

People Tracking and Identification with a Mobile Robot

Nicola Bellotto and Huosheng Hu

*Department of Computer Science, University of Essex
Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom
{nbello & hhu}@essex.ac.uk*

Abstract—In this paper we present a novel and efficient solution for tracking and identifying people with a mobile robot using multisensor data fusion. The system utilizes a laser device to detect human legs and a PTZ camera to find faces, then the relative data is fused with a sequential Unscented Kalman Filter to perform real-time tracking. A metric based on the Bhattacharyya coefficient for color histogram comparison is also adopted to identify persons wearing different clothes. Finally, integrating the information coming from the tracking and the identification modules, we improve the robustness of the data association process. Some experiments with a mobile robot show the effectiveness of our approach.

Index Terms—*People tracking, sensor fusion, histogram-based identification, Unscented Kalman Filter, data association.*

I. INTRODUCTION

After the recent advancement of computing and robotics technology, human-centred or service robots will be soon ready to serve us in our home, hospital, office and everywhere. These robots are autonomous, interactive and intelligent. They should be able to be aware of the human presence and then act properly. This means, for example, finding people in the surrounding area who are willing to interact with the robot, but also keeping track of them in order to avoid possible collisions. An efficient and robust human tracking system is therefore necessary for practical applications of service robots.

In [1], stereo vision and Kalman filtering are used to track and follow a single person with a mobile robot at short distances. The system described in [2], instead, describes a laser-based solution, which implements a heuristic algorithm to detect and keep track of moving entities. In the work of [3], a robot equipped with two laser range sensors, one pointing forward and another backward, can track several people using a combination of particle filters and Joint Probabilistic Data Association (JPDA). A computational demanding solution is also implemented by [4], where a particle filter is used for the data fusion of a laser and an omnidirectional camera. In some cases the tracking system integrates also an identification module to recognize and label people. An example is presented in [5], where classic histogram intersection is used to identify people together with a standard Kalman filter for laser-based tracking. Even the robot used by [6] uses laser for tracking and vision for identification, but it works only for a single human target. More recently, [7] presented a robust

vision-based tracking and identification which makes use of a dynamic Bayesian network for handling multiple targets, although the camera's field of view is a strong limitation.

The solution presented in this paper uses multisensor data fusion techniques for tracking people with a mobile robot. The task is performed with a SICK laser, used for legs detection, and a PTZ camera, for face detection. The integration of these two devices improves the robustness of the tracking and augments the area covered by the detection system. The position of each human target is estimated with an efficient implementation of Unscented Kalman Filter and nearest neighbour data association. Our approach takes into account the limitations of the robot's hardware and is a good solution for achieving real-time performances in case of computational constraints. Vision is finally used for identifying persons using color histogram comparison. When available, identity information is also used to discriminate different targets, improving the robustness of the data association.

The paper is organized as follows. Section II introduces legs and face detection, plus a metric for color histogram comparison. Section III explains the tracking and identification systems. In Section IV, a solution for improving data association and handling multiple tracks is proposed. Section V reports some experiments and considerations. Finally, conclusions and future work are illustrated in Section VI.

II. DETECTION

A. Legs and Face Detection

Human legs can be detected using the laser sensor mounted on the robot, if this is not too high from the floor. Several implementations can be found in literature making use of such device. However, in most of the cases the legs detection is simply based on the search of local minima [5, 3], which are generally well distinguishable only in smooth environments, like corridors or empty rooms. Other solutions are based on motion detection [2, 4], missing therefore static persons and often becoming unreliable because of the difficult robot's motion compensation.

Our legs detection algorithm instead is based on the recognition of typical legs patterns [8]. These correspond to three possible postures: legs apart (LA), forward straddle (FS) and two legs together or single leg (SL). Fig. 1 shows the three patterns and a schematic representation of the legs detection algorithm. First of all, we filter the laser data in

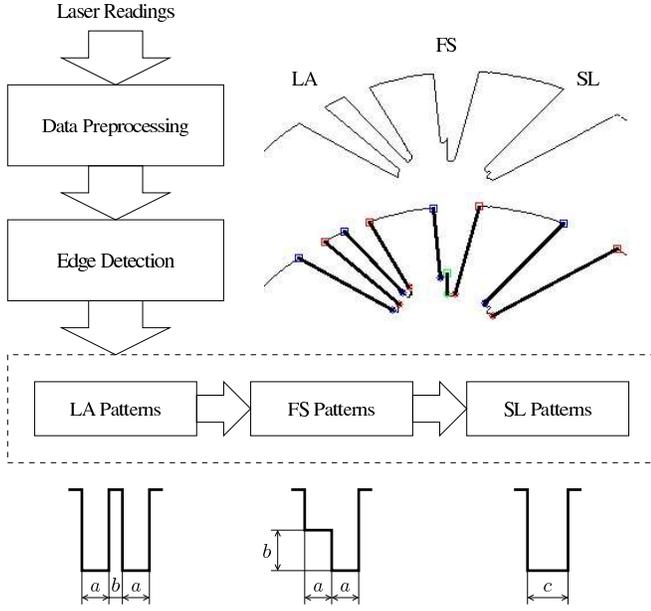


Fig. 1. Block-diagram of the legs detection algorithm. The three leg patterns are identified by sequences of vertical edges with the following constraints: leg width ($10 < a < 20$ cm), step length ($b < 40$ cm) and width of two legs together ($10 < c < 40$ cm). Patterns LA and FS are very seldom confused with other objects, while SL can be ambiguous for particular environments.

order to smooth the readings, then we detect all the edges lying on the directions of the laser scans. Finally, according to simple geometric relations and spatial constraints, we identify groups of adjacent edges which can correspond to legs. The method is quite robust even for challenging environment with clutters. Furthermore, it is computationally inexpensive and not influenced by the robot motion.

The face detection system is based on the real-time procedure of [9], which uses a cascade of classifier to extract simple but critical visual features. Besides being very fast, a nice characteristic of this algorithm is its color independence, so it is quite robust to light conditions; on the other hand, the weak point is its sensitiveness to head rotation and inclination. Finally, the direction of the face with respect to the camera, expressed with bearing α and elevation β , are calculated with the following simple transformations:

$$\alpha = \tan^{-1}\left(\frac{W/2 - u}{f}\right) \quad \beta = \tan^{-1}\left(\frac{v - H/2}{f}\right) \quad (1)$$

where (u, v) is the face's centre on an image $W \times H$ and f is the focal length in pixel units. An example of face detection is illustrated in Fig. 2.

B. Color Histograms Comparison

Different persons in the environment can be labelled according to the color histogram of their clothes, provided these are not completely identical. An efficient distance to compare color histograms is that one adopted by [10] for a

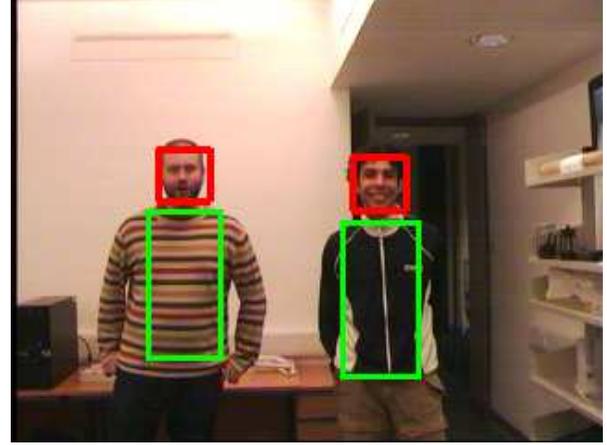


Fig. 2. Face detection and color histogram extraction. The area of the torso, from which the histogram is calculated, is 2/6 of the total human height.

mean-shift visual tracking algorithm, which is based on the sample estimate of the Bhattacharyya coefficient. Given a discrete (normalized) density of reference $\hat{\mathbf{q}} = \{\hat{q}_u\}_{u=1\dots m}$ from an m -bin histogram and another one at a given region of the image $\hat{\mathbf{p}} = \{\hat{p}_u\}_{u=1\dots m}$, the sample estimate of the Bhattacharyya coefficient is so defined:

$$\rho(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = \sum_{u=1}^m \sqrt{\hat{p}_u \hat{q}_u} \quad (2)$$

From this, the distance between the two distribution is calculated as follows:

$$d_h(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = \sqrt{1 - \rho(\hat{\mathbf{p}}, \hat{\mathbf{q}})} \quad (3)$$

Such a distance has some nice properties, among which the important fact of being scale invariant, since it uses discrete densities (instead the classic histogram intersection is not), and being normalized between 0 and 1.

The region from which we extract the color histogram of a person is the torso, since this is the only part of the body that is almost always visible from the robot's camera, either when the person is close or several meters far. As illustrated in Fig. 2, we consider the body proportions as in [6], where the torso is 2/6 of the total height (and legs are 3/6).

III. TRACKING AND IDENTIFICATION

A. State and Observation Models

When they are walking, people move often in an unpredictable way, so tracking becomes a very challenging task. Some researchers modeled human walking as a Brownian motion [11], although a constant velocity (CV) model is a better choice in case of clutters [1, 3]. Another approach consists in learning the typical human motion patterns inside a particular environment using a network of distributed sensors [5], but this is clearly an ad-hoc solution, often impossible to implement. We adopt instead a more general prediction

model [12], which is basically an extension of the CV model. The relative equations are the following:

$$\begin{cases} x_k = x_{k-1} + v_{k-1} \Delta t_k \cos \phi_{k-1} \\ y_k = y_{k-1} + v_{k-1} \Delta t_k \sin \phi_{k-1} \\ z_k = z_{k-1} + n_{k-1}^z \\ \phi_k = \phi_{k-1} + n_{k-1}^\phi \\ v_k = |v_{k-1}| + n_{k-1}^v \end{cases} \quad (4)$$

where $\Delta t_k = t_k - t_{k-1}$. Besides position (x_k, y_k) and orientation ϕ_k , a variable z_k is included for the estimation of the human height. Also, assuming a person is only walking forward, the velocity v_k is forced always positive. Noises n_{k-1}^z , n_{k-1}^ϕ and n_{k-1}^v are all zero mean Gaussians.

The measurements provided by the laser are bearing b_k and range r_k of the legs. These depend of course on the state of the robot, therefore the model includes its position and orientation (x_k^R, y_k^R, ϕ_k^R) given by the odometry, plus the following displacement of the laser device with respect of the robot's center of rotation:

$$l_k^x = x_k^R + L_x \cos \phi_k^R \quad l_k^y = y_k^R + L_x \sin \phi_k^R \quad (5)$$

The constant L_x is the distances of the laser from the robot's centre, lying on its longitudinal axis (L_y is null). The laser observation model is as follows:

$$\begin{cases} b_k = \arctan \frac{y_k - l_k^y}{x_k - l_k^x} - \phi_k^R + n_k^b \\ r_k = \sqrt{(x_k - l_k^x)^2 + (y_k - l_k^y)^2} + n_k^r \end{cases} \quad (6)$$

The noises n_k^b and n_k^r are zero mean Gaussians.

Similarly to the laser, the displacement of the camera is given by the following equations:

$$c_k^x = x_k^R + C_x \cdot \cos \phi_k^R \quad c_k^y = y_k^R + C_x \cdot \sin \phi_k^R \quad (7)$$

In addition to the robot's position, the camera model has also to take into account the pan ψ and tilt θ angles, plus its height C_z from the floor. The relative equations are therefore as follows:

$$\begin{cases} \alpha_k = \arctan \frac{y_k - c_k^y}{x_k - c_k^x} - \phi_k^R - \psi_k + n_k^\alpha \\ \beta_k = -\arctan \frac{z_k - C_z}{\sqrt{(x_k - c_k^x)^2 + (y_k - c_k^y)^2}} - \theta_k + n_k^\beta \end{cases} \quad (8)$$

Again, the noises n_k^α and n_k^β are zero mean Gaussians.

B. Sensor Fusion with UKF

For the success of a tracking algorithm, a key role is played of course by the state estimator. The Kalman filter [13] provides an efficient way to integrate different sensor data and perform the estimation. In case of linear systems with Gaussian noises, it is also proved to be optimal, while an Extended Kalman Filter can be used to provide approximate solutions in case of non-linearities. In recent years, most of

the applications for people tracking have adopted particle filters [4, 11, 3] since their performances are not constrained by linear or Gaussian assumptions. Unfortunately, in terms of computational cost, such estimators are generally quite demanding and the hardware requirements increase with the number of targets to track.

A valid solution to perform human tracking is the Unscented Kalman Filter (UKF) [14] that, instead of using a first-order linearization like in the EKF, captures mean and covariance of the probability distributions with carefully chosen weighted points, called "sigma points". Although this might sound similar to a particle filters, the two approaches differ on the fact that the sigma points are not random samples and their weights do not have to sum up to 1. Also, the number of points used by the UKF is so small to make this estimator particularly indicated for mobile robots with limited hardware resources.

In case of asynchronous and uncorrelated measurements, a Kalman filter can be updated sequentially using only the observation available at the current step [13]. So, when new information is coming just from one of the detection modules, legs or faces, only the relative observation model is used for the state correction. Moreover, when all the measurements are synchronized, a sequential update, starting from the most to the least precise sensor data, gives a better estimate for non-linear systems and it is also computationally more efficient. Under the same assumptions, the UKF can also be updated sequentially with the same benefits. When both legs and faces are detected, the filter is first updated by the laser data, which is more precise, and then by the vision.

C. Histogram Based Identification

People identification is based on a color histogram comparison as explained in II-B. First of all, if a target has been created but not yet labelled, the recognition system waits for his face to be detected, in order to estimate the region on the image where getting the histogram from. We select a window that covers approximately the human torso, then the relative histogram is compared to those ones already in memory using (3). Only the distances that fall below a certain threshold are considered and the target is labelled according to the histogram for which the distance is minimum. If none of the histograms in memory are "close" enough, the target is classified as unknown. Also, in case his face is detected, a new label is assigned to the unknown person and the relative histogram is memorized.

From several tests in real situations, under fixed light conditions and with people wearing different clothes, the histogram comparison based on (3) has proved to be very selective, with almost zero cases of misclassification, but still robust enough to recognize the same person in different postures.

IV. MULTIPLE TARGETS

A. Data Association

A fundamental part of every multitarget system is the data association, that is, the assignment of the current measurements to the proper tracks. In the last years different data association techniques have been used for people tracking, from probabilistic approaches, like JPDA [3], to multi-hypothesis algorithms [15]. In general, however, these methods are computationally quite expensive and therefore difficult to implement on powerless hardware. Instead, if the set of entities to track is not too dense, a simple solution based on Nearest Neighbour (NN) data association proved to be a good and fast alternative [5, 12].

The NN is an intuitive one-to-one association algorithm [13]. For each candidate track, the observation \mathbf{z}_k is initially predicted using the relative model. Then, a distance function d_{mn} is used to measure the similarity between every possible couple of predicted–real observation and this value is used to fill an association matrix $\mathbf{S}_{M \times N}$, where M is the number of sensor measurements and N is the number of predicted observations. Finally, the m -th measurement is chosen to update the n -th track if the relative element d_{mn} is the minimum distance (highest similarity) among all those in $\mathbf{S}_{M \times N}$; the relative m -th row and n -th column are excluded from further consideration. This last step is continuously repeated until there are no more available measurements or tracks to combine.

Before the association matrix creation, a common gating approach is adopted to avoid unlikely assignments. In practice, the gating procedure consists in excluding all the measurements \mathbf{y}_m outside a *validation region* [13]. This region is constructed around the predicted observation \mathbf{z}_n according to the relation $d(\mathbf{y}_m, \mathbf{z}_n) \leq \lambda$, where λ is a threshold and d is the Mahalanobis distance defined as follows:

$$d(\mathbf{y}_m, \mathbf{z}_n) = \sqrt{(\mathbf{y}_m - \mathbf{z}_n)^T \boldsymbol{\Sigma}_{mn}^{-1} (\mathbf{y}_m - \mathbf{z}_n)} \quad (9)$$

$\boldsymbol{\Sigma}_{mn}$ is the covariance matrix of the innovation $(\mathbf{y}_m - \mathbf{z}_n)$. The value of λ can be determined from tables of the *chi*-squared distribution ($\lambda = 3.03$ in our case).

After the validation gate, the data association proceeds with the actual NN assignment. The most common measure of similarity between real and predicted observations is based on the same distance defined in (9). However, at this point we introduced an improvement that makes use of the identity information, i.e. the color histogram of the person. Using (3), we calculate the distance $d_h(\hat{\mathbf{p}}_m, \hat{\mathbf{q}}_n)$ between the histogram $\hat{\mathbf{p}}_m$ of the region relative to the current measurement and the histogram $\hat{\mathbf{q}}_n$ of the considered track. Such a distance is compared to a threshold γ (for example $\gamma = 0.8$) and, if greater, the measurement is discarded. This is basically an additional gating, although no *chi*-squared distributions are assumed in this case. If smaller than γ instead, d_h is

multiplied to d and the result is used as a new similarity measure. The relative equations are the following:

$$d_{m,n}^* = \begin{cases} d_h(\hat{\mathbf{p}}_m, \hat{\mathbf{q}}_n) d(\mathbf{y}_m, \mathbf{z}_n), & \exists \hat{\mathbf{p}}_m, \hat{\mathbf{q}}_n \text{ and} \\ & d_h(\hat{\mathbf{p}}_m, \hat{\mathbf{q}}_n) \leq \gamma \\ d(\mathbf{y}_m, \mathbf{z}_n), & \nexists \hat{\mathbf{p}}_m \text{ or } \nexists \hat{\mathbf{q}}_n \end{cases} \quad (10)$$

In practice, whenever the histograms $\hat{\mathbf{p}}_m$ and $\hat{\mathbf{q}}_n$ exist, respectively for the measurement and the track of reference, d is weighted by d_h (with $0 \leq d_h \leq \gamma$). If $\hat{\mathbf{p}}_m$ is similar to $\hat{\mathbf{q}}_n$, the new similarity measure $d_{m,n}^*$ will be much smaller than the original one, otherwise it will tend to d . When one or both the histograms are missing, because the camera is pointing to another direction or the track is not labelled, the similarity is simply given by the Mahalanobis distance.

Using (10), two different association matrices are created, one for the legs and another one for the face detections. In the first case, the histogram $\hat{\mathbf{p}}_m$ is computed from the region corresponding to the legs' direction, if covered by the camera's field of view; its size is a function of the legs' distance and the *a priori* height of the considered track. In the second case instead, $\hat{\mathbf{p}}_m$ is computed from the region corresponding to the face's direction; its size is a function of the face's elevation and the *a priori* position of the target.

B. Creating and Removing Tracks

To create new tracks we use the laser readings discarded by the gating procedure and the assignment. In particular, only the leg patterns LA and FS, explained in Section II, are selective enough to be considered trustful for tracks creation. Faces are not taken into account at this stage as they cannot provide range information, indispensable for estimating the initial 2-D track's position. In parallel to the human tracks database, we keep then another list containing all the possible candidates. Each one of these is generated by a sequence of readings falling inside a certain region, delimited by the distance a person can cover in the interval Δt_k at a certain speed (we chose 1.5m/s). Each candidate is assigned a maximum time interval, or "lifetime": if during this interval there are enough readings falling inside its region, the candidate is promoted to human track, otherwise it is considered a false positive and removed. Finally, normal tracks are deleted from the database if not updated for more than a certain time.

V. EXPERIMENTAL RESULTS

The system has been implemented on a Pioneer robot equipped with a SICK laser and a PTZ camera. As shown in Fig. 3, the laser is located at approximately at 30cm from the floor and the camera is mounted on a special support, at about 1.5m, in order to facilitate the face detection. The on-board PC is a Pentium III 800MHz with 128MB of RAM, running Linux OS.



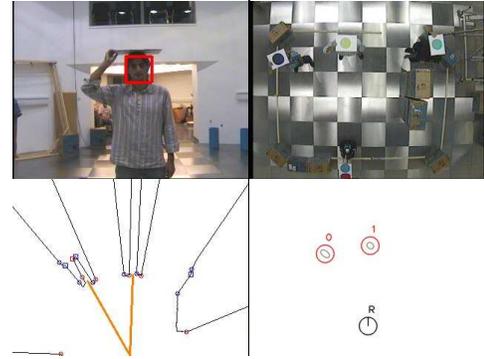
Fig. 3. Robot equipped with a SICK laser and a PTZ camera for the experiments. The camera used is the black one on the top.

The whole software has been written in C++ and runs in real-time on the robot's PC, although it is possible to use an external client, connect via wireless, for remote control and debug. The laser performs scans of 180° at 5Hz, which is also the update frequency of our program, with range and angular resolution of 1cm and 0.5° respectively. The camera has a field of view of about 49° and provides images with a resolution of 320×240 pixels at 10Hz. Every track is created after at least 3 readings within the maximum time interval of 1s and is removed if not updated for more than 2s.

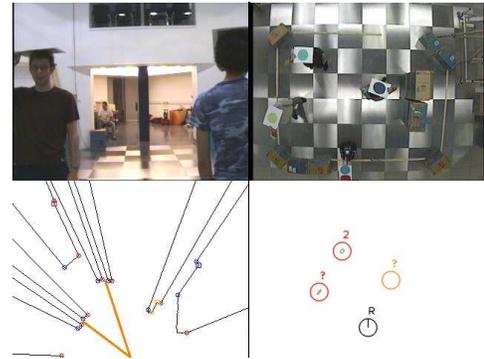
A. People Tracking and Identification

In the first experiment we wanted to test the general performance of the tracking and identification system. We set up a little environment inside the robotic arena where three people could walk in front of the robot or hide behind a cardboard wall. From a camera mounted on the ceiling we could observe the absolute positions of the three human targets, each one keeping a different color marker on the head. In this case, since the robot's camera could almost always detect at least one face just pointing straight, the pan/tilt control was kept fixed. Initially, all the persons were hiding behind the wall. Then they came out, one by one, facing to the robot's camera in order to be labelled a first time. Finally, they started moving together at normal walking speed, sometimes going back behind the wall, and sometimes performing circular paths in front of the robot. Several occlusions made the tracking task particularly difficult.

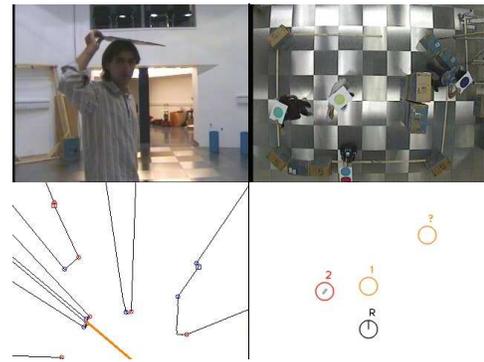
A few situations are illustrated in Fig. 4: for each of them, an image from the robot's camera and a laser scan are shown on the left, while a bird-eye-view and an estimation result show the tracks on the right. During the experiment, approximately 60s long, the estimated positions and the labels of the target were correct most of the time. In Fig. 4(b), target "0" and "1" were temporary lost after a sharp turn, so they were labelled with "?". Unfortunately, once visible again a few instants later, only target "1" was correctly re-labelled. Target "0" instead was not facing the robot and could not be recognized, so in Fig. 4(c) he was still keeping a "?" label.



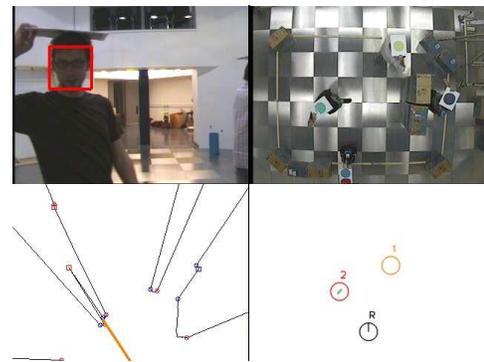
(a) Initial tracking and labelling of the people.



(b) Target "0" and "1", temporary lost, marked with "?".



(c) Only target "1" reacquires the correct label.



(d) The target behind the wall is removed after 2s.

Fig. 4. People tracking and identification. Three persons are labelled, in order, from "0" to "2" and tracked during paths in front of the robot. A sequence of four different moments, at intervals of about 5s, is shown.

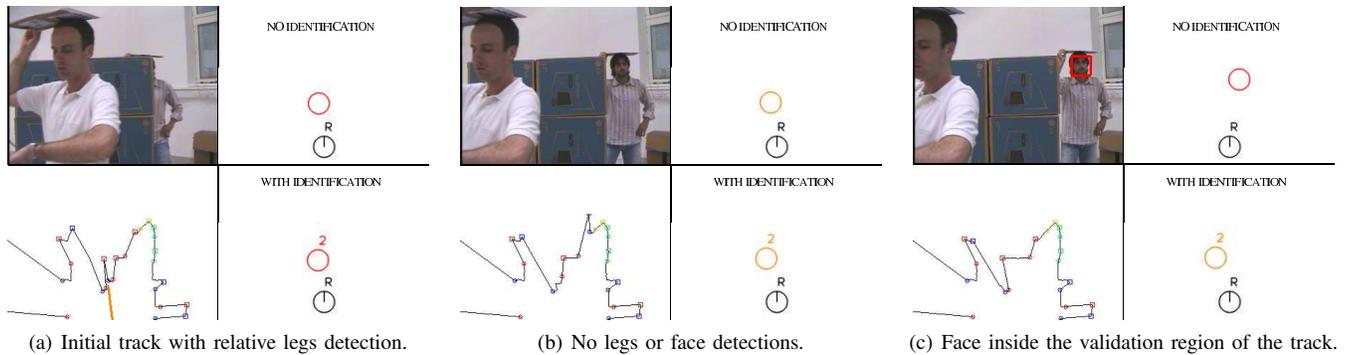


Fig. 5. A person being tracked in front of the robot. The face of a different person is detected and it can be erroneously associated to the track.

B. Data Association With and Without Identification

We show a typical case of data association problem and compare the results with and without the integration of the histogram-based identification. In Fig. 5 there is a sequence of images and laser scans taken at intervals of 0.4s during one of our experiments. In the first frame, the legs of the closest person were detected; in the second one, no detections were available; finally, in the third frame, only the face of another person was detected. Note that, in the last case, the detected face was inside the validation region of the track.

The top right part of each subfigure shows the track of the person in front of the robot without using the histogram-based identification. The track was initially correct, but in Fig. 5(c) there was an estimation error because the system used the other face to update the track. In this case, the gating process, based only on positional information, was not good enough to discard the wrong measurement. In the bottom right part, instead, which represents exactly the same situation, the gating process performed correctly thanks to the identity information. As shown by Fig. 5(c) indeed, the track estimate with identification was still correct, since the color histogram of the second person was different from the initial one and the relative face was therefore discarded.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, a novel approach for tracking and identifying people with a mobile robot is presented. At first, a sequential version of the UKF was implemented, which is computationally efficient and suitable for non-linear systems. Different human targets are then identified using color histograms, and the identity information integrated into the data association process to improve multitarget tracking. Experiments show the good performance of the proposed system.

To improve the robustness against environmental changes, we will implement an adaptive thresholding scheme for the histogram based identification. The next step of our research will be then the study of feasible real-time solutions for joint multitarget tracking and identification applied to human-robot interaction.

REFERENCES

- [1] D. Beymer and K. Konolige, "Tracking people from a mobile platform," in *IJCAI-2001 Workshop on Reasoning with Uncertainty in Robotics*, Seattle, WA, USA, 2001.
- [2] M. Lindström and J.-O. Eklundh, "Detecting and tracking moving objects from a mobile platform using a laser range scanner," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, vol. 3, Maui, HI, USA, 2001, pp. 1364–1369.
- [3] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "People tracking with mobile robots using sample-based joint probabilistic data association filters," *Int. Journal of Robotics Research*, vol. 22, no. 2, pp. 99–116, 2003.
- [4] P. Chakravarty and R. Jarvis, "Panoramic vision and laser range finder fusion for multiple person tracking," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Beijing, China, 2006, pp. 2949–2954.
- [5] M. Bennewitz, G. Cielniak, and W. Burgard, "Utilizing learned motion patterns to robustly track persons," in *Proc. of Joint IEEE Int. Workshop on VS-PETS*, Nice, France, 2003, pp. 102–109.
- [6] G. Cielniak and T. Duckett, "Person identification by mobile robots in indoor environments," in *Proc. of the IEEE Int. Workshop on Robotic Sensing (ROSE)*, Örebro, Sweden, 2003.
- [7] W. Zajdel, Z. Zivkovic, and B. J. A. Kröse, "Keeping track of humans: Have I seen this person before?" in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Barcelona, Spain, 2005, pp. 2093–2098.
- [8] N. Bellotto and H. Hu, "Multisensor based human tracking for an indoor service robot," 2007, submitted to *IJRR*.
- [9] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [10] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, South Carolina, USA, 2000, pp. 142–149.
- [11] M. Montemerlo, W. Whittaker, and S. Thrun, "Conditional particle filters for simultaneous mobile robot localization and people-tracking," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Washington DC, USA, 2002, pp. 695–701.
- [12] N. Bellotto and H. Hu, "Vision and laser data fusion for tracking people with a mobile robot," in *Proc. of the IEEE Int. Conf. on Robotics and Biomimetics (ROBIO)*, Kunming, China, 2006, pp. 7–12.
- [13] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Y. Bar-Shalom, 1995.
- [14] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [15] G. Taylor and L. Kleeman, "A multiple hypothesis walking person tracker with switched dynamic model," in *Proc. of the Australian Conf. on Robotics and Automation*, Canberra, Australia, 2004. [Online]. Available: <http://www.araa.asn.au/acra/acra2004>