

ECE 498 - ArmTrak: Paper Review

Amod Agrawal, amodka2

April 7, 2019

1 Summary

Hand gesture recognition is a fairly developed area in mobile computing and has had various solutions for high-level activity recognition like drinking, eating, and driving, etc. However, this work tries to solve a much harder problem of tracking the user's arm posture and not the activity the user is performing. This problem is considered hard because ArmTrak has only one point of (noisy) data source on the wrist and the goal is to track the motion the entire arm. The key idea that this work uses is that the arm posture follows a certain model.

Human hand posture can be decomposed as a function of the elbow location, wrist location, and wrist rotation. Since one point of the elbow is fixed, its location has 2 DoF. For a fixed elbow point, the forearm that defines the wrist location has 2 DoF as well. The wrist rotation for a fixed elbow and wrist location has 1 DoF. This means a given posture lies in five-dimensional space (5 DoF). The authors use the fusion of compass, gyroscope, and accelerometer at the smartwatch to estimate the posture, however, the error grows unboundedly. To solve this problem, the authors use the motion model of the arm since the arm motion has certain constraints and can't just move arbitrarily. This motion model of the arm along with the IMU readings are then fused using Hidden Markov Model (HMM). Authors find that the direction of the forearm is strongly coupled to the arm posture and since the orientation of the watch can give us the direction, the variable search space is immensely reduced which can then be solved using IMU values.

2 Critique & Opinion

While most of the work prior to ArmTrak has focussed on high-level activity recognition like eating, drinking, etc. This work is much more technically deep because it tries to estimate the posture of the arm itself. This posture information can then provide context and features to various machine learning models or algorithms for activity recognition or gesture recognition.

The work naturally relies on the assumption that the orientation of the watch can be precisely calculated. Orientation estimation requires unpolluted gravity readings and non-fluctuating magnetometer readings. If the device doesn't see much of the static motion, it's hard to use gravity to correct the error. Similarly, indoor environments can have magnetic fluctuations that can pollute magnetometer readings.

The work also assumed that the user is static while performing arm gestures. If the user is in motion, the accelerometer data at the wrist is polluted by the linear acceleration of the body. Decoupling body motion and hand motion is a much harder problem without using additional IMUs on the body.

Authors also show that even though the search space has significantly reduced due to the elbow point cloud, it's still pretty large for on-line estimation. There is a significant latency in the estimation which suggest that new constraints will further reduce the search space and encourages us to further build-up on this work. Finally, even though authors show results across multiple users, it's difficult to say how the arm motion models will scale up with global deployment.