Video-guided Camera Control for Target Tracking and Following

Jake Gemerek, Silvia Ferrari, Brian H. Wang, and Mark E. Campbell
Laboratory for Intelligent Systems and Controls (LISC), Cornell University

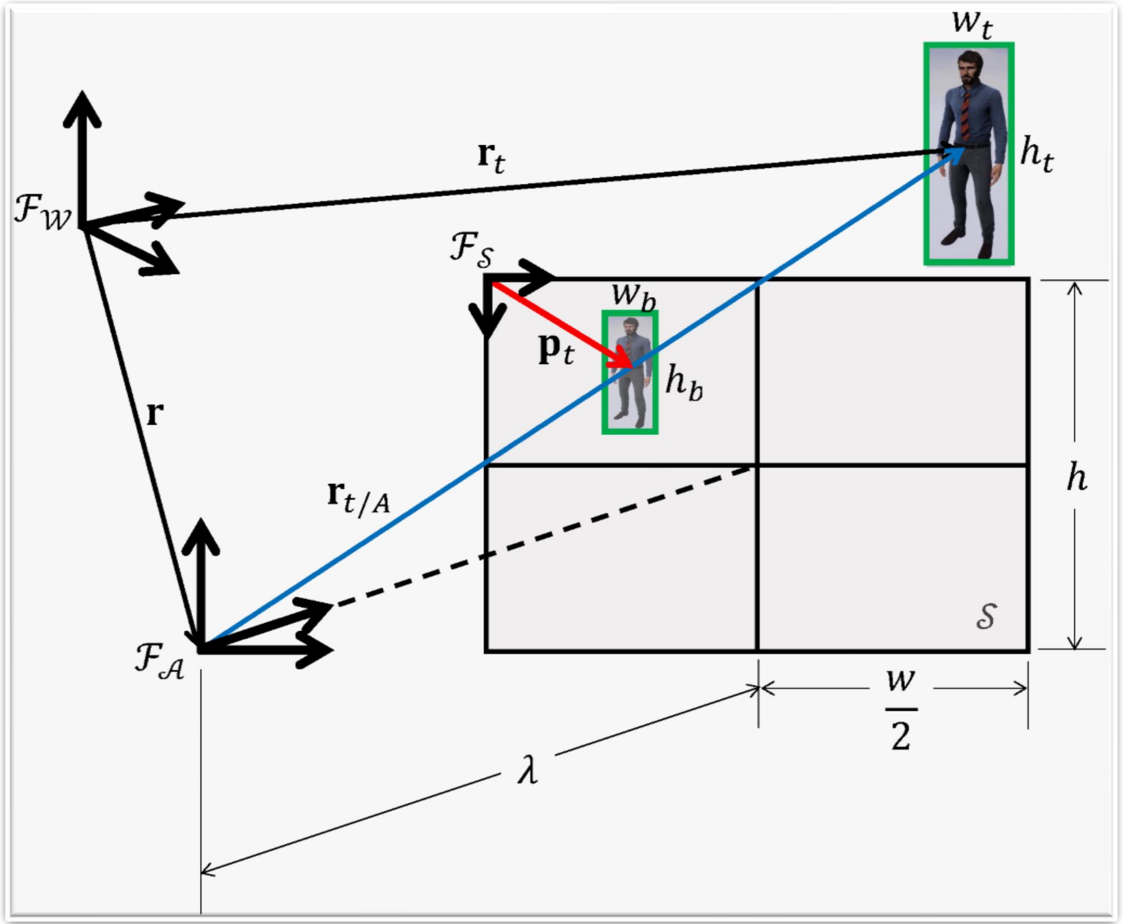
Abstract

This paper considers the problem of controlling a nonholonomic mobile ground robot equipped with an onboard camera characterized by a bounded field-of-view, tasked with detecting and following a potentially moving human target using onboard computing and video processing in real time. Computer vision algorithms have been recently shown highly effective at object detection and classification in images obtained by vision sensors. Existing methods typically assume a stationary camera and/or use pre-recorded image sequences that do not provide a causal relationship with future images. The control method developed in this paper seeks to improve the performance of the computer vision algorithms, by planning the robot/camera trajectory relative to the moving target based on the desired size and position of the target in the image plane, without the need to estimate the target's range. The method is tested and validated using a highly realistic and interactive game programming environment, known as Unreal Engine, that allows for closed-loop simulations of the robot-camera system. Results are further validated through physical experiments using a Clearpath Jackal robot equipped with a camera which is capable of following a human target for long time periods. Both simulation and experimental results show that the proposed vision-based controller is capable of stabilizing the target object size and position in the image plane for extended periods of time.

Problem Formulation

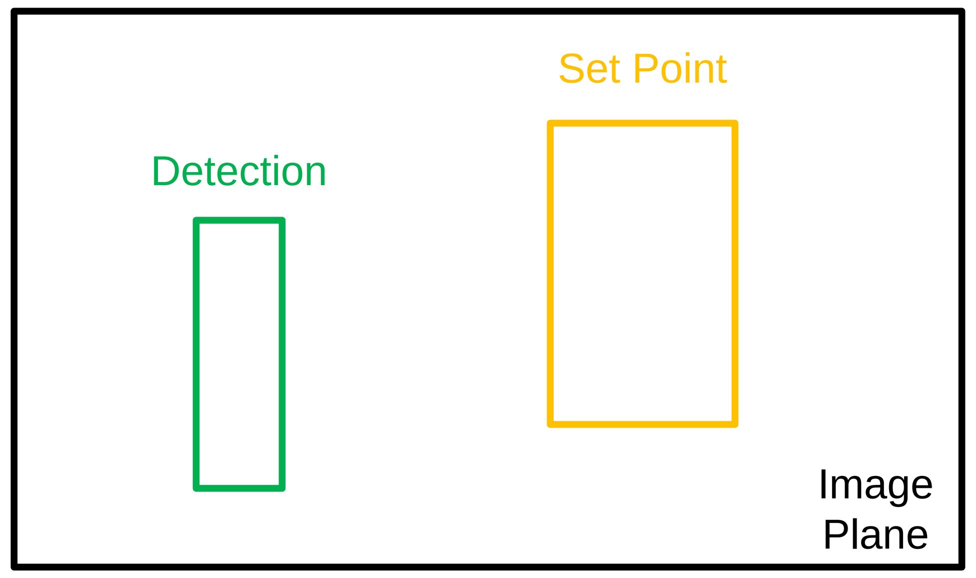
- Workspace, $\mathcal{W} \subset \mathbb{R}^3$, is populated with target human, $\mathcal{T} \subset \mathcal{W}$, and mobile ground robot, $\mathcal{A} \subset \mathcal{W}$.
  - Robot has onboard camera with FOV $\mathcal{S} \subset \mathcal{W}$.
  - Robot configuration vector: $\mathbf{q}(t) = [x(t), y(t), \theta(t)]^T$
- Robot obeys nonholonomic unicycle kinematics,

$$\dot{\mathbf{q}}(t) = \begin{bmatrix} \cos\theta(t) & 0 \\ \sin\theta(t) & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} = \mathbf{G}(\mathbf{q}(t))\mathbf{u}(t)$$



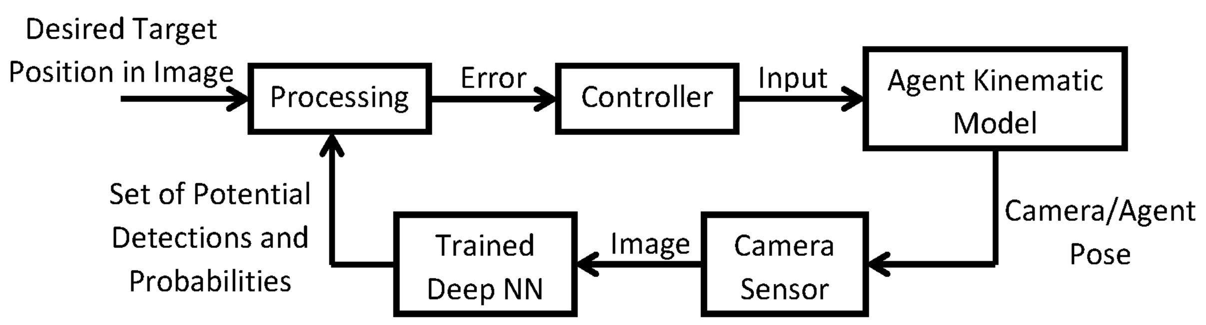
Approach

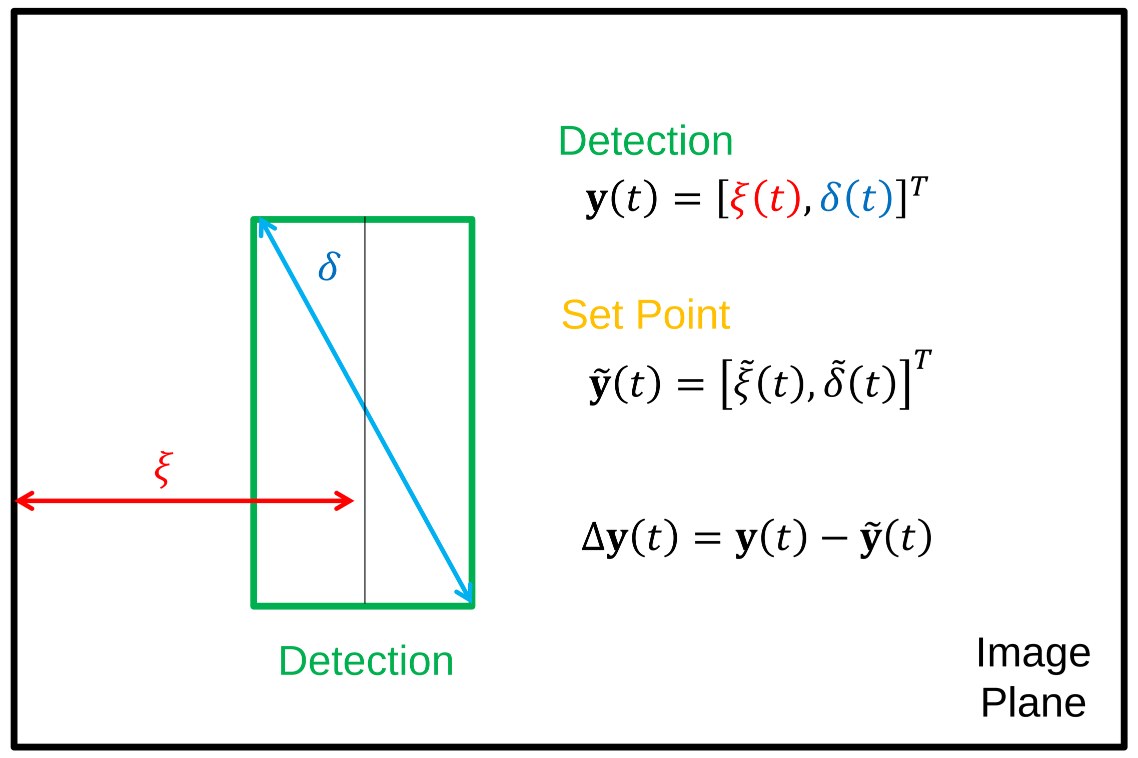
- **Control Objective**: Position the target relative to the camera FOV such that the target bounding box is driven to a desired bounding box in the image plane.



- **Perspective Projection**: Objects close to the focal point appear larger than objects far from the focal point.
  - Use the above property to control bounding box size.
  - Does not require 3D estimation.
- Control horizontal bounding box position with yaw rate input.
- Vertical bounding box position is not controllable.
- Proportional-Integral control law:

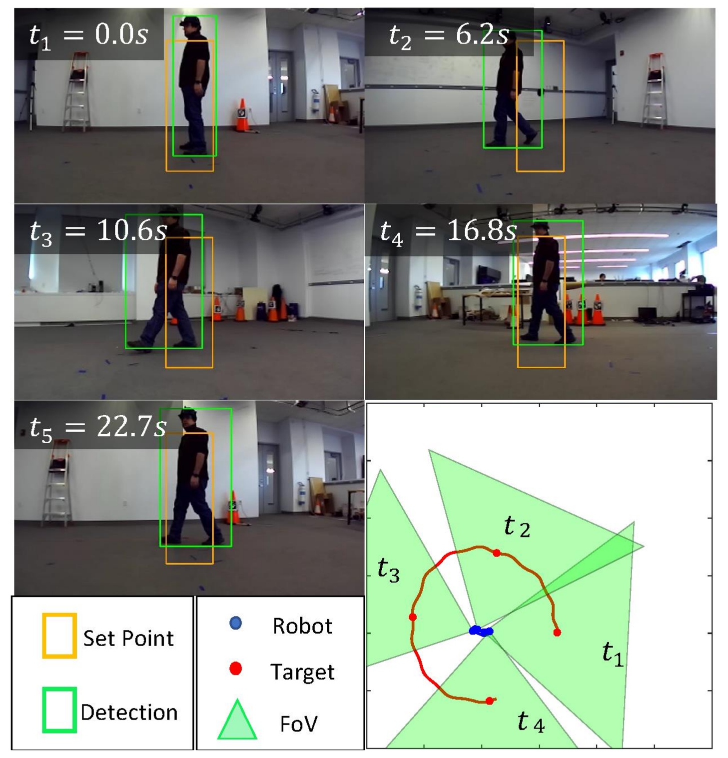
$$\mathbf{u}(t) = -\mathbf{K}_1 \Delta\mathbf{y}(t) - \mathbf{K}_2 \int_0^t \Delta\mathbf{y}(\tau) d\tau$$

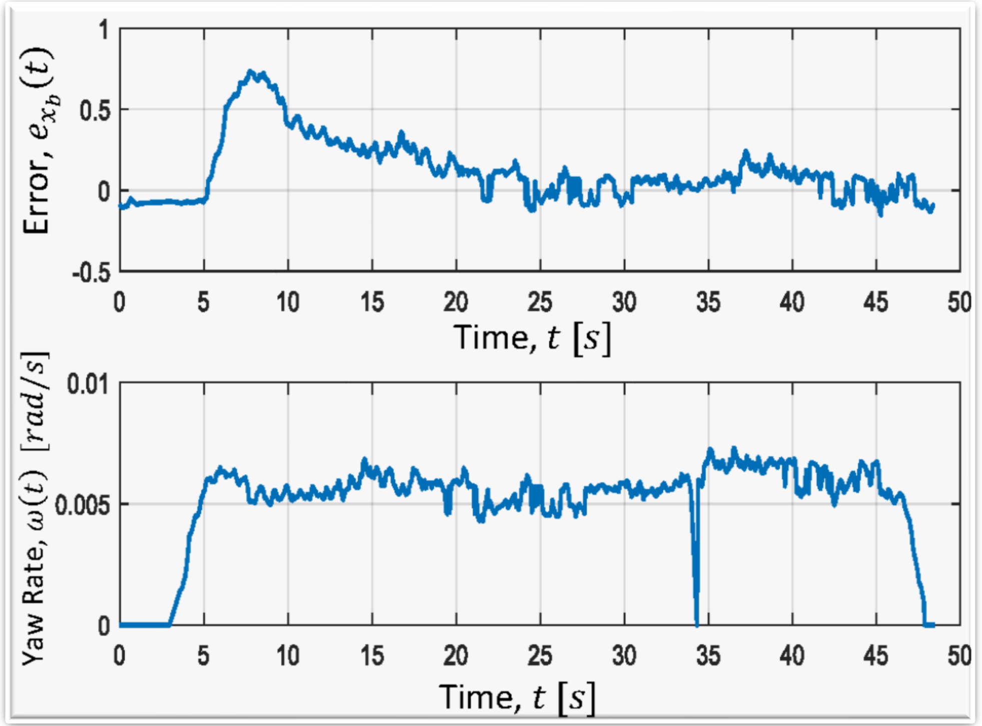




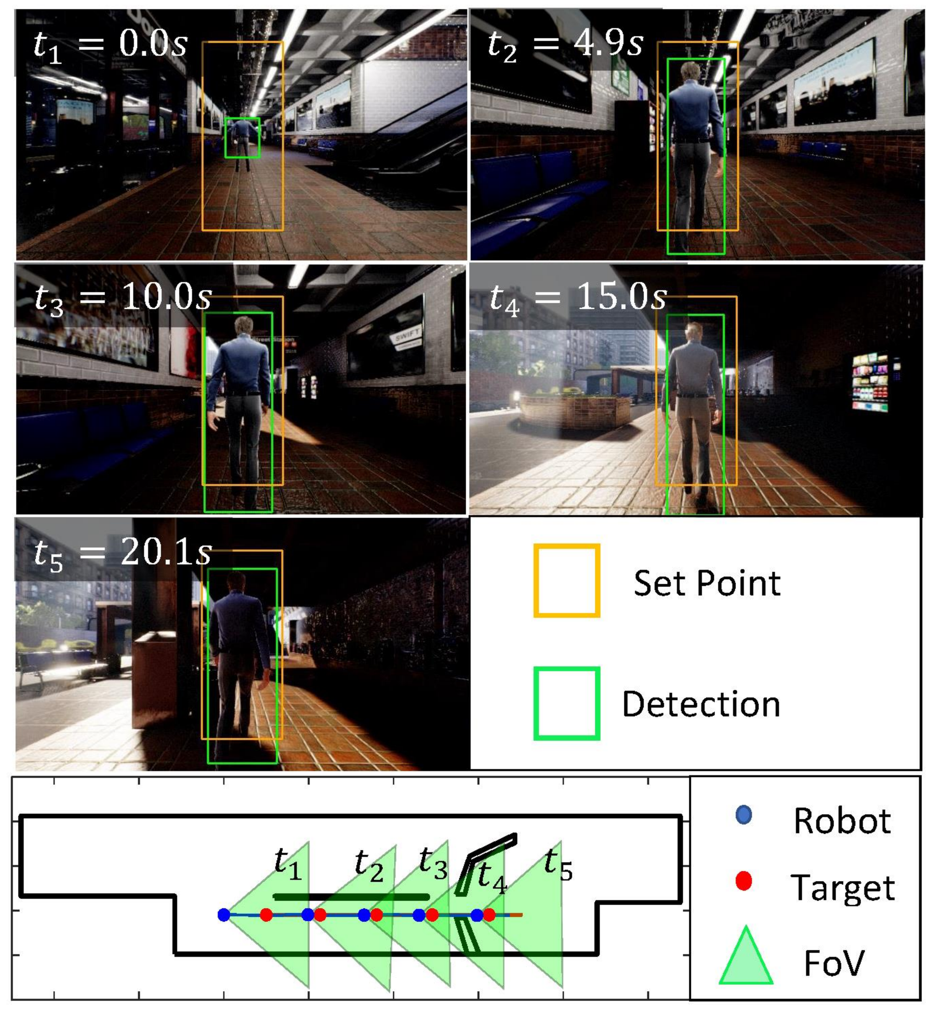
Experiments

- Step and ramp response in relative target range and heading tested in both experiments and simulation.

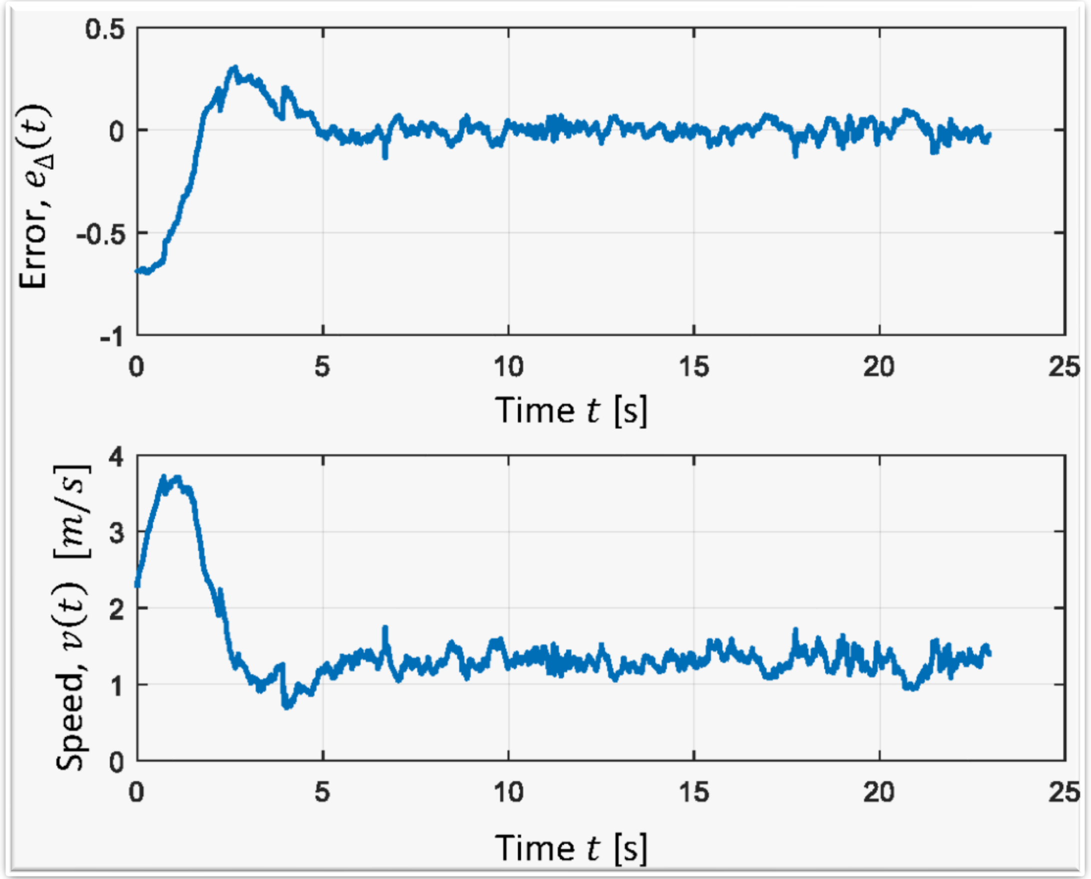




Simulations



- Simulations are programmed using the visually realistic Unreal Engine rendering engine.



Summary & Future Work

This work has provided a simple and effective method for the autonomous control of a nonholonomic mobile robot equipped with an onboard camera tasked with following a human of interest. The results show that the proposed method is capable of following a non-evasive target at typical human speeds through complex environments. There are many areas in which this work can be improved, many of which are currently being studied in the LISC. **Future work includes**:

- Static and moving obstacle avoidance using optical flow methods.
- Predictive model of human movement.
- Multi-target tracking and following in crowded environments.
- Social navigation in the presence of multiple humans.
- Dealing with missed detections and false positives.
- Dealing with occlusions and target leaving the FoV

Acknowledgement

- This research is funded by the Office of Naval Research Grant N00014-17-1-2175.

References

- [1] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and Berg, A.C. (2016). “SSD: Single Shot Multibox Detector”, *arXiv preprint.*
- [2] Wei. H., Lu, W., Zhu, P., Ferrari, S., Klein, R.H., Omidshafei, S., and How, J.P. (2014). “Camera Control for Learning Nonlinear Target Dynamics via Bayesian Nonparametric Dirichlet-Process Gaussian-Process (DPGP) Models”, *International Conference on Intelligent Robots and Systems (IROS).*