

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/229121889>

Autonomous pallet localization and picking for industrial forklifts: A robust range and look method

Article · August 2011

DOI: 10.1088/0957-0233/22/8/085502

CITATIONS

27

READS

2,537

7 authors, including:



Luca Baglivo
Eurac Research
32 PUBLICATIONS 228 CITATIONS

[SEE PROFILE](#)



Francesco Biral
Università degli Studi di Trento
106 PUBLICATIONS 1,342 CITATIONS

[SEE PROFILE](#)



Nicolas Bellomo
Technology for Propulsion and Innovation
41 PUBLICATIONS 460 CITATIONS

[SEE PROFILE](#)



Enrico Bertolazzi
Università degli Studi di Trento
92 PUBLICATIONS 1,434 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Rosetta Mission: Osiris and Giada instruments [View project](#)



Optimal control applied to human motion and physiology [View project](#)

Autonomous Pallet Localization and Picking for Industrial Forklifts. A Robust Range and Look Method

Baglivo L.¹, Biasi N.¹, Biral F.¹, Bellomo N.², Bertolazzi E.¹,
Da Lio M.¹, De Cecco M.¹

¹Department of Mechanical and Structural Engineering, University of Trento, via Mesiano 77, 38100 Trento, Italy. Phone: +39 461 28–2512

²CISAS, University of Padova, via Venezia 41, 35100 Italy

E-mail: luca.baglivo@ing.unitn.it

Abstract. A combined double-sensor architecture, laser and camera, and a new algorithm named RLPF, are presented as a solution to the problem of identifying and localizing a pallet the position and angle of which are *a priori* known with large uncertainty. Solving this task for autonomous robot forklifts is of great value for logistics industry. The state-of-the-art is described to show how our approach overcomes the limitations of using either laser ranging or vision. An extensive experimental campaign and uncertainty analysis are presented. For the docking task, new dynamic non linear path planning which takes into account vehicle dynamics is proposed.

Submitted to: *Meas. Sci. Technol.*

Keywords— pallet localization, autonomous forklift, logistics automation

1. Introduction

In the field of industrial research, there is increasing interest in extending logistics automation to unstructured environments. In practise, this means that robots have to carry out several specific human tasks without infrastructures, e.g. docking to a pallet which pose is not predetermined. Achieving of this challenging goal would have as its main result the complete materials handling automation including the trailer loading/unloading phase. The main advantages are the use of less structured and thus much cheaper stocking areas, less batch damage and less risk of injury to personnel. In this scenario, industrial autonomous vehicles must be able to 'understand' their surroundings and to cope with uncertainty. Completely or partially *a priori* unknown environments introduce more uncertainty and less repetitiveness in the robotic handling tasks, so that reliable but cost-effective sensors must be employed together with robust real-time algorithms. Conversely, faulty situations may also occur in highly repetitive and deterministic tasks when, for example, a pallet lies in an unexpected position and a blind pick-up would cause a system fault and possible damage. This paper proposes the use of a vision and a range sensor combination for novel applications to autonomous forklifts with a robust sensory processing algorithm to detect, localize, and pick up pallets in semi-structured environments. We follow the same approach of [1] in defining the problem: the term 'semi-structured' environments refers to a predetermined region hosting pallets placed with high uncertainty in an area with irregular pattern. This means that the robot only knows in which region it may find the pallet but not where it is precisely. The task assigned to the robot is, therefore, to detect pallets in a uncertain loading area, compute their position and angle, choose one pallet according to a specified criterion and, once a safe path has been autonomously computed, engage it with a very low failure rate ($< 0.1\%$). There are several works on pallet detection, localization, loading and unloading. They can be subdivided into two groups according to the two main sensor types: camera (monocular or stereo) and 2D or 3D time-of-flight Laser Range Finder (LRF). The state-of-the-art is discussed in more details in the next section. The proposed approach uses both kinds of sensor, camera and 2D LRF to overcome the main limitations of each of them. While the 2D range information is gained from the LRF and is used to detect and localize the pallet accurately according to both 1D and 2D pallet models, the camera provides complementary information along the third dimension. The colour approach used in a previous work [2] has been replaced by a more reliable model-based kernel to confirm the pallet detection in case of poor detection by the LRF. This method allows more robustness, as extensive experimental results show.

A critical task in the trajectory planning phase is to engage the pallet once it has been localized. An *ad hoc* real-time iterative trajectory planning strategy has been developed and simulated. It takes into account non-linear vehicle dynamics and is solved

in a vehicle local reference system, relative to which both pallet position and attitude are expressed. This scheme is robust to model parameters inaccuracies and vehicle and pallet state uncertainty. More details may be found in Section 6.

After the following comprehensive state-of-the-art and the motivation for the work, this paper reports the previous work on which the present one is based (Section 2). The pallet finder algorithm is described in Sections 3, 4, and 5. Lastly, experimental results and the localization algorithm uncertainty are given in Section 7.

1.1. State-of-the-art

Vision approaches exploit large amounts of information and are mainly based on the detection of natural or artificial features and/or morphological and colorimetric segmentation. Some of them use artificial features on pallets [3] and tines [4]; others extract geometric information from natural pallets [5, 6, 7], color information [8] or both [9], or use structured light [10]. 2D LRFs [11, 12] provide direct range measurements and have the important advantage of accuracy almost independent of range. However, in some cases 2D range does not provide sufficient understanding of the surroundings [13], especially in cluttered environments or ones with similar multiple pallets. 3D LRF obtained by rotating a 2D LRF [1] does not allow closed-loop trajectory control, measurements must be made done while the robot is still. Time-of-flight infrared cameras allow good object segmentation even in cases of partial occlusion [14], but pallet location accuracy and reliability versus operating distance has not yet been characterized. More details about the state-of-the-art follow.

A vision-based pallet detection and localization system mounted on a retrofitted manual forklift was first proposed by [8]. They employed a natural-feature approach based on the identification of two pallet slots and estimation of their centroids position in calibrated images. The system requires high prior knowledge of pallet pose, is active only in the final load phase, when the pallet is at a short distance from the forklift, with the tines, but not the vehicle, moving for the engagement. A pre-industrialized retrofitted autonomous forklift capable of stacking racks and picking pallets placed with limited uncertainty is described in [7]. It is based on detection of specific reference lines for concurrent camera calibration and identification, and allows stacking of well-illuminated racks and localization of pallets in front of the vehicle and close to it. [5] proposed a more complex scheme based on hierarchical visual features like regions, lines and corners using both raw and template-based detection. The main problems are complexity in decision trees and weak scale invariance. Another more recent and simpler approach proposes 3D position and attitude computation of pallets by extrinsic calibration of the camera relatively to the vehicle and the floor and completely relies on the upper and bottom edges of the front pallet face [6], for this reason it can not be considered robust. Colour-based segmentation [9, 2] turned out to be failure-prone due to ambiguity deriving from similar background objects or, for instance, from an unexpected wrapping layer covering part of the region of interest. The structured light

method, well suited for 3D reconstruction of close objects, has also been attempted for pallet pose measurement [10]. The main problem is that its accuracy quickly decreases with distance, apart from the greater complexity of the system with respect to LRFs. The issue of robust detection and picking a load at longer distances have been addressed by using artificial features placed on the object. In specific cases, like in [3], the task can justify the application of special targets on the load to be picked. In the generic pallet case, the application of fiducials [4] can solve most of the illumination and calibration problems but it is not a general solution since it requires artificial targets to be applied to each pallet, and it is also difficult to guarantee the targets integrity over time. The same consideration is valid for fiducials applied to tines [4] for a camera space closed loop engagement.

Bypassing the visual projective world, direct range measurements with LRFs can be used for pallet detection and localization but suffer from matching problems in the case of multiple targets [13]. Proposed methods are heuristic or model-based. The authors of [12] developed an outdoor autonomous forklift engaging pallets both from the ground and from truck trailers. They used a fast linear program for segment detection, applied to pre-filtered points selected by man gestures on a PDA showing an image from a camera mounted on the forklift. The pallet is identified by classification of detected segments belonging to its front face the position of which is then computed. A model fitting method based on a energy function minimization was proposed by [11]. Both approaches (and [11] and [12]) require a quite good prior knowledge of pallet pose and do not analyse their performances in two important cases: undesired objects nearby the pallet and multiple pallets in the working field of view of the sensor.

1.2. Motivation of the work

This work forms part of the AGILE European project. A partial objective of the project is to achieve an industrial implementation of the pallet localization tasks for autonomous forklifts. The above state-of-the-art shows that this task cannot be considered to be robustly solved. 3D time-of-flight cameras, overcoming problems of both 2D ambiguity and dynamic measurements problems, represents a key choice. Unfortunately, the accuracy of commercial sensors (a MESA SR4000 was used in this case) is still too low for this task. Cameras in stereo configuration can also provide 3D measurement but suffer from poor robustness due to matching disturbed by foreground and background objects. Instead, 2D range measurements, instead, can be provided with good accuracy by the safety LRF already mounted on the autonomous forklift. They suffer from the problems mentioned in the previous section. We thus propose using both range and vision data for the pallet search. Specifically, while LRF data are used for a metric position estimation of pallets, the camera is used only for consensus of pallet identification by means of pattern recognition methods to solve cases of ambiguity like those illustrated in the following section. This configuration only requires *soft* extrinsic calibration of camera and LRF and no intrinsic camera calibration because images are not used for pallet

location in the same LRF 3D space, but only for pallet identification in the image space, where the LRF points are also projected. In addition, lighting conditions have less influence only on object identification with the image, compared with full visual 3D location, using for instance an inverse perspective method: the latter requires much better illumination.

The literature, except for [12], does not report either extensive experiments, which are fundamental for industrial implementation, or test results with deviations from nominal conditions, like multiple pallet search. Our work includes the accuracy of the method obtained from extensive testing, also with treatment of ambiguous cases with more than one pallet and nearby disturbing objects, which is closer to a real scenario. It starts from a previous work [2] relying on LRF and color camera, overcoming important limitations, as described below.

2. Previous work

This method for pallet pose estimation starts from that proposed in a previous work [2], relying on LRF and colour camera. Identification of the position and orientation of a pallet was based on best fit, found with a genetic optimization, between the scan data acquired with LRF and a pallet, 2D model made up of line segments representing the horizontal section of supporting blocks (Fig. 1) Pose estimation may fail with this

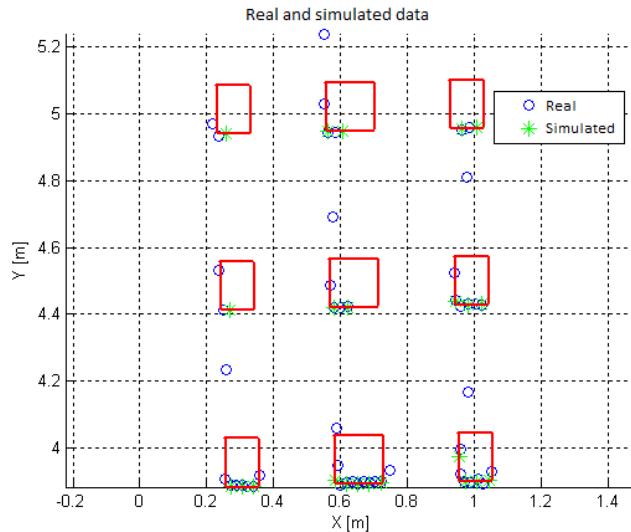


Figure 1. Pallet model as horizontal section of supporting blocks, fitted to the LRF points. Also plotted: simulated scan of the estimated pallet.

method (due to local minima in combination with poor prior knowledge of the initial position of the pallet) if there are outsider objects near the pallet, when two pallets are side by side, or when the pallet is symmetrically displaced over the main axis of the LRF. In some cases, these failures are solved by re-seeding of the genetic algorithm

until convergence, normally within two steps. To solve these ambiguous situations, a color camera is mounted near the LRF, mounted so that its field of view and that of the LRF overlap in the region of interest. A calibration procedure [15] and knowledge of the internal parameters of the camera yield the relative position and orientation of both LRF and camera. It is then possible to combine the data between the two instruments – for example, by projecting on to the image plane the LRF measured points belonging to the overlap field and assigning to them a color and a region of similar neighboring pixels in the image (check Section 4). A filter based on the color properties of the pallet is then applied: only pixels with RGB values similar to the chromatic properties of the pallet are selected. Region segmentation of candidate pallets starts from the LRF projected points, which are used as ‘seeds’ of a standard region-growing algorithm.

This method has two main drawbacks. The first is the uncertain, although upper-bounded, computation time of the genetic optimization starting from a random initial guess. Further re-seeding may also be required, and in any case does not grant convergence in ambiguous cases as described above. The second weak point is the selection of pallet colour for statistical filtering of image pixels. This is indeed a difficult operation, due to variable pallet appearance (RGB value). Changes in lighting conditions and variable pallet aging do not allow to define a narrow RGB value interval which represents pallet color. A direct consequence is that the probability of including background parts in the pallet segmented region increases. These two problems were solved, respectively by improved pallet pose initialization in the LRF data and template matching for the image processing part.

3. Pallet Localization with range data

Without loss of generality, pallet pose estimation can be simplified and speeded up by identifying an initial coarse pallet position and estimating attitude from a 1D model, and using it as an initial guess for refinement through 2D model fitting optimization with full data. We refer to such an operation as *initialization*. To initialize the pallet pose, we search the LRF scan for three collinear straight segments representing the frontal blocks of pallets as seen from a horizontal section (Fig. 2) This operation, called

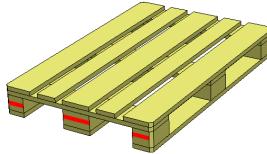


Figure 2. The line segments represents the LRF measurements relative to the pallet front section. They are searched for in the initialization phase.

pixelization, is achieved by transforming the scanned 2D Cartesian LRF space into a binary image, creating an image representing a discretized space section of the field

of view of the LRF. Image pixels assume value of 1 when they contain at least one measured point, otherwise 0 (Fig. 3). Pixelization is carried out to exploit standard image correlation techniques for model matching. The position of the front blocks of

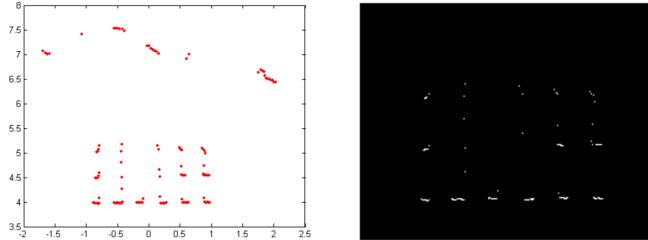


Figure 3. Left: LRF scan of two pallets side by side. Right: pixelized LRF scan, enlargement on region of interest.

the pallet can be identified, matching the image with a template binary image which shows the blocks as seen from the top (Fig. 5). Before convolving the template with the image, the angles of the candidate pallets must be identified in order to rotate the template accordingly. Candidate orientations of the pallet are computed with a Hough transform of the pixelized scan, and only the lines are kept, of which at least two segments belong to (Fig. 4). We then create the model rotated by the same angle of

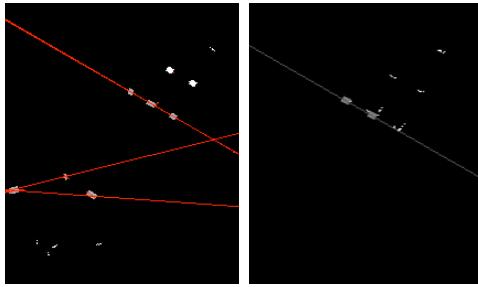


Figure 4. Hough lines in the pixelized scan. Right: selected front section line for candidate pallet

the identified line. If we detect multiple lines with different slopes, we create as many models as the number of lines. The best match position for each model is assigned as the initialization output position for pallet searching and pose refining, as described above. This procedure allows multiple pallet identification, in the LRF full visibility condition: if a pallet is found, its associated data are cancelled and the search restarts.

Starting positions are now good initial guesses for the optimization phase, so that a direct search algorithm [16] is used instead of genetic optimization. Initialization solves the problem of getting stuck in local minima and speeds up the identification process. After pose estimation refinement, as described by [13], a validation gate is triggered. A simulated scan of the pallet in the found position is simulated and each obtained point



Figure 5. First on the left is the zero angle template, the following are the templates associated with lines found in the image of Fig. 4

is associated with its nearest in the real scan. Pallet pose is accepted if the maximum percentage of coupled points (PCP – Percentage of Coupled Points) of the real scan is higher than a threshold (we choose 80%, which has to be tuned according to LRF uncertainty).

The initialization procedure can sometimes fail, with the result that other objects are classified as pallets. This may happen when the load area contains objects, such as shelves or obstacles in general, placed with the same layout as frontal pallet blocks (Fig. 6). We therefore decided to reinforce the decisional power of the *range and look*



Figure 6. Scenario with a cluster of object confused as a pallet by the LRF. Validation gate using camera image discards this false positive solution.

pallet finder (RLPF) algorithm relying on camera information. Instead of using color statistical filtering [2], we applied a procedure based on geometrical information. The PCP is computed for each initial pose from the initialization phase and refined. For those candidate poses which have a PCP higher than a certain threshold (80%), the RLPF stops successfully. Remaining candidates are analysed by interrogating the camera, and are accepted only if the image-processed data confirm the result. A logic flowchart is shown in Fig. 7. Before the image processing part of the RLPF algorithm, we illustrate LRF-camera extrinsic parameter calibration and occluded point filtering.

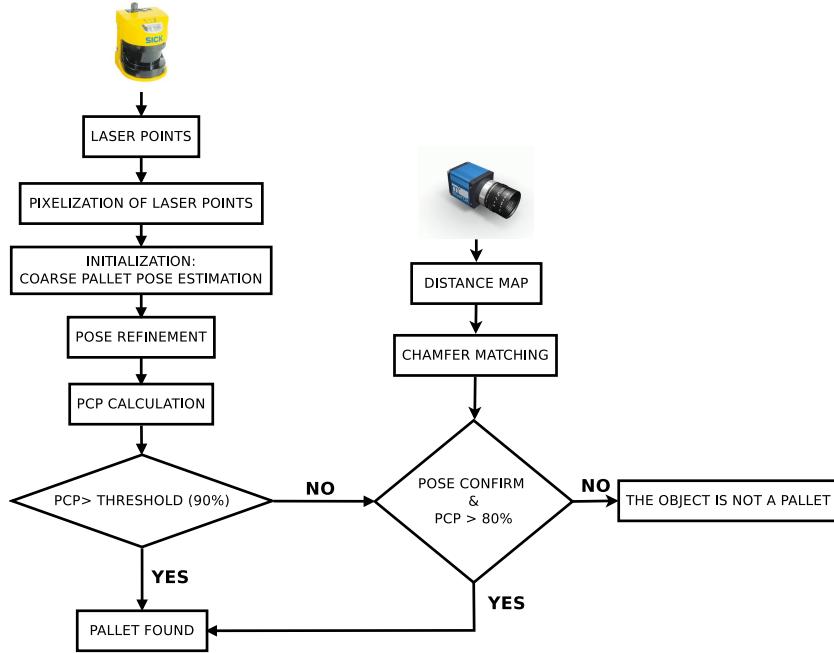


Figure 7. Flowchart of the RLPF algorithm.

4. Extrinsic calibration

Once the homogeneous matrix between LRF and camera reference system is computed, laser (infrared) measurements can be projected on the image plane Fig. 9. Calibration can be carried out with the procedure described in [15]. The ideal layout of the two instruments is zero baseline, i.e., with the optical axis of the camera parallel to the laser scan plane and its optical centre belonging to the rotation axis of the LRF mirror. The real layout is not ideal, so we apply a filter to take into account occlusions and different field of views (FOVs) between LRF and camera due to the unavoidable non-zero baseline. As shown in Fig. 8, if all the measured LRF points are projected onto the image, some points which should be hidden behind pallet blocks or are outside the camera FOV are classified as visible to the camera. To avoid such errors, we first filter out all laser points which do not belong to the camera FOV. The second criterion for filtering is based on the sequencing direction coherence of laser points, as seen from the laser source and from the camera optical centre. LRF measured points have a known angular sequencing, i.e., clockwise. In the particular case of a zero baseline, the two angular sequencing directions are always the same. If a LRF-camera baseline is different from zero, or if there are foreground objects which hide some laser points from the camera, the two sequencing directions are opposite for these points, if one sequence is taken from the laser source and the other from the camera optical centre. Checking point pair-wise the order coherence of the two sequences allows the hidden points to be filtered out. The result is effective and correct, with a not too large deviation from the zero baseline condition. The effect of the two filters is shown in Fig. 9. We refer to the

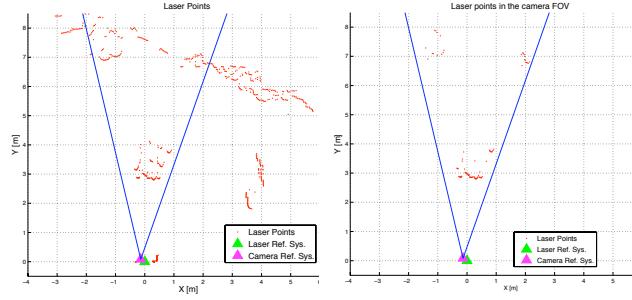


Figure 8. Left: scan plotted in LRF reference frame registered with camera reference frame. The two lines indicate the camera field of view. Right: result after application of filters that eliminate points out of camera view and occluded.

resulting image added with filtered laser points as *camera-LRF image*.

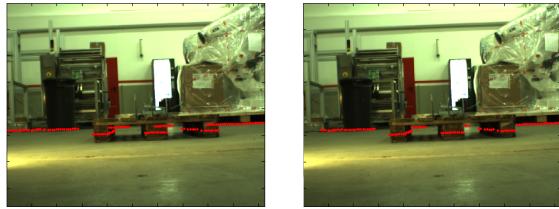


Figure 9. Same data as in Fig. 9 projected onto image as are (left) and after filters application (right)

5. Pallet localization. Camera consensus

A high PCP score turns out to be a powerful validation for pallet identification with only LRF data. Instead, lower scores (in any case greater than a fixed threshold) provide little information about whether identification was successful or not. In this case, the image is checked and the result is used as a gate for the identification output obtained from LRF data. The candidate pallet is accepted if *camera-LRF* image processing confirms the result. The validation gate evaluates the proximity measure, in image space, between the pallet position estimated from LRF points and that obtained directly from image pattern recognition. Recognition is based on geometric shape model matching [17] applied to the extracted edges. The template is made up of line segments representing the edges of the front pallet section (Fig 10). Matching template T and test image I requires computing an edge image and its Distance Transform (DT). A DT converts a binary image, like an edge image, which contains feature and non-feature pixels, into an image in which each pixel value denotes the distance to the nearest feature pixel. There are several algorithms to calculate the DT which differ from each other according to the distance metric used and the way local distances are propagated. The one we use is the Chamfer Distance, which computes an approximated Euclidean distance by integer

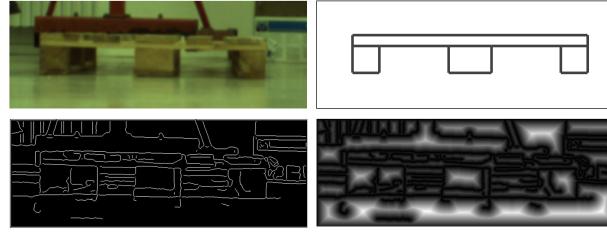


Figure 10. Clockwise from upper left corner: image acquired; pallet model edge image; extracted edges; chamfer DT image.

arithmetic [18]. After computing the DT image of I (Fig. 10), the template is convolved with it and the result is normalized by the number of egde pixels in the template. In practice, the template acts as a mask convolved with the DT image which only selects the DT values of pixels corresponding to the edge pixels of the template; then the mean of those pixels is computed. The similarity measure between I and T , for each position of T in I , is:

$$D_{\text{chamfer}}(T, I) = \frac{1}{|T|} \sum_{t \in T} d_I(t) \quad (1)$$

where $|T|$ is the number of features (pixels in our case) of the template T and $d_I(t)$ denotes the chamfer distance of the image pixel in I corresponding to the feature pixel t in T and the closest edge in I . The D_{chamfer} value is expressed according to metric, as above. The best match between T and I is the one which minimizes D_{chamfer} (Fig. 11). The advantage of using the DT image rather than the edge image is that the similarity measure is smoother as a function of the template translation parameters and therefore has better convergence performance to the solution. It also allows for some degree of dissimilarity between the template and the pallet in the image, like in the case of missing edges.

A pseudo code version of the RLPF algorithm is reported in Algorithm 1

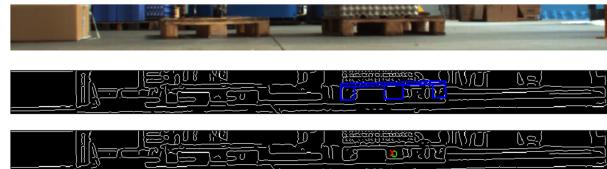


Figure 11. From top down: captured image of two pallets; edge image with DT solution after chamfer matching; LRF finds pallet position, which is projected onto image (cross) and DT solution (circle) in 2D image space. In this case, the found pallet is accepted because the two solutions are close enough to each other. The offset between the two is due to several sources of noise, mainly LRF measurement uncertainty, extrinsic parameter uncertainty, image noise on edge extraction, and DT integer metric.

Algorithm 1 PALLET LOCALIZER

```

1: acquire SCAN
2: pixelize SCAN
3: find LINES
4: for all LINES do
5:   create rotated model
6:   convolve model on pixelized scan
7:   if model match successful then
8:     return palletpose( $i$ ) {fill candidate poses set}
9:   end if
10: end for
11: if palletpose is not empty then
12:   select best pallet {user-defined criterion}
13:    $X_{pallet} \leftarrow$  best pallet pose
14:   find all points belonging to  $X_{pallet}$ 
15: else
16:   goto 1:
17: end if
18: refine  $X_{pallet}$  {apply local minimization }
19: compute PCP
20: if  $PCP > 90\%$  then
21:   print Solution found:  $X_{pallet}$ 
22: else
23:   if  $PCP > 80\%$  then
24:     camera consensus
25:     if camera consensus=true then
26:       print Solution found:  $X_{pallet}$ 
27:     else
28:       print  $X_{pallet}$  rejected
29:     end if
30:   else
31:     print  $X_{pallet}$  rejected
32:   end if
33: end if
34: goto 1:

```

6. Dynamic planning for pallet picking

The outcome of the pallet location module is pallet position, and orientation in the AGV reference frame. The maneuver for accurately picking the pallet is formulated as an optimal control problem which calculates both the trajectory to move the AGV from its initial position to the pallet location and the steer angle pattern necessary to produce the optimal trajectory. The optimal control problem is defined in a reference system, which is aligned with the actual position of the AGV and which has its origin located in the center of the rear wheel, with the x -axis pointing forward and y -axis pointing leftward, as in Fig.12. The optimal control problem is formulated as the minimization of a weighted sum of two cost functions defining the shortest and fastest maneuver as follows:

$$J_c(\cdot) = w_T T(\zeta_f) + w_S s(\zeta) + \int_0^1 \|v_{\delta_o}(\zeta)\|^2 d\zeta \quad (2)$$

in which time is parameterized with variable $\zeta \in [0, 1]$ as follows: $t(\zeta) = \zeta T(\zeta)$. Variable $T(\zeta)$, which is constant, corresponds to the minimum time, and the variable $s(\zeta)$, defined by Eq. 4, is the distance travelled by the rear wheel. The cost function (Eq. 2) basically describes a multi-objective problem in which the weighting factors w_T and w_S respectively give more emphasis to the minimum time or the minimum travelled distance. The last term implements the requirements to minimize the energy to steer the vehicle.

$$\frac{1}{T(\zeta)} \frac{d}{d\zeta} s(\zeta) = \sqrt{u(\zeta)^2 + v(\zeta)^2} \quad (3)$$

$$\frac{d}{d\zeta} T(\zeta) = 0 \quad (4)$$

$$\frac{1}{T(\zeta)} \frac{d}{d\zeta} \mathbf{z}(\zeta) = \mathbf{f}(\mathbf{z}(\zeta), \delta(\zeta)) \quad (5)$$

The cost function in Eq. 2 must also comply with the dynamic equations which represent the AGV dynamic behavior and described by Eq. 5. Vector $\mathbf{z}(\zeta)$ includes the state variables such as the cartesian coordinates which define the vehicle position and attitude $x(\zeta), y(\zeta), \psi(\zeta)$, the longitudinal and lateral velocities $u(\zeta)$ and $v(\zeta)$ and the yaw rate $\psi_{dot}(\zeta)$, and lateral forces $FY_r(\zeta)$, $FY_{fr}(\zeta)$ and $FY_{fl}(\zeta)$ of the tires. The input is the steer angle of the rear wheel $\delta(\zeta)$, the dynamic characteristics of which are described by the following differential equations included in Eq. 5:

$$\frac{1}{T(\zeta)} \frac{d}{d\zeta} \delta(\zeta) = v_\delta(\zeta) \quad (6)$$

$$\frac{\tau_\delta}{T(\zeta)} \frac{d}{d\zeta} v_\delta(\zeta) + v_\delta(\zeta) = v_{\delta_o}(\zeta) \quad (7)$$

where $v_{\delta_o}(\zeta)$ is the optimal control and *i.e.* the steer angle rate. Equation 7 briefly describes the rear wheel dynamics where the time constant τ_δ is experimentally estimated. In other words it represents the effect of wheel inertia and steering electric motor characteristics. The tire forces are saturated according to the Magic Formula of

Pacejka [19]. In addition, the minimization problem is subject to boundary conditions to impose actual AGV initial states and final target position and orientation of the AGV (which corresponds to pallet position).

$$B(\mathbf{z}(\zeta_i), \mathbf{z}(\zeta_f)) = 0 \quad (8)$$

Lastly, minimization must comply with the following constraint:

$$|\delta(\zeta)| \leq \delta_{MAX} \quad (9)$$

The above described optimal control problem is solved with an indirect approach described in [20]. The Two–Point Boundary Value Problem is then discretized and the resulting non–linear system of equations is solved as in [20]. The solver is real–time capable and robust and allows the trajectory and the control to be recomputed as the AGV is moving toward the pallet. The computed steer angle history is used as set points for the electric drive’s low level controller, until a new set is available from the optimizer. This scheme corresponds to a NonLinear Receding Horizon scheme that is able to cope with the inaccuracies of the estimation of the AGV state and dynamic model, which can cope with the inaccuracies of the estimation of the AGV state and dynamic model. As pallet uncertainty is expected to decrease with pallet distance from the vehicle, dynamic planning allows for improvement of trajectory error compensation and hence for precise picking. Figure 13 shows two maneuvers related to two different initial orientations of the AGV. Those maneuvers are continuously updated on–line as the vehicle approaches the pallet.

7. Experimental results

To validate RLPF pallet pose estimation 150, varioius relative positions and angles between pallet and LRF camera in varying scenarios, industrial or laboratory environment, with cluttered background and one or more pallets, were tested. The pallets were placed at a distance of about 4 meters from the camera–LRF system, with an angle varying from -15° to $+15^\circ$, the 0° angle occurring when the line segment from LRF source to pallet center was orthogonal to the pallet front plane.

To verify the capability of the camera to improve identification robustness, in some scenarios we introduced objects with the same geometry and layout as pallet base blocks – a situation in which the initialization phase based on 2D LRF data may fail (see example in Fig. 6). LRF data representing the cluster of disturbing objects are detected as a possible pallet but with a lower PCP threshold. Therefore, camera–LRF image analysis is triggered but does not provide a positive consensus because of the mismatch of edges in relation to the template, preventing the algorithm from producing a false positive. Note that, in this application, a very low failure rate is tolerated, but it is relative to missing detections, whereas the detection of false positivesshould be as low as possible, in order to avoid damage and stopping the transport system. Also evaluated was the presence of casual objects near pallet blocks, in order to simulate obstacles or objects falling from the pallet load (Fig. 14). In these conditions, the PCP

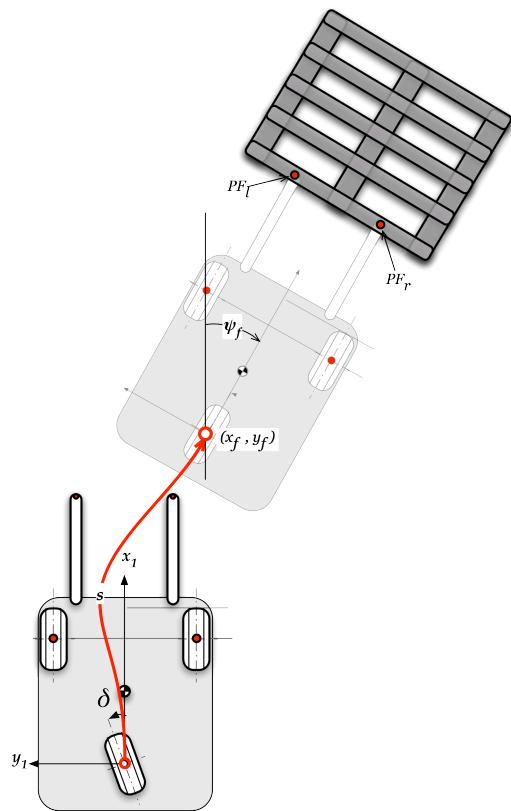


Figure 12. Initial position of AGV and target position for pallet picking defined with respect to initial position reference frame.

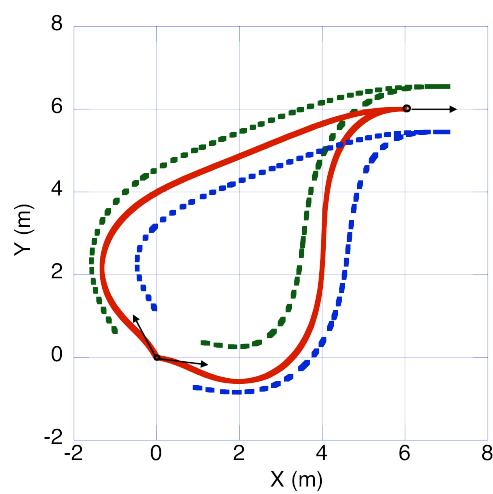


Figure 13. Optimal paths for two initial different orientations of AGV. Continuous lines: rear wheel path; dashed lines: paths of front wheels

is under the threshold only if the obstacle is very close to the blocks, because the object occludes the rear blocks of the pallet. However, this is preventive miss detection which, indirectly, does indicate the presence of undesired objects in the loading area. One



Figure 14. A possible scenario with an object which has fallen near a pallet. LRF may miss pallet detection in such cases, although the forklift is prevented from docking.

problem experienced with the camera consensus phase is environmental illumination. We tested various lighting conditions (Fig. 15), observing 5 cases of missed identification of pallet caused by poor or saturating light. This sometimes introduces unsatisfying edge detection on the image, reducing the features to be matched with the template and thus causing identification to fail. All the 145 remaining cases were solved correctly.



Figure 15. A sample of various test scenarios in industrial (upper row) and laboratory (lower row) environments with several pallet distances and angles, type of lighting, and one or two pallets in the camera field of view.

7.1. Performance analysis

The uncertainty of pallet pose calculated by the RLPF algorithm was estimated by considering the main influence which pallet parameters could have on pose estimation: distance from LRF source; face angle; supporting blocks width; color. Because the ground truth position and angle of the pallet with respect to the LRF are difficult to measure precisely with a reference measuring system, we computed RLPF uncertainty by point-to-point motion estimation, keeping the pallet fixed and moving the LRF camera. The displacement and rotation calculated from the two relative pallet poses estimated by the RLPF were then compared with reference ones. The LRF camera was mounted on a SCARA robot and moved along a line in 4 steps of $141.4 \pm 0.1 \text{ mm}$ and $\pm 0.02^\circ$ motion

linearity without rotating the LRF camera. The pallet pose computed from the initial position was taken as the reference from which to calculate the following four steps of displacements and rotations, and the errors were computed. Their covariance, together with mean error, was assumed as RLPF uncertainty. Table 1 lists results in term of standard deviations s , average e_{mean} and maximum errors e_{max} for two pallet dimensions (standard maximum tolerance EPAL and normal in-tolerance EPAL) and two colors for normal size (natural wood painted blue). For each of the three case combinations, 3 different distances and angles were tested: [2, 3, 4] m and [0, 7.5, 15] deg, for a total of 45 RLPF outputs each (5 LRF camera positions per pair of pallet distances and angles). s , e_{mean} and e_{max} are relative to each set of 45 measurements. No correlation between pallet dimension or color and RLPF measurement errors was detected. The maximum error over 135 pose measurements was 10.8 mm for displacement and 0.78 deg for pallet attitude.

	$X[\text{mm}]$			$Y[\text{mm}]$			$Z[\text{mm}]$		
	s_x	e_{mean}	e_{max}	s_y	e_{mean}	e_{max}	s_θ	e_{mean}	e_{max}
Standard max format	3.9	-0.7	-10.4	2.2	1.2	6.7	0.3	0.006	-0.78
Nomal wood	3.0	-2.0	-9.2	2.9	0.8	5.9	0.28	0.10	0.76
Normal blue	2.1	-0.9	4.9	3.0	1.4	7.4	0.27	0.12	0.67

Table 1. RLPF experimental uncertainty

8. Conclusions

This work presents a multi-step, multi-sensory strategy called RLPF, which uses an LRF for uncertain pallet position estimation and camera LRF image processing for consensus in cases of ambiguous or weak detection. Extensive experimental tests show that the algorithm can be implemented in industrial autonomous vehicles as a stand-alone module, for use in extending logistics automation toward unstructured environments. Critical cases with disturbing nearby objects and difficult lighting were analysed. Tests provide a miss detection rate of about 1% and zero wrong detections in about 300 trials, and the estimated uncertainty is compliant with safe and functional AGV picking task requirements for standard pallet formats. Until now, no-light working has not been implemented, but can be solved with artificial lighting mounted on vehicles or infrared vision. For the docking task, a new adaptive non-linear control which takes vehicle dynamics into account is proposed.

9. Acknowledgments

The authors would like to thank Fusionsystems GmbH and Digipack s.r.l. as partners in the AGILE European project, within which this work was carried out.

10. Bibliography

- [1] Roger Bostelman, Hong Tsai, and Tommy Chang. Visualization of pallets. In *Intelligent Robots and Computer Vision XXIV: Algorithms, Techniques, and Active Vision*, volume 6384, October 2006.
- [2] Luca Baglivo, Nicolas Bellomo, Enrico Marcuzzi, Marco Pertile, Enrico Bertolazzi, and Mariolino De Cecco. Pallet pose estimation with lidar and vision for autonomous forklifts. In *Information Control Problems in Manufacturing*, volume 13, pages 616–21, V.A. Trapeznikov Institute of Control Sciences, Russia, 2009.
- [3] Cédric Pradalier, Ashley Tews, and Jonathan Roberts. Vision-based operations of a large industrial vehicle: Autonomous hot metal carrier. *J. Field Robot.*, 25:243–267, April 2008.
- [4] Michael Seelinger and John-David Yoder. Automatic visual guidance of a forklift engaging a pallet. *Robotics and Autonomous Systems*, 54(12):1026 – 1038, 2006.
- [5] Rita Cucchiara, Massimo Piccardi, and Andrea Prati. Focus based feature extraction for pallets recognition. In *BMVC*, 2000.
- [6] Sungmin Byun and Minhwon Kim. Real-time positioning and orienting of pallets based on monocular vision. In *Proceedings of the 2008 20th IEEE International Conference on Tools with Artificial Intelligence - Volume 02*, pages 505–508, Washington, DC, USA, 2008. IEEE Computer Society.
- [7] Alonzo Kelly. Model-based object pose refinement for terrestrial and space autonomy. In *6th International Symposium on Artificial Intelligence, Robotics and Automation in Space (ISAIRAS 01)*, Montreal, Quebec, Canada, June 2001.
- [8] G. Garibotto, S. Masciangelo, M. Ilic, and P. Bassino. Service robotics in logistic automation: Robolift: vision based autonomous navigation of a conventional fork-lift for pallet handling. In *Advanced Robotics, 1997. ICAR '97. Proceedings., 8th International Conference on*, pages 781 –786, July 1997.
- [9] J. Pagès, X. Armangué, J. Salvi, J. Freixenet, and J. Martí. Computer vision system for autonomous forklift vehicles in industrial environments. In *The 9th. Mediterranean Conference on Control and Automation, MED'2001*, Dubrovnik, Croatia, June 2001.
- [10] J. Nygards, T. Hogstrom, and A. Wernersson. Docking to pallets with feedback from a sheet-of-light range camera. In *Intelligent Robots and Systems, 2000. (IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on*, volume 3, pages 1853 –1859 vol.3, 2000.
- [11] D. Lecking, O. Wulf, and B. Wagner. Variable pallet pick-up for automatic guided vehicles in industrial environments. In *Emerging Technologies and Factory Automation, 2006. ETFA '06. IEEE Conference on*, pages 1169 –1174, 2006.
- [12] M.R. Walter, S. Karaman, E. Frazzoli, and S. Teller. Closed-loop pallet engagement in unstructured environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, October 2010. Accepted, To Appear.
- [13] L. Baglivo, N. Bellomo, G. Miori, E. Marcuzzi, M. Pertile, and M. De Cecco. An object localization and reaching method for wheeled mobile robots using laser rangefinder. In *Intelligent Systems, 2008. IS '08. 4th International IEEE Conference*, volume 1, pages 5–6 –5–11, 2008.
- [14] Hyojoo Son, Changwan Kim, and Kwangnam Choi. Rapid 3d object detection and modeling using range data from 3d range imaging camera for heavy equipment operation. *Automation in Construction*, 19(7):898 – 906, 2010.
- [15] Qilong Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, volume 3, pages 2301–2306, 2004.
- [16] M. Wright. Direct search methods: Once scorned, now respectable, 1995.
- [17] Dariu Gavrila. Pedestrian detection from a moving vehicle. In *Proceedings of the 6th European Conference on Computer Vision-Part II*, ECCV '00, pages 37–49, London, UK, 2000. Springer-Verlag.

- [18] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching: two new techniques for image matching. In *Proceedings of the 5th international joint conference on Artificial intelligence - Volume 2*, pages 659–663, San Francisco, CA, USA, 1977. Morgan Kaufmann Publishers Inc.
- [19] Hans Pacejka. *Tyre and Vehicle Dynamics*. Elsevier, 2005.
- [20] E. Bertolazzi, F. Biral, and M. Da Lio. Real-time motion planning for multibody systems. *MULTIBODY SYSTEM DYNAMICS*, 17:119–139, 2007.