A probabilistic exclusion principle for tracking multiple objects

John MacCormick, Andrew Blake
University of Oxford
www.robots.ox.ac.uk/˜ab/vdg.html

Abstract

Tracking multiple targets whose models are indistinguishable is a challenging problem. Simply instantiating several independent 1-body trackers is not an adequate solution, because the independent trackers can coalesce onto the best-fitting target. This paper presents an observation density for tracking which solves this problem by exhibiting a probabilistic exclusion principle. *Exclusion arises naturally from a systematic derivation of the observation density, without relying on heuristics. Another important contribution of the paper is the presentation of* partitioned sampling*, a new sampling method for multiple object tracking. Partitioned sampling avoids the high computational load associated with fully coupled trackers, while retaining the desirable properties of coupling.*

1. Introduction

This paper proposes a mathematically rigorous methodology for tracking multiple objects. The fundamental problem to be addressed is demonstrated in figure 1. Two instantiations of the same tracking algorithm, with different initial conditions, are used to track two targets simultaneously. When one target passes close to the other, both tracking algorithms are attracted to the single target which best fits the head-and-shoulders model being used. One might think of avoiding this problem in a number of ways: interpreting the targets as "blobs" which merge and split again [8, 9], enforcing a minimum separation between targets [14], or incorporating enough 3D geometrical information to distinguish the targets [11]. However, each of these solutions can be unattractive. A blob interpretation does not maintain the identity of the targets, and is difficult to implement for moving backgrounds and for targets which are not easily segmented. A minimum separation relies on heuristics and fails if the targets overlap. Incorporating 3D information is impossible without detailed scene modeling.

So it seems we must instead address the fundamental problem: that the observation model used to interpret image measurements permits two targets to occupy the same point in configuration space too easily. More specifically, a single piece of image data (such as an edgel, or a colour blob), must not independently reinforce mutually exclusive hypotheses. What is needed is a "probabilistic exclusion principle", and an observation model exhibiting this behaviour is described in this paper. The formal model will initially be derived for "wire frame" targets — objects which have detectable boundaries but which do not occlude each other. We then describe how occlusion reasoning about solid objects can be incorporated naturally into the same framework. The most interesting feature of this approach is that it works even when the targets are *indistinguishable given the available information*. This is of both theoretical and practical interest.

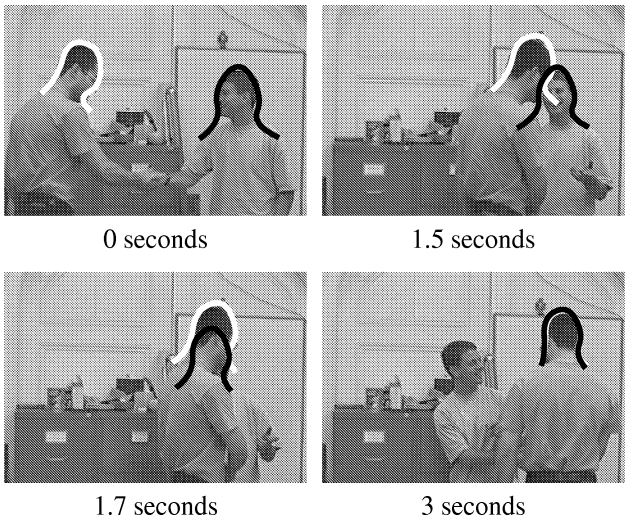


Figure 1. With an observation model designed for one target, two trackers initialised in distinct configurations eventually lock on to the one target which bests fits the model. The objective is to derive an observation model which does not permit the presence of two targets to be inferred from measurements of only one.

Many visual tracking systems for multiple objects have been developed. One standard technique is the probabilistic data association filter (PDAF) [1], and other successful

examples include [8, 9, 13, 14]. These generally employ a combination of blob identification and background subtraction; both techniques are complementary to the method proposed here. In particular, our exclusion principle does not allow two targets to merge when their configurations become similar; instead, the model continues to interpret the data in terms of two targets. As will be seen, it is a natural consequence of the methodology that the probability distribution for an obscured target diffuses until it is reinforced by further data. Furthermore, the method works for unknown and constantly changing backgrounds. The authors of [14] proposed a promising method for combining colour blob and edge information, and incorporated an exclusion principle by using a joint PDAF. However, their algorithm for fusing edgel information enforced an arbitrary minimum separation between targets. Gordon [7] employs a similar multi-target tracking methodology to this paper but with a rather different observation model and no explicit exclusion principle.

One of the difficulties with tracking multiple objects is the high dimensionality of the joint configuration space. Section 5 introduces a method known as *partitioned sampling* which diminishes the computational burden associated with the increased dimensionality of multi-target spaces.

2. The observation model

The target objects in this paper are described by their outlines, which are modeled as B-splines. We will call any such outline a *contour*. The space of contours which can correspond to a target or set of targets is called the *shape space*[3], and is parameterised as a low-dimensional vector space $\mathcal{X}$. The space $\mathcal{X}$ generally has 5–50 dimensions. A configuration $\mathbf{x} \in \mathcal{X}$ is measured by the method of figure 2, obtaining a list of image coordinates $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots \mathbf{z}_M)$. A component of $\mathbf{Z}$ is itself a vector $\mathbf{z}_i$ consisting of the measurements made along fixed *measurement lines* (see the figure) of the configuration $\mathbf{x}$. So $\mathbf{Z}$ is a function of both the image and the configuration.

Consider just one measurement line, of length L, positioned in an image known to contain two target objects. A one-dimensional edge detector is applied to this line, and some features are detected at image coordinates $z_1, z_2, \ldots z_n$. For a given target configuration $\mathbf{x}$, there are three possibilities to consider: the measurement line may intersect $c = 0, 1$ or 2 of the targets. The probability densities for each case are denoted $p_c(n; \mathbf{z})$. To calculate the p_c, several concrete assumptions are adopted:

- $c = 0$ ("random background clutter"): The probability of obtaining n features is $b(n)$, learnt from randomly placed measurement lines in typical images. The positions of the n features $\mathbf{z} = (z_1, z_2, \ldots z_n)$ are drawn from the uniform distribution on the measurement line.

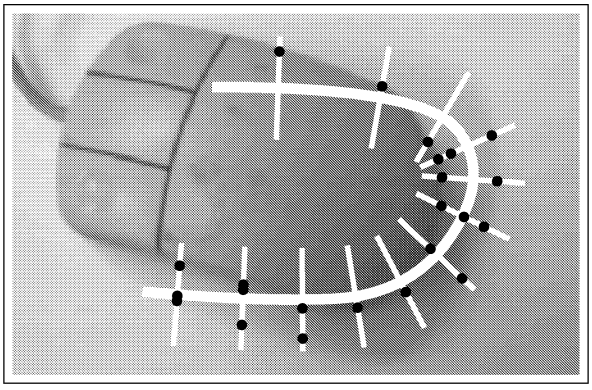


Figure 2. Measurement methodology. The thick white line is $\mathbf{x}$ — a mouse-shaped contour in some hypothesised configuration. The thin lines are *measurement lines*, along which a one-dimensional feature detector is applied. Black dots show the output of the feature detector, which in this case responds to rapid changes in intensity — one-dimensional edges.

- $c = 1$ ("single target"): One of the n features corresponds to the target boundary, whose hypothesised position on the measurement line is denoted ν. If the boundary feature is z_i, then z_i is assumed to be drawn from a fixed probability distribution $\mathcal{G}(z_i|\nu)$. In this paper $\mathcal{G}(z_i|\nu)$ is a Gaussian centred on ν with variance σ^2 (we take $\sigma = 7$ pixels in the examples later). The remaining $n-1$ features are assumed to be drawn from the random background clutter distribution described above.
- $c = 2$ ("two targets"): Two of the n features, say z_i, z_j, correspond to target boundaries at hypothesised positions ν_1, ν_2. They are drawn from $\mathcal{G}(z_i|\nu_1), \mathcal{G}(z_j|\nu_2)$ respectively with, importantly, $i \neq j$. In other words, any edge feature can correspond to at most one target boundary. It is this assumption which leads to the enforcement of a probabilistic exclusion principle described later on. Again the remaining $n-2$ features are drawn from the background distribution.

It is clear that this model can be generalised to higher values of c, but for clarity only the cases $c = 0, 1, 2$ are considered here. With the above assumptions, some simple combinatorial arguments show the appropriate pdfs are:

$$p_0(n; z_1, \ldots z_n) = b(n)/L^n \tag{1}$$

$$p_1(n; z_1, \ldots z_n|\nu) = b(n-1) \sum_{k=1}^{n} \mathcal{G}(z_k|\nu)/nL^{n-1}$$

$$p_2(n; z_1, \ldots z_n|\nu_1, \nu_2) = b(n-2) \sum_{i \neq j} \frac{\mathcal{G}(z_i|\nu_1)\mathcal{G}(z_j|\nu_2)}{L^{n-2}n(n-1)}$$

However, one more detail is required for this to be an acceptably realistic model: occasionally the target object's

boundary is not detected, because the background and target happen to have similar grey-scale values. It is therefore assumed that when $c = 1$ there is a small fixed probability q_{01} of the edge detector failing to detect the target boundary, and $q_{11} = 1 - q_{01}$ that it will succeed. This is precisely analogous to the non-detection probabilities used in PDAFs [1]. Similarly, when $c = 2$, there are fixed probabilities q_{02}, q_{12}, q_{22} that $0, 1, 2$ target boundaries are detected successfully. Thus we can define overall pdfs for the cases $c = 0, 1, 2$ as follows:

$$\begin{aligned}\tilde{p}_0(\cdot) &= p_0(\cdot)\\ \tilde{p}_1(\cdot|\nu) &= q_{01}p_0(\cdot) + q_{11}p_1(\cdot|\nu)\\ \tilde{p}_2(\cdot|\nu_1,\nu_2) &= q_{02}p_0(\cdot) + q_{12}(p_1(\cdot|\nu_1) + p_1(\cdot|\nu_2))/2\\ &\quad + q_{22}p_2(\cdot|\nu_1,\nu_2)\end{aligned}$$

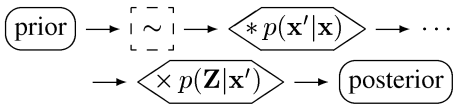
Typical graphs of the last two functions are shown in figure 3. Here and later we take $q_{01} = 0.1, q_{11} = 0.9, q_{02} = 0.05, q_{12} = 0.2, q_{22} = 0.75$; these values were determined by trial and error on simple examples.

The above discussion was framed in terms of a single measurement line, but for any given hypothesised configuration $\mathbf{x}$, the measurements $\mathbf{Z}$ will arise from say M distinct measurement lines. Let $c(i)$ be the number of target boundaries crossing the ith measurement line for a given configuration $\mathbf{x}$, and let $\boldsymbol{\nu}^{(i)}$ be the coordinates of these intersections. By making the approximation that outputs on distinct measurement lines are statistically independent, we define the *exclusive likelihood function* as

$$\mathcal{P}(\mathbf{Z}|\mathbf{x}) = \prod_{i=1}^{M} \tilde{p}_{c(i)}(\mathbf{z}^{(i)}|\boldsymbol{\nu}^{(i)}). \tag{2}$$

3. Tracking multiple wire frames

Tracking is performed in this paper by the Condensation algorithm [10], which is capable of dealing with complex likelihood functions such as (1). A Condensation tracker represents the state of a system at time t by a weighted set of *samples* or *particles* $s_1^t, \ldots s_N^t$ whose weights are $\pi_1^t, \ldots \pi_N^t$. The set is propagated to the next time step by (i) resampling: sample N times with replacement from the set of particles s_i^t, according to the weights π_i^t — this produces a set $s_1^{t\prime}, \ldots s_N^{t\prime}$; (ii) dynamical propagation: sample from $p(\mathbf{x}^{t+1}|\mathbf{x}^t = s_i^{t\prime})$ to choose each s_i^{t+1}; and (iii) measurement: examine the image to obtain the features $\mathbf{Z}^{t+1}$, then assign each of the new particles a weight $\pi_i^{t+1} \propto p(\mathbf{Z}^{t+1}|\mathbf{x}^{t+1} = s^{t+1})$. The three transformations of the particle set in any time step can be conveniently summarised diagrammatically:



(3)

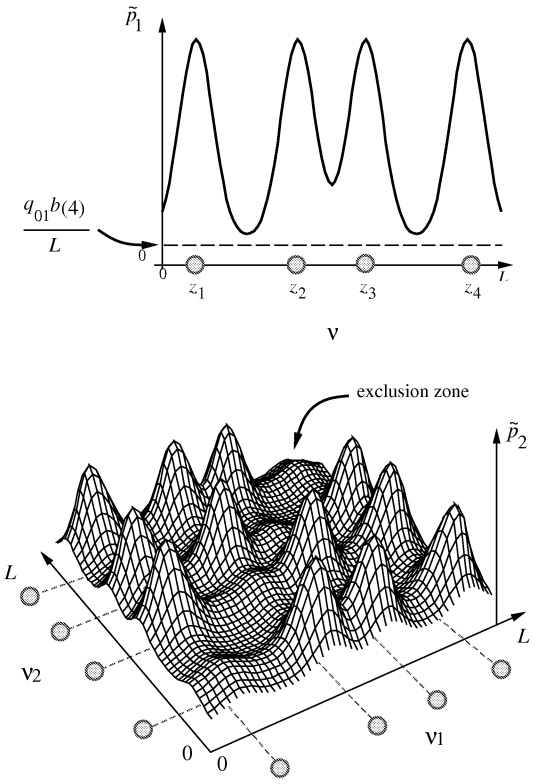


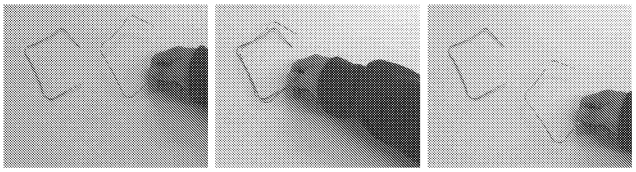
Figure 3. 1- and 2-target likelihood functions for a single measurement line. Top: $\tilde{p}_1(n; \mathbf{z}|\nu)$ graphed with respect to ν. The likelihood peaks near the 4 measurements z_i (shown as shaded circles). Bottom: $\tilde{p}_2(n; \mathbf{z}|\nu_1, \nu_2)$ graphed with respect to ν_1, ν_2. The likelihood peaks near pairs of measurements z_i, z_j (shaded circles and dotted lines), but a probabilistic exclusion principle operates: because the sum in the definition of p_2 excludes $i = j$, the probability peaks are much smaller on the line $\nu_1 = \nu_2$.

The $\sim$ symbol represents resampling as described above, the $*$ is application of some dynamics and the $\times$ represents multiplication (i.e. re-weighting) by the measurement density. Of course, to demonstrate the exclusion principle we use the likelihood function $\mathcal{P}(\mathbf{Z}|\mathbf{x})$ as the measurement density. Note that $\mathcal{P}$ as defined in (2) is not valid for opaque objects, since the model expects to observe all boundaries, even when occluded. However, it is valid for wire frame objects, so an experiment on wire frames was performed. As a control for the experiment, we need a likelihood $\mathcal{P}'$, similar to $\mathcal{P}$, but which does not incorporate an exclusion principle. Naming the two targets A and B, and writing $c_A(i)$ for the number of intersections of A with line i, let $\nu_A^{(i)}$ be the coordinates of these crossings and define the *1-body density*

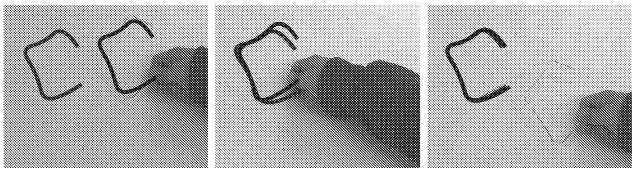
$$\mathcal{P}_A(\mathbf{Z}|\mathbf{x}) = \prod_{i=1}^{M} \tilde{p}_{c_A(i)}(\mathbf{z}^{(i)}|\boldsymbol{\nu}_A^{(i)}), \tag{4}$$

and similarly for $\mathcal{P}_B$. We take $\mathcal{P}' = \mathcal{P}_A\mathcal{P}_B$, so the posteriors for A and B given $\mathbf{Z}$ are treated as independent. A typi-

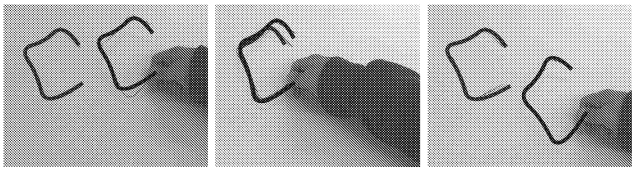
cal graph of $\mathcal{P}'$ for just one measurement line would be similar to figure 3 but with four additional peaks down the line $\nu_1 = \nu_2$. Figure 4 shows the results: as expected, $\mathcal{P}$ successfully maintains exclusion between the targets whereas $\mathcal{P}'$ does not.



(a) *raw sequence of moving wire frames*



(b) *tracking using naïve joint observation density*



(c) *tracking using probabilistic exclusion principle*

Figure 4. The exclusion principle operating on a wire-frame example. (a) Three stills from a sequence of two pieces of wire with similar shapes. Note that for several frames in the middle of the sequence, the two wires have very similar configurations. (b) Results using the likelihood $\mathcal{P}'$, which does not incorporate an exclusion principle. When the configurations become similar, both targets settle on the best-fitting wire. (c) Successful tracking using the exclusion principle likelihood $\mathcal{P}$.

4. Tracking multiple opaque objects

The wire-frame model can be adapted for use with solid objects. The method uses the mixed state Condensation tracker of [10], combined with a "2.1D" model of the targets. The basic idea of a mixed state Condensation tracker is that each particle carries a discrete label in addition to the continuous parameters describing its configuration. The extended state is defined to be

$$\mathbf{X} = (\mathbf{x}, y), \mathbf{x} \in \mathbb{R}^d, y \in \{1, \dots N_s\},$$

where y is a discrete variable labeling the current model, and $\mathbf{x}$ is a vector of continuous parameters specifying the configuration of the targets. In the two-object case, $\mathbf{x} = (\mathbf{x}^A, \mathbf{x}^B)$ and y can take one of two values: $y = 1$ if A is nearer the camera than B, and $y = 2$ if B is nearer than A. This is what we mean by a 2.1D model: the only 3D geometric aspect to be inferred is whether target A can occluded target B or vice versa.

The process density can then be decomposed as follows:

$$p(\mathbf{X}_t|\mathbf{X}_{t-1}) = P(y_t|\mathbf{x}_t, \mathbf{X}_{t-1})\, p(\mathbf{x}_t|\mathbf{x}_{t-1}),$$

and if $y_{t-1} = j$ and $y_t = i$ this can be written more explicitly as

$$p(\mathbf{X}_t|\mathbf{X}_{t-1}) = T_{ij}(\mathbf{x}_t, \mathbf{x}_{t-1}) p(\mathbf{x}_t|\mathbf{x}_{t-1}).$$

where T_{ij} is a transition matrix and p is a density specifying the continuous dynamics for a particle. Here it is appropriate for $T_{ij}(\mathbf{x}_t, \mathbf{x}_{t-1})$ to be independent of $\mathbf{x}_{t-1}$. If $\mathbf{x}_t^A$ and $\mathbf{x}_t^B$ overlap then the occlusion relationship cannot change in the the current time-step and so we take $T_{ij}(\mathbf{x}_t)$ to be the identity matrix. If $\mathbf{x}_t^A$ and $\mathbf{x}_t^B$ do not overlap then there is a small, fixed probability value of y will change, represented by taking $T_{ij}(\mathbf{x}_t) = \begin{pmatrix} 1-\delta & \delta \\ \delta & 1-\delta \end{pmatrix}$ with $0 < \delta \ll 1$. The examples later take $\delta = 0.01$, which corresponds to a time constant of 2.0 seconds for a given discrete state.

The mixed state Condensation tracker presented here incorporates a significant difference to that of [10] — the observation density $p(\mathbf{Z}_t|\mathbf{X}_t)$ depends not only on $\mathbf{x}_t$ but also on the discrete state y_t. The multi-target likelihood function (2) is used, but now the intersection counts $c(i)$ are calculated using the discrete variable y and the 2.1D geometry to determine if a given boundary feature should be visible or not, as in figure 5. To emphasise this we can write $c(i, y)$ for the number of *visible* target boundaries intersecting the ith measurement line of a configuration $(\mathbf{x}, y)$; the coordinates of the visible boundaries on the ith line are written $\boldsymbol{\nu}^{(i,y)}$. Then a likelihood incorporating occlusion reasoning can be defined as in (2), but replacing $c(i)$ by $c(i, y)$:

$$\mathcal{P}_{\text{occl}}(\mathbf{Z}|\mathbf{x}) = \prod_{i=1}^{M} \tilde{p}_{c(i,y)}(\mathbf{z}^{(i)}|\boldsymbol{\nu}^{(i,y)}). \qquad (5)$$

The likelihood $\mathcal{P}_{\text{occl}}$ performs well in experiments. Figure 6 shows a typical sequence involving occlusion. The configuration space has 16 dimensions: 8 key-frames from principal components analysis of templates [2, 5], for each of 2 targets). Tracking is performed with $N = 2000$ particles, and predictive dynamics in the form of Brownian motion with an amplitude matched to the speed of a walking person. Note how the occluded contours diffuse at 0.7 seconds. Because of the exclusion principle they coalesce again only when some evidence from the correct target is observed. The undesirable tracking behaviour of figure 1 has been corrected.

As a canonical tracking challenge, the same multiple target methodology was applied to the "leaf sequence" used in [10]. Two leaves were tracked, using an affine shape space

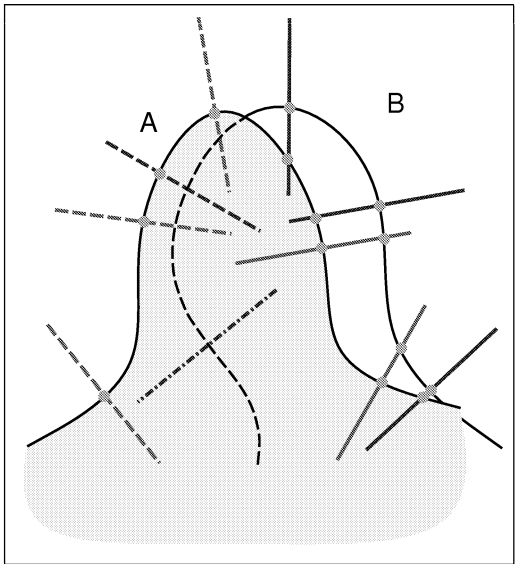


Figure 5. Intersection counts calculated from 2.1D geometry. In this diagram, $y = 1$, meaning the shaded area is occluded by target A. Visible intersections of measurement lines and target boundaries are shown as solid circles. The solid lines have intersection number $c = 2$, dotted have $c = 1$ and dashed $c = 0$. These are the c-values used in (5).

and $N = 4000$ samples with learnt dynamics. The need for 4000 samples is reduced to 750 by the partitioned sampling method described in the next section. Tracking is successful despite occlusions; some stills are shown in figure 7.

5. Partitioned sampling for Condensation

A potential limitation of the Condensation algorithm is that if the state space has many dimensions, then the number of particles required to model a distribution can be very large indeed. This is of particular concern when tracking multiple objects, since the number of dimensions in the state space is proportional to the number of objects. Fortunately, "partitioned sampling" significantly reduces this curse of dimensionality. It is the statistical analogue of a hierarchical search: the intuition is that it should be more efficient to search *first* for whichever target is unoccluded, and only then to search for another target which may lie behind.

We must omit some details here, but the basic algorithm is as follows. Assume the dynamics for each target can be decomposed as $p(\mathbf{x}''|\mathbf{x}) = \int_{\mathbf{x}'} p_B(\mathbf{x}''|\mathbf{x}')p_A(\mathbf{x}'|\mathbf{x})$, where p_A are the dynamics for target A and similarly for B. Also recall the one-body densities $\mathcal{P}_A, \mathcal{P}_B$ defined by (4). Then one time-step of the partitioned sampling algorithm is defined by the diagram in figure 8. In figure 8, the $\sim \mathcal{P}_A$ symbol means "resample according to weights $\propto \mathcal{P}_A$, then re-weight the selected particles with weights $\propto 1/\mathcal{P}_A$." It can be shown this operation does not alter the underlying distribution represented. Hence we can prove figure (8) is a

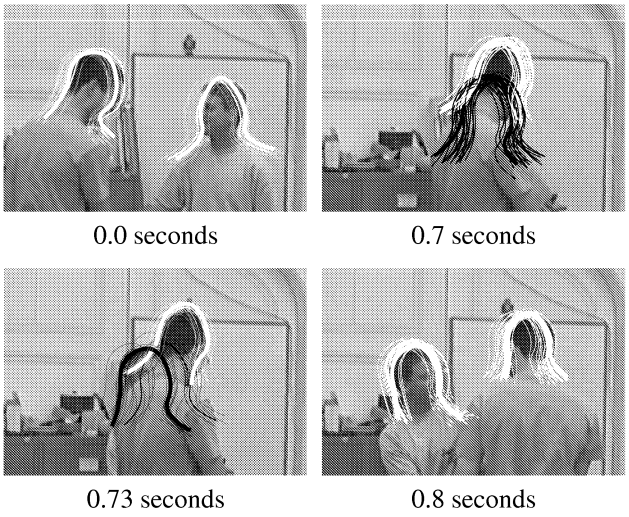


Figure 6. Successful tracking with a density incorporating occlusion reasoning 20 of the 2000 particles are shown in each frame, with widths proportional to their probabilities. Each particle consists initially of two white contours: one initialised on target A and one on target B. A contour is drawn in black if its value of y implies that it is partially occluded.

Figure 7. Tracking multiple leaves, in moving clutter and with occlusions. The black contour shows a correctly inferred occlusion.

valid simulation of (3) by simply deleting the $\sim \mathcal{P}_A, \sim \mathcal{P}_B$ steps and fusing the two paths which now have identical effects on the particles.

Particles for which $y = 1$ follow the top path, which positions the A-components first (near peaks in $\mathcal{P}_A$), since these particles believe A is unoccluded. Particles for which $y = 2$ follow the bottom path, since they believe B is unoccluded. The final result is that many more particles survive the resampling process, compared to the non-partitioned process, and the posterior is represented more accurately. Figures 9 and 10 demonstrate this on a toy example with a 2-dimensional configuration space.

Evaluating the performance of particle filters such as Condensation is a difficult problem [4, 6, 12]. To compare the two schemes (figure 8) and (equation 3) we use Doucet's [6] *estimated effective sample size* $\hat{N}$ defined for a set of

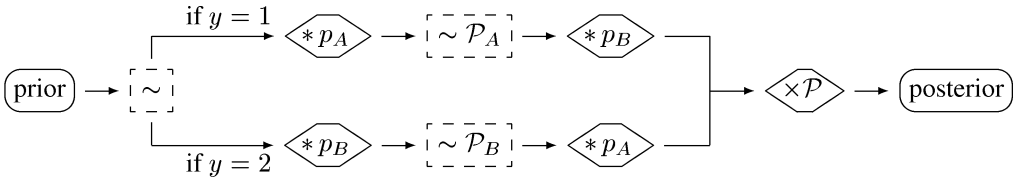


Figure 8. Diagrammatic representation of partitioned sampling. Particles with different discrete states y follow different paths through the diagram.

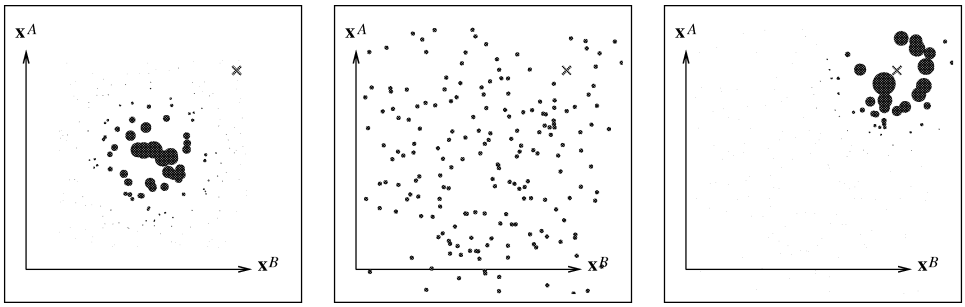


Figure 9. Conventional (i.e. non-partitioned) Condensation. The true position of the target in this 2-dimensional configuration space is shown as a cross; particles representing a probability distribution are shown as circles whose areas are proportional to their weights. Each step shown is one stage in the condensation diagram (3). (a) The prior. (b) After the dynamics have been applied. (c) After reweighting by the posterior. The posterior is centred at approximately the correct position, but this representation of the posterior is not very accurate because relatively few particles have significant weights. In technical terms, the effective sample size (6) is low. Superior results are achieved using partitioned sampling (figure 10).

particles with weights $\pi_1, \ldots \pi_N$ as

$$\hat{N} = \left(\sum_{i=1}^{N} \pi_i^2 \right)^{-1} \tag{6}$$

Intuitively, this corresponds to the number of "useful" particles: if all have the same weight $1/N$ then $\hat{N} = N$, whereas if all but one of the weights are negligible we have $\hat{N} = 1$. Any other distribution of the weights falls between these two extremes. Figure 11 compares $\hat{N}$ for the conventional ("unpartitioned") and partitioned methods. It is clear that partitioned sampling achieves much higher values of $\hat{N}$ than unpartitioned sampling and that we can therefore expect much better tracking performance for the same computational expense. We can show this is indeed the case in a practical example: figure 12 shows stills from a certain sequence tracked by each method. With partitioned sampling, and $N = 750$ particles, the tracking succeeds. However, despite using 4 times as many particles, unpartitioned sampling fails to track on the same sequence.

6. Conclusion

An exclusion principle for tracking multiple, indistinguishable targets has been introduced, which prevents a single piece of image data independently contributing to similar hypotheses for different targets. In its raw form, the model is valid only for wire-frame objects. However, by extending the tracking methodology to permit discrete states for describing the world in 2.1 dimensions, the same type of model can be used to track solid objects. Moreover, the approach requires only a simple model of the targets and no knowledge whatsoever of the background, which may itself be moving non-rigidly. A second contribution of the paper is to introduce partitioned sampling: a method of using particle filters with multiple objects, without incurring excessive additional computational cost for the extra dimensions.

So far the probabilistic exclusion principle has been developed for only the specific type of edge-based measurements described here. However, the fundamental idea is that any single measurement should reinforce multiple hypotheses coherently; it is hoped this can be used to guide the implementation of exclusion principles for more general observation processes.

Acknowledgements We are grateful for the support of the Royal Society of London (AB) and the EU (JM).

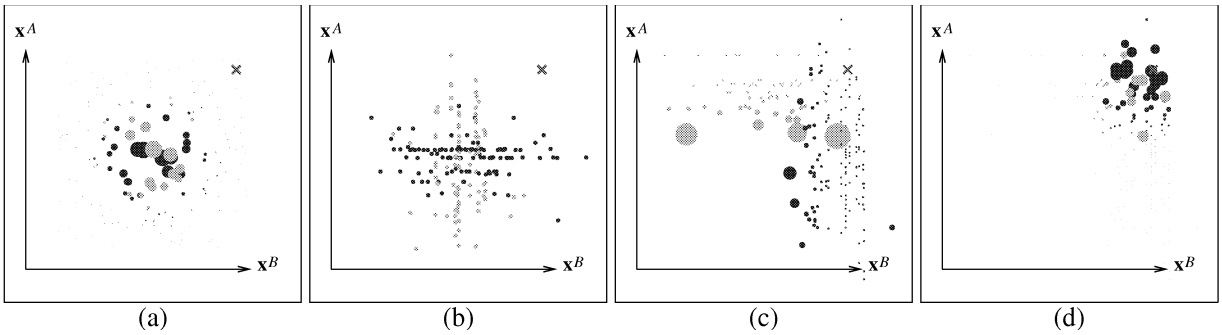


Figure 10. Partitioned sampling. Each step shows a stage from the condensation diagram figure 8. The 2-dimensional configuration has been augmented with a binary variable y, shown as black ($y = 1$) or grey ($y = 2$), and the value of this variable determines which branch is taken in figure 8. (a) The prior. (b) Dynamics have been applied in the $\mathbf{x}^A$-direction for black particles and the $\mathbf{x}^B$-direction for grey particles. (c) The weighted resampling operation has been performed. (d) The remaining dynamics are applied, and then the particles are re-weighted by the posterior. Note how fine-grained the sample set for the posterior is, compared with the final set from conventional sampling in figure 9. In other words, this representation of the posterior has a higher effective sample size (6) than that in figure 9.

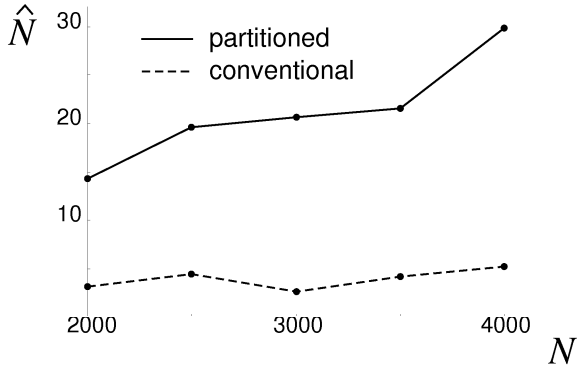


Figure 11. Estimated effective sample size $\hat{N}$ for partitioned and conventional (unpartitioned) sampling methods. The graph shows the average value of $\hat{N}$ following a 10-frame sequence tracking two leaves. Note the superior performance of the partitioned sampling method.

Figure 12. Unpartitioned sampling can fail when partitioned sampling does not, even if more particles are used. The final frame from a tracked sequence is shown: with unpartitioned sampling, the tracking fails despite using 4 times as many particles as the partitioned method.

References

[1] Y. Bar-Shalom and T. Fortmann. *Tracking and Data Association*. Academic Press, 1988.

[2] A. Baumberg and D. Hogg. Learning flexible models from image sequences. In J.-O. Eklundh, editor, *Proc. 3rd European Conf. Computer Vision*, pages 299–308. Springer-Verlag, 1994.

[3] A. Blake and M. Isard. *Active contours*. Springer, 1998.

[4] J. Carpenter, P. Clifford, and P. Fearnhead. An improved particle filter for non-linear problems. Technical report, Dept. of Statistics, University of Oxford, 1997.

[5] T. Cootes and C. Taylor. Active shape models. In *Proc. British Machine Vision Conf.*, pages 265–275, 1992.

[6] A. Doucet. On sequential simulation-based methods for Bayesian filtering. Technical report, Dept. of Eng., University of Cambridge, 1998.

[7] N. Gordon. A hybrid bootstrap filter for target tracking in clutter. *IEEE Trans. Aero. Elec. Systems*, 33:353–358, 1997.

[8] I. Haritaoglu, D. Harwood, and L. Davis. w^4s: A real-time system for detecting and tracking people in 2.5D. In *Proc. 5th European Conf. Computer Vision*, volume 1, pages 877–892, Freiburg, Germany, June 1998. Springer Verlag.

[9] S. Intille, J. Davis, and A. Bobick. Real-time closed-world tracking. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 697–703, 1997.

[10] M. Isard and A. Blake. A mixed-state Condensation tracker with automatic model switching. In *Proc. 6th Int. Conf. on Computer Vision*, pages 107–112, 1998.

[11] D. Koller, J. Weber, and J. Malik. Robust multiple car tracking with occlusion reasoning. In *Proc. 3rd European Conf. Computer Vision*, pages 189–196. Springer-Verlag, 1994.

[12] J. Liu and R. Chen. Sequential monte carlo methods for dynamic systems. *J. Amer. Statist. Assoc.*, 93, 1998. In press.

[13] N. Paragios and R. Deriche. A PDE-based level-set approach for detection and tracking of moving objects. In *Proc. 6th International Conf. Computer Vision*, pages 1139–45, 1998.

[14] C. Rasmussen and G. Hager. Joint probabilistic techniques for tracking multi-part objects. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 16–21, 1998.