People Detection from Laser Range Finder Data

João José Estêvão Nisa
joao.nisa@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2019

Abstract

The integration of robots in our daily lives became an acquired reality over the last few years. Connected to the quality of this integration is the perception that those robots have about the environments where they are placed. Therefore, having an understanding of their surroundings, in a social setting, is a synonym of the necessity of identifying people that are in the environment under analysis. This dissertation aims that purpose, People Detection, more specifically, their legs. Contrary to some work already done in this context, the working principle of the method presented in this dissertation seeks, above all, to estimate the disposition of the legs during the various stages of a normal walk. Those dispositions, performed by any passer-by, are represented by geometric patterns that slide along all the data collected by a Laser Range Finder. The main goal of this sliding process is to check if there is some correspondence between the patterns and the data collected which reflects the presence of any pedestrian. The developed solution and the central objective of this work ended up reveling positive results in both test scenarios.
Keywords: Detection, People, Legs, MOnarCH, Lasers, Patterns

1. Introduction

The field of social robots is very recent, however, that doesn't stop the major accomplishments made so far resulting in the development of astounding robots.

Besides the possibility of this type of robots ending up with a different purpose of what was projected at the beginning, the majority of the developments made so far aim for cooperation improvement between those robots and their users.

In the Humanoid Robots field - where the robot used for the development of this work fits - it's important to point the most recent and notable machines: *ASIMO* and *PEPPER*. *ASIMO* was first introduced in 2000 and since then has been targeted with multiple improvements that allowed him to be - at least so far -, one of the most advanced robots in the world. *PEPPER*, was introduced afterwards and was applied in some industry fields.

Apart from some features already available in the robots mentioned above such as, the ability to recognize objects as well as human emotions, establishing dialogue with people and understanding the environment that surrounds them, they present a common feature: social interaction. To reach that peculiar capacity, there's an important step that must be understood: people detection and/or social dynamics.

Another point worth mentioning, that is far from being less important than the previous, it's the motivation source of this dissertation which is the belief that this robots can be important in the recovery and treatment of children for example. This is specially relevant since that according to Dra. Filomena Pereira (director of the Pediatric Department of the Portuguese Institute of Oncology) "*everything that is ludic and promoter of a bigger well-being and of a major quality of life, is therapeutic as well*".

This work was guided by the intention of developing an innovative methodology that detects people and understands their dynamics as well as the capacity that this specific work has by combining two aspects that we particularly appreciate (Robots and Health care).

2. State-of-the-Art

As an introduction to the research made in the scope of this particular chapter, it is important to compare the strengths and the weaknesses of the two main modalities of *People Detection*: the *Detection of People based on Laser Range Data* and *based on Images*.

It is important to state that laser data has a remarkable precision concerning the measurement of long distances, however, they provide a restrictive perspective of the environment under analysis. This matter is solved with a capability that images don't

offer: the impermeability of laser detectors to the change of lightning and climate conditions.

The performance of *Detection based on Images* is particularly affected by lightning or climate changes (contrasting with what happens with laser data). The field of view of this modality is narrower when compared to the field of view of lasers.

2.1. Detection Based on Images

This detection approach is strongly connected to two different image characteristics: the gradient and the perception of depth.

Therefore and anchored in this characteristics (depth and gradient) there are two primordial methods. They are the *Histogram of Oriented Depths* [8] and the *Histogram of Oriented Gradients* [1, 8, 18]. The first method aims to fragment the image in multiple uniform cells, those cells are then attributed to gradient orientations. Those orientations define multiple histograms each one belonging to a cell that comes from the fragmentation process that are classified by a Support Vector Machine in a subsequent phase. The second method is an upgrade of the first one but relative to the depth values.

To close the Oriented Histograms context, there is the *Combo-HOD* [1]. This method aims, as the name indicates, to combine the advantages of the two previous methods.

It is important to highlight the *Sliding-Windows*, which are multi-scale windows that digitalize the image in regions of interest and in different scales. These windows are very similar to the *Bounding-Box*. The implementation of this method arises from the restriction of the field of analysis as well as an increasing efficiency, in a computation perspective, that this method allows. For example, it is enough to take into account that, through this windows and regions of interest (that result from the implementation of those windows) there is no need to analyze the whole picture.

Fast Depth-Based Region Proposals Generation [19] which uses a number of templates, those templates have the ability to identify the regions of interest.

There is also the *Fast Region-Based Convolutional Neural Network* [19] that generates a vector composed by probabilistic values that correspond to people detection.

To finish, it is important to mention detectors that use a set of geometric characteristics that correspond to parts of the human body, as an example of this there are the *PRD* [9], which seeks to identify a set of geometric characteristics innate to the human physiognomy in the multiple readings performed, and also *SIFT* [1] detectors that, with a classification process made with a certain point cloud, present the interest points of that cloud.

2.2. Detection Based on Laser Range Data

The Detection Based on Laser Data is preferentially used due to the efficiency and precision that are inherent to the use of those laser devices.

One of the methods that is used consists mainly on a vertical disposition of parallel laser layers [13, 5, 14]. The layers, arranged at different heights, will allow the detection of different parts of the pedestrian body. Usually, the data acquired is submitted to a fusion process between the different lasers that form a layer and, after that, the different layers are fused as well.

Complementary to the objective of detecting people is the one that detects and also distinguishes walking aids [20] (wheelchairs, walkers, etc.) of a pedestrian.

There are also methods that lay their classification process in sets of features that prevent, simultaneously, the fact that the Humans are not always moving. Therefore, the best alternative is to appeal to geometric features that are designed and extracted to detect circular shapes with *Inscribed Angle Variance*[21], for example.

In this type of methodologies, the features used, can be divided in four groups: geometrical, statistical, spacial relation and nearest neighbors [7].

A completely different perspective, relative to those mentioned so far, is defended by methodologies that seek the projection of the acquired data in a 3D plane [18].

To complete what was already described in this mechanism, it's important to mention the *DROW* Detector [3] as well that, for every laser point, predicts the class and the center of the nearest object.

2.3. Main Processes

There are, on the Detection Based on Laser Data processes - independently of their scope -, procedures that are inherent to those, as well as, *Segmentation* [11, 6, 12, 2, 20, 16] and *Classification* [11, 12, 2, 17, 15].

2.3.1 Segmentation

The Segmentation criteria is based on the distance between two consecutive points and it's known as Jump Distance. Therefore, points belong to the same segment if and only if the distance between them is lower than a certain threshold. It's the constant comparison between the threshold and the two consecutive points distance, that enables the possibility of reporting breakpoints. This threshold can be adaptative by the geometric rule designated as *Adaptative Breakpoint Detector*.

Still in this context and to treat models with non-linear processes, there is the *KF-Based Breakpoint*

Detector [4] where *Kalman Filters* are used.

2.3.2 Classification

Usually, the Classification process is the one that precedes the recognition of segments as part or non-part of humans, in this particular case. Adaboost [11, 12, 2, 10, 20, 17] is the algorithm that gathers all the potentialities mentioned above and is also the one where the choice normally falls when we talk about Classification processes.

3. Methodology

This chapter aims to present all the work developed from the software point of view, always taking into account the hardware that will be implemented. For this, and in the first instance, all the hardware involved in the work is addressed through a general presentation of the robot used - which is intended to confer the ability to detect people -, the components involved in it and also the sensors. From the software point of view, two processes that were fundamental in the distinction of laser points belonging to and not belonging to human legs are addressed (since this distinction is quite complex to make by projecting range points, according to acquisition order). Finally, the architecture of the developed detector and the processes it comprises are explained.

3.1. MOnarCH

With a design inspired on what children consider to be a generic robot, MOnarCH is the practical implementation of this work. MOnarCH has two types of robots under the same platform: one more sophisticated that targets social interactions with people/children (also known by the SO abbreviation) and the other one, more simple, that is used to increase the robot's perception about the environment that surrounds him (known by PO abbreviation), as well as assist your navigation, location and performance of some low level interaction elements. This particular robot also has four Mecanum wheels supplied by four independent motors, motors that increase the handiness and the performance of the platform. The four batteries that this MOnarCH has (two of them to feed the SO and the PO, one to supply the remaining electronic components and the last one to supply the DC motors), when entirely charged, grant to the robot an autonomy of three hours.

3.1.1 Components

The robot's platform includes the following high-level components:

- Two depth cameras equipped with a microphone (in this specific case, two Kinects);
- Three servo motors to operate in two robotic arms and a head;
- 10,1" Touch screen;
- LEDs;
- Capacitive Cells;
- Optional components like RFID Reader, StarGazer and UWB;
- Pico Projector.

3.1.2 Sensors

To have the maximum perception of the environment under analysis and to guarantee his own physical integraty, the robot uses a set of sensors specialized in distinctive areas.

- **Navigation** level, the robot uses a codifier to monitor the velocity of the motors, as well as, the inertia sensor to determine the orientation and also lasers, to detect obstacles and the environment geometry;
- **Perception** and **Interaction**, to this matter is used a depth camera to track and recognize human gestures, microphones and a RFID reader with the purpose of identifying people that carry an identification card;
- **Environment**, this android has a sensitive capacity relative to the environment in which is placed, through lightning and temperature sensors.

3.2. Auxiliar Processes

During the experimental process - and more specifically, during the development of the People / Legs detector - the need arises (for self-perception) to be able to distinguish between segments that belong to legs and those which, on the other hand, do not. This distinction allows us to consider which geometric properties are important that the detector should take into account when operating. That said, and bearing in mind that during the development phase it is important to consider large amounts of data, two processes need to be developed, one for converting range data into two-dimensional data and one for grouping or dividing the date into segments (Segmentation).

3.3. Development of the Detector

This work distinguishes itself with the fact that the developed detector doesnt need a labelling process to execute the classification of the data collected. This method is very similar to a process that already exists in **Detection Based on Images** field, the Sliding-Windows. However, instead of floating

size windows, there are multiple reference patterns - that correspond to different positions that legs can have during a normal walk - with whom laser data is compared.

3.3.1 System Architecture

The action of the detector comprises the following topics (as the **Figure 1** illustrates):

1. First, one of the reference patterns is created (since there are six in total). After this first step - and according to the size of the pattern in question - the various analysis intervals used to submit all laser range data captured by the robot to a process of the same nature are created. It is important to mention that there is the necessity to adapt the dimensions of those intervals since, one reference pattern that occupies, generally, 10 positions of the 721 positions in total, has to be compared with another 10, as it wouldn't make sense to compare these 10 positions with 20;

2. After the first step, if the analysis interval gathers the consensus of all geometric and spacial features taken into account by this detector, is determined the fitting percentage of the interval of analysis to the reference pattern under analysis;

3. Through multiple experiments, it was possible to verify which fitting percentage value should be considered as an interval of analysis as a positive detection;

4. After gathering all the positive detections, those are evaluated by two classifiers, one of them adapted to dynamic detections and the other one adapted to static ones.

3.3.2 Reference Patterns

In order to arrive to the Patterns (**Figure 2**) in question, the use of Auxiliary Processes and the analysis of what results from them are resorted to in some datasets of laser data recorded for this purpose. We can infer, that the patterns in question are not only the positions relative to legs during the walking process with a regular velocity, but also those whose disposition offers more robustness (at a Recall level) and safeness (at a precision level) to the detection process. This means that small patterns as well as patterns whose representation does not reveal any similarity between data that is bound and not bound to Conversion and Segmentation processes, more easily induce the detector in error.

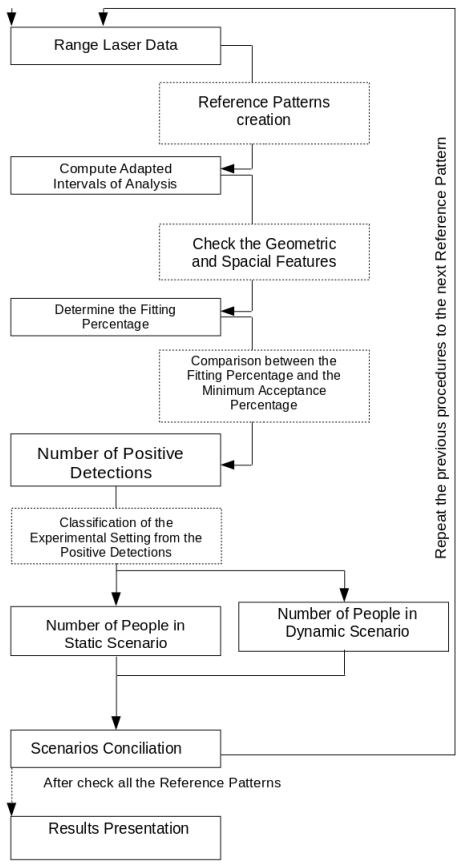
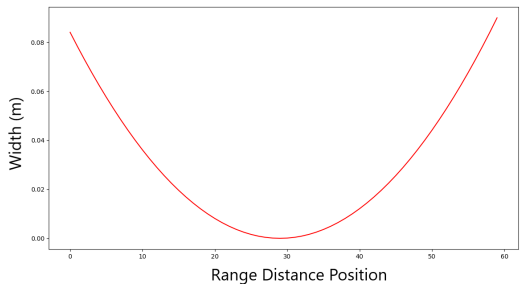
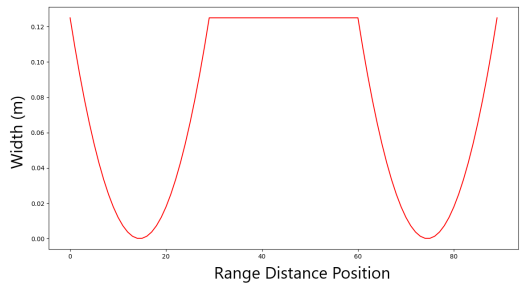


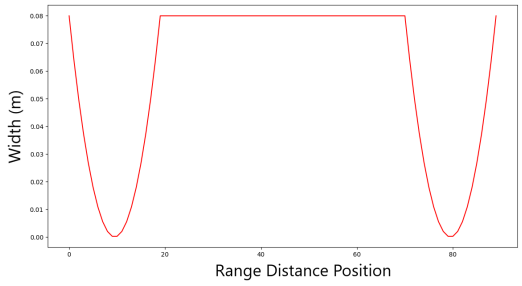
Figure 1: People/Leg Detector Architecture.



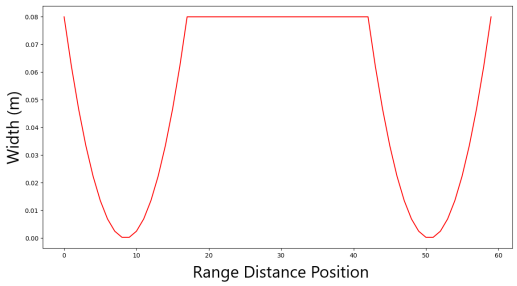
(a) Reference Pattern 1



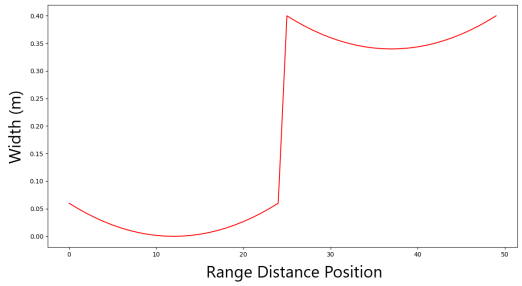
(b) Reference Pattern 2



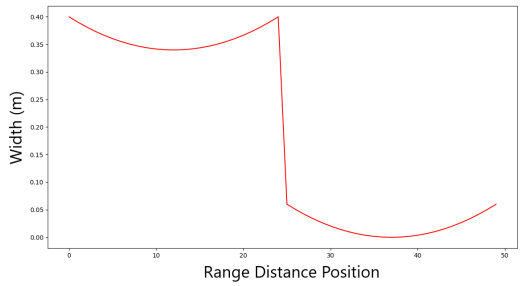
(c) Reference Pattern 3



(d) Reference Pattern 4



(e) Reference Pattern 5



(f) Reference Pattern 6

Figure 2: Reference Patterns

These errors are equivalent to the existence of structures, in given scenarios, that are very similar to legs when they are other things, which means that we are in front of a False Positive that decreases the value of Precision.

This patterns are the outcome of a mathematical procedure. The legs are, in every pattern considered, curves that correspond to parabolas (quadratic functions whose expression has the following equation, $f(x) = ax^2 + bx + c$), that come from an interpolation process between the minimum point with the two points whose ordinate value corresponds to the maximum value of the curve.

3.3.3 Geometric and Spacial Features

To improve the detector - in order to avoid the largest number of False Positives possible -, and make him more efficient from a computational point of view, some Geometric and Spacial Features were built.

This Features seeks, to avoid spacial measurements that go against what is normal at a width leg level and length (minimum and maximum) of the step or even the linearity level of a certain segment.

Therefore, it's relevant to point the Features taken into account:

- Reference Pattern 1
  1. Exclude segments that have more than 20 cm of width (width in this case, is the distance between the maximum point and the minimum point, as depicted in **Figure 3**);
  2. Exclude segments that have less than 7 cm;

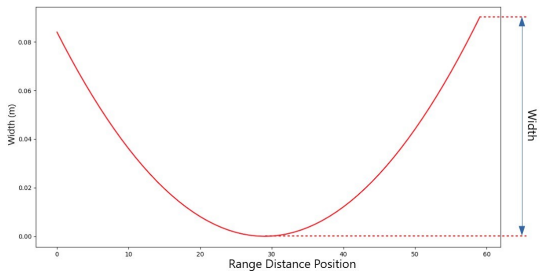


Figure 3: Width of a Segment.

- Reference Pattern 2
  1. Exclude segments that have more than 12.5 cm;
- Reference Pattern 3 and 4
  1. Exclude segments whose range value is the same for 7 or more points, in other words, segments that are, in certain part, a line;
- Reference Pattern 5 and 6
  1. Exclude segments which prevent a step from having more than 70 cm of length

(the length, in this specific case, corresponds to the difference between the midpoint of a leg and a midpoint of the other, like it is possible to see in **Figure 4**);

2. Exclude segments that prevent a step from having less than 15 cm of length.

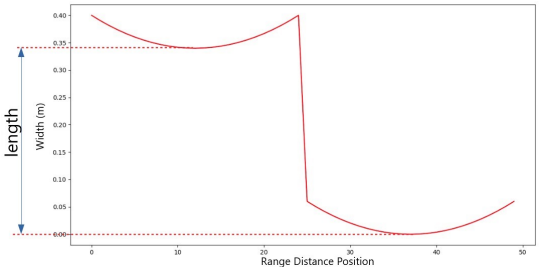


Figure 4: Length of a Segment.

3.3.4 Comparison between the Reference Patterns and the Data Obtained

In detection processes there is usually a step that dictates the presence or absence of what we want to detect, and associated to that step is always a comparison. This comparison aims at two elements that usually consist of the acquired data and what we intend to detect.

In this case, the comparison process begins shortly after the different reading distances have been broken down into several intervals of an appropriate size to the reference patterns against which all collected distances will be compared to. For example, if the reference pattern target of the comparison process occupies 70 read positions, then the 721 distances are divided into 651 groups of size 70 so that there is consistency between the size of the elements to be compared. There are 651 groups that, when analyzed, are equivalent to sliding the reference pattern across all the collected data since the same pattern "visits" all groups separately to find a match.

To improve this process, Geometrical and Spatial Features have been established to avoid comparisons with data that reveal distances that do not match either the inherent dimensions of the anatomy of any human leg.

Upon approval of all Features, a fitting percentage is assigned to the data group under analysis. At the same time, a minimum acceptance percentage from which the subset of data is considered as a positive match was also defined.

That said, it can be inferred that this method is, albeit in a different context, very similar to the Sliding-Windows that stand out in Image Based Detection.

3.3.5 Fitting Percentage

To reach a capable model to compute the fitting percentage between the reference pattern and a segment composed by n range distances captured by the laser, an implementation of three distinctive models, was made.

The models taken into account during this process were:

- Mean Value;
- Variance;
- Correlation.

To solve the problems raised by the first two models (relative to the codomain and mean values dependence), it is possible to conclude the method that best fits this problem is the *Correlation*, which considers the linear association between the reference pattern and the segment under analysis.

The implementation of this method, was possible due to the `numpy` library that outputs the correlation matrix.

$$R_{xy} = \begin{pmatrix} r_{xx} & r_{xy} \\ r_{yx} & r_{yy} \end{pmatrix} \tag{1}$$

r_{xx} and r_{yy} are equal to 1 in every case scenario due to the correspondence of random variable X with itself and due to the correspondence of Y with itself, respectively. For that reason, the value adopted to compute the fitting percentage is either r_{xy} or r_{yx} (since they're the same).

Afterwards, it is only missing how the detector takes the perception that the surface where the laser beams collide, is in fact, a leg and, for that matter, a person.

To do that it is defined a minimum acceptance fitting percentage, which means that every fitting percentage above this one is relative to a person. For example to the reference pattern 6 it was defined an acceptance fitting percentage of 90% so, every segment that revealed a higher percentage, is considered as a person.

3.3.6 Experimental Setting

By now, the detector should be able to give reasonable detection results, however there are two relevant properties to take into account:

- The people counting process, since that until now the detector only has the perception of a walker, but not the ability to know if those detections are relative to one or more people;
- Ability to differentiate how many detections are relative to dynamic people and, on the

other hand, how many are relative to static people since both can be inserted in the same scenario.

To answer these problems, the detector's operation is divided into two strands: one directed only to dynamic scenarios (i.e. moving passers-by), and the other directed to static scenarios only.

Experimental Setting - Static Strand

This strand is responsible for ensuring two aspects: (1) what is the minimum number of occurrences that corresponds to the static occupation of a person in a given position and (2), what is the distance between each position that allows to establish the difference between two static people inserted in the same scenario.

In order to address the first aspect, a data analysis was performed in order to determine the minimum value of occurrences from which it is considered that the number of detections made in a certain position correspond to a static person.

The second aspect arises because, when the number of occurrences leads the system to infer that in position pos_i there is a static person and if in pos_{i+1} there's also a number of occurrences that takes the system to draw the same type of conclusions then, it must assessed that the static detections found at both positions correspond to the same static passer-by and not to two separate ones. This aspect was solved by a data analysis process with which it was sought to determine from what distance, verified between two consecutive static positions, it is considered that the detections made correspond to two or more static people and not to only one.

Experimental Setting - Dynamic Strand

The number of dynamic people present in the scenario is determined according to the update frequency of the variable that keeps the best fitting percentage verified so far.

For this purpose, the system defines several equally spaced intervals where the minimum and maximum limits correspond to the number of updates that the best correlation value undergoes.

Considering that a person corresponds to the updating of this value a maximum of 8 times (value resulting from a Data Analysis process) in 2 seconds, the calculation of a larger number of people meets the illustrated dynamics present in **Figure 5**.

As pictured in this figure, that dynamic it's only applicable if the update number is as least one. Otherwise, the number of dynamic people is zero.

$$\text{Número de Pessoas Dinâmicas} = \begin{cases} 1, & \text{if } best\_fit\_perc\_actualizations\_num \text{ is in [1, 8]} \\ 2, & \text{if } best\_fit\_perc\_actualizations\_num \text{ is in [9, 17]} \\ \ldots & \\ n, & \text{if } best\_fit\_perc\_actualizations\_num \text{ is in [9 * n, 9 * (n + 1) - 1} \end{cases}$$

Figure 5: Dynamic Number of People Counting Scheme.

Experimental Setting - Dynamic and Static Conciliation

The conciliation of both scenarios mentioned above is able to combine the different results obtained in each of them because: (1) it would not be beneficial if the results were divided by the two existing experimental settings; and also because (2) it must be prevented that two positions in both scenarios are not summed together, since the same person would be contributing twice to the counting of people in the evaluated scenario.

Therefore, and also taking into account that in both scenarios the positions of the detections are properly kept, the process that establishes their conciliation involves verifying these common positions and excluding them before the computation of the number of dynamic persons in the scenario in question.

At the end of the verification mentioned above, the results obtained by both sides are summed up in order to obtain the global detection result.

4. Results

This section aims to present the behavior of the developed detector (which for presentation purposes was renamed MOnarCH Detector) in practical context as well as the points where it proves to be particularly capable and those in which revealed some weaknesses.

Since it is relevant to understand what position the MOnarCH Detector takes in relation to other research in this specific area, in both mentioned cases the results obtained by this work are compared with the Leg Detector package (`http://wiki.ros.org/leg\_detector`) and also the Leg Tracker package (`https://github.com/angusleigh/leg\_tracker`).

- Leg Detector: a reference amongst the various systems in place for people detection using laser data. Its working principle is limited to the use of several Geometric Features and their application in a machine learning classifier. If so, the detected legs resulting from the classification process are then linked and followed through a mechanism based on Kalman Filters. The range data collected by the laser serves as input to the aforementioned classifier, the latter being trained for data of this nature and to be able to distinguish legs between them;

- Leg Tracker: an update to the Leg Detector since it presents similar characteristics to the previous one however with some evolutions. This system has distinct capabilities that eventually come together for the central purpose of detecting and following people. This action is achieved by segmenting laser detections. At the same time, the confidence levels of segmentation performed in order to reduce the number of False Positives are taken into account. Through a method that involves neuronal networks as well as Kalman Filters, it becomes possible to follow any and all passers-by perceived by the lasers used in this process.

4.1. Assessment Situation in Controlled Environment

In order to favor certain detection situations, eight evaluation experiments were set up. In half of them we tried to control the dynamics exerted by the targeted passers-by, while in the other half, there was no pressure to exert a predefined path or positions.

All experiments took place at the Institute for Systems and Robotics (ISR) in the North Tower of Instituto Superior Técnico.

4.1.1 Experimental Results

The results obtained by the aforementioned libraries and the MOnarCH detector in the various experiments described above are shown in Table 1.

Table 1: Results obtained by the MOnarCH detector and the two comparison libraries.

Experience		LD	LT	MD
1	Recall	66.8%	0.0%	100.0%
	Precision	5.4%	0.0%	100.0%
2	Recall	56.7%	0.0%	100.0%
	Precision	5.2%	0,0%	100.0%
3	Recall	47.1%	22.1%	90.0%
	Precision	1.8%	89.4%	81.8%
4	Recall	24.2%	53.4%	74.1%
	Precision	2.1%	98.5%	90.9%
5	Recall	31.2%	66.2%	100.0%
	Precision	6.7%	96.9%	90.0%
6	Recall	24,0%	14.8%	98.7%
	Precision	7.5%	99.8%	100.0%
7	Recall	30.3%	21.9%	87.5%
	Precision	7.6%	100.0%	96.6%
8	Recall	41.1%	50.6%	63.3%
	Precision	11.2%	100.0%	98.0%

Note: LD stands for Leg Detector, LT stands for Leg Tracker and MD stands for MOnarCH Detector.

Based on the results presented, the following conclusions can be drawn:

- The leg detector method has a high number of False Positives that explain the displayed precision values. Taking into account Experiment 1, the content of which is the front facing of one person in front of the MOnarCH during the entire experiment, the leg detector reports an average of 12.94 people for each detection result determined. This is equivalent to considering that most of the segments present in this experiment belong to legs.
  Regarding the recall values, it can be concluded that they are better (25.5% on average) in situations where the experimental scenario is mostly static. That said, and by comparing the methods taken into account for this income statement, it can be concluded that the leg detector has the worst performance;

- The leg tracker method apparently ignores people who are in a static experimental setting. This conclusion comes from the experimental results obtained in the first two experiments, whose content only handles static scenarios. In these two cases, both Recall and Precision are 0%, which perfectly illustrates the behavior initially mentioned.
  On the other hand, in situations with a higher dynamic component the leg tracker exhibits a different behavior. Considering experience 5, which targets a person in a dynamic scenario in its entirety, it is possible to verify that the performance of this method improves considerably (especially when comparing these results with those obtained in experiments 1 and 2);

- The MOnarCH detector turns out to be impeccable for static experimental scenarios as illustrated by the results of experiments 1 and 2.
  In experiment 3 the results reveal a Recall superior to accuracy, which is not very common. This is explained by the presence of a structure in the scenario where the experiment took place, which closely resembles one of the six Reference Standards this detector takes into account, thus contributing to the increase of False Positives.
  It is also important to consider the results obtained in the last experiment, since this is where the value corresponding to Recall drops considerably compared to the values obtained in previous experiments. This is because one of the three people in a static experimental setting prints an arrangement on their legs that does not resemble any of the reference patterns highlighted by the detector.
  Lastly, the results obtained by the detector MOnarCH ended up as the best among the three selected methods for this specific evaluation.

4.2. **Real Environment Assessment Situation**
This experiment aims to integrate the robot into a real environment in which there is no control over the dynamics exerted by passers-by inserted in this scenario. The importance of these tests is that they reveal the ability that the MOnarCH detector demonstrates when integrating into real environments.

Regarding the dynamics of the scenario it should be noted that this experiment targets the college lobby before the first morning class (which starts at 8am).

4.2.1 Experimental Results

In this evaluation the results obtained are the following:

- Recall: 67,74%
- Precision: 81,82%

Overall, these results are mainly due to the following set of factors:

- Although the controlled environment study did not reveal it, the MOnarCH detector has some difficulties in dealing with large distances between its lasers and people;
- During this assessment there was a structure that has had some impact on the value of Precision. This is because it is very similar in shape to Reference Pattern 1;
- The existence of groups of people walking sideways (resulting in occlusion) also has a direct impact on Recall's value;
- Some of the passers-by were moving faster than normal. This factor interferes with laser readings that do not have any agreement with the Reference Patterns highlighted for this process;
- In order to finalize and justify the value of Recall, it should be noted that this is due to the arrangement of legs by passers-by when they are static, and which do not resemble any of the Standards set aside for this detector.

5. Conclusions

This paper aims to study the adaptation of a method commonly used in Image Data Based Detection to Laser Data Based Detection. In this case, instead of multi-scale detection windows that slide across the image in order to detect people, we have Reference Patterns adapted to the dispositions that legs make during a normal walk.

In Controlled Evaluation Environment, where results are expected to be superior, the detector that emerges from this work achieves an average Recall of 89.2% and an average accuracy of 94.6%, thus outperforming other detectors who were tested for the same scenarios. In Real Environment, its performance declines slightly due to the limitations raised in the previous section obtaining a Recall of 67.72% and an Accuracy of 81.82%.

Given these limitations, this work may have multiple implementations, especially in terms of monitoring social dynamics (following the recommendations raised in the next section), as well as streamlining the process of integrating this type of robots in a social context, which is one of main purposes of current robotics.

Future Work

Although this work yields remarkable experimental results, there are still some aspects that need to be taken into account so that it can be improved and thus make the detector more robust. Some points of possible improvement might be:

- The excessive delay. The MOnarCH detector reveals a delay that makes it impossible to implement in real time applications.This factor can be addressed by implementing some Python accelerators (which was the programming language used in this paper), such as PyPy Project or Numba which are tools recognized for achieving remarkable acceleration results;
- The limited detection found in dynamic experimental scenarios. Taking into account the various experimental results obtained, it is possible to verify that the detector performance has been decaying with regard to the detection of people in a dynamic experimental setting. Possibly, this is due to the lack of adequate Reference Patterns for this specific point, therefore, a study more focused on people dynamics is required in order to be able to fill this defect;
- The interpretation of social contexts. This means that it would be interesting to expand this detector to a Machine Learning method that, based on the number of people detected, could infer whether or not the space where the robot is located, is crowded to a point that makes its circulation impossible. This way, it would be possible to streamline their social integration;
- Detection with MOnarCH in motion. In this work a static detector was developed so it would also be important to widen its scope when the robot is in motion. This change would lead to the reprogramming of some steps

understood by the MOnarCH detector architecture and would always have to take into account that scanning times have a direct influence on the collected data.

References

[1] Human detection using depth information by kinect.

[2] K. O. Arras, . M. Mozos, and W. Burgard. Using boosted features for the detection of people in 2d range data. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 3402–3407.

[3] L. Beyer, A. Hermans, T. Linder, K. O. Arras, and B. Leibe. Deep person detection in two-dimensional range data. 3(3):2726–2733.

[4] G. A. Borges and M. J. Aldon. Line extraction in 2d range images for mobile robotics. 40(3):267–297.

[5] A. Carballo, A. Ohya, and S. . Ichi Yuta. Multiple people detection from a mobile robot using double layered laser range finders.

[6] W. Chung, H. Kim, Y. Yoo, C. B. Moon, and J. Park. The detection and following of human legs through inductive approaches for a mobile robot with a single laser range finder. 59(8):3156–3166.

[7] J. Cui, H. Zha, H. Zhao, and R. Shibasaki. Laser-based detection and tracking of multiple people in crowds. 106(2):300–312.

[8] IEEE Staff and IEEE Staff. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems.*

[9] Institute of Electrical and Electronics Engineers. *2017 IEEE International Conference on Multimedia and Expo (ICME) : Hong Kong, 10-14 July 2017.*

[10] H. Kheyruri and D. Frey. Comparison of people detection techniques from 2d laser range data.

[11] D. Li, L. Li, Y. Li, F. Yang, and X. Zuo. A multi-type features method for leg detection in 2-d laser range data. 18(4):1675–1684.

[12] D. Li, L. Li, M. Zhou, and X. Zuo. Robust features for leg detection in 2d laser range data. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, volume 42, pages 417–424. International Society for Photogrammetry and Remote Sensing.

[13] O. Martinez Mozos, R. Kurazume, and T. Hasegawa. Multi-layer people detection using 2d range data.

[14] O. M. Mozos, R. Kurazume, and T. Hasegawa. Multi-part people detection using 2d range data. 2(1):31–40.

[15] M. Mucientes and A. Bugarn. People detection with quantified fuzzy temporal rules. In *IEEE International Conference on Fuzzy Systems.*

[16] C. Premebida and U. J. Nunes. Segmentation and geometric primitives extraction from 2d laser range data for mobile robot applications.

[17] D. L. Rizzini and S. Caselli. Experimental evaluation of a people detection algorithm in dynamic environments.

[18] L. Spinello and R. Siegwart. Human detection using multimodal and multidimensional features. In *Proceedings - IEEE International Conference on Robotics and Automation*, pages 3264–3269.

[19] A. Vasquez, M. Kollmitz, A. Eitel, and W. Burgard. Deep detection of people and their mobility aids for a hospital robot. In *2017 European Conference on Mobile Robots, ECMR 2017.* Institute of Electrical and Electronics Engineers Inc.

[20] C. Weinrich, T. Wengefeld, C. Schroeter, and H. M. Gross. People detection and distinction of their walking aids in 2d laser range data based on generic distance-invariant features. In *IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions*, pages 767–773. Institute of Electrical and Electronics Engineers Inc.

[21] J. Xavier, M. Pacheco, D. Castro, A. Ruanot, and U. Nunes. Fast line, arc/circle and leg detection from laser scan data in a player driver. In *Proceedings - IEEE International Conference on Robotics and Automation*, volume 2005, pages 3930–3935.