Bayesian Multi-camera Surveillance

Vera Kettnaker Ramin Zabih
Computer Science Department
Cornell University
{kettnake, rdz}@cs.cornell.edu

Abstract

The task of multi-camera surveillance is to reconstruct the paths taken by all moving objects that are temporarily visible from multiple non-overlapping cameras. We present a Bayesian formalization of this task, where the optimal solution is the set of object paths with the highest posterior probability given the observed data. We show how to efficiently approximate the maximum a posteriori solution by linear programming, and present initial experimental results.

1 Multi-camera surveillance

Video surveillance in a large or complex environment requires the use of multiple cameras. In this paper we address a particular task that we call *multi-camera surveillance* (MCS). The multi-camera surveillance task arises in an environment with moving objects that is monitored by multiple non-overlapping cameras, such as an office building with pedestrians, or a set of highways. The task is to reconstruct the paths taken by all objects that were visible during the observation period, despite the fact that a moving object can be temporarily out of view of any camera.

We will assume that objects moving through the monitored environment are likely to pass several cameras, and that their movement is constrained to follow certain paths. We are given the topology of these allowable paths as input, together with information about transition probabilities and transition times. We assume that these transition models are supplied as part of the input, although it would be easy to estimate them as part of the surveillance system.

We first run a motion detection and tracking algorithm on each video stream. The tracking algorithm returns one *observation interval* for each passer-by, i.e. the collection of all per-frame views of that person, annotated by the time interval and the camera location at which the observation was made.

The solution of the MCS task will consist of a set of links between observation intervals, where each link connects two successive appearances of the same object. We probabilistically model the fact that observation intervals of the same person should look similar and that all transitions and transition times should be plausible. Besides ensuring that each chain is plausible on its own, we also model how likely the hypothesized number of chains is with respect to the environment's traffic statistics. An important global constraint stems from the fact that an object can only be in one location at a time: if the motion segmentation algorithm works correctly and the cameras have non-overlapping view fields, then the links of a correct solution will form non-overlapping chains. This reduces the number of possible hypotheses considerably.

It seems difficult to directly determine which set of mutually-exclusive chains is a posteriori most likely. Yet under a few independence assumptions, we can transform a scaled version of the posterior probability in such a way that its maximum can be found by solving a linear program. Moreover, the linear program formulation of the problem naturally encodes the global constraint that the chains should not overlap. This approach is a modification and extension of Poore's linear program formulation of probabilistic data association tasks in radar tracking [12].

We will begin by presenting a Bayesian formalization of the problem of reconstructing the set of object paths given the data. Section 3 will explain how to transform the maximum a posteriori (MAP) estimation problem into a linear program. After describing some related work in section 4, we present experimental results from a system with four cameras that monitors a research lab. Section 6 describes extensions of the system that relax some current assumptions.

2 Bayesian formalization

An individual observation interval contains two different types of information. It tells us that something moved through the monitored area of a specific camera at a specific time, and it contains information about the visual appearance of the observed object during a short period of time. Since the motion segmentation algorithm may make mistakes even in determining how many objects were visible at each time, a full hypothesis has to state first where and when how many objects passed through monitored areas, and where and how many passing incidents were detected by the motion segmentation algorithm (the *incident structure*). Secondly, the hypothesis has to state which observations are successive appearances of the same object (the links).The hypothesis $\Omega = \Omega_{is} \cap \Omega_{li}$ is therefore composed of the incident structure hypothesis Ω_{is} and the link hypothesis Ω_{li}. Similarly, the observation $O = O_{is} \cap O_{app}$ consists of the observed incident structure O_{is} as well as the observed visual appearance of objects O_{app}. Bayes formula yields

$$P(\Omega|O) = \frac{P(O_{app}|\Omega, O_{is})P(O_{is}|\Omega)P(\Omega_{li}|\Omega_{is})P(\Omega_{is})}{P(O)}$$

In order to simplify the hypothesis space, we will assume that $P(O_{is}|\Omega)$ vanishes unless $O_{is} = \Omega_{is}$. This assumption means that the motion segmentation is correct in terms of the times and locations of passing incidents. We will write P' to refer to probabilities that are implicitly dependent on the incident structure Ω_{is} or the observed incident structure O_{is}.

Each object moving through the environment passes several cameras, causing a chain of 'passing incidents' that are recorded on video. We will refer to these real-life incidents mostly in terms of their position in hypothesized chains, and will write $C_i = (c_{i,1}, \ldots, c_{i,l(i)})$ to denote the hypothesis that the incidents $c_{i,1}, \ldots, c_{i,l(i)}$ were performed by the same object and form the ith chain. We will write $o_{i,j}$ to refer to the per-frame views of the observation interval associated with incident $c_{i,j}$, and we will write O_i to refer to all visual appearance data of chain C_i.

The prior is composed of three main terms:[1] the probability $P(trans(c_{i,j}, c_{i,j+1}))$ of the time length and locations of each transition, the probability $P(len(C_i))$ of a certain chain length, and the probability $P(new(\Omega_{li}, l))$ of the hypothesized frequency with which new people enter the environment at location l, which regulates the overall number of chains.

$$P'(\Omega_{li}) = \prod_{chain\ i} \prod_{link\ j} P(trans(c_{i,j}, c_{i,j+1})) \cdot \prod_{chain\ i} P(len(C_i)) \cdot \prod_{loc\ l} P(new(\Omega_{li}, l))$$

We model transitions between locations as Markov, but allow arbitrary transition time densities, so that the movement of a single object is modeled as a semi-Markov process [7].[2] Such a model can be graphically represented as a stochastic state automaton, i.e. a directed graph, where the nodes correspond to camera locations. The links represent possible transitions between connected camera locations and are annotated by a model of the transition time and the probability that an object visible in the first location will become visible next in the second camera location.

3 Transformation of the MAP estimation problem into a Linear Program

Since it is unclear how to maximize the posterior directly, we maximize instead the ratio of the posterior over the posterior of a reference hypothesis Ω^0, which states that all passing incidents were caused by different objects.

$$\frac{P(\Omega|O)}{P(\Omega^0|O)} = \frac{P'(O_{app}|\Omega_{li})P'(\Omega_{li})}{P'(O_{app}|\Omega^0_{li})P'(\Omega^0_{li})}$$

This ratio will be decomposed first into terms that refer to one chain each, and then further into terms for each chain link. These chain link terms will serve as coefficients of a linear program, whose solution will maximize the posterior probability while obeying the constraint that the chains be mutually exclusive.

Decomposition. We can assume that an object's visual appearance does not depend upon any other object, thus the likelihood decomposes into chains

$$P'(O_{app}|\Omega_{li}) = \prod_{chain\ i} P(O_i|C_i).$$

Most factors of the prior already refer to one chain each, but the number of new chains is a property inherent to the hypothesis as a whole. However, if the frequency of new objects is modeled as a Poisson process, the ratio of the time-recursive formulations of $P(\Omega|O)$ and $P(\Omega_0|O)$ can be shown to decompose into chain terms as follows:

$$\frac{P(\Omega|O)}{P(\Omega^0|O)} = \prod_{chain\ i=1}^{n} \frac{P(O_i|C_i)P(trans(C_i))P(len(C_i))\lambda_{loc(c_{i,1})}}{\prod_{j=1}^{l(i)} P(o_{i,j}|c_{i,j})P(len(c_{i,j}))\lambda_{loc(c_{i,j})}}$$

where j ranges over the observation intervals of the ith chain, and $\lambda_{loc(i,j)}$ is the mean of the per-frame Poisson probability density function for new appearances at the location of incident $c_{i,j}$. The proof has to

[1]The actual prior we use models more data aspects. They have been ommitted here to simplify the expressions.

[2]In section 6, we will explain how to implement a higher-order Markov model.

be omitted due to space constraints, but is an adaptation of Poore's proof [12] to transition graphs instead of motion in euclidean space, to more general distributions for the transition times, and to a reference hypothesis appropriate for the MCS task.

The decomposition into chains would lead to a linear program with as many variables as there are different chain hypotheses. It is however possible to reduce the complexity of the MCS task much further by decomposing the chain terms into per-link terms. In radar tracking, a decomposition into per-link terms is inappropriate because pairwise closeness of radar signals in successive frames does not capture the notion of a trajectory. In the MCS task however, the observation intervals are rich in appearance information and permit us to assess whether two observation intervals show successive appearances of the same object.

The decomposition into per-link terms makes two modeling assumptions and uses an approximation of the likelihood. It requires Markov transition probabilities and a decomposable model of chain length such as the geometric density function. If we denote by $o_1, o_2, \ldots o_z$ all (per frame) observations in a chain, then the likelihood that all these observations stem from the same object can be computed by $\prod_{h=1}^{z} P(o_h|o_1, \ldots o_{h-1})$, where $P(o_h|o_1, \ldots o_{h-1})$ computes the probability that a sample observation o_h is from the same distribution as all the previous views. We approximate this by comparing each view with only a small number of recent observations:

$$P(o_{i,j,k}|o_{i,1}, \ldots, o_{i,j-1}, o_{i,j,1}, \ldots, o_{i,j,k-1}) \approx \begin{cases} P(o_{i,1,1}) & \text{if } j = 1 \ \&\ k = 1 \\ P(o_{i,j,k}|o_{i,j,1}, \ldots o_{i,j,k-1}) & \text{if } k \neq 1 \\ P(o_{i,j,1}|o_{i,j-1}) & \text{if } k = 1 \ \&\ j > 1 \end{cases}$$

where i, j and k index chains, links, and frames, respectively, and $P(o_{i,1,1})$ is a non-informative prior over the observation space. The last case in the above expression handles the first observation in an observaton interval that is not the first incident in a chain: this first observation is matched against a model of appearance estimated from the whole previous observation interval. With this independence assumption, many terms that are common to both hypotheses cancel out, yielding

$$\frac{P(\Omega_{li}|O_{app})}{P(\Omega_{li}^0|O_{app})} \approx \prod_i \prod_{j=2}^{l(i)} \frac{P(o_{i,j,1}|o_{i,j-1}) P(trans(c_{i,j-1}, c_{i,j}))(1 - P_\chi)}{P(o_{i,j,1}) \cdot P_\chi \cdot \lambda_{loc(i,j)}}, \quad (1)$$

where P_χ (exit probability) is the parameter of the geometrically distributed chain length probability.

Transformation into a Linear Program. Above, we transformed the ratio of posteriors into a product of terms. Each of the terms refers only to a hypothesized transition (link) or its two endpoints. The best solution will be that set of links which maximizes the corresponding product of link terms. By taking the negated logarithm, the maximization of the product turns into the minimization of a sum. This makes it possible to express the maximization of the posterior under the constraint of mutual exclusivity of the chains as a linear program. More specifically, it becomes a weighted assignment problem for which very efficient algorithms exist, for example the Munkres algorithm [4] that we currently use to compute a solution. The input to the Munkres algorithm is a matrix whose elements are the negated logarithm of product terms of expression (1), one element for every possible link between two incidents and between each incident and the virtual incident 'NEW'.

Focus sets. The size of this matrix is proportional to the square of the number of observations. However, only a small fraction of the matrix elements have to be actually computed because most links can never be part of the optimal solution. These are all links between observation intervals o_A and o_B for which the hypothesis of a link between them is a priori less likely than that of the hypothesis that o_A is a new object. More precisely, these are those links for which

$$\frac{P^+ \cdot P(trans(o_B, o_A)) \cdot (1 - P_\chi)}{P(o_{A,1}) \cdot P_\chi \cdot \lambda_{loc(A)}} \quad (2)$$

is smaller than 1. Here,

$$P^+ = \max_{o \in M}(\max(P(\cdot|o)))$$

is the upper bound on the visual match probability of any possible observation matched with any of the previously seen observations $o \in M$. Since we use parametric distributions for the visual match probability, the inner maximum can be determined analytically for each of the observations already in the modelbase M, and P^+ is updated when a new observation is made.

The remaining terms in expression (2) only depend on the locations of the two observations and their relative temporal distance. In particular, the transition probability is composed of a spatial transition probability and a probability of transition times. This means that if searching for plausible previous occurrences of the person in incident o_A, we only have to consider those previous incidents whose ending times fall into the time window[3] of those ending times that make expression 2 larger than or equal to 1.

We call the set of all match candidates that pass this criterion the *focus set* of a new observation interval, because the subsequent matching process can focus on these candidates only without loss of correctness. If one stores the previous monitoring incidents ordered by location and ending times, it suffices to compute one time window per possible preceding location. This time window can then be used to prune the match candidates that are necessarily less plausible than a new object entering the scene.

Online processing. The algorithm as described above uses batch processing, which is unreasonable if the system is used for continuous monitoring. However, we can prove that there is no online algorithm that returns the same answer as the batch algorithm for all inputs. More specifically, for all k it is possible to construct a matrix that could have arisen from a tracking situation, and for which it holds that none of the assignments of the optimal solution for the submatrix containing the first $k-1$ monitoring incidents is part of the optimal solution for the matrix containing the first k monitoring incidents.

Yet inputs with very long-term effects seem infrequent, and simulations suggest that approximate solutions with very few wrong links can be obtained by the following modification of the algorithm: In order to assign a new monitoring incident A to its most likely previous occurrence (or NEW), we consider the submatrix containing A and all those monitoring incidents recorded after A whose focus set contains at least one element of A's focus set. Note that by an argument similar to that of the focus set time windows, one can determine the time one has to wait for 'contesting' incidents of A. Once the submatrix is complete, the optimal assignment for A is computed and the assigned incident marked as taken. Then an analogous submatrix is constructed for the next monitoring incident.

This means that each assignment considers a certain lookahead so as to preclude the possibility that a premature assignment drastically limits the choice of reasonable matches for future monitoring incidents. Possible conflicts with past monitoring incidents are handled because their assignments have already considered the conflict and made the assignment accordingly. This online algorithm can be made arbitrarily correct by including not only the set of possible contestants into the submatrix, but also the set of contestants of the contestants, and so on. The novelty of this online algorithm does not lie in its use of time windows, but in the dynamic choice of the time windows such as to include all direct contestants, secondary contestants and so on by means of the focus sets.

[3]The computation of the time windows requires the inversion of probability pdfs, which can approximated very fast by a table lookup. We only need one table because we compute (transition dependent) walking time probabilities from (transition independent) walking speed probabilities

4 Related work

Cox [5] appears to be the first to use probabilistic formalizations of the radar tracking community for computer vision tasks. However, he did not exploit the visual characteristics of observations (i.e., tracking features) but only used their incident structure and left the probabilistic formalization of the radar task unchanged. Huttenlocher [9] devised a tracker that could lock back onto tracking targets after they went temporarily out of the field of view. However, his visual matching method assumes smaller changes in appearance than we do (one camera vs. multiple cameras with different viewing angles). He also does not impose a prior on matchings between observations, because he assumes an environment without a spatio-temporal structure such as the one imposed by the corridors. Exploiting such structure will however allow our system to scale. Berkeley's traffic monitoring system [10] tracks cars and performs occlusion reasoning for a single video stream. The occlusion reasoning method could not be extended to handle disappearances of cars between multiple cameras.

Recently, a number of multi-camera monitoring systems have appeared in the literature. Olson and Brill [11] built an indoor monitoring system that creates a graph representing the per-frame movement and interaction of objects in a single video stream. Although their system architecture assumes multiple cameras, no analysis across cameras is performed. Boyd *et al.* [2] presented an architecture designed for multiple sensors observing a dynamically changing environment. However, their cameras overlap and the view fields are transformed into one contiguous view field. The system is designed to perform tasks that involve techniques with project-update cycles, such as Kalman tracking or HMMs. However, although their architecture is quite general, it is difficult to apply it to tasks such as ours where observed objects are invisible for extended periods of time. Grimson *et al.* [6] have built another multi-camera system that assumes overlapping camera fields: they envision observing activities by a set of cameras that are scattered in an environment and that determine automatically how to map their local view fields into one global view field. They then learn classes of observed behavior.

Huang and Russell's system [8] performs a task sim-

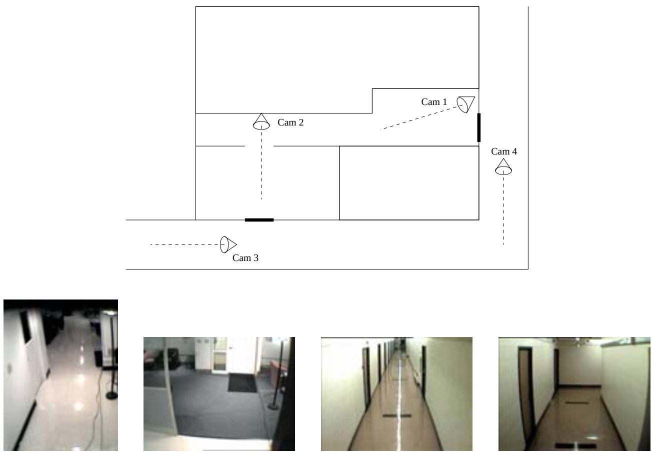

Figure 1: Floor plan of the camera setup and background snapshots from the 4 cameras.

ilar to ours: they monitor a highway at two consecutive locations and try to find matching cars. They concentrate on appearance constraints, but also transform their problem into a weighted assignment problem. They start from different premises than our derivation which leads to different link weights and to a different structure of the weighted assignment problem. Their solution is confined to setups where cameras are placed alongside a single path so that the movement of the objects is deterministic, with the exception of objects entering and exiting the environment. Our solution is much more general by allowing arbitrary corridor systems in which moving objects can choose paths. Therefore, our system is able to reconstruct the paths of all objects through an environment, which is interesting for some tasks. For example, traffic planners might want to optimize traffic light controls such that traffic flow is least interrupted for the most popular routes through a city. Huang and Russell also describe a heuristic online algorithm that trades off matching confidence with solution coverage. Unlike the online algorithm described above, they do not use a temporal lookahead, which could cause their algorithm to make premature decisions.

5 Experimental results

In order to evaluate the system, we set up a small surveillance system of 4 cameras in and around a research lab. The floor plan is depicted in figure 1, together with background snapshots from the 4 cameras. Our data contained strong reflections and shadows of the pedestrians on the corridor floors. In order to eliminate most of the background from the segmented pedestrians, we employ both a background subtraction scheme and a recent dense motion algorithm that maximizes the area of coherently moving, similar pixels [3]. The latter tends to group background pixels with faint reflections with the rest of the non-moving background. We find all coherent patches of moving pixels that surpass a certain size and track them as long as they are visible by employing a projection scheme similar to [1].

The collection of image regions corresponding to such a track is then mapped into a coarse partition of the HSV color space. We empirically designed this colorspace to distinguish between popular clothing colors such as beige, offwhite, or denim, while being coarse enough to be robust to lighting changes due to shadows. We then count for each bin how large an area of the tracked object is covered by this color and cluster the count vectors of each observation interval. The counts in each color bin across a cluster are modeled as a poisson distributed variable. This very simple scheme results in relatively robust probabilities of visual similarity.

Instead of modeling walking times for each transition, we use a single frame-quantized gamma pdf to model walking speeds. This reduces walking time model construction to measuring the distances between camera view fields. Penalty distances had to be added for transitions that involve opening of regular doors (exiting the lab) and doors with a security card lock (entering the lab).

We conducted an experiment of about 8 minutes, where two subjects walked separately and together as many paths through the system as they could think of, always changing clothes in between different paths so as to impersonate different people. Since the experiment was conducted on a summer morning, only three additional people walked through our setup. The experiment resulted in a total of 28 observation intervals from 14 true tracks. We count the tracks that two people walked together as one track because the basic tracker consistently merged the two people together and therefore also into one observation interval. The next version of the system will include a more sophisticated motion segmentation algorithm to reduce the frequency of such merges.

Figure (2) shows observation intervals and the correct observation links from a subsequence of the experiment. Overall, 28 links had to be estimated, because the system determines for each incident either a preceding incident or links the incident to 'NEW'. Our initial results are quite promising: only two out of the 28 incidents were assigned to an incorrect predecessor. In both cases, the transition times of the suggested links were likely, and the clothing of the correct and wrong matches had similar color and differed only in the pants' length.

However, the data also contains two cases in which the same person appears again after an unnaturally long disappearance time, but is not recognized as previously seen by the system: in the first case, a person unrelated to the experiment crossed the hallway and disappeared into a room from which he reappeared after a few minutes to cross the hallway again. Neither the crossing behavior nor the disappearance in rooms is modeled in our current system, and therefore the system labeled both appearances of this person as 'NEW'.

The other case of a long disappearance time was constructed deliberately: one of the subjects paused on a very short stretch of hallway for several seconds so as to simulate a pedestrian that would stop to chat with another pedestrian (which violates the modeling assumption that the person would just walk through the hallway). In this case, the system also labeled the second appearance as 'NEW'.

It would be interesting to extend the system in a way that would detect such special cases from the fact that such exit/new events would occur for two or more pedestrians at the same time, namely for the people who talk to each other. For this experiment, the batch version and the online version with a lookahead that includes only the direct contestants yield the same solution.

These first results were obtained in difficult lighting situations and with a very weak representation of visual appearance, as well as significant segmentation errors.[4] But they nonetheless suggest that our approach performs well. Our focus sets led to reasonable time windows and ensured that each observation only had to be compared with a very limited number of other observations. The average size of the focus sets in this experiment was 1.6, while without the focus sets we would have needed to compare an observation with an average of 13.5 other observations.

6 Extensions

There are three obvious extensions that generalize the current model.

Handling segmentation errors. Throughout the paper, we have assumed that the motion segmentation algorithm works correctly, at least in terms of the number, location, and time of the incidents it reports. However, in practice observations of two objects can be merged into one if they are too close and partially occlude each other. In these cases, we can still express a solution in terms of links between observation incidents by relaxing the constraint that the chains must be mutually exclusive. This can be achieved by allowing each observation incident to appear in an arbitrary number of chains (instead of in at most one), and to add a penalty term for chain crossings to the objective function. The resulting problem is still a linear program and can be solved efficiently, but in order to make true chain crossings reasonably likely, we will have to define the matching probability in a way that allows partial matches of observations without leading to too many false positives.

If the amount of occlusion is too large, an object will remain invisible. This can probably be modeled by introducing a detection probability into the expression of the posterior, as is common in the radar tracking community.

Higher order Markov models. In the prior, we used a (first-order) Markov model for transition probabilities. If we assume instead that the location an object goes next is dependent on the current location and the previous locations, we obtain a multidimensional assignment problem. This problem is NP-complete, but can be rapidly approximated by Lagrangian relaxation [13].

However, it may be easier instead to add the primary motion direction of an object in the camera image as another parameter of the transition model. Such an extended transition model would give U-turns a low probability, for example.

Handling overlapping cameras. If two cameras overlap, we can replace them by a single virtual camera with a larger field of view. This requires mosaicing together the images, which can be done with standard techniques such as [14].

7 Conclusions

This paper introduced the multi-camera surveillance task and a Bayesian formalization. We showed how the MAP solution can be found under some additional independence assumptions by transforming the problem into a compact linear program. We demonstrated the viability of our approach with results from an 8 minute experiment with 4 cameras, for which nearly all links were correctly reconstructed.

Acknowledgements

We thank Yuri Boykov for in-depth discussions and Carlos Saavedra for helping out with the experiment. This research has been supported by a grant from Microsoft. The second author has been supported by DARPA under contract DAAL01-97-K-0104.

[4] The segmentation errors were due to strong reflections and shadows on the hallway floor.

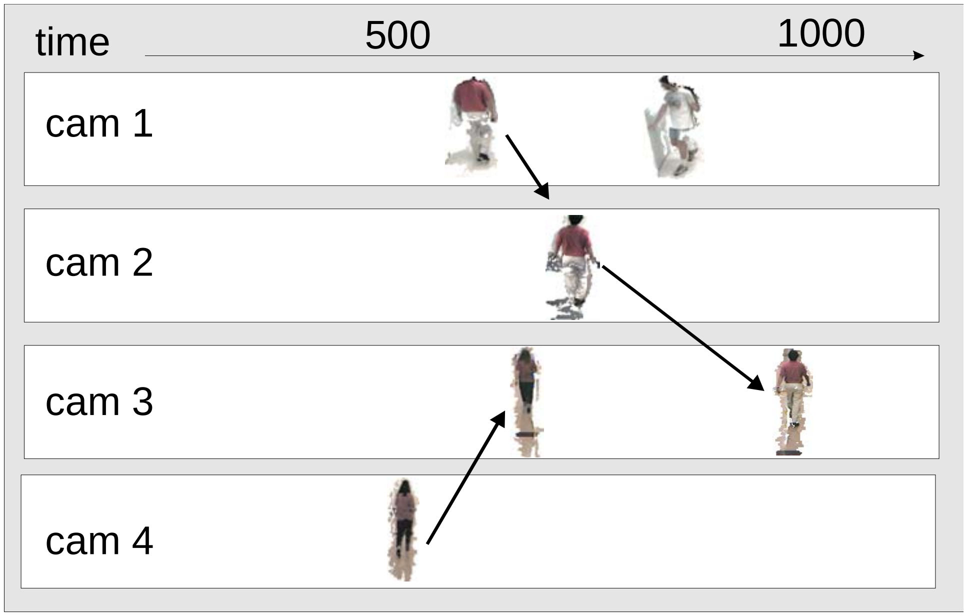

Figure 2: An example subsequence of the experimental 8 min sequence. Passing incidents are represented by the observation in the middle of the interval.

References

[1] P. Bouthemy and E. François. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *IJCV*, 10(2):157–182, 1993.

[2] J. Boyd, E. Hunter, P. Kelly, L. Tai, C. Phillips, and R. Jain. MPI-video infrastructure for dynamic environments. In *IEEE Conf. on Multimedia Computing and Systems*, pp. 249–254, 1998.

[3] Y. Boykov, O. Veksler, and R. Zabih. A variable window approach to early vision. *IEEE PAMI*, 20(12):1283–1294, 1998.

[4] F. Burgeois and J.-C. Lasalle. An extension of the Munkres algorithm for the assignment problem to rectangular matrices. *Comm. of the ACM*, 14:802–806, 1971.

[5] I.J. Cox and S.L. Hingorani. An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *IEEE PAMI*, 18(2):138–150, 1996.

[6] W.E.L. Grimson, C. Stauffer, R. Romano, and L. Lee. Using adaptive tracking to classify and monitor activities in a site. In *CVPR*, pp. 22–29, 1998.

[7] R.A. Howard. *Dynamic Probabilistic Systems.* Wiley, 1971.

[8] T. Huang and S. Russell. Object identification in a Bayesian context. In *IJCAI*, pp. 1276–1282, 1997.

[9] D. Huttenlocher, J. Noh, and W. Rucklidge. Tracking nonrigid objects in complex scenes. In *ICCV*, pp. 93–101, 1993.

[10] D. Koller, J. Weber, and J. Malik. Robust multiple car tracking with occlusion reasoning. Technical Report UCB/CSD-93-780, University of California at Berkeley, EECS Dept., 1993.

[11] T. Olson and F. Brill. Moving object detection and event recognition algorithms for smart cameras. In *DARPA Image Understanding Workshop*, pp. 159–175, 1997.

[12] A.B. Poore. Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking. *Computat. Optimization and Applications*, 3:27–57, 1994.

[13] A.B. Poore and A.J. Robertson. A new Lagrangian relaxation based algorithm for a class of multidimensional assignment problems. *Computational Optimization and Applications*, 8:129–150, 1997.

[14] R. Szeliski. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*, pp. 22–30, March 1996.