VIDEO UNDERSTANDING FOR GROUP BEHAVIOR ANALYSIS

Garate C., Bremond F. - INRIA Stars Team Francia
{carolina.garate, francois.bremond} @inria.fr

Abstract

The automatic video interpretation in the cognitive vision field has become an important research topic for real life application during the last years. One clear example is Video Surveillance. The main goal is to recognize the behaviors of a group of people (2-5 persons) involved in a scene depicted by a video sequence. This problem could be solved with the implementation of an automatic system to recognize in real time. So far, there are different research topics related to crowd or isolated individuals, but only a few works have addressed the recognition of group behavior. The current approach uses recent advances in group tracking and behavior recognition to process large amounts of video surveillance data from an underground railway station and perform a statistical analysis. The most important advantages of our approach are the robustness to process long videos and the capacity to recognize several and different events at once.

Video Processing

• The approach considers a process chain of 5 stages:

Segmentation → Blob Detection → Physical Objects Tracking → Group Tracking → Event Detection

Group Tracking

- Our technique is based on the one proposed in Zaidenberg et al. (2012). *A generic framework for video understanding applied to group behavior recognition*.
- Input algorithm: set of trajectories of the physical objects.
- Mean-Shift clustering algorithm to find similar trajectories.
- Group characterized through three features:
  - intra-object distance average.
  - average standard deviations of speed.
  - average standard deviations of direction.

$$GroupIncoherence = \omega1 \cdot distanceAvg + \omega2 \cdot speedStdDev + \omega3 \cdot directionStdDev$$

- Minimizing groupIncoherence value: lower value the group coherence is higher
- Update of a group:
  - associate cluster with probable group of most individuals.
  - add object if its probable father was in the group associated with the cluster.
- Global Tracking:
  - re-acquisition method: link current group Ids with previously lost IDs in a time window.

Behavior Recognition

Knowledge Application

ScReK proposes a simple formal declarative language.

- Constraints are based on operators:
  - Simple: equality, negation, >, …
  - Temporal: before, during, …
  - Customized: definition of new operators.

Type — Name

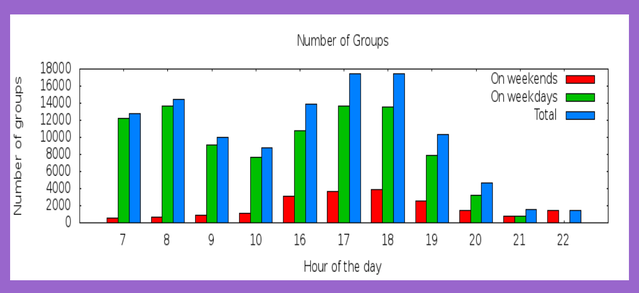
CompositeEvent (Group_Converting_Crowd ,
PhysicalObjects ((g1 : Group), (z1 : Zone), (z2 : Zone))
Components ((c1 : Group_Changes_Zone (g1, z1, z2)))
Constraints ((z2->Name = EntranceTurnstilesZone)
(variation(g1->AvgIntraDist) <= DistForCrowd))
Alarm ((Level : URGENT)))

- List of sub-scenarios to define the complex scenario
- List of objects Involved in the Scenario model
- Level of importance Of the scenario
- operators

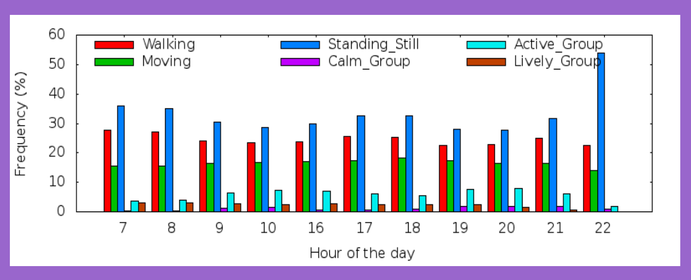
Event Recognition Algorithm

- Optimization step
  - Optimal scenario models are composed of:
    - * At most two components.
    - * At most one temporal constraint.
  - Scenario model tree: rules defining how recognized simple scenarios trigger the recognition of complex scenarios.
- Recognition process
  - Instantiate and recognize all primitive states according to detected objects.
  - Recognize complex events according to scenario model tree.
  - Compare recognized scenarios at frame t with previous ones.
  - Update or create a new recognized scenarios.

Group Detection Results

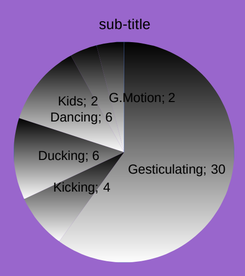


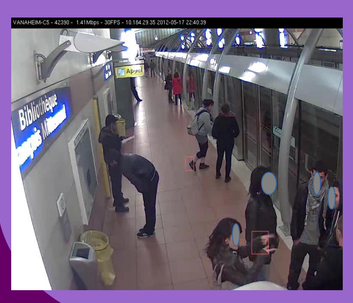
Group Event Recognition Results

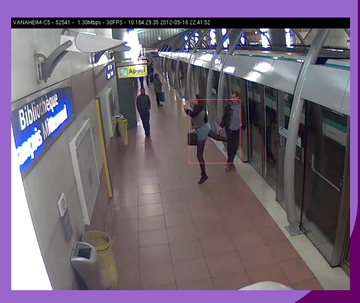


Sudden Movements Results

• 50 annotated sudden movements







Conclusion

- Real time in long surveillance videos.
- Recognize multiple and different events at the same time.
- Robustness regardless video length.
- Statistically results for the large amount of videos processed

Future work:

- Add sudden movements for groups
- Add probabilities to the event recognition phase → predict dangerous and criminal events