Information Fusion 21 (2015) 68–81

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Collaborative detection of repetitive behavior by multiple uncalibrated cameras

Tommi Määttä [a,*], Aki Härmä [b], Hamid Aghajan [c,d], Henk Corporaal [a]

[a] Electrical Engineering Department, Eindhoven Technical University, Netherlands
[b] Philips Research, Eindhoven, Netherlands
[c] Ambient Intelligence Research (AIR) Lab, Stanford University, USA
[d] Image Processing and Interpretation Research (IPI) Group, Gent University, Belgium

ARTICLE INFO

Article history:
Available online 18 February 2014

Keywords:
Repetitive behavior
Transition modeling
Multi-view
Fusion

ABSTRACT

In smart environments, the embedded sensing systems should intelligently adapt to the behavior of the users. Many interesting types of behavior are characterized by repetition of actions such as certain activities or movements. A generic methodology to detect and classify repetitions that may occur at different scales is introduced in this paper. The proposed method is called Action History Matrices (AHM). The properties of AHM for detecting repetitive movement behavior are demonstrated in analyzing four customer behavior classes in a shop environment observed by multiple uncalibrated cameras.

Two different datasets, video recordings in the shop environment and motion path simulations, are created and used in the experiments. The AHM-based system achieves an accuracy of 97% with most suitable scale and naive Bayesian classifier on the single-view simulated movement data. In addition, the performance of two fusion levels and three fusion methods are compared with AHM method on the multi-view recordings. In our results, fusion at the decision-level offers consistently better accuracy than feature-level, and the coverage-based view-selection fusion method (51%) marginally outperforms the majority method. The upper limit with the recorded data for accuracy by view-selection is found to be 75%.

1. Introduction and related work

Vision networks using multiple cameras offer means to cover vast areas and to observe people in various environments. A single camera might sometimes provide enough information to detect behavioral patterns with a reasonable accuracy. However, by having other views that simultaneously provide data from other viewpoints, the accuracy in decision making can be expected to improve. The amount of improvement depends on the fusion mechanisms that combine the data. The decisions on the activities and behavior can be used to control various services, e.g., in smart homes or shops, or to detect important events, e.g., in care homes or hospitals.

1.1. Behavior modeling

Behavior is often defined as actions or reactions of a person in response to external or internal stimuli [1]. *External* stimuli relates to interaction of a person with the environment while *internal* stimuli relates to the task objective and the emotional and physiological state of the person. In this sense behavior resembles a computer program that is executed as a response to the input. The stimulus–response model is a simplified mechanistic model of human behavior but is useful in motivating the approach taken in this paper. This paper focuses on the development of a behavior recognition system, which estimates the internal and external stimuli that drive the behavior. The generic behavior system is illustrated in Fig. 1.

In an ideal situation a stimulus would always lead to a deterministic behavior. However, in the real world observed behavior also has the stochastic component which depends on many internal and external factors. One may roughly model the stochastic component of behavior as a response to a noise component that is added to the stimuli. Even though a person is executing a certain behavior, or a deterministic program, the person may sometimes be thoughtless, get accidentally lost, or just be clumsy in their interactions with other objects/persons in the scene. In the behavior recognition system, see the lower part of Fig. 1, there is another noise component which is related to the resolution, distortions, and additive noises in the environment and the sensors. The design of a behavior

* Corresponding author. Tel.: +31 (0)6 17573216.
E-mail addresses: tommi.maatta@loft-nedsense.com (T. Määttä), aki.harma@philips.com (A. Härmä), aghajan@stanford.edu (H. Aghajan), h.corporaal@tue.nl (H. Corporaal).

http://dx.doi.org/10.1016/j.inffus.2014.01.010
1566-2535/

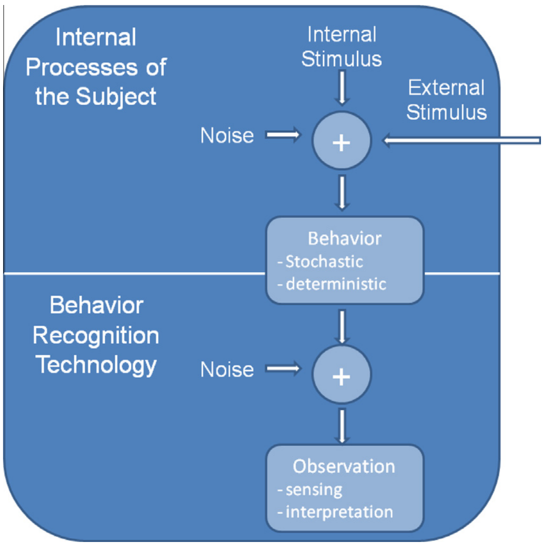

Fig. 1. The definitions and relations of behavior and the technology monitoring it.

recognition system should be optimized such that it eliminates both noise components from the final interpretation of the observation, and focuses on the detection of the deterministic part.

The basis for a behavior recognition system is a mathematical model of behavior. In psychology, mathematical models of behavior have often been associated with the *behavioristic* [2] research tradition best known for the famous works of Pavlov. Most of the early mathematical models were related to the regulation of behavior, for example, through reinforcement or punishment [3], or input–output relations in reaction to external stimuli, see [4,5]. Some of more recent work includes the use of differential equations [6], dynamic programming [7], dynamic time warping [8], and sequential pattern recognition methods [9] for complex behavior modeling.

Repeated patterns are most often modeled using graphical models [10] such as Markov models, which translate the sequential behavior into state transition probabilities. The graphical models often require large training data which may not always be available. The deterministic part of behavior can be considered as a program that the subject is executing to complete a certain task [6,7]. However, there may be many actual programs a subject can execute to complete a task [11]. A graphical model consisting of several state transitions cannot necessarily handle this large variability in programs to complete the same task.

1.2. From behavior patterns to generic repetition

The literature on automatic detection of activity patterns mostly focuses on rigid definitions of the patterns. In [12], the activity patterns are discovered by first finding the frequent sets of actions, followed by a search of topology and time relations, and finalized by identifying additional conditions, such as time of day and other contextual information. Benabbas et al. [13] use the global motion patterns based on optical flow fields to find the most usual patterns and their magnitudes, and further use these in modeling and detecting usual group behavior scenarios such as group dispersion and evacuation. Most often, motion trajectories are used in detection of unusual patterns of motion. The common approach is to learn the usual motion patterns and use them in separating away unusual patterns, like was done in [14] by training and applying a hierarchical classifier.

Repetition of actions can be considered common when people are searching for an item. Therefore, it is important that a smart monitoring system is able to detect those repetitions. However, in a rich environment the action repetition can take many different forms and cannot necessarily be detected based on a comparison to usual/unusual motion trajectories or activity patterns. Focus of this paper is on the detection of action repetition *per se* without a need to make a clear distinction between the different types of repetitions.

An action may take place at the global or a local level. Movement that is generated by a person moving through an environment is referred to as *global* motion. *Local* motion is defined as motion of the limbs, body and head when subject is, e.g., interacting or adjusting himself. In case of both global and local motion, the behavior is represented by sequential execution of different actions. The main tool for discovering the repetitive action patterns is the analysis of the transitions between the actions. The main application context in this paper is the behavior of a customer in a shopping environment. In this context the most interesting behaviors are those cases in which customers start to repeat their actions when they are confronted by a situation in which they are uncertain how to proceed. For example, a customer might stay longer periods within the same area, return to previously visited area, or alternate between previously visited areas. It would be desired that a smart environment can detect such behavior and trigger means to assist the customer.

The AHM method for describing repetitive behavior is defined in Section 2. The computer vision algorithms for person tracking and dynamic motion path creation are presented in Section 3. In Section 4 we introduce the features, models and model structures used in classifying the repetitive patterns. The two types of experimental data, recorded and synthetic, together with the required data preprocessing steps are shown in Section 5. In Section 6 we present the classification and fusion results. Section 7 contains a discussion of the results and the main conclusions of this work.

2. Behavior by transitions: Action History Matrices

Behavior modeling based on, e.g., global movement trajectories of customers, would require unrealistic amounts of annotated training data. These global trajectories can be significantly different from person to person even for the same behavior pattern. The proposed method was designed primarily to address these problems. At the same time the approach allows the inclusion of some insights about the environment into the model. It is assumed that the essential characteristics of behavior (i.e., the program) related to an internal stimuli can be observed from the statistics of the state transitions.

Considering a shopping example where a person, who cannot find the right product category, alternates between the activities of walking around the shop and stopping to scan the selection in different parts of the shop. However, a subject who has difficulty in choosing between two or more products may walk between a small number of locations and alternate between walking and close examination of the products. In both cases the observed complex behavior consists of relatively simple state alterations which can be successfully detected from a camera image.

2.1. AHM – definition

A transition matrix is a statistical representation of transitions between different states of the subject during an observation period. Transition matrices have been used earlier in behavior modeling [15], but this paper aims at a somewhat broader definition of a methodology, enabling various types of data, such as activity and movement information, to be collected in a matrix form. The transitions are accumulated in a matrix, denoted here as