the MDL penalty function (Rissanen, 1983). In the present work, Θ denotes the $K(1+N+N^2)$-dimensional vector containing all the parameters of the given Gaussian mixture:

$$\Theta = ((\theta_1, \pi_1), \ldots, (\theta_K, \pi_K)), \tag{4}$$

where

$$\theta_k = (\mu_k, \Sigma_k). \tag{5}$$

This is a vector containing all the coordinates of the means μ_k and all the entries of the covariance matrix Σ_k. The conditions in (3) ensure that f is indeed a density function.

In Algorithm 1, three steps are crucial to correctly estimate the Gaussian Mixture Model. Firstly, an initial model is estimated, function *Initialize_model*. Good results were reported using random initialization (McLachlan and Peel, 2000). Thus, in the present approach, the initialization is obtained using random generation based on the follow equations:

$$\begin{aligned}
\pi_k^{(0)} &= \frac{1}{K_{max}}, \\
\mu_k^{(0)} &= p_n \rightarrow n = \lfloor (k-1)(N-1)/(K_{max}-1) \rfloor + 1, \\
\Sigma_k^{(0)} &= \frac{1}{N} \sum_{n=1}^{N} p_n p_n^T.
\end{aligned}$$

Secondly, the Expectation Maximization algorithm is estimated using the standard approach (Figueiredo and Jain, 2002), but using as penalization criteria the MDL penalty function (Rissanen, 1983). Equation 6 shows this criteria, assuming that $L = \frac{K}{2}\left(M^2 + 3M + 2\right) - 1$ and Y is the point cloud with N points used to estimate the GMM.

$$MDL(K, \theta) = -\log p(p|K, \theta) + \frac{1}{2} L \log(N \cdot M). \tag{6}$$

The last function considered is the *Resize*. Its aim is to find two Gaussians l and m that minimize the MDL criteria. Equation 7 shows the distance metric to be optimized in order to find out the best pair of Gaussians to be joined, wherein $\Sigma_{(l,m)}$ is the covariance matrix resulting from the union of Gaussians l and m. Finally, the result obtained from the algorithm is the best GMM θ^*, with size K^* and having MDL^* as penalty criteria.

$$d(l, m) = \frac{N\pi_l}{2} \log\left(\frac{|\Sigma_{(l,m)}|}{|\Sigma_l|}\right) + \frac{N\pi_m}{2} \log\left(\frac{|\Sigma_{(l,m)}|}{|\Sigma_m|}\right). \tag{7}$$

Fig. 5-a illustrates clouds of 3-D points describing a synthetically generated ideal corridor, where one object has been inserted. The GMM that is obtained from this cloud of points is depicted in Figs. 5b-c. It is described as a mixture of few 3D Gaussians (for 3D data, $M = 3$), being each one associated to the clusters of points identified in the scene: walls, ceiling and the novelty.

Algorithm 1 Gaussian Mixture Models estimation algorithm - Input: Point cloud Y and integer K_{max}

$\theta^{(0)} \leftarrow Initialize\_model(K_{max})$
$k = K_{max}$
$[MDL^{(1)}, \theta^{(1)}] \leftarrow EM\_algorithm(k, Y, \theta^{(0)})$
$MDL^* \leftarrow MDL^{(1)}$
$\theta^* \leftarrow \theta^{(1)}$
$K^* \leftarrow k$
$i = 1$
while $(k > 0)$ **do**
  $\theta_{(l,m)}^{(i)} \leftarrow Resize(\theta^{(i)})$
  $k \leftarrow k - 1$
  $[MDL^{(i+1)}, \theta^{(i+1)}] \leftarrow EM\_algorithm(k, Y, \theta_{(l,m)}^{(i)})$
  $i \leftarrow i + 1$
  if $(MDL^* > MDL^{(i)})$ **then**
    $MDL^* \leftarrow MDL^{(i)}$
    $\theta^* \leftarrow \theta^{(i)}$
    $K^* \leftarrow k$
  end if
end while

2.4. Computing changes

Once both datasets being compared are modeled in the GMMs mathematical space, they have to be processed with the aim of estimating changes, *i.e.* detecting and segmenting novelties. This section presents two alternative approaches that can be used to accomplish this goal. The first one, which was proposed in Núñez et al. (2009) and Drews Jr et al. (2010), follows a greedy algorithm based on Earth-Mover's Distance (EMD), which measures the distance between two distributions. Being simple, this approach only takes into account the similarity between two GMMs (absolute constraints). A second approach was recently proposed by the authors in Núñez et al. (2010), which overcomes this limitation by also considering relative contraints. In this later approach, the problem is formulated as a maximum clique problem in an undirected graph. In the GMM mathematical space, the novelties detection algorithms presented in this paper allows simultaneously to detect new objects in the scene or something that has been removed from it.

2.4.1. EMD-based algorithm

The *Earth Mover's Distance* (EMD) was proposed by Tomasi et al. (1998) as a metric for measuring distance between two distributions of points in space for which a distance between points is given.

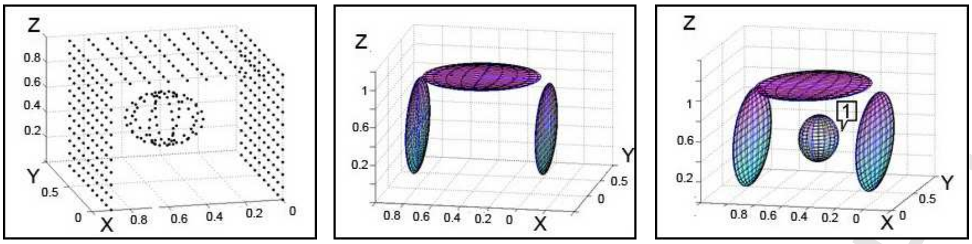
Let GMM $\Theta = ((\theta_1, \pi_1), \ldots, (\theta_n, \pi_n))$ and GMM $\Gamma = ((\gamma_1, \pi_1), \ldots, (\gamma_m, \pi_m))$ be associated with two 3D point clouds at different instant times. Let θ_i and γ_j be Gaussian functions and π_i and π_j their associated weights, respectively. In order to identify the novelties in the environment, the two GMMs are modeled as weighted points $(\theta_i, \pi_i)_\Theta$ and $(\gamma_j, \pi_j)_\Gamma$. Thus, the distance between the two GMMs is computed as

$$d_{GMM}(\Theta, \Gamma) = \mathrm{EMD}\left(\{(\theta_{1\ldots n}, \pi_{1\ldots n})\}, \{(\gamma_{1\ldots m}, \pi_{1\ldots m})\}\right). \tag{8}$$

Equation (8) can be used as a quantitative metric for detecting changes in the environment. In the most obvious way, the

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure 5: GMM computation for a synthetically generated corridor: (a) cloud of points representing an ideal 3D corridor where a new object was placed inside; (b) GMM representing the corridor map without the object placed inside; (c) GMM associated to the cloud of points (a). The Gaussian function representing the novelty is labeled by '1'.

problem of change detection could be tackled with EMD by defining a threshold U_{th} which represents the maximum value beyond which it is assumed that a novelty exists, *i.e.* a significant distance between the two GMMs exists. However, using a fixed threshold constitutes a naive approach, as it would be very difficult to tune U_{th}. Therefore, we propose a greedy algorithm that overcomes this limitation. An example of the application of this technique can be seen in Fig. 5, where GMMs associated with clusters of 3D points are shown. After applying the EMD-based novelty detection algorithm to these two Gaussian Mixtures, a novelty is detected in the maps (marked as '1' in Fig. 5-c).

The overall structure of the EMD-based algorithm is outlined in Algorithm 2. The method achieves more than just detecting changes, since the novelty is segmented and the set of points associated to it is retrieved using the posterior probability. In each iteration, the algorithm selects a Gaussian $x(\mu, \Sigma)$ from Θ with the greatest quantified change d_{GMM}, computed by the function *GreedySelectGMM*. Furthermore, this function returns both the d_{GMM} and the new set Π. It works by computing EMD between Γ and the new sets. These new sets are generated by removing one Gaussian at a time from Θ. The best Gaussian is removed from the initial mixture Θ and is also included in the new Gaussian mixture model Π. The distance d_{GMM} is compared iteratively with the previous EMD distance. The algorithm returns a list of sets of points S. Each set represents the segmented region by one Gaussian, using the posterior probabilities computed by the function *ChoosePtsfromGaussian* that has as arguments a point cloud P used for generating the novelty GMM and a Gaussian x. If $S = \{\emptyset\}$, the algorithm assumes that there are no novelties, implying that the two GMMs are similar. Moreover, the posterior probability allows the system to identify the topological relation between the segmented regions. This kind of information could be useful both for recognition and for identification, providing the means to build a semantic representation of the environment.

2.4.2. Structural matching algorithm

The EMD-based algorithm described in Sec. 2.4.1 is strongly dependent on the number of Gaussians. Furthermore, that algorithm only deals with absolute constraints, *i.e.* the Euclidean distance between the mean vectors of Gaussian functions. However, valid pairwise associations between Gaussian

Algorithm 2 Novelty Detection algorithm - Input: GMMs Θ and Γ

```
d_GMM ← EMDdistance(Θ, Γ)
Π ← ∅
repeat
    d_GMM_old ← d_GMM
    [x(μ, Σ), Π, d_GMM] ← GreedySelectGMM(Θ, Γ)
until (d_GMM_old < d_GMM)
Π ← Π − x(μ, Σ)
S ← {∅}
for all x(μ, Σ) ∈ Π do
    S ← S ∪ ChoosePtsfromGaussian(P, x(μ, Σ))
end for
return S
```

functions of two GMMs, *i.e.* relative constraints, are also worth to be considered in order to attain a more robust novelty detection. With this purpose, Núñez et al. (2010) have recently proposed an alternative novelty detection and segmentation algorithm which takes into consideration both absolute and relative constraints.

In this alternative approach, the matching problem is formulated as a graph-theoretic data association problem. Thus, the fundamental data structure of this step is a correspondence graph (Barrow and Burstall, 1976) representing valid pairwise associations between Gaussian mixtures, as it is depicted in Fig.6. Cliques in this graph indicate mutual associations compatibility and, by performing a maximum clique search, the joint compatible association set is emanated from the best matchings of Gaussian mixtures that can be found. In the proposed approach, the correspondence graph is built through the application of both *absolute* and *relative* constraints. The nodes of the graph indicate individual association compatibility and are determined by absolute constraints. Conversely, edges of the correspondence graph indicate joint compatibility of the connected nodes and are determined by relative constraints.

As before, let Θ and Γ be two GMMs associated with two 3D point clouds of the same section of an environment, acquired at different times. Moreover, let $\theta_i(\mu_i, \Sigma_i) \in \Theta$ and $\gamma_j(\mu_j, \Sigma_j) \in \Gamma$ be Gaussian functions, where (μ_k, Σ_k) is a vector containing all the coordinates of the means μ_k and all the entries of the covariance matrix Σ_k. The method to compute the correspondence

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

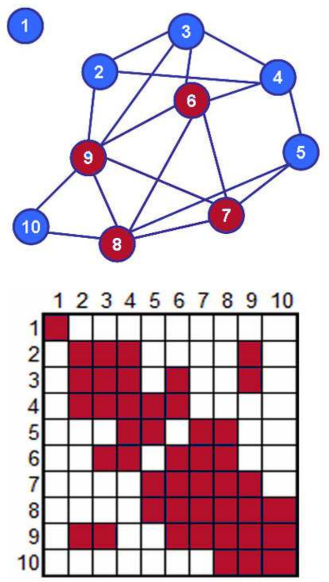


Figure 6: In the proposed approach, nodes represent tentative matchings when considered individually. Edges indicate compatible pairwise associations between Gaussians. A clique is a set of mutually consistent associations, *e.g.* the clique marked in red implies that the matching may coexist.

graph comprises of two major steps (Núñez et al., 2010): the definition of the nodes of the correspondence graph and the definition of the edges of the correspondence graph. These steps are detailed below. Afterwards, it is also described how the correspondence graph can be used to reliably detect novelties.

- Definition of the nodes of the correspondence graph. The nodes of the graph represent tentative matchings of Gaussian distributions from two GMMs, Θ and Γ, after applying an absolute constraint. Let $|\Theta| = n$ and $|\Gamma| = m$ be the number of Gaussian functions belonging to the two GMMs, respectively. Firstly, the algorithm generates the $n \times m$ matrix T_t for all pairwise combinations, by calculating the distance between the two Gaussian functions:

$$d_{\theta_i,\gamma_j} = max(d_{\mu_{ij}}, d_{\Sigma_{ij}}), \quad (9)$$

where $d_{\mu_{ij}}$ is the Euclidean distance between the two Gaussian functions using the cartesian coordinates of the mean vector, and $d_{\Sigma_{ij}}$ is the distance between the covariance matrices associated with the Gaussian functions. According to Forstner and Moonen (1999), this distance can be defined as:

$$d_{\Sigma_{ij}} = \sqrt{\sum_{k=1}^{N} ln^2 \lambda_k(\Sigma_i, \Sigma_j)}, \quad (10)$$

where λ represents the generalized eigenvalues of Σ_i and Σ_j, and N is the matrices' dimensionality.

The matrix entry associated to the matching of two similar Gaussian functions presents a small value. On the other hand, large values of T_t correspond to dissimilar features. Pairwise matched features, whose matrix values are smaller than a fixed threshold U_T^t, constitute the set of tentative matches. Thus, graph nodes are defined as the set of all possible combinations of these pairwise descriptors. Fig. 7 shows the process achieved in order to obtain the nodes of the graph for a simple real example. Fig. 7-a illustrates two sets of Gaussians associated to two different instant of time (Θ and Γ). This kind of representation for the Mixture of Gaussians is common along this paper. The ellipsoids are associated to each Gaussian (mean and covariance matrix) and they are drawn using different colors and with 1.5 standard deviation as diameter. The use of different colors is only for improving the visualization of the mixture. In the figure, firstly, the algorithm calculates the distance between the pair of Gaussian θ_{1_a} and γ_{1_b}. If this distance is lower than the threshold, the algorithm generates the node $(1_a, 1_b)$ (Fig. 7-b). The algorithm calculates the distances between the pair of Gaussian in a similar way (Figs. 7c-e).

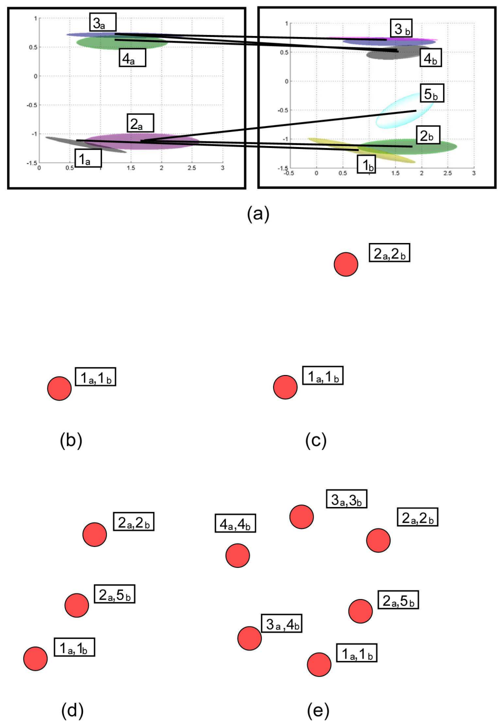


Figure 7: Nodes represent tentative matchings when considered individually: (a) two set of Gaussians associated with two different instant times (Θ and Γ); (b)-(e) the algorithm calculates the distance between the pair of Gaussian and generate the nodes if the distance is lower than the threshold.

- Definition of the arcs of the correspondence graph. For all pairwise combinations of matchings in T_t, the relative constraint matrix R_t is computed through the use of relative constraints on the mathematical space of GMMs. A

ACCEPTED MANUSCRIPT

pair of matched Gaussian functions (θ^i, γ^i) and (θ^j, γ^j) is consistent *iff* they satisfy the relative constraint:

$$max(\omega_{d_\mu}, \omega_{d_\Sigma}) \leq U_R^t, \tag{11}$$

with

$$\begin{aligned} \omega_{d_\mu} &= \sqrt{|(d^{\Theta}_{\mu_{ij}})^2 - (d^{\Gamma}_{\mu_{ij}})^2|} \; and \\ \omega_{d_\Sigma} &= \sqrt{|(d^{\Theta}_{\Sigma_{ij}})^2 - (d^{\Gamma}_{\Sigma_{ij}})^2|}, \end{aligned} \tag{12}$$

where U_R^t is a threshold defined by the user. Thus, the corresponding entry in the relative constraint matrix R_t contains a '1' if the constraint is satisfied (an edge is added to the graph), and '0' otherwise. Fig. 8 shows an example of the arc definition algorithm for the examples described in Fig. 7. For instance, in Fig. 8-a, the relative constraint between ($1a$, $1b$) and ($2a$, $2b$) matches, and then node ($1a$, $1b$) is connected to node ($2a$, $2b$) through an edge Fig. 8-b. The process is repeated for all the pairwise of Gaussian (Figs. 8c-d). The relative constraint between ($2a$, $5b$) only matches with ($2a$, $2b$).

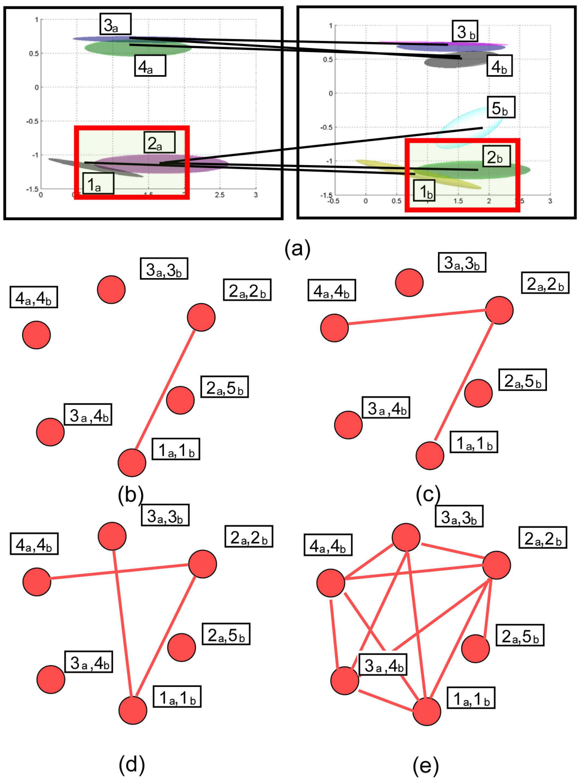


Figure 8: Nodes represent tentative matchings when considered individually. Edges indicate compatible pairwise associations between Gaussians. A clique is a set of mutually consistent associations, *e.g.* the clique marked in red implies that the matching may coexist.

- Maximum clique detection and novelty identification. The set of mutually consistent matches which provides the largest clique is calculated. This is equivalent to finding the maximum clique on a graph with adjacency matrix R_t (see Núñez et al. (2010)). After the computation of the maximum clique of the correspondence graph, a set of mutually compatible associations is obtained, *i.e.* a set of matched Gaussian functions (see red lines in Fig. 9). In this way, the algorithm takes into account structural relationships to detect correct associations, which result in 3D points in the environment that are not associated with changes in the robot's surroundings. Thus, the set of Gaussian functions in Θ that are not included in the clique represents detected novelties. In Fig. 9, the only node which is not included in the clique, ($2a$, $5b$), represents the novelty that is detected in this example within the robot's workspace.

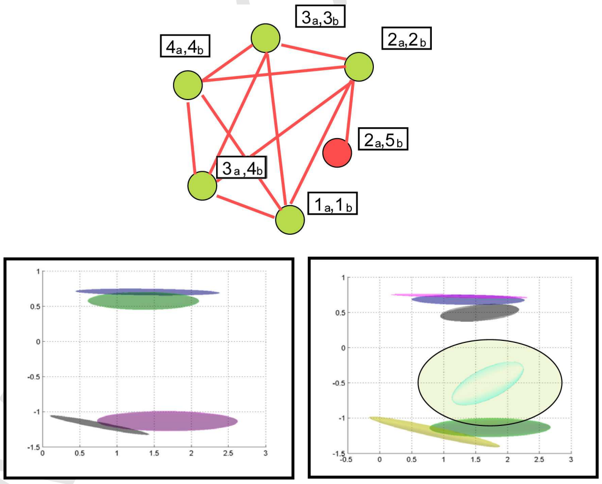


Figure 9: A clique is a set of mutually consistent associations, *e.g.* the clique marked in red implies that the matching may coexist.

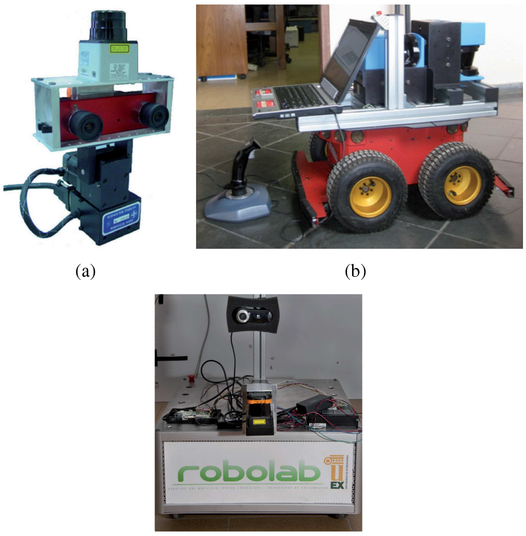
3. Experiments

The proposed methods have been evaluated using real data. The algorithms were developed in C++ and the benchmark tests were performed on a notebook with a 2.0GHz AMD Turion X2 Ultra CPU and 3Gb RAM running using GNU/Linux Ubuntu 10.0. Real data has been acquired using three different platforms.

The first experimental platform was static comprised of a Hokuyo URG-04LX laser range finder mounted on a Directed Perception PTU-D46 pan-tilt. A set of experiments was carried out with this platform in three different environments. For any of these environments, the experiment comprised three steps. Firstly, a 3D map was acquired to obtain a representation of the environment. Afterwards, a novelty was introduced. Finally, in order to obtain statistically significant results, the experiments were repeated ten times for each test area. The same procedure was used in two different mobile robotic platforms. First, a Pioneer P3-AT with two SICK LMS-200 mounted orthogonally was used in the experiments. The robot is located using a

ACCEPTED MANUSCRIPT

SLAM algorithm (DP-SLAM) (Eliazar and Parr, 2003) with the information acquired from the frontal laser and wheels odometry. The localization information from SLAM together with the second laser allows the robot to generate 3D maps. In the other mobile robotic platform, a 3D capture system mounted on a Robex robot was used (Gutiérrez et al., 2011). This sensor consists of a Hokuyo URG-30LX laser rotated by a step motor, where its resolution is configurable by the active perception system that the robot is equipped with. Fig. 10 illustrates the three platforms used in our experiments. Fig. 10-a shows the static sensor platform and Fig. 10-b shows the mobile platform equipped with two orthogonal laser range finders. Finally, Fig. 10-c illustrates the robot Robex and its low-cost 3D sensor based on a 2D Hokuyo laser range finders [1].



(a) (b) (c)

Figure 10: Experimental platforms: (a) platform using Hokuyo URG-04LX laser range finder and a pan-tilt unit; (b) mobile platform using Pioneer robot and two SICK laser orthogonally mounted; and (c) robot Robex and its 3D perception system based on a Hokuyo URG-30LX laser finder.

3.1. Quantitative Assessment

In order to validate the proposed method, contingency tables were built relating the system response to ground truth information. The ground truth data was manually segmented using a dataset edition tool developed by the authors, where the points that represent the changes are selected. Statistical significance of the association between the ground truth (points related to the novelty) and the novelty detection algorithm response was established as in Vieira Neto and Nehmzow (2008). They use three different statistical indicators based on χ^2 analysis of the contingency table. These metrics are used as a reference to evaluate the GMM-based method proposed herein.

[1]A simple USB camera is used for acquiring RGB information of each 3D point, but it is was not used in the experiments described here.

The Cramer's V ($0 \leq V \leq 1$) and the uncertainty coefficient U ($0 \leq U \leq 1$) are used to quantify the strength of the association, and they are computed as proposed in Vieira Neto and Nehmzow (2008). Smaller values for these statistics indicate weaker associations, and values closer to one represent stronger association.

Another metric that is used is the index of agreement κ, which is used to assess the agreement between actual novelty status and the algorithm response, in way similar to V and U. However, the κ statistic may yield negative values. The valid range for κ is $\kappa \in [-1, 1]$, where negative values represent the level of disagreement between the system response and the ground truth. Table 1 explains the relation between the κ intervals and the corresponding level of agreement.

Table 1: κ intervals and corresponding levels of agreement between ground truth and novelty filter response (Vieira Neto and Nehmzow, 2008).

Interval	Level of Agreement
$\kappa \leq 0.1$	No
$0.1 < \kappa \leq 0.4$	Weak
$0.4 < \kappa \leq 0.6$	Clear
$0.6 < \kappa \leq 0.8$	Strong
$0.8 < \kappa \leq 1.0$	Almost Complete

3.2. Estimation of parameters

The proposed method requires choosing values for a set of parameters. These parameters are:

1. The threshold value σ_{max}, which determines the maximum surface variation for splitting a cluster of points in the preprocessing stage. In the experiments conducted here this value was set to 0.1. It is empirically selected for a typical commercial laser data density. However this is not a critical parameter, since its value implies only the number of points used in the next stages.
2. The maximum number of Gaussian functions in the Gaussian Mixture Model, K_{max}. This parameter depends on the environment and the size of the map. In this article, this parameter is changed from 10 to 30 in order to validate its effect on the change detection algorithms results.
3. There are two thresholds used in the generation of the correspondence graph. First, the minimum distance between two Gaussian functions for being considered a tentative match, U_T^t (absolute constraint). Second, the threshold value U_R^t which is used in the definition of the arcs of the graph. It represents the minimum value in the relation between two pairs of tentative matches for a match to be considered as coherent. The benchmark used to set them correctly was similar for the two stages. This step is based on the work of Blanco et al. (2010). Optimal thresholds are calculated by minimizing the probability P_{err} of misclassifying a association as a valid (v) or an invalid (w)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

candidate. It is described as:

$$\begin{aligned} P_{err}(U_T^t, U_R^t) &= \\ P(\omega)P_{err}(U_T^t, U_R^t \mid \omega) &+ P(v)P_{err}(U_T^t, U_R^t \mid v) \\ &= P(\omega)P(d_{ij} < U_T^t, \delta_{ij} < U_R^t \mid \omega) \\ &+ P(v)\left[1 - P(d_{ij} < U_T^t, \delta_{ij} < U_R^t \mid v\right], \end{aligned}$$

where a misclassification will occur if: (i) a distance d_{ij} is less than both thresholds U_T^t and U_R^t, and it was a wrong correspondence; or (ii) a valid pairing does not generate values that are larger than the thresholds U_T^t and U_R^t.
Considering no *a priori* information about the probability of being in a valid or invalid association, that is P(v) = P(ω) = 1/2, the method evaluates the joint conditional densities $P(d_{ij}, \delta_{ij} \| v)$ and $P(d_{ij}, \delta_{ij} \| \omega)$ from histograms according to a set of 3D maps where the ground truth is known ($N = 40$).
Finally, the values $U_T^t = 3.5$ and $U_R^t = 1.0$ were selected. These two values allow the algorithm to reduce the number of false positives and to improve the precision at the end of the change detection process.

3.3. Experimental setup with 3D mapping

Fig. 3.3 depicts three experiments in an indoor environment located at Institute of Systems and Robotics – Coimbra, where the platform depicted in Fig. 10-a was used. In each one of these experiments, there were two acquisitions: in the first one, an initial 3D map of the test site was processed. In the second one, a new 3D map of the same test site was processed after introducing a novelty. These new maps are shown in Fig. 3.3-b. They are an opened door, a person in the corridor and a closed door inside the room, respectively. This is illustrated by a picture from the test site in Fig. 3.3-a. Fig. 3.3-c depicts GMM representations from the new 3D map, where the point clouds are shown in blue dots. The Gaussians that represent the segmented novelties in the GMMs space is shown in Fig. 3.3-d. These results demonstrate the ability of GMMs to model the environment with different complexities and the proposed algorithm for detecting changes in them. For all of these experiments, parameter K_{max} was set to 20.

In a second experiment, the robotic platform depicted in Fig. 10-b was used to test the proposed novelty detection and segmentation method in an office building located in the Computer Science Department of the Federal University of Minas Gerais, Brazil, as it is shown in Fig. 11. For the experiments depicted in Fig. 11, three different novelties were included in order to evaluate the results of the algorithm: a cylinder, a person and a box (see Fig. 11-a). Fig. 11-b illustrates the 3D laser data acquired with the novelties by the robot, after the pre-processing stage. The GMMs associated to these 3D maps is shown in Fig. 11-c, and the actual novelty is marked in the figure. Results of the proposed method are drawn in Fig. 11-d. As it is shown in the figure, the Gaussian functions associated to the novelties introduced in the environment were successfully extracted with the proposed method.

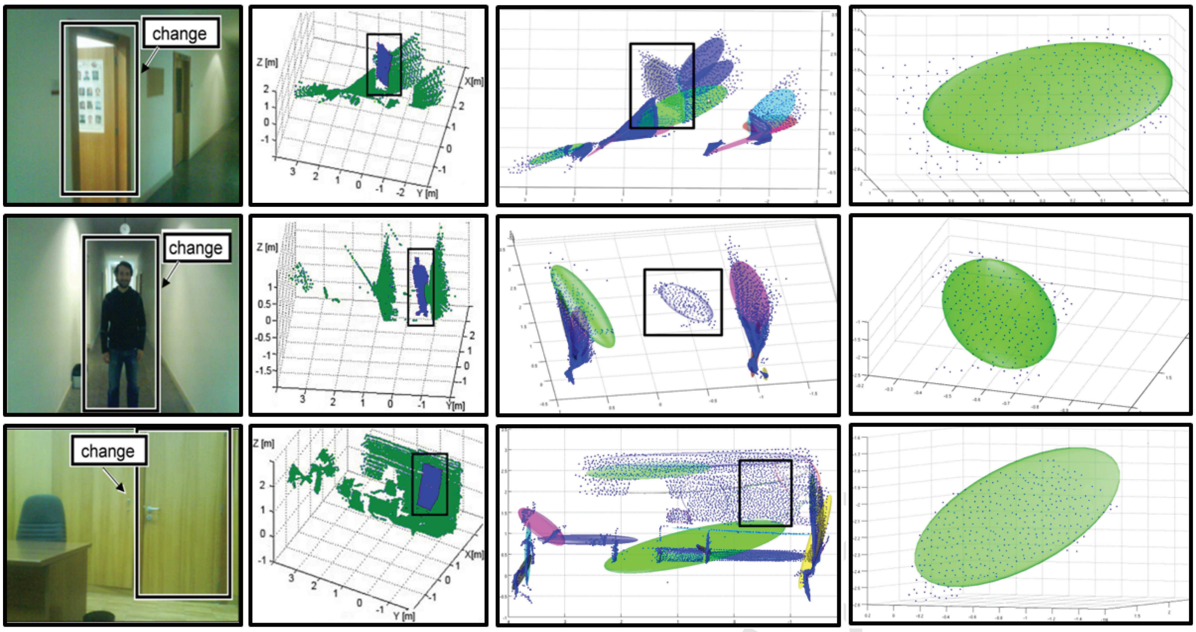
3.4. Comparison with Loménie's method - Gaussian Mixture Models vs. K-Means Clustering

In this section, the use of Gaussian Mixture Models (GMMs) as a representation model of 3D maps is validated. Moreover, it is compared with the technique for partitioning 3D maps described in Loménie (2004), which mainly consists of a specific K-means algorithm denoted as UFP-ONC (Gath and Geva, 1989). Results obtained with the two methods (GMMs and UFP-ONC) are compared in order to deal with the change detection problem. Thus, in order to detect these changes, one important feature is the capability to previously segment the environment in different parts associated to changes or non-changes. Furthermore, another important aspect is related with the computation cost of these. Thus, three results are calculated for each dataset. First, the results from Loménie's method using the automatic choose of the best representation, considering the maximum number of partition as 9. Second, due to the poor quality of these results, the 3D map partition is enforced to be divided into 9 clusters. Finally, the same results were obtained using our approach for clustering using GMMs, with $K_{max} = 9$. Loménie's method uses fuzzy inference in order to select the points related to each cluster. For our experiments, $S = 0.2$ is used to associate all points to the cluster. For larger values, the method does not make association between points and a cluster.

The datasets used in this comparison are three indoor experiments. They have approximately 30,000 points and are illustrated by Fig. 3.3-a. Results are shown in Fig. 12, wherein the segments are drawn in different colors. Fig. 12a-c illustrates Loménie's results considering the automatic best solution between 1 and 9 clusters. The method is fast, but it did not generate good segmentation. As it is shown in Fig. 12a-c, it selected three clusters in the first case and two in second and third experiments. The time spent is about 2 minutes for a Java implementation for all of them. In Fig. 12d-f the number of clusters was forced to be equal to 9. It generates similar results to the GMMs, but the time consumption is much greater. Results after using GMM segmentation are shown in Fig. 12g-i. The computation time of Loménie's method is about one hour. Conversely, the GMM-based method spent 1 minute, therefore it was faster than the first approach applied to Loménie's method but presenting similar results to the second one.

The results shown in Fig. 12-f and in Fig. 12-i do not allow detecting the change in the scene (*i.e.* the door), because it clusters the door and the wall in the same segment. Due to the large computational time of Loménie's method, it is unfeasible to increase the number of clusters in order to evaluate it. Our methodology computes GMMs with $K_{max} = 20$ and obtains a good representation in 60 seconds, as depicted in Fig. 13. This elapsed time is approximately the same time spent to acquire the maps by the robot using the described acquisition system. Similar to Fig. 12, it shows the representation based on mixture of Gaussians from the dataset by color division. One interesting issue is related to the possibility of detecting the door in the dataset with a good precision (marked in the figure as a white box), due to the good approximation by a Gaussian.

ACCEPTED MANUSCRIPT



(a) (b) (c) (d)

Figure 11: Experiments in indoor environments: (a) test sites – a door in the corridor has been opened (top), a person appears in the corridor (middle) and the room's door has been closed (bottom); (b) real observations from Hokuyo laser range sensor (novelties manually segmented in blue);(c) the GMM representations of the maps containing the novelties (black box); (d) Gaussian functions and sets of points representing the segmented novelties.

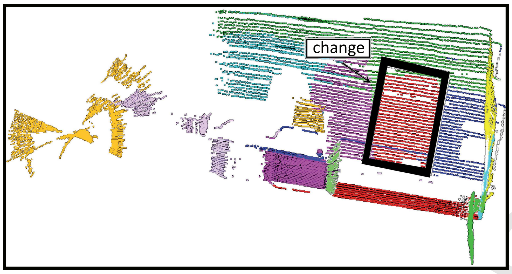


Figure 14: Results of the environment segmentation based on Gaussian Mixture Models using the described approach with $K_{max} = 20$, considering the maps illustrated in Fig.3.3-a (bottom). It allows detecting changes in the environment after applying the algorithm proposed in this article (box delimits the position of the door).

3.5. Evaluation of the robustness

In this section, both methods for detecting changes in the environment using the mathematical space of Gaussian functions described in this article have been compared and their results analyzed in terms of robustness and computational cost.

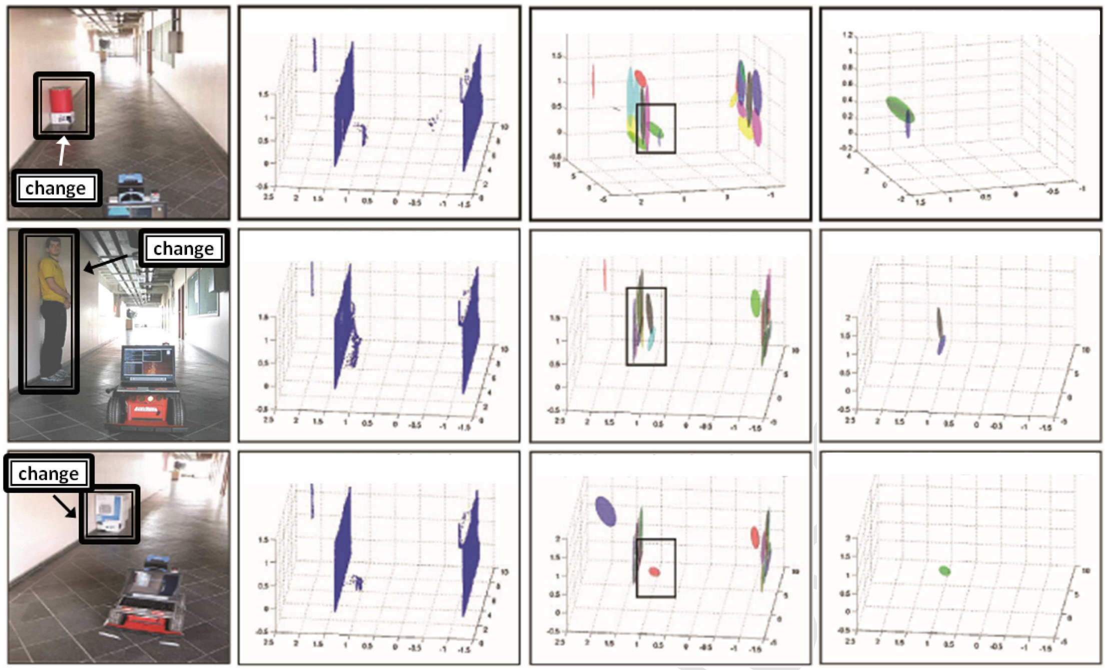
With this aim, sets of 3D laser range data collected by Pioneer robot (see Fig. 10-b) have been used. These datasets are shown in Fig. 11. For each 3D point, the GMM was calculated. Afterwards, each change detection algorithm was executed and the contingency tables (Table 1) were built. The results are shown in Table 2. In this table, "cylinder", "person" and "box" correspond to the first, second and third row in Fig. 11, respectively.

Table 2 shows the results obtained with the maps shown in Fig. 11, where the number of Gaussians selected by the methods is equal to 16. Both change detection algorithms have a low level of association strength (Cramer's V and uncertainty coefficient U less than 0.6) in both cases, the cylinder and the box, respectively. Moreover, EMD-based algorithm shown slightly better results than Structural matching one. This weakness of associations is due to the 3D map acquisition process: the ceiling suffers with parallax problem depending on the robot's path during the data acquisition stage. This produces a poor segmentation result using GMMs. Fig. 14 illustrates the GMM obtained using $K_{max} = 16$ for one of the experiments.

As it is shown in the first two columns in Table 2, EMD-based method has advantages over the structural matching algorithm. It is interesting to point out that changes in the robot scene are detected even in difficult conditions of the selected dataset. As it is shown in Fig. 14, the ceiling is segmented using the most of the Gaussians. However, the proposed method allows the detection of changes in a discriminative way, *i.e.* the changes are represented by different Gaussians. It is possible to see that both methods were able to detect changes with a few outliers in two cases (cylinder and box).

In order to improve the results, 3D points with Z coordinate larger than 1.5 meters were removed from the dataset, *i.e.* a distance constraint was used for removing points associated to the ceiling. In this way, the results are expressive and identical due to the segmentation obtained by GMMs, allowing both methods detecting correctly changes. The last two columns in Table 2 show the results obtained in this case, where the algorithm es-

ACCEPTED MANUSCRIPT



(a) (b) (c) (d)

Figure 12: Experiments using robotic 3D maps. Induced novelties were a cylinder (a-top), a person (a-middle), and a box (a-bottom). In (b), the acquired 3D maps are shown. The GMM representations of the maps containing the novelties are depicted in (c). The segmented novelties are depiced in (d).

timates GMMs using $K_{max} = 20$ and the number of Gaussians selected by the proposed algorithm is 12. The change detection is not ideal ($\kappa = 1$) because Gaussians that represent the changes do not segment all points in the ground truth.

Table 2: Performance comparison for the experiments with different novelty detection algorithms, considering full area and area of interest in the dataset. All results correspond to statistically significant correlation between the system response and the actual novelty status (χ^2 analysis, $p \leq 0.01$).

	Complete Dataset		Area of Interest	
	EMD	Str. Matching	EMD	Str. Matching
Cylinder	V=0.570 U=0.381 κ=0.254	V=0.474 U=0.238 κ=0.237	V=0.98 U=0.987 κ=0.987	V=0.98 U=0.987 κ=0.987
Person	V=0.991 U=0.996 κ=0.996	V=0.967 U=0.883 κ=0.848	V=0.983 U=0.992 κ=0.992	V=0.983 U=0.992 κ=0.992
Box	V=0.621 U=0.571 κ=0.503	V=0.543 U=0.479 κ=0.465	V=0.968 U=0.991 κ=0.991	V=0.968 U=0.991 κ=0.991

3.6. Sensitivity to Different Conditions and Parameters

In this section, the sensitivity of the proposed detection method to different situations and parameters is evaluated. Specifically, the approach is evaluated varying the size of the object that represents the change in the scene. Besides, the robot's point of view is modified in the data acquisition stage. For each situation, the number of Gaussians K was varied between 8 and 20.

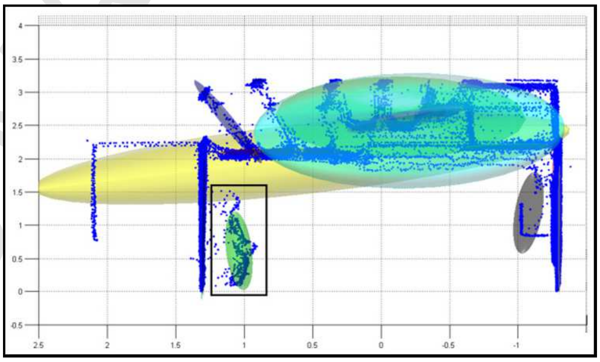
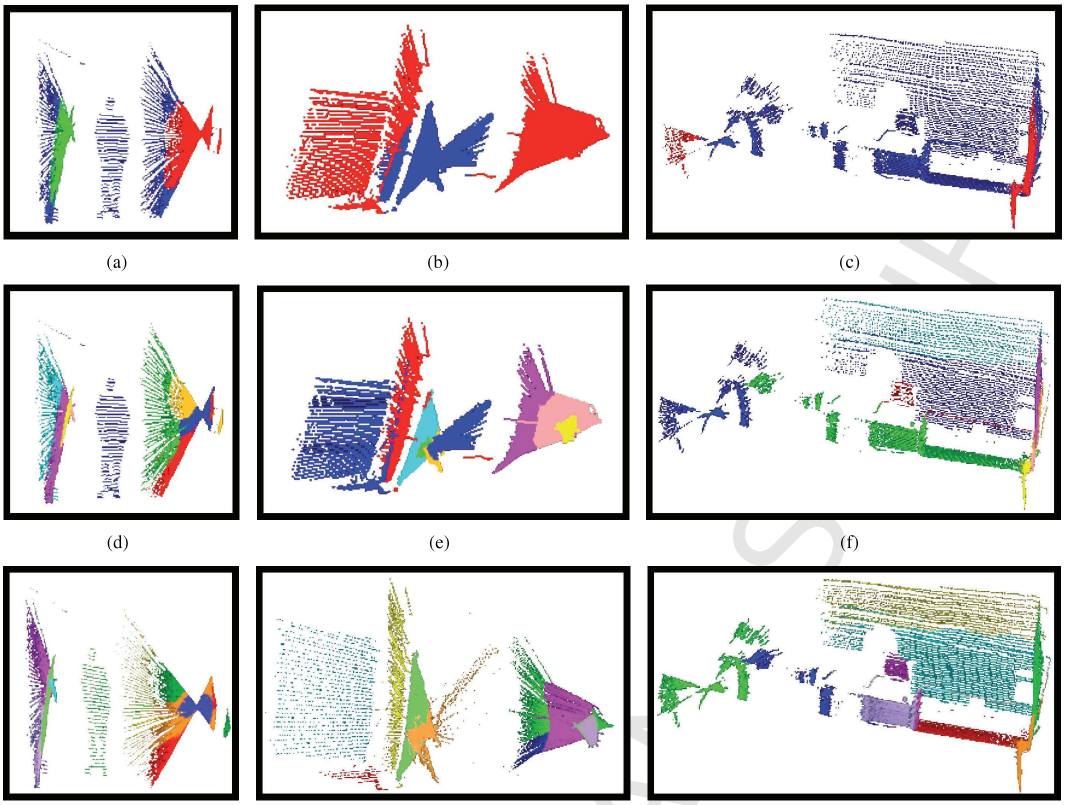


Figure 15: Frontal view of the Gaussian matematical space obtained from the complete datasets shown in Fig. 11-b using $K_{max} = 16$. The black rectangle is used to indicate the change.

Firstly, the robotic platform depicted in Fig. 10-c was used to evaluate the method in an office environment with boxes of different sizes, but in a similar position with respect to the robot. Twelve maps were acquired with the robot. Fig. 15 shows the real scene used for two of these experiments (Fig. 15-a) and the results calculated by the algorithm (15-d). Intermediate steps (*i.e.* 3D maps after acquisition stage and GMM representations) are illustrated in Fig. 15-b and Fig. 15-c. For these experiments, the number of Gaussian selected by our method is equal to 8 and 12, respectively. A comparative study using the

ACCEPTED MANUSCRIPT



(a) (b) (c) (d) (e) (f) (g) (h) (i)

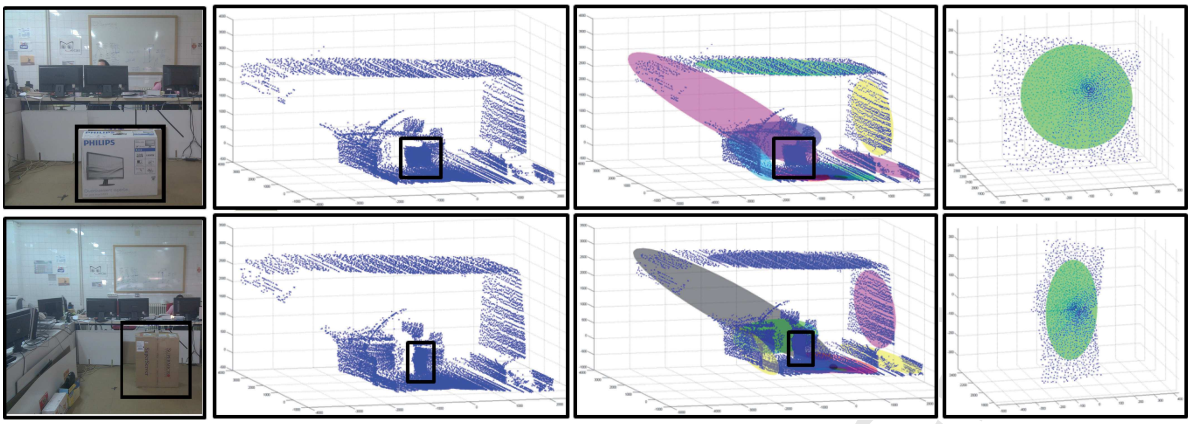
Figure 13: Comparison between K-means-based method and Gaussian Mixture Models in real datasets, in order to evaluate the ability of these methods to detect changes in 3D maps: (a-c) results using Loménie's method, considering the best result between 1 and 9 clusters; (d-f) results from Loménie's method but forcing 9 clusters; (g-i) GMMs results using the approach proposed herein with K_{max} = 9.

metrics presented in Sec. 3.1 was done. The average values V, U and κ are shown in Fig. 17-a and Fig. 17-b for the greedy EMD-based and the structural matching methods, respectively. Both methods were able to detect changes in the scene (see Table 1), though the number of outliers is significant. For both methods, the best change detection is attained using $K = 8$ and $K = 12$ Gaussians. For other values of K, the approach is able to detect changes but with some outliers. On the other hand, greedy-EMD algorithm and structural matching has similar average results for all values of K. Afterwards, in order to evaluate the influence of the robot's point of view with respect to the position of the change, a new experiment was carried out. Similarly to previous experiments, the map including the novelty (a human being) was acquired 12 times from different point of views (robot's localization is considered to be solved). In Fig. 16-a, the real scene of the environment is illustrated for two different experiments. Then, 3D maps are shown in Fig. 16-b. Fig. 16-c illustrates the GMM associated to the map. Finally, Gaussians and the points that represent the segmented novelties in the GMMs space are shown in Fig. 15-d. The same metrics were used to evaluate the performance of the algorithms. Results are shown in Fig. 17-c and Fig. 17-d for greedy-EMD and structural matching algorithms, respectively. Both methods are able to detect changes for different points of view with similar average values (*i.e.* U, V and κ). However, the quality of the detection is sensitive to the number of Gaussian and the results are largely dependent on the segmentation algorithm.

4. Conclusion and Future Works

This article described two methods to detect changes in 3D real environments for robot navigation. Real data acquired by laser scanners is preprocessed in order to reduce the size of the point clouds. *Gaussian Mixture Models* (GMMs) are used to obtain a new representation of the point clouds. It was validated and compared with a state-of-art algorithm in order to deal with real situations, where a good representation of 3D point cloud (maps) is required. A novel greedy algorithm based on Earth Mover's Distance metric and a structural matching algorithm are employed to quantify the existence of a novelty in the scene.

Results of the proposed algorithm demonstrate the reliability of the method in several real scenarios. Furthermore, the proposed techniques were compared in terms of robustness, accuracy and sensitivity. The methods are evaluated using the χ^2 analysis, as proposed by (Vieira Neto and Nehmzow, 2008).



(a) (b) (c) (d)

Figure 16: Experiments with two boxes of different sizes: (a) test sites – an office environment with a box (novelty) in two different sizes; (b) real observations from the Robex platform (novelties are manually segmented in black); (c) the GMM representation of the maps containing the novelties (black box); (d) Gaussian functions and point set associated to the segmented novelties.

The results of χ^2 analysis shown a small advantage of the EMD-based algorithm over structural matching algorithm in very difficult datasets. The proposed method may be easily extended to detect things removed from the scene, which is achieved by simply using the current map as the reference input.

The applicability in mobile robots was evaluated showing the capabilities of the method to be applied in real robots located by odometry and laser-based SLAM. In the present work, the robot localization is an important constraint but has smaller impact in the performance of the method in relation to point-to-point approaches. The features in the proposed method is based in GMM, thus two independent information are important: shape and location. Therefore, it is robust to small localization error. However, it could be interesting to take into account this problem when the localization error is significant.

The techniques presented in this article open new opportunities to develop automatic processes for surveillance of infrastructures, by proposing a new and robust approach for searching and detecting changes within large amounts of data.

Future work will be focused on the use of the current novelty detection algorithm in real field robotic applications, like surveillance or exploration of dangerous environments, where the presence and modeling of novelties could be important. Also, an extension of the Gaussian Mixture Models estimation method will be developed to work iteratively, so that data can be captured and processed online by the mobile robot. One possible approach is to use the GMM method with online learning (Kristan et al., 2008; Declercq and Piater, 2008). Furthermore, an efficient implementation in GPU of the GMM estimation, the computational bottleneck of the proposed framework, is being carried out by the authors. Another important improvement is to include a registration module in the system. It will allow the system to deal with over fitting volumes, and to avoid the strict need for having maps expressed in the same reference coordinates frame.

Acknowledgments

This work was partially supported by: EU projects IRPS (FP6-IST-45048) and PROMETHEUS (FP7-ICT-2007-1-214901); by the MICINN AIB2010PT-00149 project; by the Junta de Extremadura IB10062 project; by CAPES, FAPERGS, FAPEMIG, and CNPQ, Brazil; and by Portuguese Foundation for Science and Technology (FCT) and FEDER funds.

The authors would like to thank to the colleagues in C3-FURG, DCC-UFMG, ISR-UC, and to Prof. Nicolas Loménie who made available the source code of his method to be used in our tests.

References

Andreasson, H., Magnusson, M., Lilienthal, A.J., 2007. Has something changed here? autonomous difference detection for security patrol robots, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3429–3435.

Barrow, H.G., Burstall, R.M., 1976. Subgraph isomorphism, matching relational structures and maximal cliques. Information Processing Letter 4, 83–84.

Blanco, J., González, J., Férnandez-Madrigal, J., 2010. An Experimental Comparison of Image Feature Detectors and Descriptors applied to Grid Map Matching. Technical Report. University of Málaga, Málaga, Spain.

Chandola, V., Banerjee, A., Kumar, V., 2009. Anomaly detection: A survey. ACM Computing Surveys (CSUR) 41, 1–58.

Comon, P., 1994. Independent component analysis, a new concept? Signal Processing 36, 287–314.

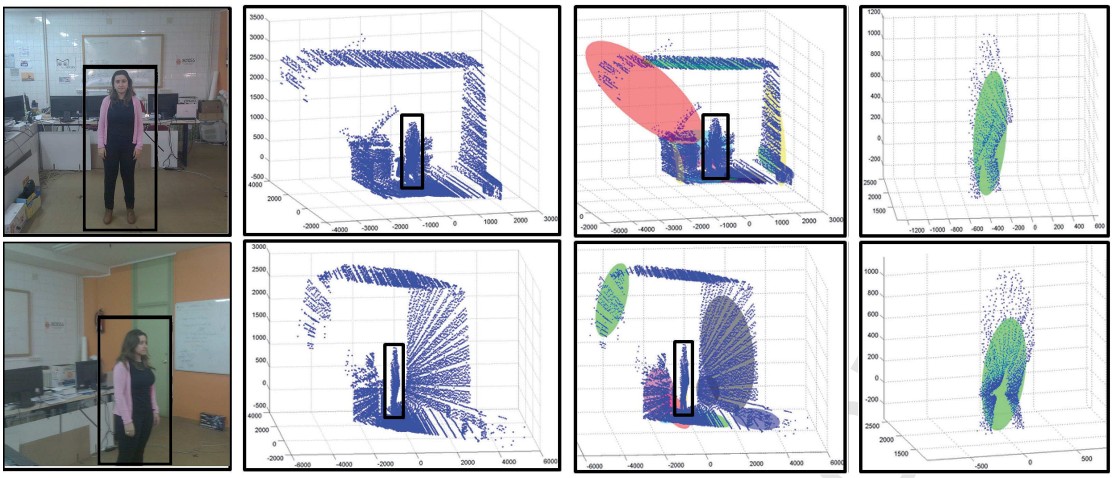
Declercq, A., Piater, J.H., 2008. Online learning of Gaussian mixture models - a two-level approach, in: Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, pp. 605–611.

Drews Jr, P., Núñez, P., Rocha, R., Campos, M., Dias, J., 2010. Novelty detection and 3D shape retrieval using superquadrics and multi-scale sampling for autonomous mobile robot, in: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Anchorage, Alaska, USA. pp. 3635–3640.

Eliazar, A., Parr, R., 2003. DP-SLAM: Fast, robust simultaneous localization and mapping without predetermined landmarks, in: Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI-03), pp. 1135–1142.

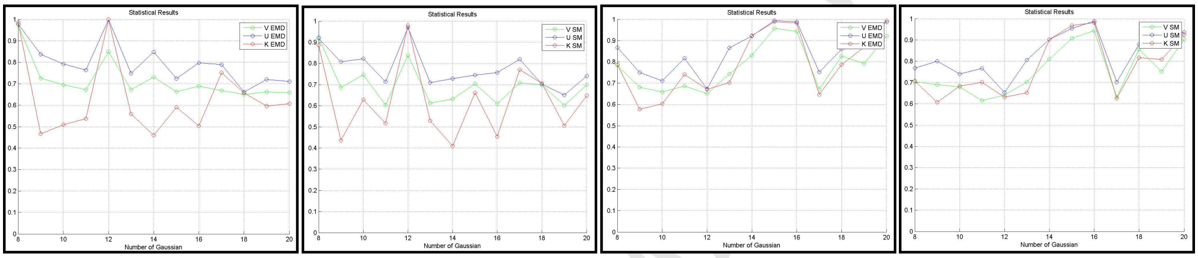
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



(a) (b) (c) (d)

Figure 17: Experiments with two different view points: (a) test sites – an office environment with a person as a change; (b) real observations from the third platform (novelties manually segmented in black); (c) GMM representations of the maps containing the novelties (black box); (d) Gaussian functions and point set representing the segmented novelties.



(a) (b) (c) (d)

Figure 18: Statistical Results (average values) using κ metric in two datasets: (a-b) box as a change shown in Fig. 15, using greedy EMD-based and structural matching algorithms, respectively; (c-d) person as a change shown in Fig. 16, using greedy EMD-based and structural matching algorithms, respectively.

Figueiredo, M., Jain, A., 2002. Unsupervised learning of finite mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23, 381–396.

Forstner, W., Moonen, B., 1999. A metric for covariance matrices. Technical Report. Dep. of Geodesy and Geoinformatics, University of Stuttgart, Germany.

Fraley, C., Raftery, A.E., 1998. How many clusters? which clustering method? answers via model-based cluster analysis. The Computer Journal 41, 578–588.

Gath, I., Geva, A., 1989. Unsupervised optimal fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, 773–781.

Goldberger, J., Gordon, S., Greenspan, H., 2003. An efficient image similarity measure based on approximations of KL-divergence between two Gaussian mixtures, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 487 – 493.

Gutiérrez, M., Martineta, E., Sánchez, A., Rodríguez, R., Núñez, P., 2011. A cost-efficient 3d sensing system for autonomous mobile robots, in: Proc. of XII Workshop of Physical Agents 2011 (WAF'2011), Albacente, Spain.

Helén, M., Virtanen, T., 2007. Query by example of audio signals using Euclidean distance between Gaussian mixture models, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007), pp. 225–228.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer Series in Statistics, Springer-Verlag Inc., New York, EUA. 2nd edition.

Kristan, M., Skocaj, D., Leonardis, A., 2008. Incremental learning with Gaussian mixture models, in: Proceedings fo the Computer Vision Winter Workshop (CVWW), pp. 25–32.

Lai, K., Fox, D., 2009. 3D laser scan classification using web data and domain adaptation, in: Proceedings of the Robotics: Science and Systems (RSS), pp. 1–8.

Loménie, N., 2004. A generic methodology for partitioning unorganised 3d point clouds for robotic vision, in: Proceedings fot the Canadian Conference Computer and Robot Vision, pp. 64–71.

Markou, M., Singh, S., 2003. Novelty detection: A review - part 1: Statistical approaches. Signal Processing 83, 2481–2497.

Marsland, S., Nehmzow, U., Shapiro, J., 2002. Environment-specific novelty detection, in: Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior on From Animals to Animats (ICSAB), pp. 36–45.

McLachlan, G.J., Peel, D., 2000. Finite mixture models. John Wiley & Sons.

Mufti, F., Mahony, R., Heinzmann, J., 2012. Robust estimation of planar surfaces using spatio-temporal ransac for applications in autonomous vehicle navigation. Robotics and Autonomous Systems 60, 16–28.

Núñez, P., Drews Jr, P. Bandera, A., Rocha, R., Campos, M., Dias, J., 2010. Change detection in 3D environments based on Gaussian Mixture Model and robust structural matching for autonomous robotic applications, in: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2633–2638.

Núñez, P., Drews Jr, P., Rocha, R., Campos, M., Dias, J., 2009. Novelty detection and 3D shape retrieval based on Gaussian Mixture Models for autonomous surveillance robotics, in: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS), pp. 4724–4730.

Pauly, M., Gross, M., Kobbelt, L.P., 2002. Efficient simplification of point-sampled surfaces, in: Proceeding of IEEE Visualization (VIS), pp. 163–170.

ACCEPTED MANUSCRIPT

Premebida, C., Ludwig, O., Nunes, U., 2009. LIDAR and vision-based pedestrian detection system. Journal of Field Robotics 26, 696–711.

Reynolds, D.A., Rose, R.C., 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Transactions on Speech and Audio Processing 3, 72–83.

Rissanen, J., 1983. A universal prior for integers and estimation by minimum description length. The Annals of Statistics 11, 416–431.

Rubner, Y., Tomasi, C., Guibas, L.J., 2000. The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision 40, 99–121.

Rusu, R., Marton, Z., Blodow, N., Dolha, M., Beetz, M., 2008. Towards 3D point cloud based object maps for household environments. Robotics and Autonomous Systems 56, 927–941.

Sofman, B., Bagnell, J.A., Stentz, A., 2010. Anytime online novelty detection for vehicle safeguarding, in: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 1247–1254.

Sofman, B., Neuman, B., Stentz, A., Bagnell, J.A., 2011. Anytime online novelty and change detection for mobile robots. Journal of Field Robotics 28, 589–618.

Stamos, I., Liu, L., Chen, C., Wolberg, G., Yu, G., Zokai, S., 2008. Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes. International Journal of Computer Vision 78, 237–260.

Tarassenko, L., Hayton, P., Cerneaz, N., Brady, M., 1995. Novelty detection for the identification of masses in mammograms, in: Proceedings of International Conference on Artificial Neural Networks (ICANN), pp. 442–447.

Thrun, S., Fox, D., Burgard, W., Dellaert, F., 2001. Robust Monte Carlo localization for mobile robots. Artificial Intelligence 128, 99–141.

Tomasi, C., Rubner, Y., Guives, L., 1998. A metric for distributions with applications to image databases, in: Proceedings of International Conference on Computer Vision (ICCV), pp. 59–66.

Vieira Neto, H., Nehmzow, U., 2007. Visual novelty detection with automatic scale selection. Robotic and Autonomous Systems 55, 693–701.

Vieira Neto, H., Nehmzow, U., 2008. Visual novelty detection for autonomous inspection robots, in: Takahashi, Y. (Ed.), Service Robot Applications. I-Tech Education and Publishing, Vienna, Austria, pp. 309–330.

Yu, L., Liu, H., 2004. Efficient feature selection via analysis of relevance and redundancy. Journal of Machine Learning Research 5, 1205–1224.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biography of each author**

Paulo Lilles J. Drews Jr is M.Sc. in Computer Science at Federal University of Minas Gerais(UFMG), Brazil. His master thesis was focused in change detection and shape retrieval in 3D Maps. He has graduated in Computer Engineer in 2007 at Federal University of Rio Grande (FURG), Brazil. His main research interests are computer vision, image processing, machine learning and robotics. He was a researcher at the ISR Coimbra's Mobile Robotics Lab, collaborating on Prometheus project and on the IRPS project. Currently, he is assistant professor at FURG and a PhD student in Computer Science at Federal University of Minas Gerais.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biography of each author**

Pedro Núñez is Assistant Professor associated to Tecnología de los Computadores y las Comunicaciones
department, at University of Extremadura (Spain). He is Telecommunications Engineer since 2003, and he obtained his PhD Telecommunications degree in 2008, at the University of Malaga (Spain). During these years, he developed his research as member of the group of Ingeniería de Sistemas Integrados (ISIS) at University of Malaga. Between 2008 and 2010, he was a Post-Doc researcher at Institute of Systems and Robotics of University of Coimbra (Portugal). In 2007, he started to work as professor at University of Extremadura, whitin the Robolab (Laboratorio de Robótica y Visión Artificial) group. Currently, he is leader of different research projects related to social robots (APSUBA and MI-BOT).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biography of each author**

Rui P. Rocha is an Assistant Professor with the Department of Electrical Engineering and a Senior Researcher at the Institute of Systems and Robotics (ISR), both at the University of Coimbra, Portugal. He received the Engineering degree, the M.Sc. degree, and the Ph.D. degree in Electrical and Computer Engineering from University of Porto in 1996, 1999 and 2006, respectively. His main research topics are cooperative robotics, 3-D map building, distributed control, cybernetic transportation systems, and multi-agent systems. He has been involved on several FP6 and FP7 European funded projects developed
in consortium for the past few years.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biography of each author**

Mario Fernando Montenegro Campos is a Professor of Computer Vision and Robotics in the Department of Computer Science at the Federal University of Minas Gerais(UFMG), Brazil. He holds B.S. degrees in Engineering, and M.S. in Computer Science, all from the UFMG, and a Ph.D. in Computer and Information Science from the University of Pennsylvania. His research interests include cooperative robotics, robot vision, sensor information processing. His main contributions are in haptics, multirobot cooperation, aerial robotics and robot vision. He is the founder and director of the Vision and Robotics Lab(VeRLab). He has been a Distinguished Lecturer in the IEEE RAS.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biography of each author**

Jorge Dias has a Ph.D. degree on Electrical Engineering at University of Coimbra, specialization in Control and Instrumentation, November 1994. Jorge Dias conducts his research activities at the Institute of Systems and Robotics at University of Coimbra. His research area is Computer Vision and Robotics, with activities and contributions on the field since 1984. He has several publications on Scientific Reports, Conferences, Journals and Book Chapters. Jorge Dias teaches several engineering courses at the Electrical Engineering and Computer Science Department, University of Coimbra. He is responsible for courses on Computer Vision, Robotics, Industrial Automation, Microprocessors and Digital Systems.

ACCEPTED MANUSCRIPT

***Photo of each author**

***Photo of each author**

***Photo of each author**

ACCEPTED MANUSCRIPT

***Photo of each author**

***Photo of each author**