# Learning Moving Objects in a Multi-Target Tracking Scenario for Mobile Robots that use Laser Range Measurements

Polychronis Kondaxakis, Haris Baltzakis, Panos Trahanias
Institute of Computer Science
Foundation for Research and Technology – Hellas (FORTH)
E-mail:{konda, xmpalt, trahania}@ics.forth.gr

*Abstract*— **This paper addresses the problem of real-time moving-object detection, classification and tracking in populated and dynamic environments. In this scenario, a mobile robot uses 2D laser range data to recognize, track and navigate moving targets. Most previous approaches either rely on pre-defined data features or off-line training of a classifier for specific data sets, thus eliminating the possibility to detect and track different-shaped moving objects. We propose a novel and adaptive technique where potential moving objects are classified and learned in real-time using a Fuzzy ART neural network algorithm. Experimental results indicate that our method can effectively distinguish and track moving targets in cluttered indoor environments, while at the same time learning their shape.**

## I. INTRODUCTION

TARGET tracking in mobile robotics is an essential attribute but at the same time it presents to the designer with a very complicated problem. It usually requires a combination of sensors and algorithms in order, for a moving observer, to detect and track multiple moving targets in cluttered and dynamic indoor environments. The problem becomes even more complicated if hardware augmentation and sensor fusion (e.g. cameras with laser scanner) is not one of the designer's options.

In this paper we consider a multi-target tracking scenario where a mobile robot utilizes 2D laser range measurements to maneuver possible moving objects in its space. To achieve this task, the deployed algorithm has to successfully extract and initiate tracks for only moving targets. The key idea behind this work is to filter the raw sensor data through a number of stages to facilitate a robust moving object extraction and learning algorithm. There are three main steps on the moving data detection procedure. The background map attenuation stage, the track initiation process and finally, the moving-object's shape classification and learning. The last stage is the main contribution of this paper. It achieves not only accurate track initiation, but at the same time the robot learns in real-time the shape of any moving target in its environment.

Previous approaches that use laser data such as [1] and [2], detect moving objects on the mapping level by employing multiple probabilistic grid maps thus quantizing in 2D the robot's environment. The performance of grid-map architectures was extensively tested in [3]. However, in general, grid-map techniques demand a large amount of processing power and suffer from quantization errors. These quantization outliers become even more intense when the robot observer is moving freely in its space. An alternative method that does not involve grid-map techniques was reported in [4]. Although the idea of this approach is similar to the background attenuation stage of our algorithm, it lacks the very critical track initiation and classification levels of our approach.

Track initiation techniques have been reported in [5], [6], [7] and [8]. The first two approaches ([5], [6]), formulate the track initiation problem as multiple composite hypothesis testing using the maximum likelihood estimator with probabilistic data association (ML-PDA). In [5], the hypotheses are based on the minimum description length (MDL) criterion, and in [6] a new method called optimal gating is proposed. Furthermore, the authors in [7] propose a track initiation method based on Hough Transform (HT), and in [8], enumeration strategies (kd-trees) are utilized to link together observations from different time steps and initiate tracks. Acknowledging that a track initiation algorithm could provide significant moving target filtering in dynamic and cluttered environments, the proposed algorithm is augmented by the method reported in [3].

Finally, geometric feature analysis is a fundamental technique for detecting moving objects using 2D laser data. Available methods so far, approach the problem of detecting geometric features such as lines, circles, legs by manual design and threshold hand-tuning [9]. Moreover, the extracted features are compared to objects stored in a database to recognize targets [10]. Alternatively, a more adaptive approach can be found in [11] and [12] where the authors train, in off-line mode, the AdaBoost classifier to identify peoples' legs. However, this off-line training prohibits the system to learn new moving objects, meaning that in case of a different shaped moving target appears (another robot), the detection mechanism fails.

Our approach combines features from all the above techniques producing an innovative three-stage moving object detection algorithm. Furthermore, by augmenting a Fuzzy ART neural network classifier [13] with object

This work is supported by the FP6-Information Society Technologies program, European Commission INDIGO research project (Contract No. IST-045388).

categorization functions, we achieve on-line system training. This expands the available system with all-moving object detection and learning characteristic.

Conclusively, the filtered laser scanner measurements that represent moving targets are then become available to a Joint Probabilistic Data Association with Interacting Multiple Model (JPDA-IMM) algorithm, which performs the actual trajectory tracking.

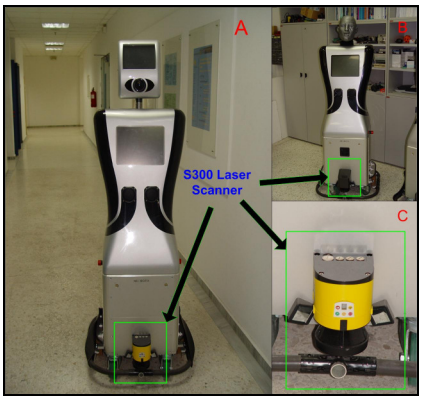


Fig. 1. (A) The NEOBOTIX ME-470 mobile robot platform main configuration. (B) An alternative configuration augmented with an interactive robotic head. (C) The SICK S300 laser range scanner.

## II. Stationary Background Attenuation

### A. Laser Data Collection

This research was conducted using a SICK S300 Laser Range Scanner adjusted on the base of a NEOBOTIX ME-470 mobile robot platform (Fig. 1). The scanner has an angular resolution of 0.5° and, although being capable to provide 270° scans, in our experiments it was configured to exploit only the 180° central span since the sides were blocked by the front face of the robot (Fig. 1A). Therefore, we utilize 361 laser measurements with 30-70*mm* range accuracy independent of the actual range. Optimally, this scanner can provide 12.5 frames per second (80*ms* minimum response time) with a maximum range of 30*m*. Moreover, as a result of the indoor operational specifications in our experiments, the laser was configured to operate at a maximum of 9.6*m* distance. We represent a measurement frame at time index $k$ by a vector $\mathbf{Z}_{laser}(k)=\{\mathbf{z}_1(k),\mathbf{z}_2(k),\ldots,\mathbf{z}_{361}(k)\}$, where $\mathbf{z}_l(k)$ with $l=1\ldots361$, is a vector composed by distance and bearing readings, $\mathbf{z}_l(k)=[r_l(k),\varphi_l]^T$. The bearing component is obtained from the fixed vector-array, $\mathbf{\Phi}=\{(\varphi_1=0^\circ),(\varphi_2=0.5^\circ)\ldots(\varphi_{361}=180^\circ)\}$.

### B. 1D Range Mapping and Non-Static Object Detection

Stationary background-map attenuation presents the first step on extracting moving objects in a dynamic environment. The proposed methodology finds non-static objects by detecting occupation of previously unoccupied space. Here, instead of dividing the world in a 2D occupancy grid map, a vector array, which is composed by 361 elements, accumulates background range measurements. This 1D representation of robot's background space offers a powerful and at the same time a less demanding (in terms of computational resources) approach.

More specifically, the robot does not store an absolute map of its surroundings, as it was the case in [3]. It performs a one-to-one measurement frame alignment and comparison from a previous time step $\mathbf{Z}_{laser}(k-1)$ to the current measurement frame $\mathbf{Z}_{laser}(k)$, using its odometry and an Iterative Closest Point (ICP) algorithm [14].

The two-frame $(\mathbf{Z}_{laser}(k-1),\mathbf{Z}_{laser}(k))$ alignment and comparison procedure is as follows:

**1.** From odometry, we obtain the robot pose $\mathbf{X}_R(k)=[x_R(k),y_R(k),\theta_R(k)]^T$ (position and orientation). It has been assumed that pose is acquired simultaneously at every laser measurement frame. The range scan points of the previous measurement frame, $\mathbf{Z}_{laser}(k-1)$ are transformed in the current coordinate frame of measurement vector, $\mathbf{Z}_{laser}(k)$ using (1).

$$\mathbf{z}_l^*(k|k-1)=\begin{bmatrix}x_l^*(k|k-1)\\ y_l^*(k|k-1)\end{bmatrix}= \begin{bmatrix}r_l(k-1)\cos(\varphi_l+\theta_R(k)-\theta_R(k-1))+x_R(k)-x_R(k-1)\\ r_l(k-1)\sin(\varphi_l+\theta_R(k)-\theta_R(k-1))+y_R(k)-y_R(k-1)\end{bmatrix} \quad (1)$$

with $l=1\ldots361$, $\mathbf{Z}_{laser}^*(k|k-1)=\{\mathbf{z}_1^*(k|k-1),\ldots,\mathbf{z}_{361}^*(k|k-1)\}$ [1].

**2.** To counteract additional translation and rotation point-set differences between vectors $\mathbf{Z}_{laser}(k)$ and $\mathbf{Z}_{laser}^*(k|k-1)$, an Iterative Closest Point (ICP) algorithm is utilized. A closed-form solution of the ICP algorithm can be found in [14]. This particular version is augmented by kd-trees enumeration technique. Therefore, by providing increased throughput in data selection, the augmented ICP becomes ideal for real-time implementations. Its output is a translation ($\mathbf{T}_{2\times1}$) and rotation ($\mathbf{R}_{2\times2}$) matrix.

**3.** The background map (Fig. 2(A)) is represented by a vector $\mathbf{V}_{back}(k)=\{\boldsymbol{\rho}_1(k),\ldots,\boldsymbol{\rho}_{361}(k)\}$ where $\boldsymbol{\rho}_l(k)$ is a vector composed by maximum distance readings and the fixed bearing component, $\boldsymbol{\rho}_l(k)=[r_{lback}(k),\varphi_l]^T$.

For moving-object acquisition and background-map update, we propagate the background vector from time index $k$-1 to the current step. This is achieved by applying odometric (same as (1) by substituting $r_l(k-1)=r_{lback}(k-1)$) and ICP (3) transformations to its components.

$$\boldsymbol{\rho}_l^*(k|k-1)=\begin{bmatrix}r_{lback}^*(k|k-1)\\ \varphi_{lback}^*(k|k-1)\end{bmatrix}=\text{cart2pol}\left\{\left(\mathbf{R}_{2\times2}\times\begin{bmatrix}x_{lback}^*(k|k-1)\\ y_{lback}^*(k|k-1)\end{bmatrix}\right)+\mathbf{T}_{2\times1}\right\} \quad (2)$$

[1] The asterisk (*) on any vector variable designates that the vector has gone through some short of geometric transformation due to robot's odometry.

where $\mathbf{T}_{2\times1}$ and $\mathbf{R}_{2\times2}$ are obtained from part **2**.

Next, the background vector is reinitialized and its transformed members, which are calculated by (2), are aligned into a new background array:

$$\mathbf{V}^*_{\left[\text{floor}\left(\left(\varphi^*_{l_{back}}(k|k-1)\times 361\right)/\pi\right)+1\right]_{back}}(k|k-1)=\begin{bmatrix} r^*_{l_{back}}(k|k-1) \\ \varphi_{\left[\text{floor}\left(\left(\varphi^*_{l_{back}}(k|k-1)\times 361\right)/\pi\right)+1\right]} \end{bmatrix} \quad (3)$$

with $l=1\ldots361$ and $0^\circ \le \varphi^*_{l_{back}}(k|k-1) \le 180^\circ$.

Given the current measurement frame $\mathbf{Z}_{laser}(k)$ (Fig. 2(B)), the elements of the background map are updated as follows:

$$\boldsymbol{\rho}_l(k)=\begin{bmatrix} r_{l_{back}}(k) \\ \varphi_l \end{bmatrix}=$$

$$\begin{cases} [r_l(k),\varphi_l]^T \text{ if } r^*_{l_{back}}(k|k-1)+\delta+\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} > r_l(k) \ge r^*_{l_{back}}(k|k-1)-\delta-\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} \\ [r_l(k),\varphi_l]^T \text{ if } r^*_{l_{back}}(k|k-1)+\delta+\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} < r_l(k) \\ [r^*_{l_{back}}(k|k-1),\varphi_l]^T \text{ if } r^*_{l_{back}}(k|k-1)-\delta-\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} > r_l(k) \end{cases} \quad (4)$$

where $\delta = 0.4m$ is a range tolerance threshold value and $\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}}$ is an adaptive variance component obtained from ten sequential background range values for each $l$. This component provides additional uncertainty when the robotic observer is moving, by dynamically adjusting $\delta$.

Moving points in absolute robot coordinates are given by:

$$\mathbf{z}^*_l(k)=\begin{bmatrix} x^*_l(k) \\ y^*_l(k) \end{bmatrix}=\begin{bmatrix} r_l(k)\cos(\varphi_l+\theta_R(k))+x_R(k) \\ r_l(k)\sin(\varphi_l+\theta_R(k))+y_R(k) \end{bmatrix} \quad (5)$$

$$\text{if } \; r^*_{l_{back}}(k|k-1)-\delta-\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} > r_l(k)$$

Therefore, the remaining moving-object points (Fig. 2(B)) are represented as $\mathbf{Z}^*_{mov}(k)=\{\mathbf{z}^*_1(k),\mathbf{z}^*_2(k),\ldots,\mathbf{z}^*_s(k)\}$ with $\mathbf{Z}^*_{mov}(k)\subseteq\mathbf{Z}^*_{laser}(k|k-1)$ and $s\le361$.

### C. *Keeping Background Map Consistency*

To this point, the proposed stationary background attenuation method presents a simple but yet efficient non-static object detection algorithm. However, when the robot is moving the alignment between the background vector $\mathbf{V}_{back}(k-1)$ and $\mathbf{Z}_{laser}(k)$, is not perfect due to odometric errors. The ICP counterbalances these errors, keeping background-map integrity, but still outliers cannot be totally eliminated.

An innovative approach refreshes the map when needed, providing the appropriate consistency and robustness. The idea is to generate an additional array vector, which will be called "background-weight array" and is given by: $\mathbf{W}_{weight}(k)=\{w_1(k),\ldots,w_{361}(k)\}$. These weights accommodate a consistency belief of the background-map's elements and can take values from 1 to 20. Each weight component (out of 361) is assigned to a particular location in the background vector array (also composed by 361 locations) and they increase/decrease linearly according to (6).

$$w_l(k)=$$

$$\begin{cases} w_l(k-1)+1<20 \text{ if } r^*_{l_{back}}(k|k-1)+\delta+\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} > r_l(k) \ge r^*_{l_{back}}(k|k-1)-\delta-\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} \\ 20 \text{ if } r^*_{l_{back}}(k|k-1)+\delta+\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} < r_l(k) \\ w_l(k-1)-1\ge1 \text{ if } r^*_{l_{back}}(k|k-1)-\delta-\sigma_{\eta_{back}\{(k-1)\cdots(k-11)\}} > r_l(k) \end{cases} \quad (6)$$

When a weight value decreases to 1, the algorithm substitutes at the background vector the corresponding range-data component with the latest range reading ($r_l(k)$). Therefore, this technique renews the background map prohibiting further deterioration.

It is worth noticing that weights representing occluded background areas from verified targets, maintain the same values ($w_{l_{occluded}}(k)=w_{l_{occluded}}(k-1)$).

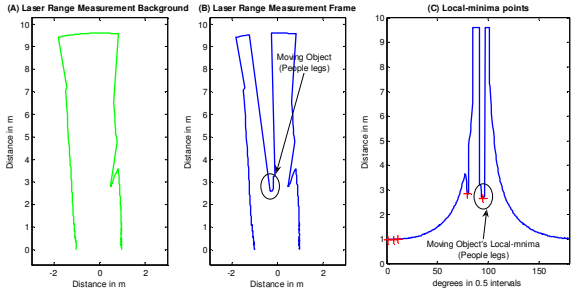


Fig. 2. (A) A background map of a corridor. (B) The same corridor, as described by a laser measurement frame including a moving target. (C) Extracted local-minima points from corridor laser measurements.

## III. Data Clustering and Track Initiation

The second stage of the proposed moving object detection algorithm is the data clustering and track initiation procedure.

To initiate target tracking, the segmentation of the raw data points included in vector set $\mathbf{Z}^*_{mov}(k)$ provides clusters of scan points that are grouped together. Each cluster represents a potential moving target, human or otherwise. Here, the algorithm adopts the distance clustering procedure presented in [3]. Clusters are represented as $\mathbf{C}_{SEG}(k)=\{\mathbf{c}_1(k),\mathbf{c}_2(k),\ldots,\mathbf{c}_q(k)\}$ with $t=1\ldots q$ and $\mathbf{C}_{SEG}(k)\subseteq\mathbf{Z}^*_{mov}(k)$ where each cluster can acquire different number of vector points and is represented as $\mathbf{c}_t(k)=\{\mathbf{z}^*_1(k),\mathbf{z}^*_2(k),\ldots,\mathbf{z}^*_p(k)\}$ with $p\in\mathbf{N}$.

In comparison to [3], the proposed moving object detection technique is enhanced by a local-minima filtering algorithm. The algorithm applies a convolution operator for data smoothing in $\mathbf{Z}_{laser}(k)$ and obtains a local-minima vector $\mathbf{L}^*_{\min}(k)=\{\mathbf{l}^*_1(k),\mathbf{l}^*_2(k),\ldots,\mathbf{l}^*_{\varepsilon361}(k)\}$, $\mathbf{L}^*_{\min}(k)\subseteq\mathbf{Z}^*_{laser}(k)$. Thus, a filtered version of the moving object cluster vector $\mathbf{C}_{SEG}(k)$ is obtained as follows:

$$\tilde{\mathbf{C}}_{SEG}(k)\leftarrow\mathbf{c}_t(k) \;\text{ if }\; \mathbf{c}_t(k)\cap\mathbf{L}^*_{\min}(k),\; t=1\ldots q \quad (7)$$

Fig. 2(C) displays with a red cross, the extracted local-minima points from a data frame $\mathbf{Z}_{laser}(k)$.

A centre-of-gravity (COG) point-vector is maintained for filtered every cluster. Thus, $\overline{\mathbf{C}}_{SEG}(k)=\{\bar{\mathbf{c}}_1(k),\bar{\mathbf{c}}_2(k),\ldots,\bar{\mathbf{c}}_q(k)\}$ corresponds to the centre-of-gravity vector. Having acquired specific moving targets from data clustering, the algorithm initiates new tracks as described in [3].

## IV. On-line geometric classification

A novel contribution of this publication is the implementation of a third-stage moving-object filtering mechanism where a real-time learning system categorizes the clusters into groups. These clusters are obtained in the track-initiation procedure and represent potential moving objects. The classification is based on a geometric profile of each available cluster and it uses the same fourteen feature definitions as the ones described in [11]. To achieve on-line categorization and learning, we needed a model classifier capable of rapid stable learning of recognition categories in response to arbitrary sequences of analog input patterns.

The Fuzzy ART neural network ([13]) is a vector classifier. The input vectors are categorized depending on similarities with any previous patterns. If the similarity criterion is not satisfied, a new category is created. Thus, the classifier solves the dilemma between plasticity and stability by enabling the learning to occur only in a resonance state. In this way, the network learns new information without destroying the old one. Moreover, the Fuzzy ART neural network exploited herein is capable of learning stable recognition categories in response to both analog and binary input patterns. A more detailed explanation of the Fuzzy ART algorithm can be found in [13].

The fourteen analog geometric features (e.g. number of points, width, linearity, circularity, radius, etc.) of each cluster are inserted directly to the classifier with only minimum data normalization pre-process (the analog input components must be in-between the interval [0, 1]).

A key assumption/fact in this research is that the available clusters, for which a track has been initiated, mainly represent moving objects. There is only a minimum and not frequent presence of outlier-clusters in $\tilde{\mathbf{C}}_{SEG}(k)$ with an initiated track. A voting system assigns a belief-state to the categories learned by the classifier and distinguishes between real moving targets and outlier targets.

More specifically, whenever a moving object is verified as target by either track initiation or data association (JPDA) process, its geometric profile is learned by the classifier and a category number becomes available. A voting vector, $\mathbf{B}_{vote}(k)=\{b_1(k),\ldots,b_r(k)\}$ with $r$ = number of initiated categories, accumulates the belief-variables. These variables contain cumulative counts, which express the confidence for a category to originate from a moving target. They are increased linearly (from $b_{1_{MIN}}(k)=0$ to $b_{1_{MAX}}(k)=50$) every time the Fuzzy ART algorithm classifies an input-cluster into one of the available categories.

Clusters that, due to Fuzzy ART network classification, belong to dominant categories are recognized as moving targets. A category is defined as dominant if, at time index *k*, its associated belief-variable is inside a threshold margin. For example, we say that category 2 is dominant and represents a moving object if and only if the following condition is validated:

$$b_{higher}(k)\geq b_2(k)\geq b_{higher}(k)-15 \qquad (8)$$

where $b_{higher}(k)$ is the belief-variable from $\mathbf{B}_{vote}(k)$ with the higher accumulated value at time index *k*.

Last but not least, a forgetting factor is applied to the voting vector meaning that as time progresses, the counts of every belief-variable decrease linearly and at a constant rate until they reach a zero value (one count per ten algorithm iterations). Thus, this technique dynamically updates the available moving object knowledge, eliminating at the same time possible outliers. A number of initiated categories with their associated belief-variables is displayed in figure 4.

## V. The tracking Algorithm

The three-stage non-stationary object detection process, as described in previous sections, filters out the row laser data providing the tracking algorithm only with the initiated tracks of moving objects. As were the case in [3], this work also deploys a JPDA-IMM hybrid filter to estimate the trajectories of the acquired targets. A thoroughly presentation of the JPDA-IMM algorithm is provided in [15]. The algorithm utilizes two linear dynamic motion models: (a) the Constant Velocity (CV) model and (b) the Coordinated Turn (CT) model with constant angular velocity rate.

## VI. Results

To illustrate the performance of the proposed architecture, a series of tests were conducted using real laser data collected from a SICK S300 laser scanner as described in section II-*A*. The mobile robot (Fig. 1(A)), is configured to run at a maximum translational speed of $0.3\,m/\sec$ and a maximum rotational speed of $10\,\deg/\sec$. Based on [3], the JPDA-IMM estimator was tuned to detect targets with initial velocities that are mutch smaller than the abiove-mentioned robot's.

For this experiment, the Fuzzy ART neural network was configured to learn up to ten categories. In this way, it can grow to encompass new data according to the vigilance parameter, which defaults to 0.85. By choosing a larger

vigilance parameter (e.g. closer to one), we can narrow the algorithm's morphological generalization for each category. Therefore, the output of the ART algorithm a distinctive probabilistic profile (Fig. 3(A), (B)), which is unique for each moving object class. The bias is set to $10^{-6}$, with a maximum of 100 epochs and the learning rate is set to 1.0 (fast-learning).

A tracking scenario that lasted several minutes included a two people and a robot (Fig. 1(B)) as potential targets. Their laser data signatures are displayed in figures 3 (D) and (C) respectively. According to this scenario's sequence, the robotic platform displayed in Fig. 1(A) moved freely along a laboratory corridor. The algorithm runs for 1200 iterations (3min) delivering 854 frames where laser data were successfully associated to tracks and learned by the classifier. It is worth mentioning that during this experiment, the algorithm detected no false positives providing 100% of tracking accuracy. However, there were some isolated situations where moving targets could not be correctly detected and thus the algorithm had some difficuklty initialing a new track. This is due to the regular appearance of one of the moving objects (the robot) during a specific segment of the sequence. In that case, its distinctive profile dominated the voting-vector $\mathbf{B}_{vote}(k)$, suppressing any other category information. For objects with different geometric profiles, the algorithm needed time to readjust its voting-vector and to allow new track establishment. Fig. 3 (A) and (B) provide the probabilistic signatures for the frame of the robot which is tracked in this sequence and for the peoples' legs respectively. These profiles were obtained by separately classifying each established track to one of human-legs or robot-frame categories. Having acquired different voting-vectors for each classified track (e.g. $\mathbf{B}^{d}_{Robot}(k)$ for *d*-robot tracks and $\mathbf{B}^{e}_{Legs}(k)$ for *e*-human leg tracks), we added the voting-vectors of each category and normalized them to one. The two bar-plots provide a rigorous proof that objects with different morphological characteristics are classified as different histograms by the Fuzzy ART neural network algorithm. Moreover, the obtained results support our initial assumption that moving targets appear more frequently than outliers in a dynamic environment and thus can be learned and filtered-out using a neural network classifier.

The bar plot presented in Fig. 4, is the voting-vector developed at the end of the 3-min experiment. Each bar represents an initiated category and its related belief-variable. The larger accumulated belief-variable is 47 (Category 3), which is near the upper limit variable ($b_{1_{MAX}}(k)=50$). Based on the above conclusions, we achieved to distinguish the clusters representing people's legs (smaller geometric profile) and are associated to categories 3, 7 and 8. On the other hand, a robot-cluster

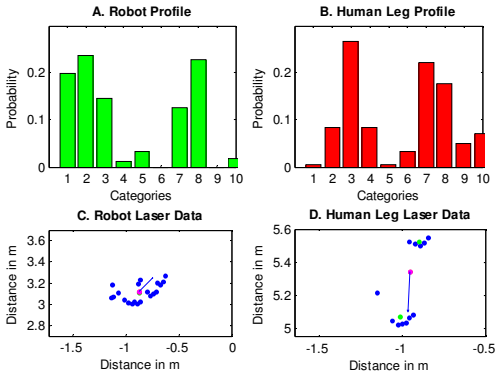


Fig. 3. (A), (B) Robot & Human Leg probabilistic profile for ten available categories. (C), (D) Robot & Human Leg laser data signatures.

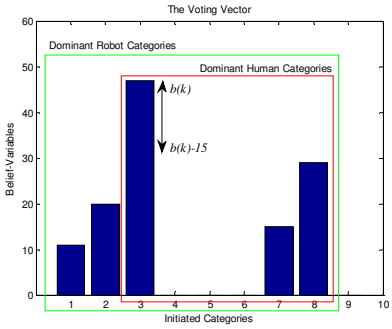


Fig. 4. The voting vector at the end of the 3-min experiment.

(larger geometric profile), is mainly described by categories 1, 2, 3, 7 and 8. Therefore, it is fair to say that in the proposed experiment, the human profile presents a distinguishable subclass of the robot profile.

Next, the tracking efficiency of the JPDA-IMM estimator is displayed in figures 5 and 6. Because the true pose vectors of the targets ("ground" truth) are not known, estimation errors cannot be evaluated. Instead, the root mean square (RMS) prediction errors are used as a performance criterion. Fig. 5(A) illustrates the RMS prediction errors (explained in Appendix). The tracker has a maximum error of 0.54*m* and a mean value of 0.112*m*. Its standard deviation error value is 0.09*m*. The Time-Average NIS test is presented in Fig. 6. To achieve low variability of the test statistic based on a single run the NIS is averaged in over twenty (*K*=20) future steps. The blue line represents the upper bound of the (one-sided) 95% probability concentration region. In cases that the filter performs correctly (which is the case in our experiments), the averaged normalized innovation squared (Appendix) is expected to fall inside that region. Note the much narrower range of the interval corresponding to larger values of active tracks (Fig. 5(B)). This illustrates the variability reduction in such repeated tracks.

## VII. CONCLUSION

This paper presents a novel approach on moving object

detection, recognition and learning using a single 2D laser range scanner. The proposed technique dynamically learns the shape of new potential moving targets and initiates a multi-target tracking mechanism for trajectory estimation. The results indicate that both the Fuzzy ART classifier and JPDA-IMM estimator are robust to dynamic changes delivering an accurate tracking system.

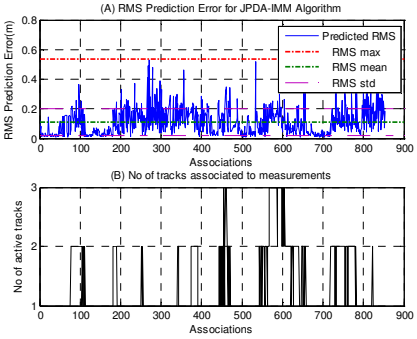


Fig. 5. (A) Predicted RMS error, (B) Number of active tracks.

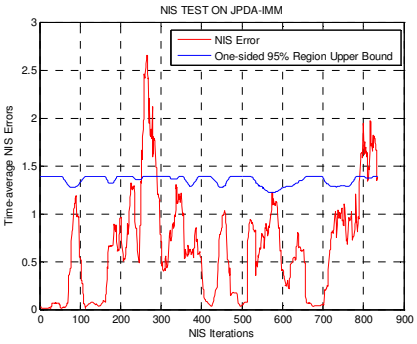


Fig. 6. The NIS test verifies JPDA-IMM algorithm's stability.

Future research will include further analysis on track categorization as well as improvements on all three filtering and classification levels, providing an even more precise knowledge on the algorithm's interdependences.

## APPENDIX

The RMS prediction error averaged over all established tracks (at a given time):

$$\sigma_{RMS}(k)=\sqrt{\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M_n^*}\sum_{t=1}^{T}\mu^t(n)\frac{\beta_m^t(n)}{c}\left[\left(v_{x_m}^t(n)\right)^2+\left(v_{y_m}^t(n)\right)^2\right]} \quad (9)$$

where $c=\sum_{m=1}^{M_n^*}\beta_m^t(n)$, $N$=no of active tracks associated to $M_n^*$ measurements, $T$=max-no of models in IMM, $\mu^t(n)$=mode probability, $\beta_m^t(n)$=marginal association probability, $v_{x/y_m}^t(n)$= innovations in the $x$, $y$ directions of track $n$ measurement $m$ and model $t$.

The Time-Average Normalized Innovation Squared (NIS) test is described by the following equations:

$$\bar{\varepsilon}_v(k)=\frac{1}{K}\frac{1}{N}\sum_{k=1}^{K}\sum_{n=1}^{N}\varepsilon_v^n(k) \quad (11)$$

$$\varepsilon_v^n(k)=\sum_{m=1}^{M_n^*}\sum_{t=1}^{T}\mu^t(n)\frac{\beta_m^t(n)}{c}\left(v_{xy}^t(n)\right)^T\left[\sum_{t=1}^{T}\mu^t(n)S^t(n)\right]^{-1}\sum_{m=1}^{M_n^*}\sum_{t=1}^{T}\mu^t(n)\frac{\beta_m^t(n)}{c}v_{xy}^t(n) \quad (12)$$

$S^t(n)$ is the innovation covariance matrix of model $t$, $K$=no of averaging time steps.

## REFERENCES


[1] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "Tracking Multiple Moving Objects With a Mobile Robot using Sample-Based Joint Probabilistic Data Association Filters." In *International Journal of Robotics Research (IJRR)*, vol. 22, no. 2, 2003, pp. 99-116.

[2] J. Almeida, A. Almeida and R. Araujo, "Tracking Multiple Moving Objects for Mobile Robotics Navigation." In *Proc. of 10th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'05)*, Catania, Italy, 19-22 September 2005, pp. 203-210.

[3] P. Kondaxakis, S. Kasderidis and P. Trahanias, "A Multi-Target Tracking Technique for Mobile Robots Using a Laser Range Scanner." In *Proc. IEEE International Conference on Robots and Systems (IROS'08)*, Nice, France, 22-26 September 2008, pp. 3370-3377.

[4] M. Lindstrom and J.-O. Eklundh, "Detecting and Tracking Objects from a Mobile Platform using a Laser Range Scanner." In *Proc. of 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS'01)*, Wailea, Hawaii, October/November 2001, Vol 3, pp.1364 – 1369.

[5] H. Chen, Y. Bar-Shalom, K. R. Pattipati and T. Kirubarajan, "MDL Approach for Multiple Low Observable Track Initiation." In *IEEE Trans. On Aerospace and Electronic Systems*, vol. 39, no. 3, July 2003, pp. 862-882.

[6] H. Chen, X. Rong Li and Y. Bar-Shalom, "On Joint Track Initiation and Parameter Estimation under Measurement Origin Uncertainty." In *IEEE Trans. On Aerospace and Electronic Systems*, vol. 40, no. 2, April 2004, pp. 675-694.

[7] J. Shu-Ling, L. Yan, H. Peng, P. Guang-Lin, P. Quan, C. Yong-Mei, "An Effective Hough Transform Based Track Initiation." In *Proc. 2006 International Conference on Machine Learning and Cybernetics*, Dalian, 13-16 August 2006, pp.3196-3200.

[8] J. M. Kubica, A. Moore, A. J. Connolly and R. Jedicke, "Fast and Robust Track Initiation Using Multiple Trees." Tech. report CMU-RI-TR-04-62, Robotics Institute, Carnegie Mellon University, November 2004.

[9] J. Xavier, M. Pacheco, D. Castro, A. Ruano and U. Nunes, "Fast Line, Arc/Circle and Leg Detection from Laser Scan Data in a Player Driver." In *Proc. of 2005 IEEE International Conference on Robotics and Automation (ICRA'05)*, Barcelona, Spain, 18-22 April 2005, pp. 3930- 3935.

[10] D. Castro, U. Nunes and A. Ruano, "Feature Extraction and for Moving Objects Tracking System in Indoor Environments." In *Proc. of 5th IFAC/EURON Symposium on Intelligent Autonomous Vehicles (IAV'04)*, Lisbon, 5-7 July 2004.

[11] K. O. Arras, O. M. Mozos and W Burgard, "Using Boosted Features for the Detection of People in 2D Range Data." In *Proc. of 2007 IEEE International Conference on Robotics and Automation (ICRA'07)*, Roma, Italy, 10-14 April 2007, pp 3402-3407.

[12] L. Spinello, R. Triebel and R. Siegwart, "Multimodal Detection and Tracking of Pedestrians in Urban Environments with Explicit Ground Plane Extraction." In *Proc. IEEE International Conference on Robots and Systems (IROS'08)*, Nice, France, 22-26 September 2008, pp. 1823-1829.

[13] G. A. Carpenter, S. Grossberg and D. B. Rosen, "Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System." In *Neural Networks*, vol. 4, no. 6, 1991, pp. 759-771.

[14] F. Lu and E E Milios, "Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans." In *Proc. 1994 IEEE*

*Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, WA, June 1994, pp. 935-938.

[15] P. Kondaxakis, S. Kasderidis, P. Trahanias, "Tracking Multiple Targets from a Mobile Robot Platform using a Laser Range Scanner." In *Proc. IET on Target tracking and data fusion: Algorithms and applications*, Birmingham, UK, 15 -16 April 2008, pp. 177-184.