A Human Detection System for Proxemics Interaction

Xiao WANG, Xavier CLADY
Université Pierre et Marie Curie, UPMC-Paris6
Institut des Systèmes Intelligents et de Robotique
UMR 7222 CNRS
Paris, France
{wang, clady}@isir.upmc.fr

Consuelo GRANATA
Université Pierre et Marie Curie, UPMC-Paris6
Institut des Systèmes Intelligents et de Robotique
UMR 7222 CNRS
Robosoft, Bidart, France
consuelo.granata@robosoft.com

ABSTRACT
In this paper, we present a human detection system for a domestic robot. A 2D laser scanner based leg detector and a vision based body detector are combined using a grid fusion strategy. This approach has been evaluated on a domestic robot. Furthermore, we propose a methodology to evaluate it in relation to proxemics that could be generalized to other robot's perceptive functions.

Categories and Subject Descriptors
I.2.9 [Computing Methodologies]: Artificial Intelligence – learning, robotics, vision and scene understanding.

General Terms
Algorithms, Performance, Experimentation.

Keywords
detection, 2D laser scanner, computer vision.

1. INTRODUCTION
Detecting a human presence is a fundamental task required from a domestic robot. Another task is often associated to it: to localize the human and, in particular, to compute its relative distance from the robot. It could provide a useful information to categorized the human behavior [16] or to give socially acceptable behaviors to the robot. Indeed, social distances, also called proxemics, are known to have an important role in human-robot interaction.

This report deals with a pilot study in order to propose a new method to detect and localize humans using sensors embedded in a wheeled robot and to characterize its performances related to proxemics. We have focused our work on a laser and vision based detection which is the most common approach. Classically the range information is used to focus a vision based algorithm on a interest region of the image [14] or to reject the false alarms provided by the visual sensor [15]. Most of the recent studies realize the fusion via the tracking process [6, 18, 8, 2, 1]. We propose a grid fusion method to combine laser and vision based detectors in a more powerful detector. Furthermore a protocol is proposed in order to relate the reliability evaluation (given through ROC curves) to proxemics. Most research provide only the global ROC curve [15], or even just a table of correct positive and false negative rate [2,1,8,12,18]. Our protocol demonstrates clearly the efficiently and the complementarity of the used sensor's configuration and algorithms.

2. DETECTION ALGORITHMS
2.1 Probability Grid
We propose to merge the data from the two sensors using probability grid (sampling in cells the robot's environment). Probabilities of human presence is computed for each cell. Our grid is based on polar split of the space: each cell is identified by its polar coordinates (ρ, θ). The grid is more dense close to the robot (where we need a more accurate position estimation). The geometric relation between the camera and the laser scanner is estimated in order to project the cells of the grid in laser and vision space (cf. Fig. 1). Each cell in grid is passed through the laser and vision based detectors. The two detectors' outputs are normalized using distances of Mahalanobis, and then summed to get the final score. A threshold rule is applied on the final score.

Figure 1. Example of the grid projected in the vision space.

2.2 Laser Based Leg Detector
Some algorithms are based on the search of local minima [4]. Other approaches can only detect moving persons [9,3]. Bellotto et al. [1,2] propose a structural approach to identify typical patterns. Recently machine learning approaches have been proposed [12]. Our method is based on statistical models of the leg geometries. A frame of laser scan is split into several segments. A Gaussian model (parametrized by its mean, μ=0.13m, and its standard deviation, σ=0.03m) on leg radius is then applied to give the probability that the segment can be a leg. These values are then filled into a grid where each cell contains the probability value of the presence of a leg. The convolution on this grid with a mask taking account of another inter-leg distance Gaussian model (μ=0.23m, σ=0.04m) provides us a second grid of the conditional probability of another leg in the neighborhood.
 cells. Multiplying the two grids provides us a joint probability of legs' position.

2.3 Vision Based Body Detector

For each cell of the vision's grid, we extract a sub-image. Its size is computed assuming the people's legs are leaned on the floor. This sub-image is provided to an appearance based classifier trained to detect people. We have used a cascade of boosted classifiers quite similar to the one proposed by Viola et al. [17], but using Real AdaBoost algorithm[13], an improved version of the Adaboost algorithm. Furthermore, two types of features are employed: Haar-Like (Haar-like features or Haar features) and HoG (Histogram of Oriented Gradients). These two features are frequently used in the domain of object detection and recognition [7]. The output of the cascade, noted G, is finally fitted to a sigmoid probabilistic function [10] using the learning data base. This last process allows us to allocate a probability of human presence to each cell.

3. EXPERIMENTS

The Kompaï robot used for our experiments embeds a laser sensor Sick S300 (at 0.3m from the floor) and a camera on its head (at 1.5m from the floor). We have selected 15 persons of different heights, origins and dressed with different clothes. We have asked the participants to walk in front of the robot. We have collected 271 data (image and laser frame) in 3 different indoor environments, that represent 471 potential human observations.

![ROC curves at different spaces](image-url)

Merging the laser based legs detector and the vision based body detector works as expected: its curve (cf. the global ROC curve in Fig. 2) rises above the others. To analyze more precisely these results, we have grouped the observations following the human proxemics categorization proposed by Hall [6]. We can observe that the weakness of the leg detector observed in the global ROC curve, is partially due to the non homogeneous distribution of the data. The most of them (>50% of the potential observations) have been collected in the Public Space, where the body detector is more reliable and the legs one works badly (because the probability that a laser ray intersects a leg decreases drastically in the beginning of the Public Space). In the Personal Space, only the laser based detector works: the camera cannot see the whole body of a closed person. In the Social Space, the information provided by the two sensors have to be merged.

4. CONCLUSIONS AND FUTURE WORKS

A grid fusion method is proposed to merge the output of different detectors. Their performances are then characterized with the proxemics. The experiments demonstrate the complementarity of these sensors. This complementarity was already used experimentally in several studies [6,14,18,8,15,2,1]. But their protocols and the described results do not allow to associate the sensors (and their associated algorithms) to proxemics. In future works, we would extend this type of characterization to other functions (face detection or recognition, voice localization, speech recognition, ...). We intend to improve the HRI performances of a domestic robot by proposing the most efficient sensors and software configurations related to HR proxemics.

5. ACKNOWLEDGMENTS

This work has been supported by European Union and French National Research Agency (ANR) through AAL program (project DOMEO n°ANR-08-AALI-001-02).

6. REFERENCES


