Performance evaluation and benchmarking driving robotics R&D

Yael Edan
Department of Industrial Engineering and Management
Agricultural, Biological and Cognitive Robotics Initiative
Ben-Gurion University of the Negev

In the field of robotics which is currently booming with R&D there is a growing awareness of the importance of benchmarking and performance evaluation which I have been investigating since the late 80s. As the complexity of current robotic and intelligent systems grows, developing sound experimental approaches and benchmarking procedures is critical. Benchmarking robotics research is inherently difficult. Performance evaluation of robotic algorithms is usually based on either experiments in real environments, or by theoretical analysis. The first approach is problematic in unstructured and dynamic environments, since it is impossible to repeat experiments under identical conditions. Furthermore, typically, results are reported only for a specific robotic system and a set of self-chosen tasks. The second approach requires explicit assumptions concerning the nature of the environment, task and robot and sensory information. This approach is hard to implement since it is usually difficult to characterize performance in complicated and dynamic environments. Furthermore, its applicability is limited providing only solutions to the cases solved.

I will present how applying performance measurement and evaluation of robotic systems has advanced methodologies for improved robotic systems with special focus on industrial engineering algorithmic developments including performance analyses of grippers, robotics and sensory systems, and human-robotic systems.

Additionally, two detailed examples in which performance measurement has led to development of new powerful algorithms will be presented. Usually, performance of detection algorithms are measured through the receiver operating characteristic curve which describes the number of correct detections as a function of the number of false alarms. Our novel approach in a robotic site specific sprayer we developed is to measure performance using maximal pesticide reduction while maintaining high target detection rate. High target detection rate (foliage and grape clusters in this case) usually increases also false alarms reflected by low pesticide reduction. An algorithm that finds the optimal point which leads to maximum pesticide reduction for a certain level of target detection rate (i.e., minimum false alarms for a certain detection rate) will be presented. In another research, improved detection for robotic harvesting was achieved by incorporating different task objectives in an adaptive thresholding algorithm. The adaptive thresholding algorithm derives local 3D thresholds based on illumination conditions and takes into account a changing task objective that describes the relationships between false positive rate, true positive rate, and accuracy in the location.