Introduction to the Special Issue on Intelligent Systems for Activity Recognition

Activity recognition systems aim to understand what people (and animals) are doing by observing their movement and their environment. The emergence over the past decade of novel sensing, low-power wireless communication, fast processing and statistical algorithms has made recognition practical and useful in several fields. Notable successes include gaming, surveillance, elder care, personal fitness, sports physiology and ecological systems monitoring and protection [Philipose et al. 2004; Pollack 2005; Yang 2009]. In the past, traditional works in activity recognition shared several characteristics. They focused on activities performed one at a time in fixed instrumented areas, by individuals rather than groups, and they mostly relied on specialized sensors using kinematic, location and object-manipulation-based cues, and view data on interactive time scales. The underlying technical machinery was typically fully supervised propositional time-series analysis based on machine learning and data mining.

The recent emergence of mobile sensing and online community activities, along with growing interest in social computing, has highlighted activities of new kinds at different scales [Zhang et al. 2010]. In particular, multiple simultaneous interactions between people (e.g., based on phone logs or location data), over long periods across many people (and animals), have called for novel intelligent systems and tools to shed new light on the structure of activity. This special issue provides a selection of several key directions in this recent research.

- *Collaborative activity recommendation* considers recommending activities for a target user by finding a collection of users who have similar historical activity traces with the target one. The advantage of the collaborative approach is that it can integrate the rich knowledge from a large number of historical users with similar interests to the target one.
- *Concurrent activity recognition* considers multiple concurrent activities that involve multiple participants such as conversations. The key point for concurrent activity recognition is to jointly model the concurrent activities and capture the relationship among them such that this relationship can improve the performance, compared to modeling them separately.
- *Pair-activity recognition* considers recognizing the activities defined for a pair of people who are interacting with each other, such as chasing and fighting. Pair-activity recognition takes into account the causality of two sensor reading traces to infer the interaction of the two people from video.

DOI 10.1145/1889681.1889682 http://doi.acm.org/10.1145/1889681.1889682

—*Evaluation metric for activity recognition* is a fundamental research issue that is worth studying. Thus far, the evaluation metrics used in most activity recognition works are borrowed from other related domains like information retrieval. A specific evaluation metric is required to compare the activity sequences that comprise time shifts, insertions and deletions.

—*Inferring colocation and conversational networks using privacy-sensitive audio* is a new topic for activity recognition. It aims to find a set of privacy-sensitive features for location and networking information in separately recorded streams of audio data from mobile phones. Research in this area promises to contribute to both activity recognition and computational social science.

This issue consists of six carefully chosen peer-reviewed articles.

In "Learning Travel Recommendations from User-Generated GPS Traces," Yu Zheng and Xing Xie present a travel recommendation system for location-based activity recognition using the GPS-enabled mobile phone by mining a large number of GPS traces. They propose a generic model that recommends the target user with top interesting locations mined from all the GPS traces and a personalized model that can recognize her travel preferences by matching her historical traces with those of others. They also collected a real-world GPS trace database comprising over a hundred users over one year to validate the effectiveness of their methods.

The article "Discovering Routines from Large-Scale Human Locations using Probabilistic Topic Models" presented by Katayoun Farrahi and Daniel Gatica-Perez adopts the "reality mining" data set from MIT media lab for mining daily location-driven routines. The data set comprises 491 consecutive days of cell-ID sequences collected by 97 subjects using mobile phones. The authors propose a bag representation to encode the temporal activity words and adapt the topic model to analyze the latent structure of the location-driven daily activities. They also visualize the analysis results and give a comprehensive interpretation.

Concurrent activity recognition is discussed in "Probabilistic Models for Concurrent Chatting Activity Recognition" presented by Jane Yung-jen Hsu, Chia-chun Lian, and Wan-rong Jih. They consider the scenario that multiple concurrent conversations are taken place where multiple participants join in and each wears an audio recorder. They use the Factorial Conditional Random Fields to model the relationship among the multiple concurrent conversations. They report on improved accuracy as compared to methods that ignore latent relationships, as is done in traditional parallel conditional random fields and hidden Markov models.

Another interesting topic, pair-activity recognition, is investigated by Yue Zhou, Shuicheng Yan, Bingbing Ni, and Thomas S. Huang who present the article "Recognizing Pair-Activities by Causality Analysis." They use videos as sensors to record the trajectories of two people and analyze the strength of causality between two trajectories. The causality features are derived from the Granger Causality Test, originally used in economics. They define five typical pair-activities such as chasing and following and collect a real-world data set to validate the effectiveness of the proposed pair-activity features.

Jamie Ward, Paul Lukowicz, and Hans Gellersen propose a set of new performance metrics for activity recognition in "Performance Metrics for Activity Recognition." Compared to many previous works that directly contrast activity sequences frame by frame, the proposed metrics take into account fragmenting, insertion, deletion, and time errors. They test the metrics on three published datasets and visualize the results to yield significant new insight.

The last article of this special issue is by Wyatt et al., titled "Inferring Colocation and Conversational Networks using Privacy-Sensitive Audio and the Implications for Computational Social Science." This article explores a set of privacy-sensitive features that can be used to find colocation and conversation events in separately recorded streams of audio data, using devices such as mobile phones. An innovative aspect of the solution is that it exploits a social network of colocation and face-to-face conversation among students over a long period of time. The solution used the privacy sensitive features to infer who was speaking when, and combined those inferences with colocation inference to determine who was in conversation with whom. The article not only contributes to activity recognition, but also to computational social science.

An overriding theme of these articles suggests that as more "digital footprints" are being collected, new opportunities emerge. The articles in this special issue represent a new trend in sensor-based activity recognition, where our research focus is beginning to shift from recognizing simple activities to more complex activities and patterns, that is, from the activities of individuals to that of groups, from simple predictions of the next events to more holistic discoveries on entire behavioral sequences, and from specialized and expensive sensors to more pervasive sensors such as mobile phones. As the "digital footprints" become increasingly available, activity recognition will usher in a new generation of applications in many areas like health care, wellbeing management, public safety, city resource planning, environment monitoring, target advertising, social and culture event promotion, etc.

DAQING ZHANG
Institut TELECOM & Management SudParis, France

MATTHAI PHILIPOSE
Intel Research Laboratory at Seattle, USA

QIANG YANG
Hong Kong University of Science and Technology, Hong Kong
Guest Editors

REFERENCES

PHILIPOSE, M., FISHKIN, K. P., PERKOWITZ, M., PATTERSON, D. J., FOX, D., KAUTZ, H., AND HAHNEL, D. 2004. Inferring activities from interactions with objects. *IEEE Pervas. Comput., 3*, 4, 50–57.

POLLACK, M. E. 2005. Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment. *AI Mag. 26*, 2, 9–24.

YANG, Q. 2009. Activity recognition: Linking low-level sensors to high-level intelligence. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*. 20–25.

ZHANG, D. Q., GUO, B., LI, B., AND YU, Z. W. 2004. Extracting social and community intelligence from digital footprints: An emerging research area. In *Proceedings of the 7th International Conference on Ubiquitous Intelligence and Computing (UIC'10)*. Lecture Notes in Computer Science, Vol. 6406, 4–18.