

Research & Development - DRAFT

Evaluation of current Approaches for Situation-Awareness in Autonomous Systems

Author:

Maximilian Schöbel

Advisors:

Prof. Dr. Erwin Prassler

Prof. Dr. Paul G. Plöger

May 15th, 2015

Contents

1 Introduction

Brief but concise review of the first four "W"s.

Applications: video surveillance for security reasons or in assisted living environments for the elderly, video retrieval, human-computer interaction.

Because of large variations in used datasets and evaluation practice, approaches are discussed conceptually.

Focus on Deep Learning approaches, since conventional methods have been surveyed in many articles before (see related work section).

1.1 Situation Awareness from video data

General Definition of Situation Awareness in the context of autonomous systems.

Placement of Action Recognition among other vision-based methods, i.e. Action Prediction, Anomaly Detection, Event and Action Detection, Person/Pedestrian Detection, Gesture Recognition.

Definitions of the above methods.

General Processing Pipeline for Action Recognition: Person Detection -> Tracking -> Action Detection -> Segmentation -> Action recognition.

1.2 Survey Papers in Action Recognition (Related work)

Review of most important/recent review papers in Action Recognition with traditional and Deep Learning approaches.

1.2.1 A survey on vision-based human action recognition, Ronald Poppe (2010)

Definition of action: Uses the hierarchical classification of human motion in action primitives, actions and activities as given in Moeslund et al. (cite ??)

Action primitives are atomic movements at the limb-level.

Actions are possibly cyclic whole body movements and consist of multiple action primitives.

Activities consist of multiple actions whose subsequent execution make the movement interpretable.

Example: Action primitives: Left/right leg forward -> Action: Starting, Running, Jumping -> Activity: Jumping hurdles.

Scope: Gives a very good classification of conventional methods in human action recognition.

The discussion is split according to video representations and classification methods.

Challenges of the domain are described very well.

Deficits: No Deep Learning methods are discussed.

Datasets and benchmarks are discussed only very briefly.

1.2.2 kk

<++>

1.3 Challenges in Action Recognition

Action Recognition is a classification-task.

Intra- and inter-class variances.

Background and recording settings.

Temporal variations.

Obtaining and labeling training data.

Difference to face/gate recognition: Generalize over person characteristics.

1.4 Taxonomy

Specification of a taxonomy for conventional approaches (from survey papers).

Specification of a taxonomy for Deep Learning approaches (from survey papers if present or self-created).

2 Conventional Methods in Action Recognition (Representation-based approaches)

Brief description of conventional Methods in action Recognition, i.e. its beginning with hollistic representations and advances to aggregation-based approaches.

2.1 Global Representations

2.2 Local Representations

3 Deep Learning Methods in Action Recognition (Deep Networks based approaches)

Review of approaches that use Deep Learning.

3.1 Spatio-Temporal Networks

I.e. convolutional methods.

3.2 Multiple Stream Networks

3.3 Generative Models

Restricted Boltzmann Machine

3.4 Temporal Coherency Networks

4 Datasets and Benchmarks in Action Recognition

4.1 Review of Datasets for Human Action Classification

Review of the most important currently existing datasets, focus on newest ones (since 2013)

Reference dataset survey paper.

4.2 Alternate Benchmarks for Action Recognition algorithms

4.3 Inter-dataset approaches

5 Evaluation

What do we need, what do we have, what is best suited so far?