Multimodal Semantic Analysis of Public Transport Movements

Wolfgang Halb and Helmut Neuschmied

Institute of Information Systems,
JOANNEUM RESEARCH, Steyrergasse 17, 8010 Graz, Austria
firstname.lastname@joanneum.at

Abstract. We present a system for multimodal, semantic analysis of person movements that incorporates data from surveillance cameras, weather sensors, and third-party information providers. The interactive demonstration will show the automated creation of a survey of passenger transfer behavior at a public transport hub. Such information is vital for public transportation planning and the presented approach increases the cost-effectiveness and data accuracy as compared to traditional methods.

1 Introduction

It is very important for public transport and infrastructure providers to have exact measures of passenger movemements at public transport hubs (such as train or bus stations) and within the entire transportation network. Figures about changes of mode or type of transport (e.g., from train to bus, from private to public transport) are also valuable for public transportation planning and accurate statistics are a key factor for designing efficient and sustainable public transport [1]. Traditionally, obtaining these numbers has been expensive because it involves human observers manually counting the number of passengers over a number of days. More advanced approaches using Bluetooth hardware for wirelessly recording passenger movements [2] have the disadvantage that only passengers with a Bluetooth-enabled mobile phone can be captured. In the presented approach passenger movement data is extracted from surveillance videos and enhanced with additional information. The use of multiple cameras allows to cover larger areas and improves the detection rate. When fixed cameras are used it is possible to monitor movements over a long period of time thus enabling comprehensive statistics. The integration of data from mobile camera locations is also supported. Through a semantic analysis of video and other related data detailed information about persons' movements can be gained. The system has been designed to always ensure the privacy of passengers and persons are never identified.

2 The System in Brief

A system for multimodal, semantic analysis of person movements has been developed that incorporates data from surveillance cameras, weather sensors, and

third-party information providers. It has been developed for analysing passenger movements at public transportation hubs but can also be used for the analysis of person movements in any other setting (e.g., pedestrian movements in shopping malls). The components involved in the system are (i) visual semantic analysis, (ii) a contextual data interface, and (iii) a visualization of the combined statistics.

2.1 Visual Semantic Analysis

The visual semantic analysis is based on video input from surveillance cameras. Basic underlying principles on how the video input is processed are described in [3]. State-of-the-art object detection algorithms are used to detect people anonymously in the video input. The object detection is based on Histograms of Oriented Gradients (HOG) [4] and object trackers are used that have been refined and improved so that it is also possible to detect and track persons in crowded scenes and use multiple cameras for more accurate results. It is even possible to use video input from infrared cameras in the case of poor lighting conditions at night. As an intermediate result trajectories of person movements are available that describe the position of each person at a given point in time but they do not contain any additional semantics. This information is further processed and enhanced with a description of the persons' actions. In the user interface certain areas can be defined and linked to metadata about the area. Each time a person enters or exits an area this information is recorded. Along with metadata about the area the actual event semantics can be assessed. The area metadata contains information about the charateristics of the area, such as if the area belongs to the station or not. If an arrival/departure platform is located within the area, information about the serving bus/train lines is also included. Examples for further specialized areas are bike racks or parking lots. All information gained through the visual semantic analysis is stored in a central database for generating the passenger movement statistics. This database does not contain any sensitive information as it only contains data extracted from the visual analysis without storing any image content. It is not possible to determine a person's identity based on this information.

2.2 Contextual Data Interface

The contextual data interface is responsible for the acquisition of additional information that can enrich the detected passenger movements. Currently modules for the integration of weather information from weather sensors and bus/train schedules are implemented. The system has been designed for flexible use and thus allows including a range of further information providers. Bus/train schedules are used for determining the line a passeneger has used. In the case that multiple lines are served from a single platform it is not sufficient to know from which platform a passenger is departing for determining the used line and the additional information from the schedules is needed.

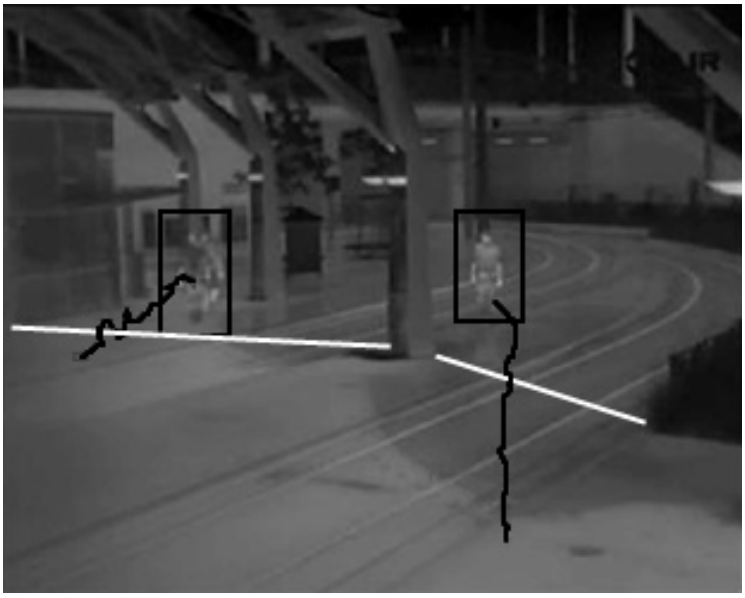

Fig. 1. Person movement trajectories.

2.3 Combined Statistics

Passenger movement statistics make use of all data that has been collected through the visual semantic analysis and the contextual data interface. Based on this information it is possible to determine the following events:

- Arrival at/departure from station with public transport (and bus/train line used)
- Entering/leaving the station and mode of transport (pedestrian, bike, car)
- Transfer between different public transport lines

Statistics can be accessed in near realtime through a web-based user interface. They contain information about the events as described above and thus present the degree of public transport line utilization. The presentation is highly customizable to show daily, hourly, or peak time movements, comparisons between different periods of time and so forth. By taking weather data into account it can be analysed if certain weather conditions (e.g., extreme temperatures, precipation, etc.) influence passenger movement behavior. This information is extremely valuable for public transportation planning to predict future passenger movement trends.

3 Demonstration

The demonstration will be highly interactive and show a complete sample usecase for a multimodal, semantic analysis of passenger movements at a public

transportation hub. For the demonstration pre-recorded surveillance camera footage from a real setup will be used. It will be shown how the visual semantic analysis can be configured to link visual regions with additional metadata about the region. The actual visual computations will be shown (see Figure 1 for an example of detected persons, their trajectories, and area crossings in a night scene, recorded with an infrared camera). The resulting statistics and the possible views will be shown as well. Through this demonstration the viewer will be able to understand the entire processing workflow and see the different possibilities for visualization of the results.

Acknowledgements The research leading to this paper was partially funded by the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT) in the 'IV2Splus - Intelligente Verkehrssysteme und Services plus' programme under project nr. 816031 (NET FLOW). We would also like to thank GRAZ AG VERKEHRSBETRIEBE (GVB) for making the video recordings possible.

References

1. David J.C. MacKay. *Sustainable Energy - Without the Hot Air*. UIT, Dec. 2008.
2. Vassilis Kostakos. Using bluetooth to capture passenger trips on public transport buses. *CoRR*, abs/0806.0874, Aug. 2008.
3. L. Paletta, S. Wiesenhofer, N. Brandle, O. Sidla, and Y. Lypetskyy. Visual surveillance system for monitoring of passenger flows at public transportation junctions. In *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*, pages 862–867, Sept. 2005.
4. Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1*, pages 886–893, Washington, DC, USA, 2005. IEEE Computer Society.