Localization of Cask and Plug Remote Handling System in ITER using multiple Video Cameras

João Ferreira[a,*], Alberto Vale[a], Isabel Ribeiro[b]

[a]*Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade Técnica de Lisboa,*
Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[b]*Laboratório de Robótica e Sistemas em Engenharia e Ciência, Instituto Superior Técnico, Universidade Técnica de Lisboa,*
Av. Rovisco Pais 1, 1049-001, Lisboa, Portugal

Abstract

The Cask and Plug Remote Handling System (CPRHS) provides the means for the remote transfer of in-vessel components and Remote Handling equipment between the Hot Cell Building and the Tokamak Building in ITER. Different CPRHS typologies will be autonomously guided following predefined trajectories. Therefore, the localization of any CPRHS in operation must be continuously known in real time to provide the feedback for the control system and also for the human supervision. This paper proposes a localization system that uses the video streaming captured by the multiple cameras already installed in the ITER scenario to estimate with precision the position and the orientation of any CPRHS. In addition, an augmented reality system can be implemented using the same video streaming and the libraries for the localization system. The proposed localization system was tested in a mock-up scenario with a scale 1:25 of the divertor level of Tokamak Building.

Keywords: ITER, Remote Handling, Cask and Plug Remote Handling System, Localization of mobile vehicles, Videos Cameras, Augmented Reality

1. Introduction

The ITER (International Thermonuclear Experimental Reactor) is a joint international research project, aiming to demonstrate the technological feasibility of fusion power as an alternative and safe power source. The Cask and Plug Remote Handling System (CPRHS) is a vehicle that provides the means for the remote transfer of (clean/activated/contaminated) in-vessel components and remote handling (RH) equipment between the Hot Cell Building (HCB) and the vacuum vessel in Tokamak Building (TB), as illustrated in Figure 1.

During the reactor's operation, the in-vessel components, such as the blankets that cover the vacuum vessel, are expected to become activated. When such components have to be removed for disposal, operations are to be carried out by the CPRHS, which is required to dock in pre-defined locations, the vacuum vessel port cells, located on the three levels of TB: divertor, equatorial and upper level. Then, the components are transported to the HCB for operations of diagnose and refurbishment or disposal of activated material. Hence, the CPRHS must dock at the docking stations through port plugs interfaces or park in parking areas at the different levels of the HCB. The CPRHS comprises three sub-systems: a cask envelope containing the load, a pallet that supports the cask envelope and the Cask Transfer System (CTS). The CTS acts as a mobile robot, provides the mobility for the CPRHS and can be decoupled from the entire system.

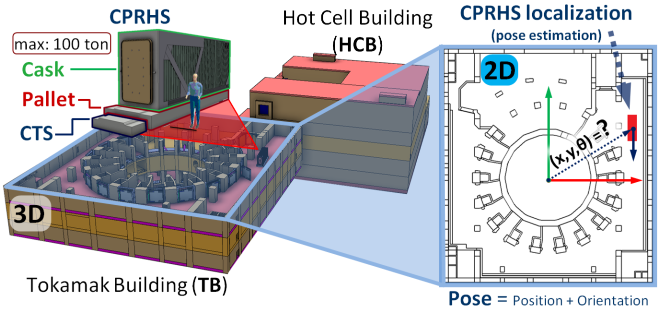

Figure 1: CPRHS and divertor level in Tokamak building of ITER.

The CPRHS is entrusted with the transportation of heavy (total weight up to 100 tons) and highly activated in-vessel components [2]. The kinematic configuration, first proposed in [1], endows it with the required flexibility to navigate autonomously or remotely controlled, in the cluttered environments of the TB and the HCB. To navigate in a cluttered environment (with less than 300 mm safety margins for CPRHS and 100 mm for CTS or for rescue casks, a particular configuration of CPRHS), it is required to know the correct pose (position and orientation) of the vehicle to follow the optimized trajectories evaluated in [3] and [4]. The localization system should estimate CPRHS pose with high level of accuracy and redundancy. Several different complementary systems could be integrated to achieve this requirements. Since the CPRHS transports activated in-vessel components, the use of on-board electronics should be kept to the minimum. The localization system components should be installed on the walls of the buildings, protected against contin-

---

[*] Corresponding author. Tel.:+351 218 419 098; fax:+351 218 417 819.
Email address: `jftferreira@ipfn.ist.utl.pt` (João Ferreira)

uous exposure. Hence, a localization system is required to estimate continuously and in real time the pose of the CPRHS/CTS when moving inside the buildings, using the sensors installed on the walls.

A previous work developed by the authors, [5], proposed a localization system using an optimized laser range finders network to estimate the pose of the CPRHS/CTS in real time with a high precision. However, the proposed system requires the installation of an additional type of sensors in the walls of the buildings. Furthermore, a second localization system is important for redundancy.

In ITER it is expected a set of video cameras inside the main buildings for supervision purposes. The video streaming captured by these cameras could also be used for the localization of CPRHS/CTS during operation.

This paper addresses a localization system using video cameras already installed on the scenario. The only additional intervention is a set of markers glued or painted on the CPRHS/CTS sides. The localization system algorithm runs in an external computer with access to the video streaming and it estimates the pose (position and orientation) of all CPRHS/CTS.

The proposed localization system is based on marker pose estimation using computer vision. This technique is a well known problem with several applications like Augmented Reality (AR) or object pose estimation. There are several software libraries such as ARtoolkit, ARtoolkitPlus, ARtag [6], which provide a ready platform for marker tracking. Wagner *et al.*, [7], developed a series of tests using ARtoolkitplus, in mobile devices to add AR to real time video. There are some experiments with mobile robot localization using cameras on the environment and markers on the vehicle, [8].

Combining the video streaming captured by the video cameras and the pose estimated by localization system is still possible to implement AR, displaying additional information on the images, e.g., the most critical points of collision or plotting the optimized trajectory to be followed.

This paper is organized as follows: Section 2 introduces the proposed localization system: the requirements, the implementation and the features. The Section 3 presents the results, followed by the conclusions and future work in Section 4.

2. The localization system

The proposed localization system aims to estimate the CPRHS or CTS (from this point forward identified simply by vehicle) pose, $X_r = [x_r \; y_r \; \theta_r]$. The proposed localization system requires access to the video streaming acquired by the cameras installed inside the ITER buildings and the installation of markers on the vehicle.

The working principle of the system and his main components are shown in Figure 2. First, a tracking algorithm identifies the markers in the video frames and obtains the marker pose in the camera reference frame. Knowing the exact position of the camera, it is possible to estimate the marker pose in the global frame and, since the marker is rigidly connected to the vehicle, know the vehicle pose as well.

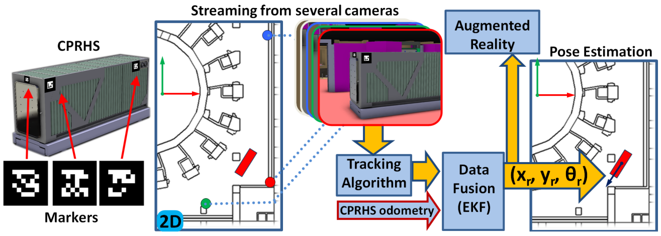

Figure 2: The main components of the proposed localization system.

The vehicle can only be observed sideways as it is nearly the same height of a building's level. To avoid occlusions several cameras and markers are necessary. To enhance the system robustness and precision, Extended Kalman Filter (EKF) can be used to fuse the poses of the several markers, estimated from the streaming of the cameras and the vehicle odometry. Knowing the poses of the vehicle and of the cameras it is possible to project AR contents in the video as an aid for remote operators.

2.1. Conditions assessment

The performance of localization system is dependent on the cameras specifications, markers and lightning conditions.

The position, Field of View (FoV) and resolution of each camera have direct impact in the system performance, as described in [9]. It is assumed that the cameras are already installed for multiple purposes inside ITER, therefore, their assessment boils down to the selection of the best cameras to include in the system. The selected cameras must maximize the area covered by the network and minimize distance between cameras and areas of interest.

The size of the markers, their contrast with the surrounding and a large Hamming distance between their patterns is crucial for a correct detection and identification, as explained in [6].

For a standard VGA camera, to detect a marker $10m$ away, the marker size must be approximately $40cm$. Given the large dimensions ($8.5 \times 2.65m$) of the vehicle, to avoid occlusions and to assure that a marker is always detected, it is advisable to install more than one marker on every side of the vehicle. Light conditions should be bright and homogeneous without specular reflections for the correct tracking of the markers.

The tracking algorithm is a critical component in the system, some techniques are detailed in [6] and [9] where edge based algorithms show clear advantages over the binary morphology based algorithms, including better detection under difficult lighting conditions.

2.2. Implementation

Let us define a transformation matrix Γ_B^A ,(1), between frames A and B, where R_B^A is a rotation matrix and t_B^A is a translation vector. T is a function that retrieves the translation and the Euler angles associated with the transformation, (e.g. $T(\Gamma_B^A) = T_B^A = [x \; y \; z \; \psi \; \varphi \; \theta]^T$ is associated with Γ_B^A).

$$p_A = \Gamma_B^A * p_B = \begin{bmatrix} R_B^A & t_B^A \\ 0_{1\times3} & 1 \end{bmatrix} * [p_x \; p_y \; p_z \; 1]_B^T \tag{1}$$

Figure 3 shows the several frames associated with the proposed localization system and the transformations between them: *1)*

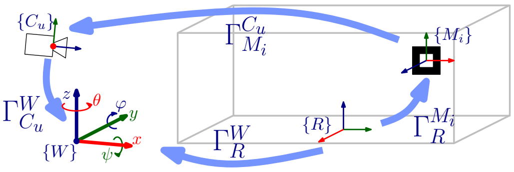

Figure 3: Localization system transformation framework.

$\Gamma^{C_u}_{M_i}$ – between marker i and camera u, it is obtained by the tracking algorithm; *2)* $\Gamma^{M_i}_{R}$ – between vehicle and the marker i, it is constant since the markers are rigidly connected to the vehicle; *3)* $\Gamma^{W}_{C_u}$ – between camera u and World, assumed to be constant, the cameras are installed on the building. It is obtained by calibration. *4)* Γ^{W}_{R} – between vehicle and World, the correct estimation of this transformation is the goal of localization system. $\Gamma^{W_u}_{R_i}$ is estimated based on the other three transformations, (2). The vehicle pose obtained through the *u-th* camera and the *i-th* marker, is given by $X^{W_u}_{R_i}$, (3).

$$\Gamma^{W_u}_{R_i} = \Gamma^{W}_{C_u} \times \Gamma^{C_u}_{M_i} \times \Gamma^{M_i}_{R} \tag{2}$$

$$D = \begin{bmatrix} 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&1&0 \end{bmatrix} , \quad X^{W_u}_{R_i} = [x_r \; y_r \; \theta_r]^T_{(u,i)} = D \times T^{W_u}_{R_i} \tag{3}$$

2.3. Localization performance

The estimation of vehicle pose is affected by process and measurements noise, the process noise accounts for wheel slippage or erroneous modeling of the kinematics. The measurement noise accounts for errors in the estimation of $\Gamma^{C_u}_{M_i}$ by the tracking algorithm, admitting that camera calibration does not add error to the estimation. The EKF fuses the information from vehicle odometry and from several cameras and markers and minimizes the error between real vehicle pose (X_r), and the estimated pose ($\hat{X}_r$). The EKF works in two major steps per iteration, prediction step and update step. At iteration k, the prediction step (4) uses the previous estimation ($\hat{X}_{r(k-1)}$), the odometry ($u_{(k-1)}$) and the known rhombic kinematic model (f), detailed in [1], to estimate the vehicle pose and an estimation covariance matrix. The update step (5) takes the measurements from the tracking algorithm (z_k) and updates the estimations of X_r and P_r, knowing the measurement model ($h(X_r)$), that relates the measurements with the vehicle pose.

$$\begin{aligned} \tilde{X}_{r(k)} &= f(\hat{X}_{r(k-1)}, u_{(k-1)}) \\ \tilde{P}_{r(k)} &= F_{(k-1)} \hat{P}_{r(k-1)} F^T_{(k-1)} + Q \end{aligned} \tag{4}$$

$$\begin{aligned} K_k &= \tilde{P}_{r(k)} H^T (H \tilde{P}_{r(k)} H^T + R_k)^{-1} \\ \hat{X}_{r(k)} &= \tilde{X}_{r(k)} + K_k (z_k - h(\tilde{X}_{r(k)})) \\ P_{r(k)} &= [I - K_k H] \tilde{P}_{r(k)} \end{aligned} \tag{5}$$

$\tilde{X}_r$ is the estimation after prediction step, $\tilde{P}$ is the covariance matrix of estimation error. F and H are the Jacobian matrices of functions f and h, that represent the kinematic model and the measurement model, respectively. Q and R are the covariance matrices of the process and measurement noises, respectively.

The measurements in this system (z_k) consist in one vector, $X^{W_u}_{R_i}$ (3), for each combination of camera with visible marker.

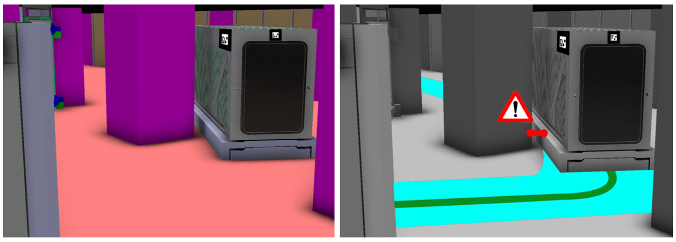

Figure 4: Original video frame captured by a camera installed in the scenario (left) and the same video frame with AR to display the occupied area and the warning of a possible clash (right).

The measurement model, h, is a simple identity function, since the system's measurements are the state itself. The matrix H (6) is the Jacobian of h, t is an identity matrix with dimension Dim, that depends on the number of cameras L and on the number of markers visible from each camera $m(u)$. The $\Sigma_{(u,i)}$ is the covariance matrix associated with the measurement noise expected to the *i-th* marker identified in the *u-th* camera. It can be a fixed value or change depending on each marker visible area and confidence level (values that can be obtained by the tracking system). The R matrix combines the covariance matrix for all measurements, that are assumed to be independent.

$$\begin{aligned} & H = I_{Dim} \quad , \quad Dim = \textstyle\sum_{u=1}^{L} m(u) \\ & \Sigma_{(u,i)} = diag(\sigma^2_x(i,j),\, \sigma^2_y(i,j),\, \sigma^2_\varphi(i,j)) \\ & R = diag(\Sigma_{11}, \ldots, \Sigma_{L\,m_L}) \end{aligned} \tag{6}$$

The integration of several cameras, with different perspectives over the vehicle, and the installation of several markers on all sides of the vehicle minimize the pose estimation error.

The localization system, based on cameras, requires calibration due to camera intrinsic parameters and installation errors. Camera calibration can be accomplished by adding some calibrating markers on known positions in the scenario. To avoid extra markers on the buildings, Rekleitis *et al.*, in [8], proposed a system that localizes a vehicle and self calibrates using the markers on the vehicle.

Although it is not a very precise localization system, it is a great option to provide redundancy and robustness to an overall localization system that integrates several complementary technologies (e.g. laser range finders). This system has the advantages of recognizing the markers, which could be helpful to identify vehicles or even operate with several vehicles simultaneously. The output of the system can be used to introduce AR, which is an asset for remote operation.

The main feature of AR is displaying virtual aids in a video frame to help the remote handling operation. Knowing the geometric transformation between the scenario and cameras referential, it is possible to add virtual elements such as warnings of clashes, predicted trajectory for the vehicle and occupied volumes, as depicted in Figure 4.

3. Experimental Results

The results were obtained with a mock-up of the CPRHS in a scenario similar to the divertor level in TB with a scale of

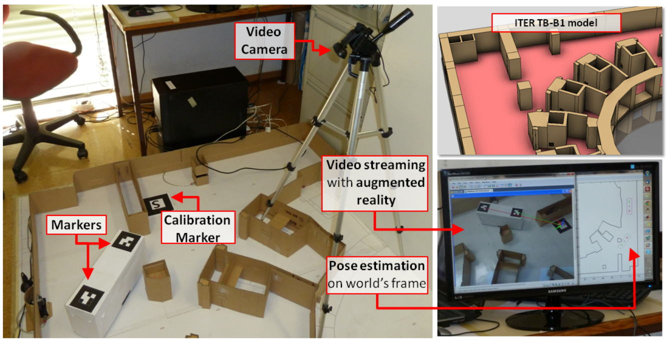

Figure 5: Experimental setup (left), divertor level (top right) and a screenshot of the real-time pose estimation with AR (bottom right).

1:25, as shown in Figure 5. The software library used to track the markers was the ARTK, based on binary morphology algorithms with a fixed threshold. The CPRHS is autonomously guided. It was used a single camera (C_1), with 800×600 resolution, and two markers installed on the top of the prototype, a front and a rear marker. The camera was calibrated offline, for its intrinsic parameters, using the calibration tool included in the ARTK. The Calibration Marker (CM), shown in Figure 5, is used to estimate the extrinsic parameters, allowing the camera to be moved between different tests.

The matrix $\Gamma_{C_1}^{W}$ (7), depends exclusively on $\Gamma_{CM}^{C_1}$, the transformation between the CM and the camera, obtained by ARTK, and on Γ_{CM}^{W}, that, in this case, is a simple offset from the world's origin (7). I_3 is the identity matrix with dimension 3 and t_{CM}^{W} is the vector between the center of the CM and the world's origin. The transformations between markers and camera are $\Gamma_{M_F}^{C_1}$ and $\Gamma_{M_R}^{C_1}$ for front and rear markers, respectively. These are obtained on-line by the ARTK at a maximum rate of 20 Hz. The transformations between the vehicle's center and each marker, $\Gamma_{R}^{M_F}$ and $\Gamma_{R}^{M_R}$, are simple translations with vectors $t_R^{M_F}$ and $t_R^{M_R}$ (8).

$$\Gamma_{CM}^{W} = \begin{bmatrix} I_3 & t_{CM}^{W} \\ 0_{1\times3} & 1 \end{bmatrix}, \quad \Gamma_{C_1}^{W} = \Gamma_{CM}^{W} \times \left(\Gamma_{CM}^{C_1}\right)^{-1} \tag{7}$$

$$t_R^{M_F} = [\,0\ -Yoff\ -Zoff\,]^T \qquad t_R^{M_R} = [\,0\ +Yoff\ -Zoff\,]^T \tag{8}$$

Being $Yoff$ and $Zoff$ the offset between vehicle's center and markers. Up to the date the implementation with EKF is not complete, but, since the measurements are the state itself, the average of the two measurements is enough to obtain a good approach of the vehicle pose $\hat{X}_r = \frac{1}{2}X_{R_F}^{W} + \frac{1}{2}X_{R_R}^{W}$. Several poses, shown in Figure 6, were measured manually (*real poses*) and compared with the *estimated poses*. The length of the localization error is always below $4cm$ and the average error along the trajectory is $1.5cm$ which are extremely good results, considering the method used. The orientation error is considerable and the system robustness is affected by lightning conditions. However, with multiple cameras, EKF and odometry, the system precision and robustness is expected to be improved, which is endorsed as a future development.

4. Conclusions and Future Work

This paper presented a localization system to estimate in real time the pose of the CPRHS or CTS operating in ITER using the video streaming captured by cameras already installed in the scenario. It is only required a set of markers glued or painted in the sides of the CPRHS or CTS to identify each vehicle. The localization system provides also the ability to include augmented reality in the video streaming, as displaying in the images the critical parts with risk of clash or plot on the floor the trajectory to be followed. The proposed localization system works as an alternative or even to integrate with other localization systems.

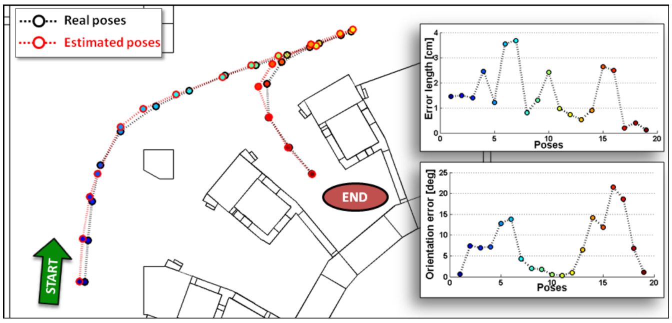

Figure 6: Pose estimation along a path in the vicinity of the entrance to a port.

The system was tested in a scale of 1:25 with the CPRHS following a pre-defined trajectory. The pose estimation along the path provided an uncertainty of less than $1.9cm$, which demonstrates the potential of this localization system.

There are open issues under development, namely the implementation of EKF with more cameras and markers, the inclusion of an edge based algorithms, the self calibration and a more detailed condition assessment for the ITER scenario.

Acknowledgments

The work was supported by FCT in the frame of the Contract of Associate Laboratories of Instituto de Plasmas e Fusão Nuclear/IST (PEst-OE/SADG/LA0010/2011) and Laboratório de Robótica e Sistemas em Engenharia e Ciências/IST (PEst-OE/EEI/LA0009/2011).

References

[1] I. Ribeiro, P. Lima, R. Aparício, R. Ferreira,"Conceptual Study on Flexible Guidance and Navigation for ITER Remote Handling Transport Casks", Proc. of the 17th IEEE/NPSS Symp. on Fusion Engineering, pp. 969-972, San Diego, USA, 1997.

[2] C. Gutiérrez, C. Damiani, M. Irving, J. Friconneau, A. Tesini, I. Ribeiro, A. Vale, "ITER Transfer Cask System: status of design, issues and future developments", Proc. of the 9th Int. Symp. on Fus. Nuclear Energy, 2009.

[3] D. Fonte, F. Valente, A. Vale, I. Ribeiro, "A Motion Planning Methodology for Rhombic-like Vehicles for ITER Remote Handling Operations", Proc. of the 7th IFAC Symp. on Intellig. Autonomous Vehicles, Italy, 2010.

[4] D. Fonte, F. Valente, A. Vale, I. Ribeiro, "Path Optimization of Rhombic-Like Vehicles: An Approach Based on Rigid Body Dynamic", Proc. of the 15th IEEE Int. Conf. on Advanced Robotics, 106-111, Estonia, 2011.

[5] J. Ferreira, A. Vale, R. Ventura, "Optimizing Range Finder Sensor Network Coverage in Indoor Environment" Proc. of the 7th IFAC Symp. on Intelligent Autonomous Vehicles, Italy, 2010.

[6] M. Fiala, "Artag, a fiducial marker system using digital techniques", Proc. of IEEE Comp. Society Conf., vol. 2, pp. 590-596, San Diego, USA, 2005.

[7] Wagner, D. and Schmalstieg, D., "Artoolkitplus for pose tracking on mobile devices", Proc. of 12th CVWW, pp. 139-146, Austria, 2007.

[8] I. Rekleitis, D. Meger and G. Dudek, "Simultaneous planning, localization, and mapping in a camera sensor network", Proc. of IEEE Conf. Robotics and Automation, vol. 54, pp. 921-932, Atlanta, USA, 1993.

[9] De Kler, T., "Integration of the ARToolKitPlus optical tracker into the Personal Space Station", Sec. Comp. Science, Univ. of Amsterdam, Netherlands, 2007.