# Multisensor Integration for Human-Robot Interaction

**Nicola Bellotto and Huosheng Hu**
Department of Computer Science, University of Essex
Wivenhoe Park, Colchester CO4 3SQ, UK
Email: {nbello, hhu}@essex.ac.uk

***Abstract*— In this paper we present a solution to human-robot interaction using a combination of visual and laser range information. Human legs are extracted from a laser scan and, at the same time, faces are detected from the camera's image. The information is integrated in a detection procedure that returns direction and distance of the surrounding people. This is eventually used by a mobile robot to approach and start interaction with humans. Unlike other similar applications, our solution works well in real-time even under limited computational resources. Experimental results show good performances of our system.**

***Index Terms*— human-robot interaction, multisensor integration, people detection.**

## I. Introduction

Recently many researchers have focused their work on the social aspect of mobile robotics, in which human detection and tracking is an essential prerequisite for human-robot interaction (HRI). Indeed, the first step for a social robot to start any kind of interaction with people is being able to detect and approach them.

Many different solutions to resolve such problem are reported, including the use of artificial marks and light-emitting devices [1], [2]. Although these systems are quite efficient, the drawback is that the human to be tracked has to carry an "attractor" device. On the other hand, many applications use the robot's on-board camera for people detection, often concentrating on the face [3], [4] or other regions of the human body [5]. Others make use of laser range sensors, identifying people as moving objects [6], [7]. Some researchers have worked on"multi-modal" systems, in which visual data is combined with laser range data [8], [9], [10], [11] or thermal images data [12] to enhance human detection and tracking performance of the system. In several cases, speech recognition is also integrated only with vision [13] or with both vision and laser data [14], [15], [16].

It is clear that in all the advanced approaches, or at least most of them, image and video processing are essential components. In particular, it is very important to develop cost-effective face detection algorithms to achieve real-time performances [17], [18], [19], [20].

In this paper, vision and laser data are combined to detect face and legs respectively. Compared to other similar methods [8], [9], [11], [15], we achieved significant improvements. Our application is able to continuously detect faces and legs even when both the robot and a person are moving. Moreover, the need of computational power is dramatically reduced and we do not make use of any additional hardware dedicated to image processing. The system is implemented on our mobile robot ATLAS, which is provided with a SICK laser range sensor and a PTZ camera. ATLAS is an interactive robot that acts as a tour-guide inside the County Hall building in London.

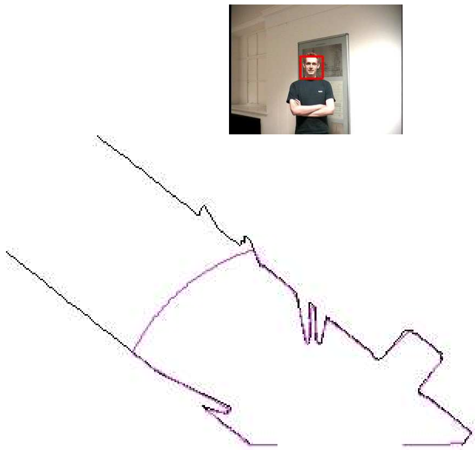

Fig. 1. Laser scan. The black line is the original laser output, the purple line is the output after filtering. In the middle we can notice the legs of a person (picture above).

The rest of the paper is organized as follows. Sections II, III and IV illustrate our approach, including some information about the practical implementation. Some experimental results are presented in Section V and a brief summary and future extensions are given in Section VI.

## II. Legs detection

The laser range sensor provides a 180° scan of the environment at approximately forty centimeters from the floor and with a half degree resolution. It is known that laser readings are very accurate and the error is in the order of millimeters. In Fig. 1 we can observe the output of a laser scan in presence of a person.

The laser range data can be represented as a function on a XY graph in which the abscissa is the angle and the ordinate is the measured distance. In the simplest case, the characteristic pattern of two legs is

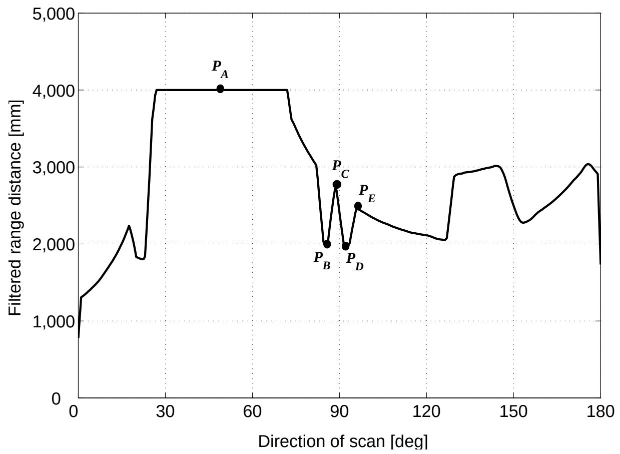


Fig. 2. Typical legs' pattern, identified by the sequence of points $P_A \rightarrow P_B \rightarrow P_C \rightarrow P_D \rightarrow P_E$.

constituted by a sequence of alternate maximums and minimums of the distance function following this order: $max \rightarrow min \rightarrow max \rightarrow min \rightarrow max$ (an example is shown in Fig. 2). The algorithm used to detect possible legs is based on this simple concept and can be divided in the following three steps:

- First of all, some noise is removed from the laser data with a simple moving-window filter. A threshold is also used to limit the maximal range. The purple line in Fig. 1 shows a typical result.
- Then, all the minimums and maximums of the distance function are extracted, memorizing at the same time those particular sequences that could identify two legs, as shown in Fig. 2.
- Finally, some basic rules are applied to discard sequences that are not legs patterns. Basically, these rules take into account the inner distance between human legs (e.g. a legs aperture of 3 m would be definitely impossible!).

More specifically, here is a detailed description of these steps. First, the implementation of the moving-window filter, which smooths the laser data, can be described as follows. If $\Delta\theta$ is the angle step of the laser scan and $d(i \cdot \Delta\theta) \equiv d_i$ is the distance measured along the direction $i \cdot \Delta\theta$, we can write the filtered value $\hat{d}_i$ as follows:

$$\hat{d}_i = \frac{1}{2N+1} \sum_{n=i-N}^{i+N} d_n \qquad (1)$$

where $2N+1$ is the size of the moving-window filter, with a window size of $5$ ($N = 2$) empirically determined. The maximal distance is limited to 4 m.

The next step is identifying sequences of minimums and maximums of the distance function, as shown in Fig. 2, which could be possible legs patterns. Such points correspond of course to the angles for which the derivative of the distance function is null. In practice, to extract such points from our samples, we consider all the sequences

TABLE I

LEGS DETECTOR ALGORITHM

{filtering}
**for** $i = 1$ to $LASER\_DATA\_SIZE$ **do**
  $\hat{d}_i \Leftarrow$ filter $d_i$ with (1)
**end for**

{pattern recognition}
$P \Leftarrow \oslash$ {set of legs patterns}
$S \Leftarrow \hat{d}_1$ {current sequence}
**for** $i = 2$ to $LASER\_DATA\_SIZE$ **do**
  **if** ($\hat{d}_i$ is a max.) $\vee$ ($\hat{d}_i$ is a min.) **then**
    **if** $\hat{d}_i$ and the last element of $S$ satisfy (3) **then**
      add $\hat{d}_i$ to $S$
      **if** $S$ contains 5 elements **then**
        add $S$ to $P$ {new candidate pattern}
      **end if**
    **else if** $\hat{d}_i$ is a max. **then**
      $S \Leftarrow \hat{d}_i$ {reset the sequence}
    **end if**
  **end if**
**end for**

{pattern selection}
**for all** $S \in P$ **do**
  **if** $S$ does not satisfy (4) **then**
    remove $S$ from $P$
  **end if**
**end for**

$S = \{\hat{d}_h, \hat{d}_{h+1}, \ldots, \hat{d}_{h+k}\}$ for which:

$$\begin{aligned} \left|\hat{d}_i - \hat{d}_{i+1}\right| < \epsilon \\ \text{with } i = h, h+1, \ldots, h+k-1 \end{aligned} \qquad (2)$$

where $k$ is the number of consecutive duplets for which is valid the expression above. Then, we simply take the point $\hat{d}_{h+k/2}$ in the middle of the sequence as minimum or maximum. Note that (2) means we consider all the segments with inclination less than $\epsilon$. Although a little rough, this approximation permits a very fast computation with a sufficient precision for our task. The value $\epsilon$ must not be too big, in which case we could miss some local minimum or maximum, and not too small, thus to avoid the effect of residual "noise" left by the previous filter. We found $\epsilon = 5$ cm to be a good compromise.

The distinction between minimum and maximum is simply deducted from the fact that one always follows the other (there cannot be two consecutive minimums!). When a maximum or minimum is extracted, the algorithm checks if this contributes to reconstruct a possible legs pattern. In practice, it tries to identify a sequence of points from $P_A$ to $P_B$, like shown in Fig. 2, respecting the following vertical constraints:

$$\begin{cases} \hat{d}_{P_A} - \hat{d}_{P_B} > L_{ext} \\ \hat{d}_{P_C} - \hat{d}_{P_B} > L_{int} \\ \hat{d}_{P_C} - \hat{d}_{P_D} > L_{int} \\ \hat{d}_{P_E} - \hat{d}_{P_D} > L_{ext} \end{cases} \qquad (3)$$

where $L_{ext} = 20$ cm and $L_{int} = 50$ cm in our application.

Finally, a last control is applied to discard improbable legs patterns, excluding all the cases where the distance

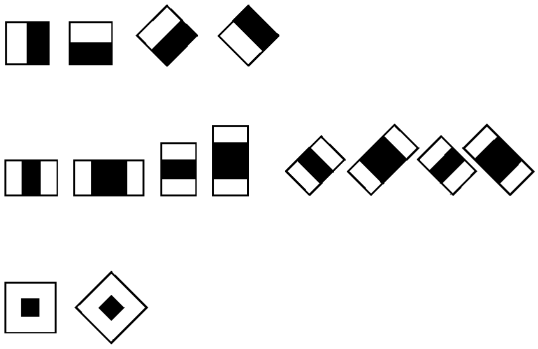

Fig. 3. Set of Haar-like features for face detection. The total amount of features contained in a $24 \times 24$ window is about 120000.

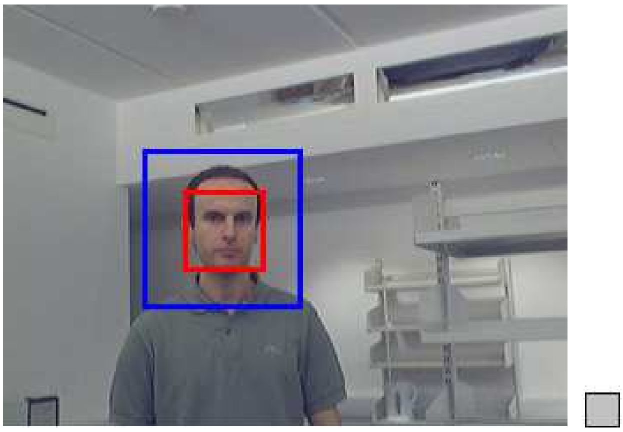

Fig. 4. At time $t = 0$, the whole image $320 \times 240$ is scanned, with a minimum face size of $20 \times 20$ pixel (small gray square on the right). The bounding box of the detected face (red square) is $45 \times 45$.

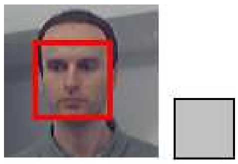

Fig. 5. At time $t = 1$, the scan covers only a sub-image $90 \times 90$ (blue square in the previous image), looking for faces not smaller than $36 \times 36$ pixel (80% of the previous face).

TABLE II
FACE DETECTOR ALGORITHM

```
if no face tracked then
   min. face size ⇐ 20 × 20
   scan the whole image
else {tracking}
   min. face size ⇐ 80% of the tracked face
   sub-image size ⇐ twice the tracked face
   scan the sub-imago centered on the tracked face
end if
```

between two adjacent legs is too big:

$$\|P_D - P_B\| < D_{step} \tag{4}$$

Considering that the laser device is about at forty centimeters from the floor, a good value for such limit is $D_{step} = 50$ cm. The pseudo-code of the procedure for legs detection is summarized in Table I.

Thanks to the precision of the SICK laser data, the legs detector permits to know the position of the surrounding people with great accuracy. Of course there are cases where it is almost impossible to distinguish legs from other objects, for example because of their position with respect to the robot (e.g. one leg covered by the other) or simply because the people is too close to a wall. In these situations the face detector, explained in the next section, plays a fundamental role.

## III. Face detection

Face detection is a very difficult task because involves many challenging issues. Some of these are pose, presence of structural components (like beards, glasses, etc.), facial expression, occlusions, image orientation and environmental conditions (e.g. light). In our application we make use of a recent object detector system [18], which is a refined version of the broadly known algorithm created by Viola & Jones [19]. Such system is already implemented and trained for face detection in *OpenCV* [21], the computer vision library we adopted. Compared to the numerous solutions illustrated in literature [17], the chosen one shows a good balance between detection performances and computational speed. It is also worthwhile specifying that this method is color independent, therefore adapt for different skins and more robust to varying light conditions.

Briefly, the face detection algorithm of [18] and [19] works as follows. Using a cascade of pre-trained classifiers, it extracts Haar-like features from subwindows of the image. The set of features is illustrated in Fig. 3. Each classifier rejects bad input samples operating a discrimination on the output samples of the previous classifier. Two important parameters that influence the speed of the algorithm are the resolution of the image and the minimum size of the sub-windows, which is the minimum size of the searched faces.

To increase the performance of our face detection module, we implemented a simple and fast face tracking algorithm with an adapting regulation of the parameters. In practice, at the beginning we scan the whole image, which is $320 \times 240$ in our case. If one or more faces are detected, we choose the closest one. At the next time step, for the scan we consider only the sub-image containing the face. We found a good solution taking a sub-image double the size of the face. At the same time, we modify the minimal size of searched faces, setting this parameter to 80% of the current face size. Fig. 4 and Fig. 5 show an example of how the tracking works.

With this method we obtain two important results:

- First, we increase significantly (~4 times) the face detection speed, of course provided some face is actually present;
- Second, we keep track of one face as long as it can

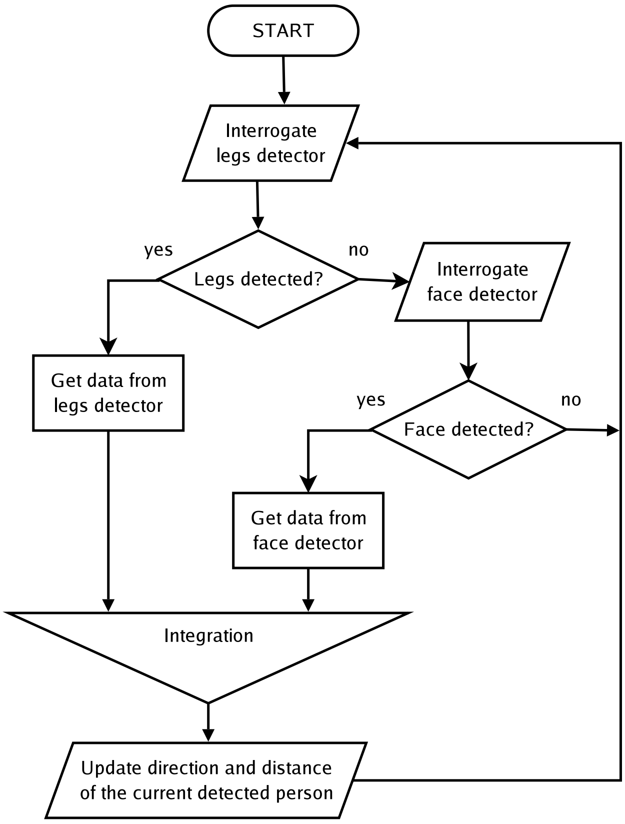


Fig. 6. Flowchart of the detection program. The information from the laser legs detector, more reliable, has the priority on the face detector.

Fig. 7. ATLAS, the interactive museum guide robot. The test environment, in this case a wide foyer, has artificial lights, large door-windows and big movie-screens.

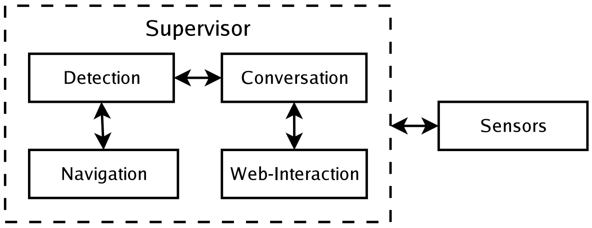


Fig. 8. Scheme of the application that integrates people detection with the other modules. Each block is a thread, all controlled by the supervisor.

be detected, avoiding cases where the face of interest changes continuously (e.g. when two faces are at the same distance and the selection method tries to chose the closest).

Table II illustrates a simplified version of the algorithm to detect and track a face. Like the legs detection module described in Section II, the face detection module can also return the position of a person with respect to the robot. Indeed, being known the field of view (FOV) of the camera, the direction of the person is simply proportional to the horizontal position of his face inside the image. The calculus of the distance instead is more complicated and, in absence of a stereo camera, is normally resolved using dynamic vision techniques, like depth from motion [22] or depth from focus [23]. However, these methods normally require the people to be nearly static, which is not our case. We chose then a more naive approach, fast and good enough for our pourposes. To each face indeed is associated a bounding box (see Fig. 4), the size of which changes with the distance. We use a simple conversion factor, determined empirically from the height of the bounding box at fixed distances of the face. Using the position of the face in the image and the size of its bounding box, we can therefore calculate roughly the location of the person with respect to the robot. Of course, while the direction is quite precise, the error of the distance is considerable. Nevertheless, in our experiments it has been proved to be reliable enough within a range of 2 m, which is the area inside where the human-robot interaction starts.

## IV. Hybrid implementation

At the current stage, the combination of the two modules, legs and face detection, is simple but efficient. In practice, we realized experimentally that the laser based legs detection is very accurate and in most of the cases is much more reliable than face detection. Moreover, the computational time needed by the legs detector is much less than that one required by the face detection module. We decided then to give priority to the information coming from the laser and use the face detection only when the former does not detect any person. Our choice is supported also by the fact that the range covered by the laser device is much wider than the camera view. While the laser covers a semicircular area with a radius of several meters, the camera view is limited to approximately $40°$. Also, even if the camera is fixed at about 1.5 m from the floor (which is an average of the people height), there are cases when a face cannot be detected because the person is too tall or too short and very close to the robot.

The flowchart of the detection program is illustrated in Fig. 6. First, the legs detector is interrogated and, if any person is found, direction and distance of the closest one are recorded. In case no legs are detected, the control passes to the face detector and if it succeeds, it provides

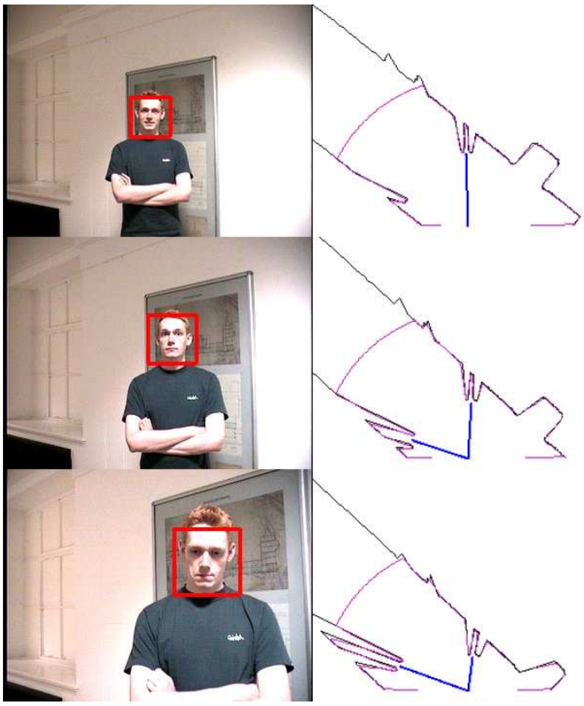

Fig. 9. Approaching a human. In the sequence we can see the detected face on the left and the relative laser legs detection on the right. ATLAS stops in front of the human (last snapshot) to start the interaction.

the position of a tracked face. Otherwise the procedure simply restarts.

Our robot ATLAS is an ActivMedia PeopleBot equipped with a SICK laser range sensor, a PTZ camera, audio system and a touch-screen (see Fig. 7). The on-board computer is a Pentium III 800 MHz with 256 MB of memory, running Linux operating system. The camera has been mounted on a special support to increase the height of its view, which is now about 1.5 m. This permits in general better performances when detecting faces.

The software has been realized with a modular approach. The legs and the face detectors have been thought as "virtual" devices that can independently return direction and distance of detected people. Such detectors are part of a library that can be expanded with additional functionalities (e.g. motion or sound detectors).

Furthermore, the detection thread is part of a more complex application that involves navigation, conversation and web-interface, thus to provide a full interactive robot guide. In Fig. 8 we can see a scheme of the whole application. Every module is an independent thread, the main one being the supervisor, which controls all the others using a FSM approach. A thread is also dedicated to the sensors, keeping the data up-to-date and synchronizing the access to the input devices.

More in detail, the detection module provides the navigation with the position of humans and leaves it the task to approach them avoiding obstacles. The detection also informs the conversation module on the presence of people, so that the latter can attract them or start an interaction when close enough. At this stage, the web-

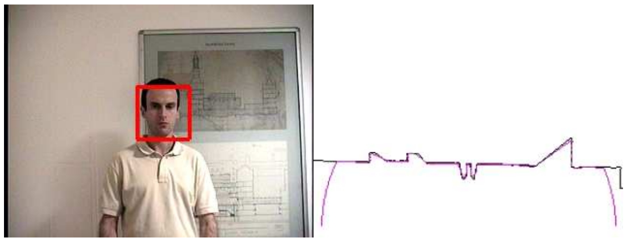

Fig. 10. Person too close to the wall. The legs pattern searched with a laser scan, on the right, can be confused with other objects in the environment. The face detector helps to resolve the confusion.

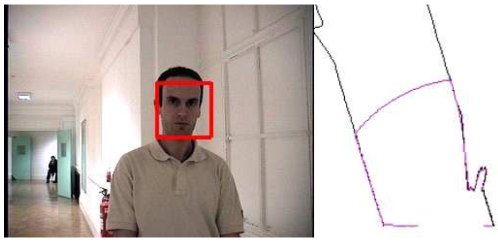

Fig. 11. Legs too close to each other. Instead of two columns, the laser shows just one big column. Even in this case, the face detector helps to resolve the ambiguity.

interface is also available to show visual information to the user or get input through the touch-screen.

## V. Experimental results

Most of the experiments have been conducted in a corridor and a foyer of the County Hall in London. The environment is a very good test-bed for interactive robots like ATLAS. First of all, the wide space permits several people to interact with the robot at the same time. Furthermore, the light condition is very challenging because varying from artificial to natural illumination. Fig. 7 shows part of the wide foyer.

### A. Example of approach

In the first experiment, we report a successful approach achieved by ATLAS using both laser and vision data. In Fig. 9 we show a sequence of three snapshots taken during a human approach. Both face and legs are continuously tracked while the robot is moving toward the person approximately at 30 cm/s. The update speed of the legs detection module, about 6 Hz, is limited only by the hardware and the Operating System. Of course, due to image processing, the speed of the face detector is lower, in average 3 Hz. When it reaches a person at a proper distance for interaction, the robot stops (last snapshot of the sequence). We can also note that, during the second and third time-steps, the laser detects another person on the left, who is not visible from the camera.

### B. Failure of the legs detector

Most of the times the laser is sufficient to detect a person. However, there are cases for which it is impossible

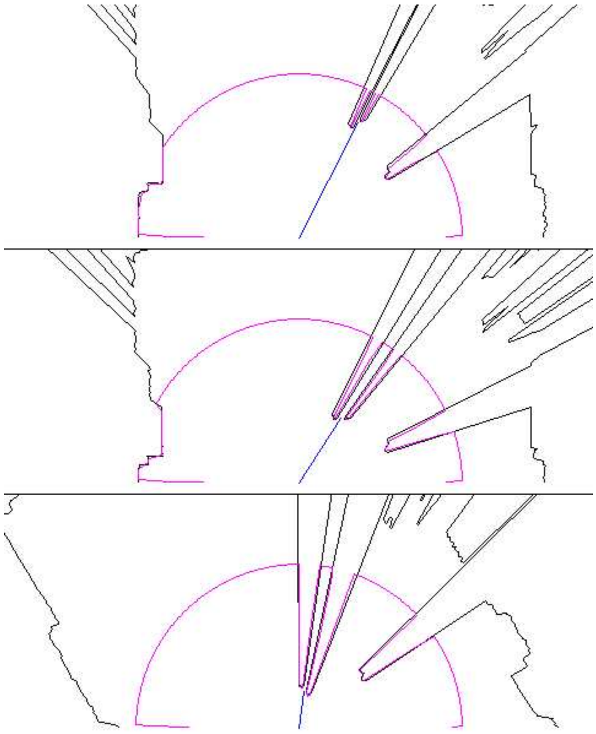

Fig. 12. Approaching a moving human. ATLAS detects a person who is walking and getting closer. Therefore it turns right and approaches him for interacting.

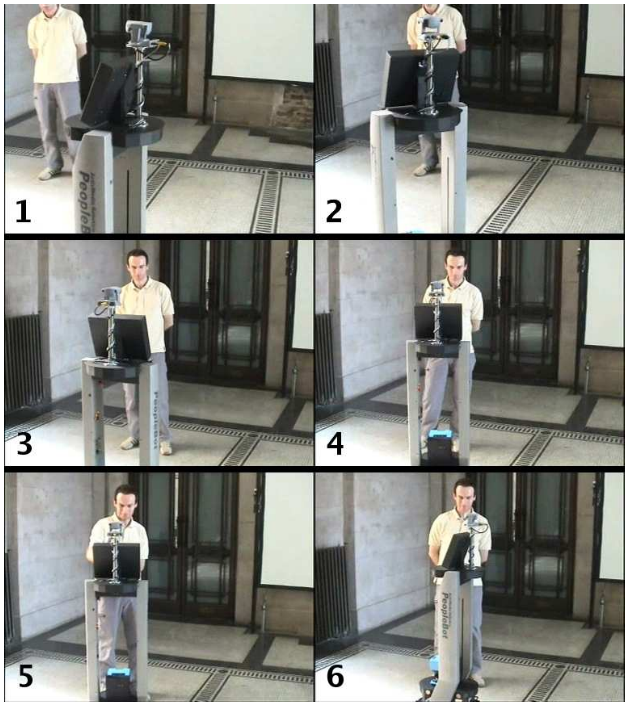


Fig. 13. Approach and interaction. In the sequence, ATLAS detects the person and approaches him (frame 1-3). After that, they interact using speech and the touch-screen (frame 4-5). Finally, when the interaction terminates, ATLAS turns and moves away (frame 6).

to distinguish and recognize legs in the environment. In this section we show a couple of typical examples where the legs detector fails but, thanks to the integration with the face detection, the robot is still able to identify a human. In Fig. 10 we can see a situation where the person to approach is too close to a wall, so that the legs pattern is not clearly identifiable. In particular, such pattern is discarded by the constraints (3) and (4) given in Section II. In Fig. 11 instead the legs are too close to each other, looking like a single column (this could easily happen also if the person is a lady with a long skirt).

### *C. Approaching a moving human*

Often people are not just statically waiting for the robot to reach them, but prefer to move towards it looking for some interaction. We show that ATLAS can actually handle this kind of situations and approach humans even when they are walking to get close. In Fig. 12 there is an example where a person, on the right side of the robot, is walking torwards it. ATLAS detects the human legs using the laser, then it turns and starts to approach the person, who in the meanwhile is getting closer and closer. Finally, they both stop in front of each other at a proper distance to start the interaction.

We must point out that, during most of the action, the face is out of the camera view, so the laser is the only device ATLAS can rely on. Eventually, when the person is approximately in front of the robot, the face detector could also help the tracking. This suggests that an additional pan-tilt control of the camera would be preferable.

### *D. Interaction*

The main task of ATLAS is to welcome the visitors entering the County Hall's foyer. This consists in detecting and approaching the people, greeting them and, if they desire, providing them with some useful information about the current exhibition. The interaction can last from a few seconds to a few minutes, depending on the interest of the user. To communicate, ATLAS shows HTML pages synchronized with speech, getting feedback from users through simple yes/no answers and with the touch-screen. In Fig. 13, we show a complete sequence of detection, approach, interaction and leave. The video with this and other performances is available online at *http://privatewww.essex.ac.uk/~nbello*.

## VI. Conclusions and future work

This paper presents a novel human detection system that integrates the information coming from a laser legs detector and an visual face detector. The way to recognize a typical human legs pattern from range information and how to speed up the face detection with tracking are explained. The results show that the system can be successfully used on a mobile robot to interact with people in real time.

In our future work we would like to use the pan-tilt movement of the camera to track people while the robot is moving. Probability methods will be investigated in order to improve the robustness and fault tolerance.

## References


[1] A. Ohya, "Human robot interaction in mobile robot applications," in *Proc. of the 2002 IEEE Int. Workshop on Robot and Human Interactive Communication*, 2002, pp. 5–10.

[2] Y. Nagumo and A. Ohya, "Human following behavior of an autonomous mobile robot using light-emitting device," in *Proc. 10th IEEE Int. Workshop on Robot and Human Communication*, Bordeaux and Paris, France, September 2001, pp. 225–230.

[3] K.-T. Song and C.-C. Chien, "Visual tracking of a moving person for a home robot," *Journal of Systems and Control Engineering*, to be published.

[4] J. N. K. Liu, M. Wang, and B. Feng, "ibotguard: an internet-based intelligent robot security system using invariant face recognition against intruder," *IEEE Trans. on Systems, Man and Cybernetics, Part C (SMC-C)*, vol. 35, no. 1, pp. 97–105, February 2005.

[5] D. Beymer and K. Konolige, "Tracking people from a mobile platform," in *Proc. of IJCAI-2001 Workshop on Reasoning with Uncertainty in Robotics*, Seattle, WA, USA, 2001. [Online]. Available: http://www.aass.oru.se/Agora/RUR01/proceedings.html

[6] M. Bennewitz, W. Burgard, and S. Thrun, "Learning motion patterns of persons for mobile service robots," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Washington, DC, USA, 2002, pp. 3601–3606.

[7] W. Burgard, P. Trahanias, D. Hähnel, M. Moors, D. Schulz, H. Baltzakis, and A. A., "Tourbot and webfair: Web-operated mobile robots for tele-presence in populated exhibitions," in *Proc. of the IROS 02 Workshop on Robots in Exhibitions*, 2002.

[8] J. Blanco, W. Burgard, R. Sanz, and J. Fernánez, "Fast face detection for mobile robots by integrating laser range data with vision," in *Proc. of the Int. Conf. on Advanced Robotics (ICAR03)*, vol. 2, Coimbra, Portugal, 2003, pp. 953–958.

[9] S. Feyrer and A. Zell, "Robust real-time pursuit of persons with a mobile robot using multisensor fusion," in *Proc. of the 6th Int. Conf. on Intelligent Autonomous Systems (IAS-6)*, Venice, Italy, 2000, pp. 710–715.

[10] M. Kleinehagenbrock, S. Lang, J. Fritsch, F. Lömker, G. A. Fink, and G. Sagerer, "Person tracking with a mobile robot based on multi-modal anchoring," in *Proc. of the IEEE Int. Workshop on Robot and Human Interactive Communic. (ROMAN)*, Berlin, Germany, 2002, pp. 423–429.

[11] M. Scheutz, J. McRaven, and G. Cserey, "Fast, reliable, adaptive, bimodal people tracking for indoor environments," in *Proc. of the 2004 IEEE/RSJ Int.l Conf. on Intelligent Robots and Systems (IROS '04)*, vol. 2, Sendai, Japan, 2004, pp. 1347–1352.

[12] G. Cielniak and T. Duckett, "People recognition by mobile robots," in *Proc. of AILS-2004, 2nd Joint SAIS/SSLS Workshop*, Lund, Sweden, 2004. [Online]. Available: http://ai.cs.lth.se/ails04/

[13] H. Asoh, N. Vlassis, Y. Motomura, F. Asano, I. Hara, S. Hayamizu, K. Ito, T. Kurita, T. Matsui, R. Bunschoten, and B. Krse, "Jijo-2: An office robot that communicates and learns," *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 46–55, 2001.

[14] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun, "Learning motion patterns of people for compliant robot motion," *The Int. Journal of Robotics Research*, vol. 24, no. 1, pp. 31–48, 2005.

[15] J. Fritsch, M. Kleinehagenbrock, S. Lang, G. A. Fink, and G. Sagerer, "Audiovisual person tracking with a mobile robot," in *Proc. Int. Conf. on Intelligent Autonomous Systems*, F. G. et al., Ed. Amsterdam: IOS Press, 2004, pp. 898–906.

[16] S. Lang, M. Kleinehagenbrock, J. Fritsch, G. A. Fink, and G. Sagerer, "Detection of communication partners from a mobile robot," in *Proc. of the 4th Workshop on Dynamic Perception*, Bochum, Germany, 2002, pp. 183–188.

[17] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 34–58, 2002.

[18] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in *Proc. the IEEE Int. Conf. on Image Processing 2002*, vol. 1, New York, USA, 2002, pp. 900–903.

[19] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features." in *CVPR (1)*, Kauai, HI, USA, 2001, pp. 511–518.

[20] M.-H. Yang, D. Roth, and N. Ahuja, "A SNoW-based face detector," in *The Conference on Advances in Neural Information Processing Systems (NIPS)*. MIT Press, 2000, pp. 855–861. [Online]. Available: http://l2r.cs.uiuc.edu/ danr/Papers/nips00.pdf

[21] OpenCV, "Open Computer Vision Library." [Online]. Available: http://sourceforge.net/projects/opencvlibrary

[22] J. Campbell and P. Pillai, "Leveraging limited autonomous mobility to frame attractive group photos," in *Proc. of the 2005 IEEE Int. Conf. on Robotics and Automation (ICRA '05)*, Barcelona, Spain, 2005, pp. 3407–3412.

[23] S. Ghidary, Y. Nakata, T. Takamori, and M. Hattori, "Human detection and localization at indoor environment by home robot," in *Proc. of IEEE Int. Conf. on Systems, Man, and Cybernetics*, vol. 2, Nashville, TN, USA, 2000, pp. 1360–1365.


**Nicola Bellotto** received his Laurea degree in Electronic Engineering from the University of Padua in Italy. He is currently a PhD candidate in robotics at the University of Essex in UK. His doctoral thesis focuses on Multisensor Data Fusion for Human-Robot Interaction. Other research interests include mobile robotics, computer vision and robot self-localization. Before joining the Human Centred Robotics (HCR) Group in Essex, he has been an active member of the Intelligent Autonomous System Laboratory in Padua and of the Centre for Hybrid Intelligent Systems at the University of Sunderland. He gained also several years of professional experience in embedded systems programming and he is currently working for Robocity Ltd. in London as research assistant.

**Huosheng Hu** received the MSc degree in industrial automation from the Central South University in China and the PhD degree in robotics from the University of Oxford in the U.K. Currently, He is a Professor in Computer Science at the University of Essex, leading the Human Centred Robotics (HCR) Group. His research interests include mobile robotics, sensors integration, data fusion, distributed computing, intelligent control, behaviour and hybrid control, cooperative robotics, tele-robotics and service robots. He has published over 180 papers in journals, books and conferences within these areas. He is currently Editor-in-chief for International Journal of Automation and Computing, and a reviewer for a number of international journals such as IEEE Transactions on Robotics, Automatic Control, Neural Networks and International Journal of Robotics Research. He is a Chartered Engineer, a senior member of IEEE, and a member of IEE, AAAI, ACM, IAS, and IASTED.