Dima Jamal Al Damen

Activity Analysis: Finding Explanations for Sets of Events

Doctor of Philosophy
September 2009

Automatic activity recognition is the computational process of analysing visual input and reasoning about detections to understand the performed events. In all but the simplest scenarios, an activity involves multiple interleaved events, some related and others independent. The activity in a car park or at a playground would typically include many events. This research assumes the possible events and any constraints between the events can be defined for the given scene. Analysing the activity should thus recognise a complete and consistent set of events; this is referred to as a global explanation of the activity. By seeking a global explanation that satisfies the activity's constraints, infeasible interpretations can be avoided, and ambiguous observations may be resolved.

An activity's events and any natural constraints are defined using a grammar formalism. Attribute Multiset Grammars (AMG) are chosen because they allow defining hierarchies, as well as attribute rules and constraints. When used for recognition, detectors are employed to gather a set of detections. Parsing the set of detections by the AMG provides a global explanation. To find the best parse tree given a set of detections, a Bayesian network models the probability distribution over the space of possible parse trees. Heuristic and exhaustive search techniques are proposed to find the maximum a posteriori global explanation.

The framework is tested for two activities: the activity in a bicycle rack, and around a building entrance. The first case study involves people locking bicycles onto a bicycle rack and picking them up later. The best global explanation for all detections gathered during the day resolves local ambiguities from occlusion or clutter. Intensive testing on 5 full days proved global analysis achieves higher recognition rates. The second case study tracks people and any objects they are carrying as they enter and exit a building entrance. A complete sequence of the person entering and exiting multiple times is recovered by the global explanation.