ESTIMATING AND INCORPORATING TARGET ORIENTATION FOR A HUMAN FOLLOWING ROBOT USING MONOCULAR VISION

Burke M* and Brink W†

* Council for Scientific and Industrial Research, South Africa - michaelburke@ieee.org

† University of Stellenbosch, South Africa - wbrink@sun.ac.za

Abstract

We present a monocular vision-based human following robot that incorporates orientation information in its control strategy. We propose that the pose of a walking person's upper body typically indicates their intended travelling direction, and find that a simple planar fit to the back of a salient human torso contains sufficient information for the purpose of inferring orientation.

Our human following robot equipped with a single camera mounted on a pan tilt unit (PTU).

1. Introduction

The ability of a mobile robot to track and follow a human is required in a wide variety of applications, particularly in service robotics.

Traditional approaches only use some form of a human's position for navigation. A more intelligent navigation strategy could be implemented if some knowledge of the intended motion of the human target was incorporated.

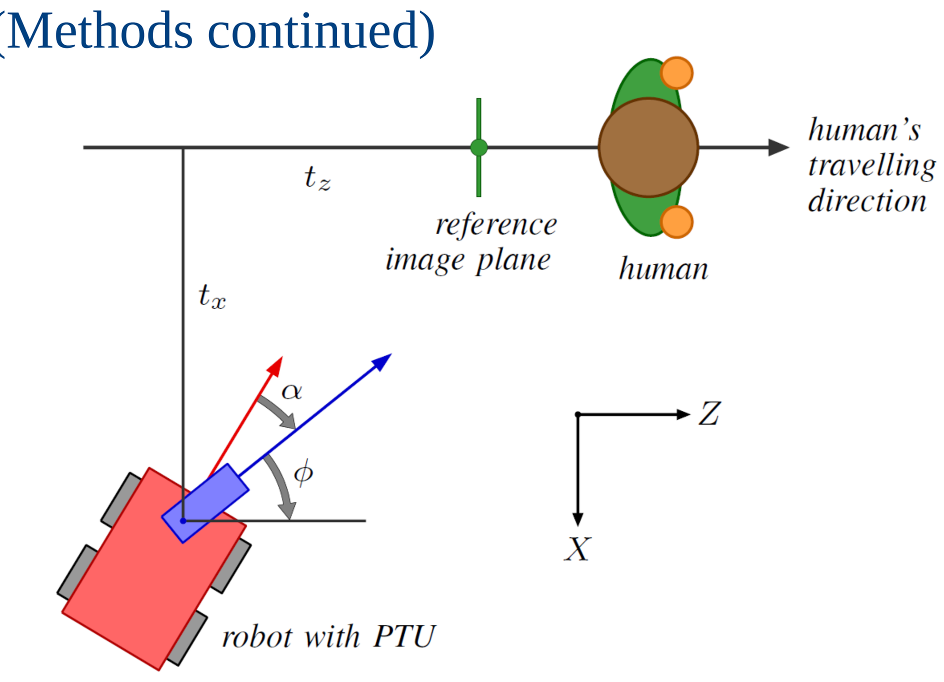
Human following navigation schemes that include orientation information require that some measure of human body pose be made. We present a direction-based approach to human following and a technique for extracting basic human pose from a single image.

2. Method

Our system operates under the assumption that the pose of a walking person's upper body typically indicates travelling direction. As we are interested only in the approximate facing direction of the human, a planar fit to the back of the torso contains sufficient information to infer travelling direction.

The human following task requires that the relative pose between robot and human be determined in real time.

(Methods continued)



The visual servo control task requires that the robot navigate so as to recreate the desired reference image.

Initially, SURF [1] feature matches relating the robot's current view of the human to a fronto-parallel image of the back of the human's torso are obtained. A planar homography mapping the features is then estimated, from which pose measurements are extracted, up to a scale [2].

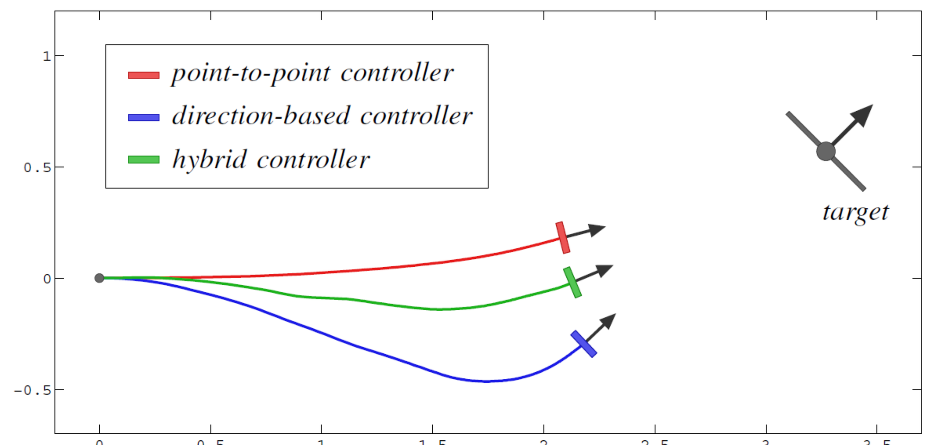
Although features detected on the back of a human torso are usually not coplanar, a robust RANSAC (Random Sampling and Consensus) based method of homography estimation [3] is able to discard errors induced by this assumption. In addition, a robust measure of homography is also required to reduce errors caused by the deformable nature of clothing.

Pose estimates are then improved greatly by including tracking using an extended Kalman filter.

Three controllers were designed for the human following system. The first merely centered the human in view, while the others made use of the extracted orientation information in their control strategies.

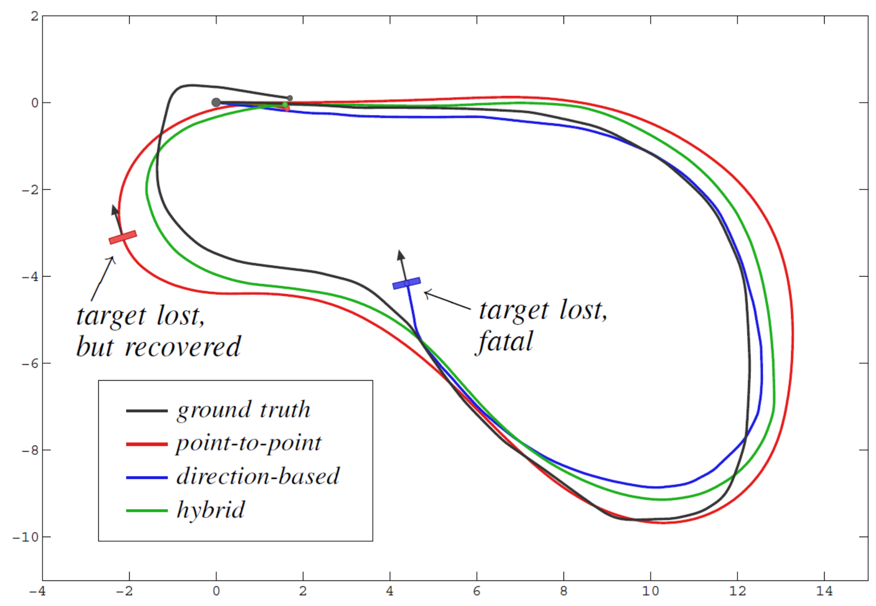
3. Results and Conclusions

Pose measurements were successfully estimated under a wide range of conditions. The robust homography estimation represents a good planar approximation to the human torso.

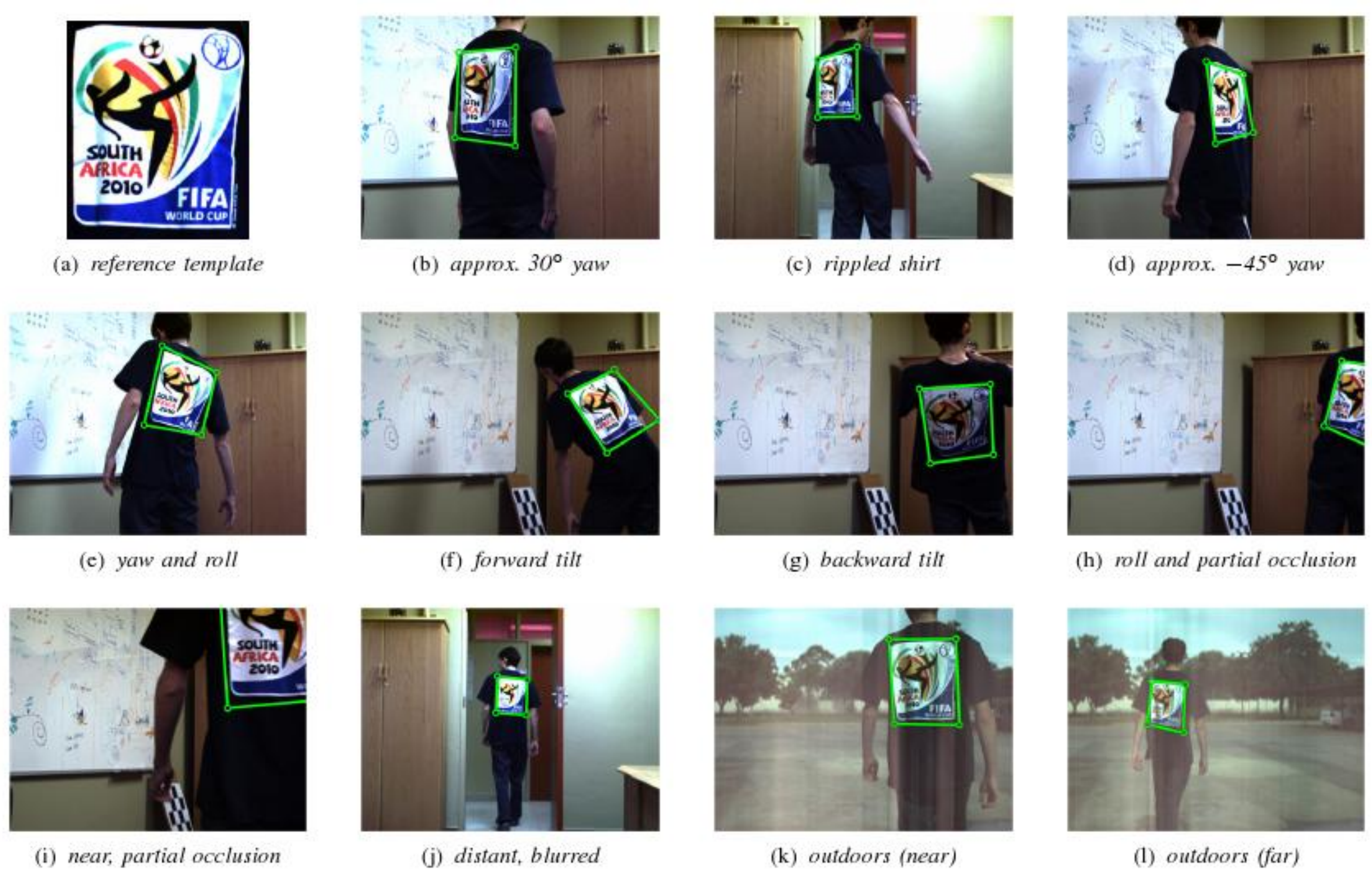


The step responses of the three controllers designed show how they differ in operation.

- The point-to-point controller discards orientation and is vulnerable to losing a sharply turning target.
- The direction-based controller corrects orientation, but is vulnerable to losing a fast target as it drives along a non-ideal trajectory.
- The hybrid controller combines the benefits of these controllers in an attempt to remedy their defects.



The hybrid controller performs best as it uses pose information to navigate so as to avoid losing sight of the target.



Results showing the efficacy of the planar fit to a human torso. Position and orientation information is extracted from the planar fit and used to control the robot.

References

[1] H. Bay et. al., "Speeded-up robust features (SURF)," Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359, 2008.

[2] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," International Journal of Pattern Recognition and Artificial Intelligence, vol. 2, pp. 485–508, 1988.

[3] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge University Press, 2004.

Acknowledgements

Thanks to MIAS for the use of their platforms and facilities. This work has been submitted to the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems.

UNIVERSITEIT STELLENBOSCH UNIVERSITY

our future through science