Towards Understanding and Modeling Individual Behavior and Group Dynamics

Michele Bezzi, Robin Groenevelt

Accenture Technology Park, 449, Route des Cretes, F-06902, Sophia Antipolis, France
{michele.bezzi, robin.groenevelt@accenture.com}

Abstract. Understanding human behavior is a very complex task. In this paper we present our initial thoughts on modeling and automatic recognition of human activities. We argue that to successfully model human behavior, we need to consider both individual behavior and group dynamics. To demonstrate these theoretical approaches, we also introduce two experimental systems for collecting experimental data of complex behaviors. The first system focuses on recognition and prediction of daily home activities in the context of elderly healthcare, and, in its current implementation, it can track one user only. To study more complex activities and group behavior, we are developing a second system to analyze everyday activity in our office.

1. Introduction

Automatic recognition and prediction of human activities from sensory observations is an extremely challenging task with a broad spectrum of applications: automated support system for video surveillance, context-aware human computer interfaces, performance evaluation systems information spreading optimization, environment design, health care, identifying social relationships, optimizing spatial layout, and identifying changes in people behavior. The complexity of human behavior in the physical world arises from the interaction between two main levels: *individual behavior* and *group dynamics*.

Individuals are *per se* complex entities: their actions depend not only on the sensory context, but also on various hard-to-measure factors such as past personal history, attention, attitudes, experiences, and emotions. Additional levels of complexity are found in human action description: a complex action (say ''drive to the office'') may be decomposed into a sequence of simpler task ("start the engine'', "drive out of the parking'', etc), or they may be part of more complex activities [1].

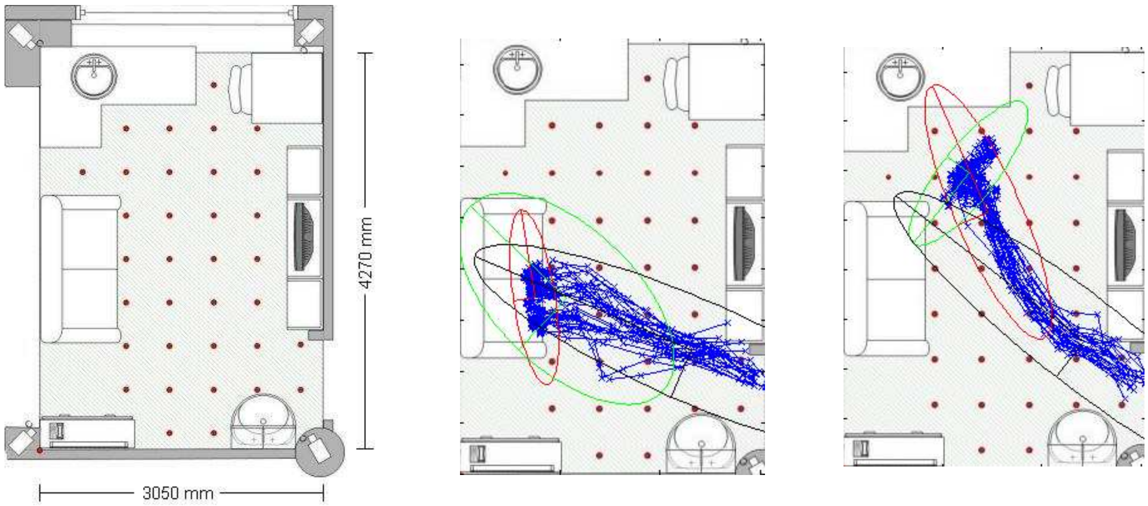
Group dynamics, often due to social interactions, are also highly complex processes. It has been found that networks of friendships or personal contacts often exhibit small world [2,3] or scale-free properties [4,5], i.e., there are many people with few connections and a few people with many connections. Taking advantage of the large use of digital communications, the properties of such networks have been extensively studied using data from emails [6,7], phone calls [8], collaborations [9], instant-messaging [10], and cross references of webpages [11]. Although this approach gives interesting results and potential applications [12,13], it does not consider physical interactions and face-to-face communications that are at the basis of human behaviors.

We are constructing a framework wherein these two levels will be integrated. This implies identifying different human individual and collective activities from sensory observations, finding a higher level description of these behaviors, and modeling these behaviors as they emerge from the interactions between the individuals. An important aspect of our study on behavior and intentions comes forth from human physical interactions, and to estimate this we will focus on the movement trajectories of people. In addition, more standard approaches based on electronic communications can be added at a later stage to provide additional information.

A successful model for understanding collective human behavior should in principle contain all these aspects, as well as a representation of the non-linear interactions among them. Although such a complete model is still far in the future, in the last decade progress has been made in this direction due to the concurrence of several causes.

First of all, recent advances in sensor technologies (i.e., RFID, cell phones, email, instant messaging, location aware devices, and energy measurement devices) along with the cheap deployment of cameras provide a large amount of data about human behavior in every day life.

Secondly, new theoretical tools are becoming available to analyze human behavior on individual and on social levels. Accordingly, the construction of models aiming at a conceptual organization of sensor data and individual based behavior into a model of human collective behavior is becoming a compelling task. Similar problems are being tackled by physicists and, more recently, by biologists in an attempt to derive macroscopic properties from microscopic



Our Laboratory | Sitting on the sofa activity | Making coffee activity

Fig. 1. Samples of trajectories for two activities (center and right panels). Observation distributions are represented by a mixture of two Gaussians (Ellipses in the figure). Only four states are shown in the figure for clarity; typical simulations are run with a 10 states Markov chain. Maximum likelihood parameter estimation was performed using an Expected-Maximization algorithm (EM, Baum-Welch).

interactions rules. A first application of such physics-derived methods to describe collective behavior has been introduced in [14].

2. Systems under development

In the rest of this paper we will describe two different systems for collecting complex experimental data of behaviors. In the first system called *Intelligent Home Services*---which has already been implemented and is in the early stages of deployment---recognition and prediction of daily home activities is made in the context of elderly healthcare. To better investigate more complex activities and inter-personal behavior due to the presence of more people, we are developing a second setting using our office as a test bed (see Sec. 2.2, *office activities*).

In the *Intelligent Home Services* system sensor data is acquired through video cameras, from which a person's location is extracted using blob extraction, Kalman filters, and heuristic methods. Video cameras have been chosen for practical reasons (price and availability) and because they provide a large amount of information about a person's activity: position, posture, attention focus, facial expression, and the location of objects. Moreover, the tracked people do not need to carry any special detectors to be monitored. The main limitations of such a system are: the complexity of image processing, the low-accuracy of automatic recognition, and the problems with simultaneous multi-person tracking. However, in a single user environment the camera system provides sufficient capabilities.

The drawback of the uncertainty of identifying individuals may be overcome using additional complementary sensors. Integration of data acquired through cameras, infrared badges, and fingerprint readers are currently under testing, and in the near future the usage of Ultra Wideband (UWB) sensors appears to be very promising. Simultaneous use of multiple noisy sensors can reduce uncertainty, improve accuracy, and increase tolerance to single sensor failures, but it also introduces the problem of how to combine these different sources of information (i.e., data fusion). We are currently investigating a probabilistic approach, based on Bayesian networks, to merge this heterogeneous information from low-level data streams. This data fusion, along with the interpretation of people's location is the focus of the study in the office environment.

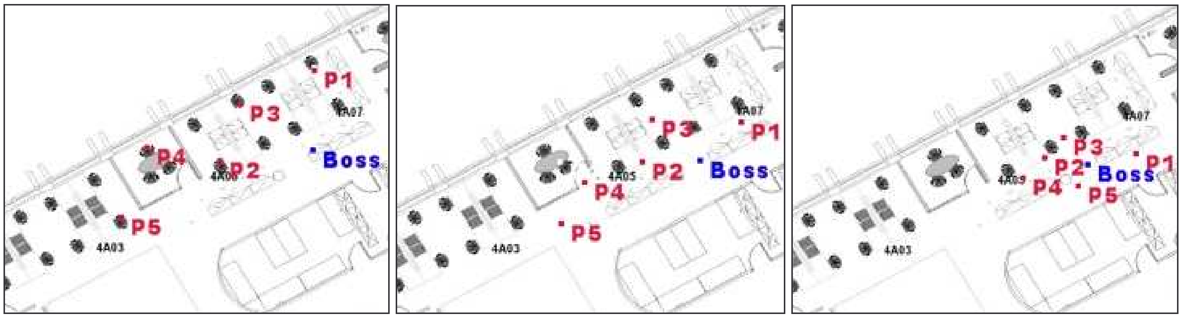


Fig. 2. Satellite view of office behavior: The arrival of *Boss* causes people *P1-P5* to come together.

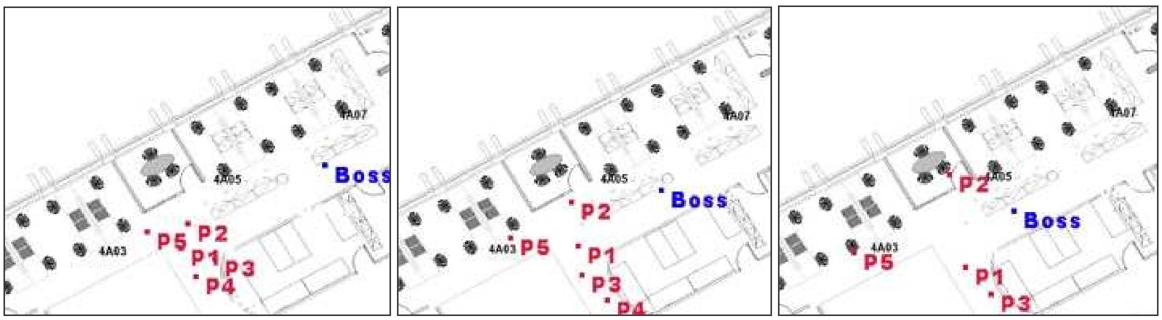


Fig. 3. Satellite view of office behavior: The arrival of *Boss* causes people *P1-P5* to scatter

2.1. Intelligent home services

Motivated by the growth of ageing populations, we developed a support system to help elderly people to live longer in their homes, reducing health cost, and improving an elderly person's quality of life. A camera-based video system has been installed in a room of our laboratory and reproduces a single room in a retirement center [15] (see Fig. 1, left panel).There were two aims to this project: (1) to develop an automatic system to monitor a person's daily activities to get insights on his moral and physical state and (2) to perform a series of cued actions whenever a person's state deteriorates and an action needs to be taken (i.e, the elderly falls and a doctor / family needs to be called).

Current implementations allow us to detect long term behavior patterns (as temporal trends in performing daily activity) and detecting events like falls. Sensory information from the video system---mainly a user's location and center point of gravity (with a precision of the order of 20 cm)---are used to identify a specific type of activity using a single layer hidden Markov model (e.g. "making coffee'' or ''sitting on the sofa'', see Fig, 1, center and right panels). After training the system, it is able to recognize various activities and perform predictions on short term behavior. Furthermore, collecting data over many sessions allow us to identify recurrent behavioral patterns (habits). The main areas for development of the current system are that it is limited by the small number of activities that can be tested in an experimental set up and that the current implementation is able to track just a single user, so no behavior related to other people or social interactions can be analyzed. Despite its simplicity, this system constitutes a valuable platform for experimenting with different strategies for human behavior recognition at an individual level, and it has been used for developing context-aware human computer interfaces [16].

2.2. Office activities

Considering the above areas for development we have chosen an office environment as a test setting for various reasons. First of all, a video-camera infrastructure is readily available in our location and the data is easily accessible. Secondly, correlations between people's behaviors are common in daily office life (see two simulated scenarios in Figs 2-3). Finally, data from the camera systems can be integrated with, or replaced by, other sensors (UWB tracking

devices, badge readers, finger print readers) and with data extensively available in electronic form (calendar, e-mail, logfiles).
The next step is the fusion of this raw-sensor data into a higher-level description of people's activities and behaviors. Potentially valuable models include graphical models [1], Bayesian networks [17,18], and context-free grammars [19]. Causality plays a crucial role in human behavior; in this respect the use of Bayesian networks can be extremely valuable, because in this framework causal dependencies can be efficiently described.
Although these models perform well for recognizing and modeling many different human behaviors, the diversity that characterizes single individual's actions should be taken in account to fully represent the richness of collective behaviors.
In other words, specific features of each individual contribute to the global behavior. From a modeling point of view, this means that individual-based models have to be introduced, where each individual is represented by its movement pattern but also by a set of features that describes individual personalities (this can be difficult to model mathematically, preliminary attempt in this direction are presented in [21]). Inter individual interaction (link in the language of social networks) can be estimated from physical interaction data, and other sources (email, instant-messaging exchange).

3. Conclusions

Automatic recognition of human behavior (both at level of single individual and of population) is critical in a number of applications, this motivated many recent studies on human activities detection and social networks. In this framework we are investigating how the two approaches (individual and social) may be integrated. This implies to identify different human individual and collective activities from sensory observations, and to find a higher level description for these behaviors, but also to model how these behaviors emerge from the interactions between the individuals. An important novel aspect of our study is modeling people physical interactions, we will focus on movement trajectories for estimating people social connections. This complements more standard approaches based on electronic communications. These models have to be tested on real life data. Thus, as a first step, we have developed a simple experimental setting where various individual daily life activities can be successfully recognized and short time activity prediction performed. A second, more complex, setting is under development to include group dynamics.

References

1. Duong, T.V.; Bui, H.H.; Phung, D.Q.; Venkatesh, S.; Activity Recognition And Abnormality Detection With The Switching Hidden Semi-Markov Model Computer Vision And Pattern Recognition, 2005. Cvpr 2005. IEEE Computer Society Conference On Volume 1, 20-25 June 2005 pp:838 - 845
2. Milgram S. The Small World Problem. Psychology Today, (1967). pp 60 – 67
3. Watts D., Strogatz. Collective Dynamics Of 'Small-World' Networks. Nature 393, P440-442, 1998
4. Albert R. ,Barabasi A.-L. Statistical Mechanics In Complex Networks. Reviews Of Modern Physics. (2002)
5. Watts D. Six Degrees: The Science Of A Connected Age. (2004).
6. Ebel, H., Mielsch L.-I., And Bornholdt, S. Scale-Free Topology Of Email Networks. Physical Review E 66, (2002).
7. Holger, E., Mielsch, L. I., Bornholdt, S., Scale-Free Topology Of E-Mail Networks, Physical Review E 66, 035103,(2002)
8. Aiello, Chung, And Lu. Proceedings Of The 32nd Acm Symposium On The Theory Of Computing. New York, (2000), P. 171.
9. Newman, M.E.J. Who Is The Best Connected Scientist? A Study Of Scientific Coauthorship Networks. Physical Review E 64, 016131, (2001).
10. Smith R. Instant Messaging As A Scale-Free Network. Cond-Mat/0206378. (2006).
11. Ripeanu, M., Iamnitchi, A., And Foster, I. Mapping The Gnutella Network: Properties Of Large Scale Peer-To-Peer Systems And Implications For System Design. Ieee Internet Computing Journal 6, 1 (2002), P50-57.
12. Wu F., Huberman B., Lada A., Tyler, J.R.. Information Flow In Social Groups. Physica A, 337:327-335, (2004)
13. Guimera, R, Uzzi, B, Spiro, J, & Amaral, L. A. N, Team Assembly Mechanisms Determine Collaboration Network Structure And Team Performance. Science 308, 697-702 (2005).
14. Helbing, D., Farkas, I., Vicsek T., Simulating Dynamical Features Of Escape Panic, Nature 407, 487-490, (2000).
15. http://panasonic.co.jp/sincere/shisetsu
16. www.accenture.com/global/services/accenture_technology_labs/r_and_i/intelligenthomeservices.htm
17. Olivier N., Rosario B., Pentland A. A Bayesian Computer Vision System For Modeling Human Interaction. Proceedings Of The First International Conference On Computer Vision Systems, P. 255 – 272. (1999).

18. Madabhushi, A., Aggarwal, J.: A Bayesian Approach To Human Activity Recognition. In: In Proc. Of The 2nd Internationalworkshop On Visual Surveillance. 25–30, (1999)
19. Ivanov, Y., Bobick, A.: Recognition Of Visual Activities And Interactions By Stochastic Parsing. Ieee Trans. On Pattern Analysis And Machine Intelligence, Tpami 22(8) 852–872 10, (2000).
20. Barabasi, A.L., Linked: The New Science Of Networks. Ny: Perseus, (2002).
21. Bagnoli, F., Berrones, A. , Franci, F. , De Gustibus Disputandum (Forecasting Opinions By By Knowledge Networks), Physica A ,332, 509518, (2004).