IDIAP RESEARCH REPORT

Understanding Metro Station Usage Using Closed Circuit TeleVision Cameras Analysis

C. Carincotte [a] M. Hick [a] X. Naturel [b]
J.-M. Odobez [b] J. Yao [b] A. Bastide [c]
B. Corbucci [d]

IDIAP–RR 08-38

June 2008

[a] Multitel asbl, Rue Pierre et Marie Curie 2, 7000 Mons, Belgium
[b] IDIAP Research Institute, Rue Marconi 19, 1920 Martigny, Switzerland.
[c] ACIC s.a./n.v. Rue René Descartes 2, 7000 Mons, Belgium
[d] ATAC SpA, Via Ostiense, 131L 0154 Roma, Italy

IDIAP Research Institute **www.idiap.ch**
Av. des Prés−Beudin 20 P.O. Box 592 1920 Martigny − Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

IDIAP Research Report 08-38

Understanding Metro Station Usage Using Closed Circuit TeleVision Cameras Analysis

C. Carincotte M. Hick X. Naturel J.-M. Odobez J. Yao
A. Bastide B. Corbucci

June 2008

Abstract. In this paper, we propose to show how video data available in standard CCTV transportation systems can represent a useful source of information for transportation infrastructure management, optimization and planning if adequately analyzed (e.g. to facilitate equipment usage understanding, to ease diagnostic and planning for system managers). More precisely, we present two algorithms allowing to estimate the number of people in a camera view and to measure the platform time-occupancy by trains. A statistical analysis of the results of each algorithm provide interesting insights regarding station usage. It is also shown that combining information from the algorithms in different views provide a finer understanding of the station usage. An end-user point of view confirms the interest of the proposed analysis.

1 Introduction

Despite the legitimacy of a number of privacy issues, Closed Circuit TeleVision (CCTV) networks are nowadays commonly present in public environments such as transportation premises, city centers or commercial establishments. In the meantime, automatic processing of video data is currently a field of activity stirring up the utmost attention in the pattern recognition community; state-of-the-art advances in this area enable the reliable extraction of surveillance-like events such as person tracking, face/object recognition, abnormal behavior and abandoned luggage detection.

Apart from surveillance and safety issues, CCTV video streams may also represent a useful source of information for urban planning and resource optimization applications. Advanced video analysis devices can indeed provide above physical sensors information and wide-area measurements that can replace, or at least complement, many conventional physical detectors. In addition, video detection performance can be easily verified and detectors are easy to reconfigure interactively. In this context, very few works address the issue of using already deployed CCTV network to perform statistics gathering (relatively unexplored problem), e.g. for maintenance/planning purposes (people counting, person classification, speed measurement...). Indeed, most studies deal with surveillance-like event detection and scenario recognition [1], allowing for example to determine whether a human is crossing the rails [2, 3], to detect overcrowding situation in the platform [4], or to produce an aggression indication [5].

In this context, our contribution is threefold. We first propose two algorithms, one to estimate the number of people in one camera view, and the other to measure the platform time-occupancy by trains. The second contribution is to provide a statistical analysis on a large dataset of the results of these detectors, to provide a better understanding of transportation stations usage from the planning and resource optimization point of view. The goal of this statistical analysis is to provide long term analysis of the station's usage, patterns of activity, and detection of abnormal events. Eventually, the joint analysis of different camera views also allows to provide a better analysis of people flow and of the station usage.

The structure of the article is as follows. Section 2 explains the algorithms for detecting people and train arrivals/departures. Section 3 provides an analysis on the results of each algorithm. Section 4 uses the combined results to provide a joint analysis. Finally, in section 5, feedbacks on the practical interest of such an analysis from the user point of view are given.

2 Video detectors

In this section, we provide a brief theoretical description of the algorithms used to perform people detection and platform occupancy measure.

2.1 Human detection

We have developed a fast method to detect humans from videos captured in surveillance applications. It is based on a cascade of LogitBoost classifiers relying on features mapped from the Riemanian manifold of region covariance matrices computed from input image features. The developed human detector relies on the approach of Tuzel *et al.* [6], which was shown to provide good detection performance for human detection in still images. It was improved by extending in several ways [7]. First, as the mapping process is slow for high dimensional feature space, we propose to select weak classifiers based on subsets of the complete image feature space. In addition, we propose to combine these sub-matrix covariance features with the means of the image features computed within the same subwindow, which are readily available from the covariance extraction process. Finally, in the context of video acquired with stationary cameras, we propose to fuse image features from the spatial and temporal domains in order to jointly learn the correlation between appearance and foreground information based on background subtraction. Our developed method can process from

5 to 20 frames/sec (for a 384x288 video) while achieving similar or better performance than existing methods.

2.2 Train arrival/departure detection

The proposed method[1] is mainly based on a tracking algorithm, which proved to be efficient in various contexts (indoor/outdoor, metro/train, camera location. . .). Furthermore, this approach does not use any background modelling estimation, thus preventing from context changes related issues, such as illumination, reflection. . . The main idea of the approach is to use trajectories from randomly distributed particles in the image to perform the train arrival/departure detection (and corresponding platform occupancy computation). Next paragraphs give an overview of the algorithm principle and stages.

The principle of the proposed method is to quickly locate moving objects in the scene, and to determine whether their motions are compatible with the requirements of the train arrival/departure (location, direction, speed. . .). To do so, particles are randomly initialized in a region of interest (i.e. "rails zone"), and tracking is activated for each particle when a defined criterion is met (roughly when motion is detected). Relevant trajectories can then be analyzed to compute useful information and eventually derive the final train arrival or departure decision.

Particle distribution Inactive particles are randomly distributed for each new image with a non-uniform rule using the calibration information, to take into account perspective over the region of interest (detection area).

Particle activation The activation criterion is based on an instantaneous motion detection, namely a thresholded frame-differencing operator. When an inactive particle is moved to a point where the frame-differencing operator is bigger than a defined threshold, the tracking for the concerned particle begins.

Particle tracking After activation, a particle continuously tracks the motion of the underlying object using a block-matching algorithm. Particles associated with uninteresting trajectories are recycled as detailed below.

Filtering of trajectories Trajectories are analyzed by computing a set of various features; linearity of track, track length, track duration, track direction, start/stop particle location. . . Trajectories are then classified as relevant ones and uninteresting ones depending on the features values. An uninteresting trajectory is then recycled while a a relevant one is kept active. Typically, trains/metros are characterized by well-defined trajectories, i.e. linear trajectories, mean direction parallel to the rails, speed linearly increasing/decreasing. . . On the other hands, passengers and tracking errors are most likely to have much more chaotic trajectories, which make them rejected.

Train arrival/departure detection All remaining trajectories are then scored; low values are attributed to trajectories of weak interest, while high scores are given to highly relevant ones. The total score is compared to a threshold to decide whether a train is potentially arriving/departing; arrival/departure time are lastly estimated using a Finite State Machine (FSM), which allows to distinguish between arrival and departure, and to compute the resulting platform time-occupancy.

[1]Intelligent Video System Software, Copyright ©2008 ACIC. For further information, please visit http://www.acic.eu/.

3 Single view analysis

In this section, the video detectors are applied on large amounts of video, and the resulting detections are analyzed individually.

3.1 Data description

(a) "Turnstiles view"

(b) "Platform view"

Figure 1: Processed camera views in Rome metro

Experiments were conducted using a dataset acquired for the CARETAKER project; videos come from 14 acquisition sessions performed in Roma metro on June 2007 (from 06.00 a.m. to 11.30 a.m.); the resolution of all sequences are PAL standard (720x576 pixels, 5 frames per second) and compressed using MPEG-4. Two cameras are used, one monitoring the platform (Fig. 1-(b)) and one monitoring the turnstiles to access the train platform (Fig. 1-(a)). In this view, several flows of people are mixing up. People can come from the left or from the top and go through the turnstiles to access the train platform, people can inquire at the desk (middle top of the image), and people that leave the platform are also seen at the bottom of the image. A small part of the platform can also be seen on the top right.

3.2 Monitoring activity in one camera view

People density estimation : We are interested in estimating the number of people over time in the turnstiles view (Fig. 1-(a)). The human detector of section 2.1 is applied every second of the video (e.g. every 5 frames). A pre-processing step smooths the signal by averaging the number of people on a 3 minutes window. Figure 2 shows this smoothed signal, where each point is thus the average number of people on a 3 min window. All the following analyses are done on this smoothed signal. Using the (smoothed) data count for all mornings, an analysis is conducted to extract and visualize the global trend of activity, and see how the evolution over one morning fit into this average trend. We specifically distinguish between week days and week-ends, to see if week-ends stand out from week days. To do so, the average and standard deviation over week days are computed at each time instant, e.g. with $c_k(t)$ the smoothed number of counts for morning k at time t and $N = 10$ the number of mornings, we compute $m(t) = \frac{1}{N}\sum_{k=1}^{N} c_k(t)$ and $\sigma^2(t) = \frac{1}{N-1}\sum_{k=1}^{N}(c_k(t) - m(t))^2$.

Data analysis : Left of figure 3 shows the week average $m(t)$ (in green), as well as its fit (blue), and the fit +/- $2\sigma(t)$ (red). These curves characterize the usual usage of the station on week days, and the envelope of 2 standard deviations (red curves) indicates the area where the behavior can be considered as 'normal'. A simple analysis shows that the average usage is more or less what we could expect: traffic is low from 6 am to 7.30 am ; it then increases steadily until 9.30 am and slowly

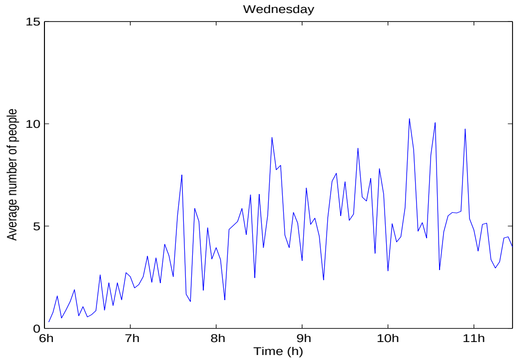


Figure 2: Average Number of people (on a 3 min window) over time on a Wednesday morning, in turnstiles view

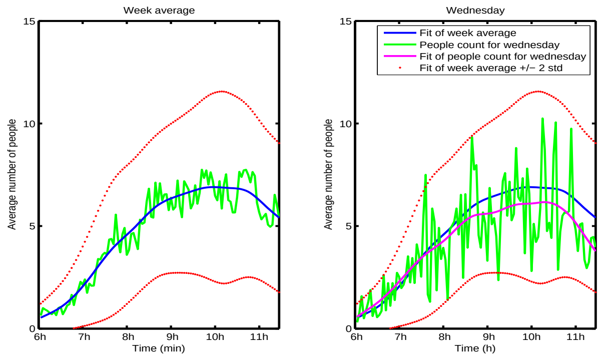


Figure 3: Left: average people density for week days, plus fit and standard deviation, Right: People density and fit on one Wednesday, compared to the week fit and std

decreases starting from 10.30 am. On the right of figure 3 is displayed an example of how a specific day (the one from Figure 2) fits into this week average, and how close its evolution (in magenta on the figure) is from the week evolution (blue and red curves are simply reported from figure on the left).

Week ends are analyzed in Figure 4, where the average of the 2 Saturdays and Sundays is plotted together with the week average. It allows to show the difference in evolution of the number of people on Saturday and Sunday mornings, and especially that it does not fit into the week average curves. Saturday appears to be a busy day; there is an early start and it also does not comply with the decrease around 10.30 am observed on week days. On the other hand Sunday is clearly less busy, and if the curve has approximately the same shape as the week days one, it is clearly shifted to the right, indicating that people are getting up late, as one might have expected.

Figure 5 shows an interesting example on a week day (Thursday). While the global trend follows the week average, an unusual event stands out well of the 2 standard deviation envelope. This peak is due to a group of obviously lost tourists, which is staying still or wandering about, for quite a long time, moving in and out of the camera view.

3.3 Monitoring train traffic

In this section, we present the evaluation and validation of the platform time-occupancy measure presented in Sec. 2.2. The accuracy of the train stop estimation was measured using annotated data (ground truth), performed manually on the whole 14 sequences.

Figure 6 firstly shows the distribution of the train occupancy-time in the station, for both ground

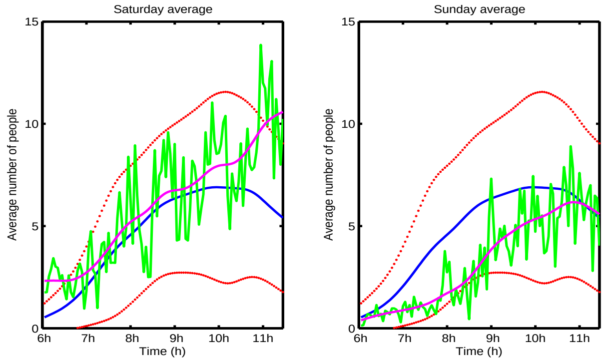


Figure 4: Saturday and Sunday averages (magenta), compared to the week average (blue).

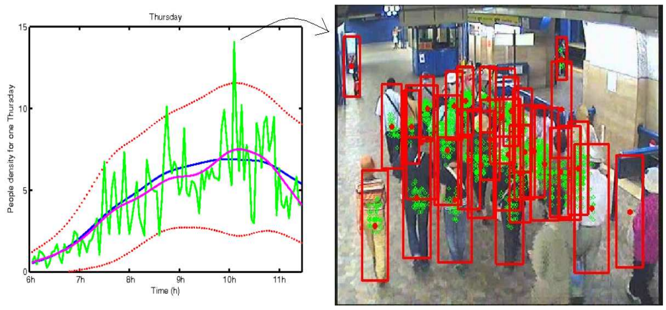


Figure 5: People density on one Thursday, compared to the week average. An unusual event can be detected around 10 am, corresponding to the frame on the left.

truth (GT) and processed data. It shows that the time-occupancy of the platform can have very different durations; while the bulk of the detections are located around 30-50s (average value of 45s), a non-negligeable number of stops are far below (10s) or far above (up to 3 min, even 6 min).

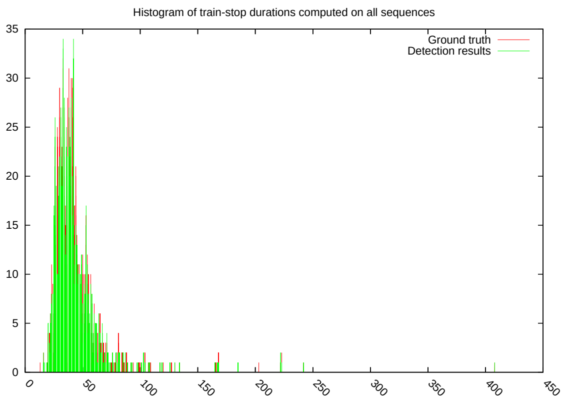


Figure 6: Histogram of the time spent by the train in the station.

In the transportation context, such outlying values can be explained by several phenomena; regulation purposes, safety checks, incidents on platform, signal failures or break-down of systems... In our case, most of these outliers come from regulation purposes and “incidents” on the platform as highlighted in Fig. 7.

Taking into account this phenomena, the detection statistics (e.g. the standard deviation of the

(a) 2007/06/10 - 100s
"quad intrusion".

(b) 2007/06/05 - 408s
"regulation purpose".

Figure 7: Screenshots corresponding to trainstop outlying values.

stops' durations) have to be analysed carefully. Tab. 1 shows the detection rates obtained during the evaluation process. As highlighted in the table, the average detection rate on the overall sequences is around 96.59% (710 detection on 735 stops in the GT), which demonstrates the effectiveness and robusteness of the proposed approach.

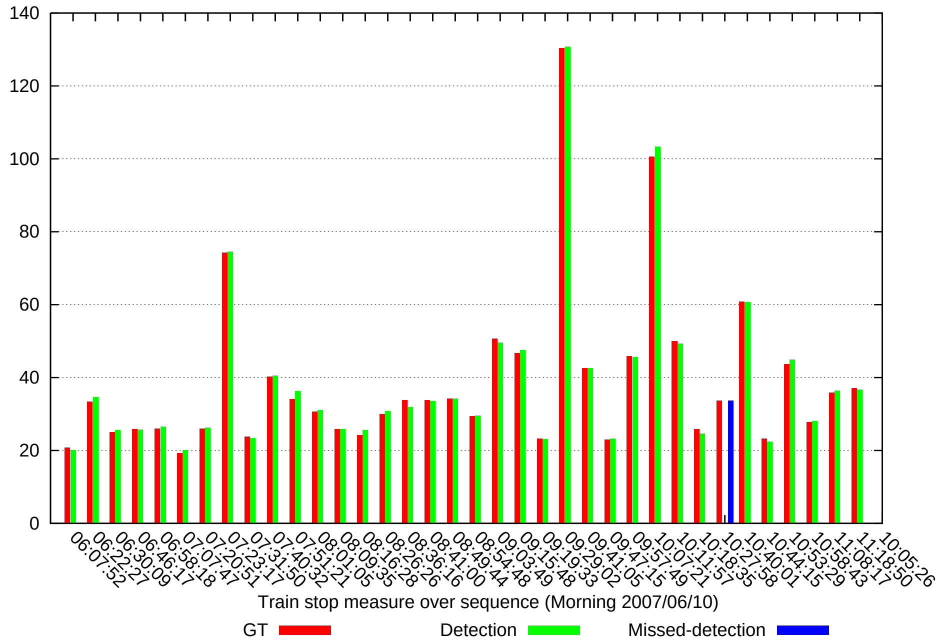
Table 1: Platform time-occupancy by metros: detection results

Day	07.06.02 (Sat)	07.06.03 (Sun)	07.06.04 (Mon)	07.06.05 (Tue)	07.06.06 (Wed)	07.0
Detection rate (nb)	100% (35/35)	97.36% (37/38)	98.38% (61/62)	96.77% (60/62)	94.44% (51/54)	
Detection rate (time)	100%	95.90%	98.20%	97.89%	97.05%	
Average stop duration (std)						
- detection -	37.14 (10.53)	42.77 (35.73)	44.41 (16.86)	53.98 (52.92)	46.63 (15.87)	45.
- ground truth -	37.21 (10.45)	43.21 (35.40)	44.34 (16.84)	53.22 (52.21)	46.08 (15.98)	45.
Mean arrival delay (s)	0.77	0.63	0.58	0.83	0.63	
Mean departure delay (s)	0.69	0.84	0.72	0.97	0.77	

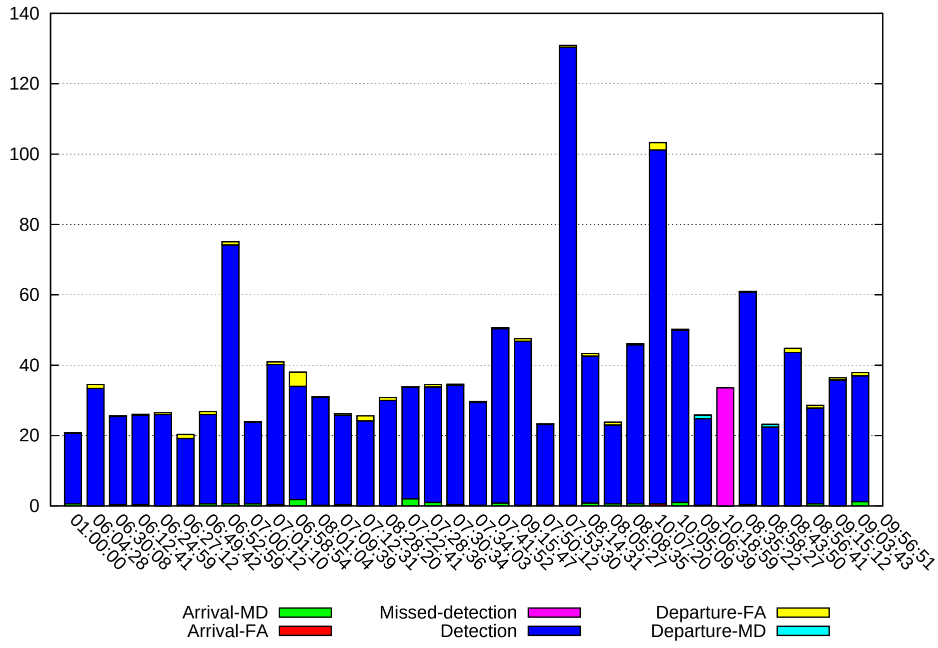
Day	07.06.09 (Sat)	07.06.10 (Sun)	07.06.11 (Mon)	07.06.12 (Tue)	07.06.13 (Wed)	07.0
Detection rate (nb)	98.03% (50/51)	97.22% (35/36)	98.18% (54/55)	98.18% (54/55)	94% (47/50)	
Detection rate (time)	97.86%	97.58%	97.79%	98.16%	95.05%	
Average stop duration (std)						
- detection -	35.13 (15.08)	38.96 (22.73)	48.37 (32.60)	48.95 (17.59)	51.04 (30.45)	42.
- ground truth -	35.09 (14.91)	38.62 (22.22)	48.54 (32.41)	48.79 (17.51)	50.35 (29.82)	41.
Mean arrival delay (s)	0.72	0.41	0.71	0.68	0.98	
Mean departure delay (s)	0.82	0.60	0.73	0.85	1.11	

Regarding the detection delays at train arrival/departure, most of them are between 0.5s and 1s which seems to be sufficient from the end-user point-of-view. Furthermore, as highlighted in Fig. 8-(b), such delays are quite negligible compared to the stop-duration itself.

From the metro operation point of view, Tab. 1 already allows to identify different behaviors depending on the days; e.g. week-days seems to have less metros and lower average detection than week-end days. So as to confirm this analysis, Fig. 9 presents cumulative plots of platform occupation



(a) Clustered histogram (GT, detection & missed-detection).



(b) Stacked histogram (arrival/departure detection delays).

Figure 8: Detection results obtained on sequence "Morning 2007/06/10".

by metros; such graph allows to reflect both trainstop duration and frequency within a single plot. It clearly exibits the fact that Saturdays and Sundays have lower platform occupation compared to the other days of the week. This conclusion is different from the one obtain in section 3.2 for Saturdays,

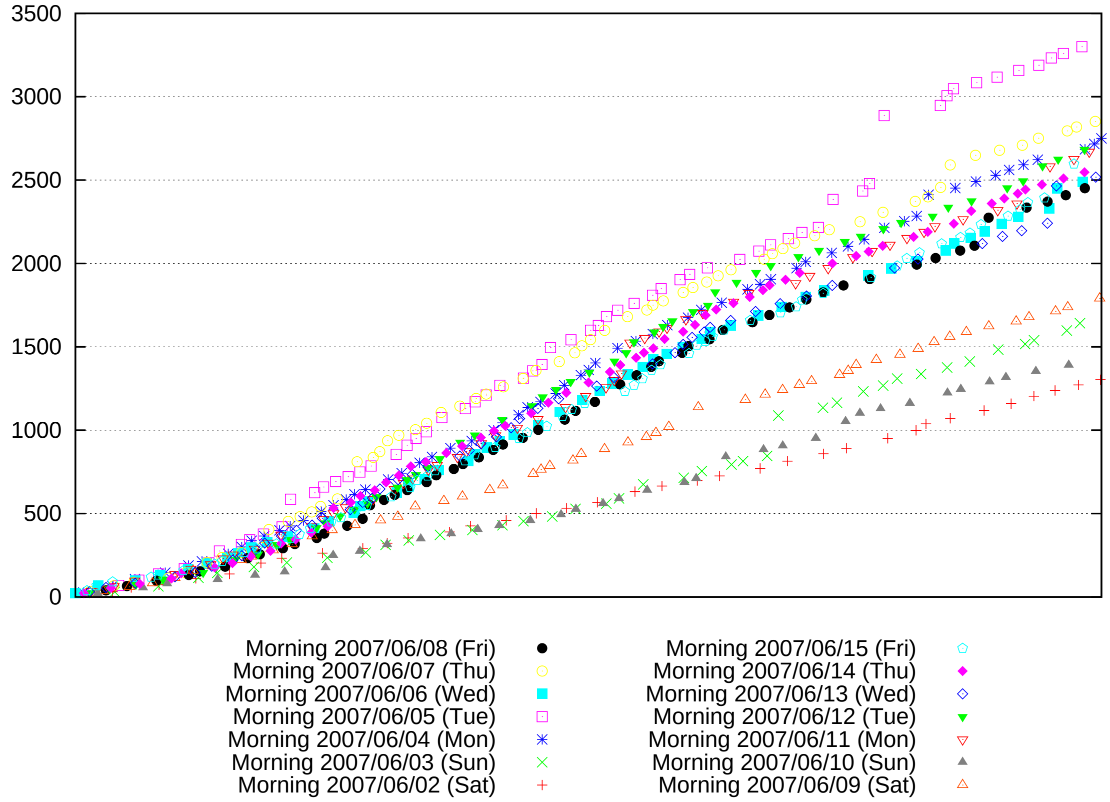


Figure 9: Plot of cumulative platform occupancy.

where we observed much more people in the turnstiles view. The next section investigates the joint usage of the two views, especially addressing this issue.

4 Multi view analysis

In this section, we investigate the joint usage of information coming from the two camera views, to provide a more in-depth analysis and solve possible ambiguities.

4.1 Understanding peaks of activities

We are looking at the relation between the number of people in the turnstiles view of Fig. 1-(a), and the train arrivals/departures, using results from sections 3.3 and 3.2. Fig. 10 shows an example of 9 train arrivals and departures over one hour, with the corresponding number of people in the turnstiles view. In this case, the signal is not averaged, i.e. the values displayed are the actual number of people at each second. It shows quite clearly that the peaks of the number of people in the turnstiles view correspond to instants where the train is present in the station. These peaks are due to people leaving the train, and passing in front of the turnstiles camera view, towards the exit. On average, it was computed that there are 7.2 persons in the turnstiles view when a train is present, whereas there are only 4.5 persons in average in the other case.

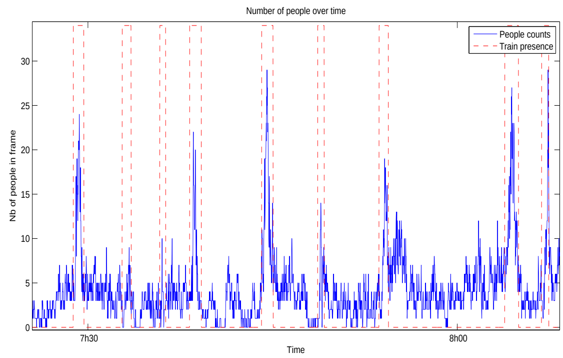


Figure 10: Detail of 9 train arrivals, plotted on the corresponding counts of people.

4.2 Understanding flows of persons

Another useful way of using the train presence information is to make the same analysis as in Fig. 3, with filtering out the instants where the train is present. More precisely, the number of people in the time slot where the train is present is replaced by the average in the period where no train is in the platform (e.g. when no flow of people is coming out of the platform). This filters out from the signal the peaks due to people coming out from the train. The obtained signal thus characterizes the usual activity in the turnstiles view, excluding the arrivals. Fig. 11 shows this filtered signal and its fit (in

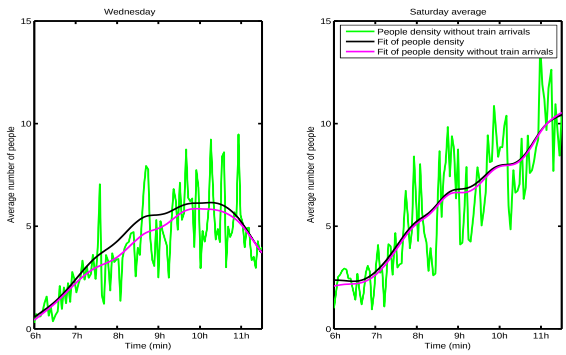


Figure 11: People density and its fit, with filtering out train arrivals. Comparison with the fit of non-filtered signal (Wednesday and Saturday)

magenta), compared with the fit of the non-filtered signal (in black). From this figure, it seems that the high variability in the number of people in the turnstiles view is not only due to train arrivals. This means that independently from train arrivals, the activity in the turnstiles view is itself very irregular. This is confirmed by the computation of the average standard deviation, which is 3.1 for the non-filtered people density signal and 2.6 for the filtered one, which is still quite high.

Another important remark is that the difference between filtered and non-filtered signals is quite noticeable for week days (one Wednesday is shown on the left of figure 11) whereas it is barely noticeable for Saturdays and Sundays (Saturdays average is shown on the right of figure 11). This means that far less people are coming out of the trains on week-ends, and thus the traffic in the turnstiles view is mainly due to people passing by or going to the platform. This is confirmed by Fig. 12, which shows the average number of people per morning, on train presence time (blue) and

Table 2: Correlation between train staying duration and cumulative number of people in turnstiles view. Critical value at 5% risk

Days	Mon	Tue	Med	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Correlation	0.75	0.43	0.77	0.40	0.60	0.44	0.31	0.58	0.44	0.76	0.20	0.57	0.51	0.31
Critical value	0.25	0.25	0.28	0.26	0.27	0.33	0.27	0.27	0.25	0.27	0.25	0.25	0.32	0.32

on train absence (green). The difference between the 2 curves is clearly larger on week days (3.1 in average) than on week-ends (1.5 in average). This clearly shows a different behaviour of users on week-ends.

In particular, the apparent contradictory results for Saturdays (more people in turnstiles view but less trains, and feedbacks from Rome operatives that Saturdays are less busy overall) are explained: the high people density is not related to train exits, but is due to people waiting or entering the platform. Two hypothesis are proposed to explain this high density. Feedbacks from the video hinted a difference of behaviour between week days users (mostly people going to work) and week-ends users (occasional travellers), which tend to stay longer in the camera view, waiting or hesitating on the way to go. The second one is that people takes different routes, and flows in this camera view are thus different. This behaviour could not have been spotted by the monitoring of the turnstiles view alone. This analysis is a very useful information that highlights a difference of behaviour and station usage on specific days.

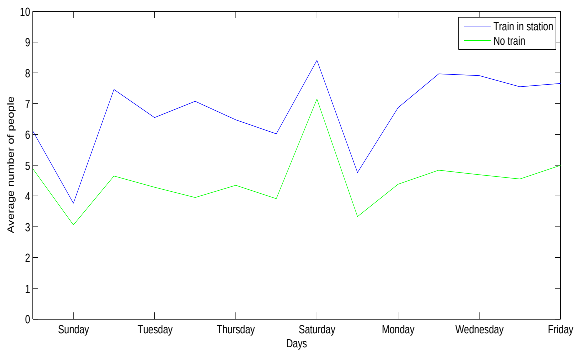


Figure 12: Average number of people per morning, train present or not

4.3 Analysing train staying duration and platform crowding

One of the interest of the metro operatives is to see whether too much people on the platform delay the departure of the train. We thus try in this section to relate the estimated number of people on the platform with the train staying duration. We are measuring the number of people in the turnstiles view (excluding instants where people leave, i.e. train present in station), which is just one entry point to the platform, the goal is thus more on catching the trend than on the exact estimation of the number of people.

To verify this hypothesis, the correlation between the train staying duration and the cumulative number of people in the turnstiles view (excluding train arrivals) is computed. The empirical correlaton coefficient is given by

$$r = \frac{1}{N-1}\sum_{k=1}^{N}\frac{(x_k-\bar{x})(y_k-\bar{y})}{\sigma_x\sigma_y}$$

where N is the number of samples, and σ are the empirical standard deviations. Outliers, e.g. trains that stay more than 100s (2 times the maximum standard deviation of stop durations) are discarded. Results in Table 2 show that most of the days exhibit a correlation, which is in most of the cases above the critical value[2], under which we cannot reject the null hypothesis (e.g. the variable are not correlated). Exceptions are Sundays, where the correlation does not seem to be significant. Fig. 13

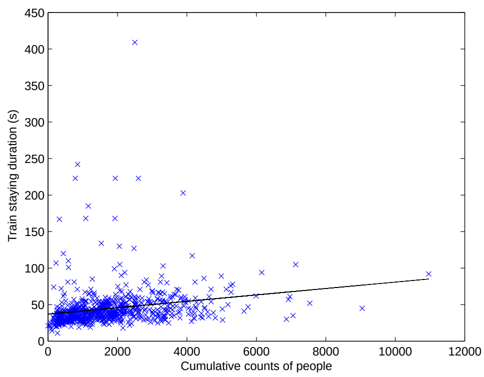


Figure 13: Relation between cumulative number of people in turnstiles view, and time spent by the train in station

shows the relation between the average number of people during waiting time (time between a train leaves and another arrives) and the time that the next train stays in station. It also confirms the results of Table 2, that there seems to be a link between the two variables, but which cannot explain the outliers. The platform crowding level can thus be an explanation for some delay and the train irregularity, but cannot explain the observed variations.

5 User evaluation

CCTV has been traditionally installed in underground environment to support the operation management with a view to enhance safety and security. Nowadays modern advance video processing technologies are opening a new role for CCTV. As Caretaker project is demonstrating, new algorithms can provide valuable information to underground and public transport operations and security managers. The possibility to automatically process on-line and massive recorded data, within the respect of privacy regulation, can indeed help the security staff which can not have possibility to process such data. The algorithms described in this paper fulfill such requirements. The estimation of number of people in a camera view can help to alert the operator when anomalies may occur e.g. when a station reach its capacity limit that request the limitation of the number of people entering the station (in Rome this alert can be risen in some station during large events). As a long-term analysis along with ticketing data fusion, it can help to provide a clear view of the trend of how the station (or its specific part) is used. The algorithm that analyse the platform time occupancy by the train can support the operations (that already have the signaling system). It can support the analysis of anomalies, with a direct reference to the related video data, to improve in the longer term the performances and regularity of the service. The most interesting thing, as shown in the previous paragraph is to combine the analysis of the two algorithms. The in depth analysis of such results could also lead to a redesign of the level of service provided by the underground, taking into also into account the transportation demand. This is clearly the first attempt to analyze such information, which is very new in the transportation world and need some time to be integrated in the usual evaluation. Furthermore the estimation of number of people in a camera view could provide more useful information if the analysis is carried out in more cameras and potentially comparing different stations.

[2]Critical values are extracted for a 5% risk, from tables in [8]

Also the analysis of the platform time occupancy by the train can provide more valuable information if extended on several station, to understand the domino effect of potential delays.

6 Conclusion

A person detector and a train arrivals/departures detector have been used on a large amount of real video data of the Rome underground. It has been shown that a statistical analysis of the results provides interesting information regarding station usage, allowing to characterize the usual behaviour of train users, distinguish trends between days of the week, and spot some unusual events. It also been shown that the joint usage of the information coming from two views can solve possible ambiguities and can provide a better understanding of the station's usage. This joint analysis can be extended to several views and to the size of a undergound network, to better understand the relationships between stations and their usage.

The exploitation of the potential of such technology, along with the comprehension and the analysis of results, could lead to a new way to exploit CCTV resources in the underground, to enhance both safety and security.

7 Acknowledegment

The work presented here is partially supported by the European Commission under the 6th Framework Program through the IST FP6-027231 CARETAKER project. For further information about the CARETAKER project, please visit `http://www.ist-caretaker.org/`.

References

[1] C. Andrea, "Event detection in underground stations using multiple heterogeneous surveillance cameras," in *Advances in visual computing: (First international symposium, ISVC 2005)*, vol. 3804, Lake Tahoe, NV, USA, December 5-7 2005, pp. 535–542.

[2] S. Oh, S. Park, and C. Lee, "A platform surveillance monitoring system using image processing for passenger safety in railway station," in *Control, Automation and Systems, 2007. ICCAS '07. International Conference on*, Seoul, 17-20 Oct. 2007, pp. 394–398.

[3] ——, "Vision based platform monitoring system for railway station safety," in *Telecommunications, 2007. ITST '07. 7th International Conference on ITS*, Sophia Antipolis, 6-8 June 2007, pp. 1–5.

[4] B. Lo and S. Velastin, "Automatic congestion detection system for underground platforms," in *Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on*, Hong Kong, China, 02-04 May 2001, pp. 158–161.

[5] W. Zajdel, J. Krijnders, T. Andringa, and D. Gavrila, "Cassandra: audio-video sensor fusion for aggression detection," in *Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on*, vol. 3804, London, UK, 5-7 Sept. 2007, pp. 200–205.

[6] O. Tuzel, F. Porikli, and P. Meer, "Human detection via classification on riemannian manifolds," in *CVPR*, 2007.

[7] J. Yao and J.-M. Odobez, "Fast human detection from videos using covariance features," IDIAP Research Institute, Tech. Rep. 07-68, 2007.

[8] G. Saporta, *Probabilités, analyse des données et Statistique.* Technip, 2006.