Predicting abnormalities in complex human-object interaction by using object affordance context

Mahmudul Hassan and Anuja Dharmaratne
School of IT, Monash University Malaysia

MONASH University
Information Technology

Abstract

Detection of anomalous video events in complex Human-Object interactions is a novel and challenging computer vision problem. Researchers have applied various contexts like spatial, temporal, sequential etc. to address this issue. This conceptual paper explores the potentiality of object's affordance as a promising context for recognizing and predicting complex human-object interaction based abnormalities in novel video sequences.

Motivation

In contrast with the views of the orthodox schools of constructive psychology, the concept of affordances (1979), has allowed us to look at the intrinsic and functional attributes of objects. As object's affordance trigger actions, the relational mapping of object class, human and the ambient scene may aid anomaly detection process in Human-Object interactions.

Applications of affordance in computer vision

From the definition of Affordance, it is evident that, it is neither subjective nor objective, it does not solely depend on the environment or the actor rather it depends on both. 'Affordance' has been used in saliently solving different aspects of Human-Object interactions (HOI) related challenges in computer vision. Such as: improving the detection process of actions and objects [1], saliently comprehend scenes, anticipate probable human activities [2] etc. Finding further potential applications of object's affordance is an extremely interesting proposition

Video based abnormality detection

The definition of anomalous actions refers to the course of actions that does not follow the expected pattern. To label a novel action abnormal, we may need to possess a profile of normality. Ironically, this 'normality profile' depends largely on the contexts. For example, when a human is lying, the 'spatial context-**where**' is very important to judge whether this is a normal action or not

Different contexts used in Computer Vision Community are:

❏ Spatial
❏ Temporal
❏ Sequential etc.

Proposed Idea

Researchers have often overlooked the context of 'object's affordance' as a potential source of anomalous human object interaction detection. Most of our daily activities do involve objects and anomalous events can occur in these actions.

For example, a knife is meant for chopping and cutting but throwing it may be an anomalous event. At the same time standing on a tool may be normal for an adult but can be potentially dangerous for a child. Even there may be abnormalities in the scene. For instance, missing grill in front of a fire place, a turned on gas hob without any utensils for longer periods etc. are potential sources of danger. We argue, that these anomalies are very hard to recognize by the current algorithms in video based anomaly detection and the intrinsic characteristics of object's affordance can serve as a great potential tool to find these human-object interaction or object-object interaction based abnormalities.

Attributes regarding Human and Objects for anomaly detection

Object related attributes
Object's affordance class, material, aspect ratio, object's state, danger index, shape, ambient object's affordance, object's group etc.

Actor related attributes
Age, gender, disability, aspect ratio, alone/group etc.

we may map and learn the mutual relational contexts between the human, object and scene in a normal setting and try to recognize anomalies in a novel event.

Proposed Approach

Probabilistic graphical models (PGM) are powerful means of reasoning in the domain of uncertainty. The Human poses and its attributes and the object and its attributes may be considered as different nodes of the PGM. With substantial amount of training videos, the structural relationships between these nodes may be learned. Inference can be done for most probable actions or for most probable joint assignments of these nodes. The inferred quantities can subsequently be used for anomaly .

Experiment

In accordance to the scope of this conceptual paper a fundamental and naive experiment has been done.

A Bayesian Network has been created with nodes corresponding to : 1(O)=Object Class (A coke Bottle, A medicine Bottle, An aerosol bottle), 2(A)=Age group of the actor (Child and adult), 3(P)=Human pose (Drinking pose and spraying pose), 4(D)=Danger Index of the object (Low and High),5(AF)=Affordance of the object (Drinking and spraying) and 6(AS)=The Action state (Normal/Abnormal).

The resulting Bayes Net is
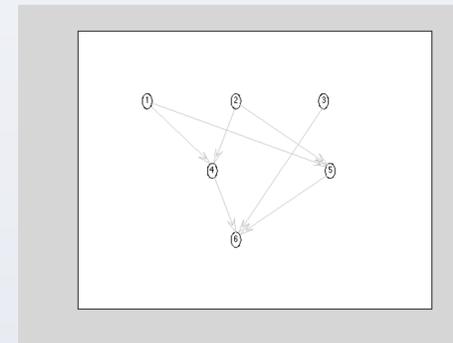
The problem was formulated as a classification problem, and to find the marginal probability

$$P(AS|O,A,P,D,AF)$$

We have considered two scenarios where we condition on the object as an aerosol spray, the age group as child. The resulting Abnormality came to

0.1550 ~ 15.50% (Normal)
0.8450~84.50% (Abnormal)

The same experiment with the conditioned age group is adult reveals a result
0.68~68% (Normal)
0.32~32%(Abnormal)

Another experiment was done where an adult has been drinking from the aerosol spray and it also provided with a desired outcome.

Contributions

❏ This research contributes a novel context for Human-Object interaction based abnormalities.

❏ Include and explores new attributes related to Human and object for more salient affordance prediction.

❏ Providing a framework for ambient affordance prediction.

Conclusion

Affordance, in usual circumstances, is a powerful feature of objects which, inherently maps the relations between the human, object and the ambient environment. This conceptual paper urges to exploit this intrinsic characteristic of object's affordance in order to address a novel challenge of Human-Object Interaction based abnormalities.

References

Gupta, A.; Davis, L.S., "Objects in Action: An Approach for Combining Action Understanding and Object Perception," Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on , vol., no., pp. 1,8, 17-22 June 2007

Hema S Koppula, Ashutosh Saxena, "Anticipating Human Activities using Object Affordances for Reactive Robotic Response",. In Robotics: Science and Systems (RSS), 2013.