

Multi-Sensor Multi-Target Tracking - Strategies for Events that Become Invisible

D.Hutber and Z.Zhang

INRIA, 2004 Route des Lucioles, B.P.93,

06902 Sophia Antipolis Cedex, France.

dhutber@sophia.inria.fr, z Zhang@sophia.inria.fr

Abstract

Within the context of multi-sensor, multi-target tracking, this paper addresses the problems of dealing with objects in the world that, for various reasons, were being observed but are no longer. A modification to the usual Mahalanobis distance metric used for matching observations to Extended Kalman Filters is proposed, and also two strategies for terminating tracks and tracking objects through blind zones are described. Experimental results obtained from a sensor-equipped vehicle navigating in road traffic are presented.

1 Introduction

Multi-sensor multi-target tracking poses an important set of problems that need to be addressed when designing, for example, multi-sensor vision systems for navigation or surveillance applications. The tracking problems have already been studied for radar systems [1, 3] in the 70's and 80's, but their application to multi-sensor vision systems is less well-understood. In particular, little attention has been paid to the disappearance of perceived objects, or *events*, and to the reasoning and strategies that can be applied to explain and deal with the subsequent problems. These subsequent problems arise from an increase in the *imprecision*, or positional uncertainty, of the event, as well as a possible decrease in its *validity* i.e. whether the perceived object still corresponds to a real object in the world, after it has ceased to be observed for some while. The work described on event disappearance in this paper is general to vision systems for both navigation and surveillance applications.

2 The Stages of an Event's Lifetime

The division of the lifetime of an event into three stages is taken from the radar literature [3]. The work described here is predicated on the use of Extended Kalman Filters (EKF's) as the basis for tracking events whose state is described by the n -dimensional state vector \mathbf{x} which has covariance matrix P . The m -dimensional observations of events are denoted by \mathbf{y} .

The principal problem to be solved in multi-target tracking is that of *data association*, i.e. which observation to associate to which event. This is more complicated when there is little history of the event, and in the presence of false alarms and missed targets it may take some time to be confident that observations correspond to a real event. This stage, during which all information about the potential event is kept internal to the tracker, is called *track initiation*. After this stage comes the *track maintenance* stage during which observations are regularly received that are associated with the event. The existence and state of the event are made known externally to the tracker (called a confirmed track file in the radar literature) and the event covariance is typically in a quasi steady-state. At some point observations are no longer associated with the event. This may be due to a variety of reasons, and a list of these for one particular class of applications is given later in this section. The event is then in the stage of *track termination*.

It is convenient to visualise the association between observations and events as being done using the mechanism of a match matrix, which is a matrix where the entries (i, j) represent the Mahalanobis distance between observation i and event j . The Mahalanobis distance between two vectors \mathbf{x}_1 and \mathbf{x}_2 with covariance matrices P_1 and P_2 respectively is defined as:

$$(\mathbf{x}_1 - \mathbf{x}_2)^T (P_1 + P_2)^{-1} (\mathbf{x}_1 - \mathbf{x}_2)$$

Algorithms for data association in the stages of track initiation and track maintenance, as well as an overall architecture for multi-sensor multi-target tracking were presented in [5]. The stage of track termination is next described in more detail.

2.1 Track Termination

As stated above, as soon as observations are no longer matched to an event, the event is in the stage of track termination. Clearly this stage overlaps with the track maintenance stage in the sense that observations are received periodically rather than continuously, and targets are detected with probability P_D which is usually less than unity.

We first consider why, for a navigation or surveillance application, an event might become invisible. There are several reasons why this can happen, listed here:

Out of Range The event goes out of the effective range of the sensors and hence is no longer seen.

Blind Zone The event goes out of the fields of view of the sensors, however it remains in a *blind zone* i.e. a region relatively important to the effectiveness of the navigation or surveillance system but not covered by any sensors.

Known Occlusion The event becomes occluded by another known event.

Manoeuvring The event manoeuvres more than the predicted maximum, and so escapes from the *gate* (matchable region in observation space) of the event.

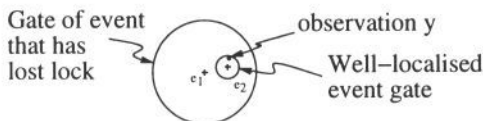


Figure 1: Observation Stealing

Not Detected The sensors no longer detect the event due to occlusion by unperceived objects, or other factors related to the sensor.

The practical consequence with the EKF algorithm of no observations being received that are matched to the event is that the gate of the event as measured by P will grow unchecked according to the maximum acceleration criterion.

3 Problems Involved with Track Termination

We are now in a position to state the problems which must be addressed by strategies for invisible events.

3.1 Observation Stealing and Oscillation

A possible result of the loss of lock between observations and events is the phenomenon of *observation stealing*. This has also been dealt with in earlier work by Zhang [7] who used a method of ‘token locking’ to prevent events with larger covariance stealing observations. In order to see how this is manifested consider the example shown in Figure 1.

Here the event e_1 has lost lock on its observations for one of the reasons described above. Event e_2 is a well-localised event that has been initiated and matches to a coherent sequence of observations. The size of the gates is indicated by the two circles, the large one is for e_1 . An observation y is matched using the Mahalanobis distance test, for which the distance for event e_1 is lower than that of event e_2 even though the Euclidean distance is much greater, causing event e_1 to ‘steal’ the observation from e_2 . A common consequence of this phenomenon is that the two events will now oscillate in size as illustrated in Figure 2.

In Figure 2 the two events e_1 and e_2 move from left to right on the page in time. At the start event e_1 has stolen the observation y_1 from the event e_2 , causing its covariance to decrease. It also matches to the two subsequent observations y_2 and y_3 . The next observation y_4 is closer in the Mahalanobis distance sense to event e_2 , since its covariance has now increased, according to the maximum acceleration law used, due to the lack of matched observations. The covariance of event e_2 now starts to decrease until observation y_7 , and so on.

This phenomenon is due to the autoscaling nature of the Mahalanobis distance and is relatively independent of the data association algorithm used (for example the JPDAF [1], and MHT algorithms e.g. [6] also suffer from this problem)

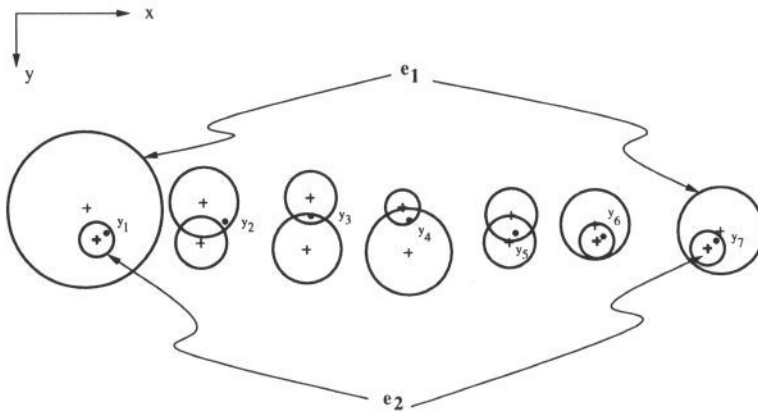


Figure 2: Observation Oscillating between Two Events

3.2 False Alarm Random Walking

With reference to Figure 1, a single false alarm appearing in the gate of event e_1 will match with the filter, and if the false alarm density is sufficiently high the event will continue to match with random false alarms indefinitely. This leads to the effect of the perceived event doing a ‘random walk’ in the observation space with potentially catastrophic effects for e.g. a navigation system.

We therefore develop an approach for the termination of tracks which takes into account the phenomena of observation stealing, oscillation and false alarm random walking.

4 A Strategy for Track Termination

The proposed method for track termination is made up of two parts; firstly the definition of a modification to the Mahalanobis distance metric to deal with observation stealing, the second a well-principled criterion for when to terminate a track.

4.1 A Modified Mahalanobis Distance

We define a modified Mahalanobis distance that is close to the normal Mahalanobis distance for small values of $|P|$ and which tends towards the Euclidean distance divided by a constant as $|P|$ gets larger. This will tend to penalise filter gates that have a large covariance, since the modified Mahalanobis distance will be larger than the (unmodified) Mahalanobis distance for large $|P|$.

A remaining problem is the determination of the constant or scale-factor upon which to base the scaled Euclidean distance. The natural scale-factor to use is that of the steady-state norm of the state covariance matrix ($\equiv |P_{ss}|$) for which no analytic solution is available in general [1] but which depends on the observation noise covariance R and the expected period of observations. Figure 3 shows the different metrics and how the scale-factor changes as a function of $|P|$.

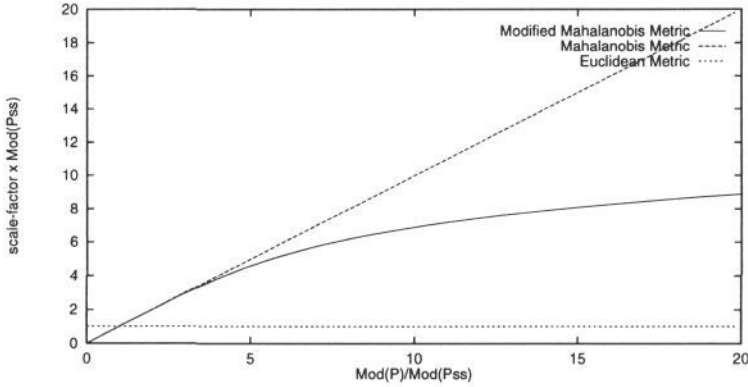


Figure 3: Graph of Distance Metrics

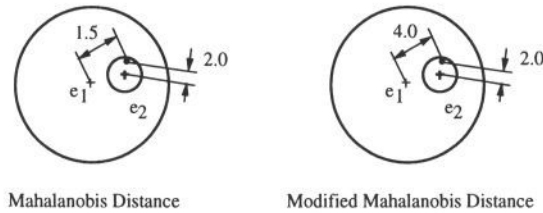


Figure 4: Unmodified and Modified Mahalanobis Distance Comparison

The scale-factor used by the Euclidean distance is constant (in the S.I. system the unit is 1 metre) and independent of $|P|$ as can be seen by the horizontal line on the graph.

The scale-factor used by the normal Mahalanobis metric is proportional to the norm of the covariance $|P|$; thus in Figure 3, as $|P|$ increases along the x-axis, the scale-factor plotted on the y-axis increases linearly.

The proposed scale-factor for the modified Mahalanobis distance for small $|P|$ is the same as the unmodified Mahalanobis distance. The normal minimum value for $|P|$ is $|P_{ss}|$. As $|P|$ increases, the scale asymptotes to a constant value $\beta|P_{ss}|$ determined by the free parameter β ($\beta = 10$ on the graph). The form of the function chosen is

$$x = k \left(\frac{y}{\beta|P_{ss}|} \right)^\alpha + \frac{y}{|P_{ss}|}$$

where α and k are other free (sensor-independent) parameters that determine how quickly the modified distance approaches the chosen asymptote $\beta|P_{ss}|$. In the results shown in Section 6, experiments indicated that their values are not critical.

We must also now distinguish between the gate size, which is related to $|P|$ and defines the search area for candidate observations, and the modified distance between an observation and a filter prediction. This difference is illustrated in Figure 4.

The gate size represented by the circles remains unchanged, indicating that the search area for candidate matches remains unchanged. This is reasonable since the gate size calculation is based on a maximum acceleration criterion which remains unchanged with the new method. The actual distances calculated using the old

and new method are shown on Figure 4, with the old scheme, the filter e_1 would steal the observation from e_1 since the Mahalanobis distance to e_1 is less. With the new scheme the distance from filter e_2 to the observation remains unchanged since the covariance of the state P_{e_2} is close to that of the steady-state. However now the distance from filter e_1 is larger with the new metric and the well-localised filter retains the observation.

4.2 A Termination Criterion

We now define firstly when to consider an event for termination, and then to define a well-founded basis for terminating an event.

An event is only considered for termination when another reason for its having disappeared cannot be found i.e. the tests for *out of range*, *blind zone* and *known occlusion* have all been found to be negative.

We note that the probability of finding a false alarm or a new event in a gate at time t_k with m-dimensional hyper-area $A_G^{t_k}$ is $(\rho_{FA} + \rho_{NT})A_G^{t_k}$, where ρ_{FA} is the expected density of false alarms and ρ_{NT} is the expected density of new events. The area of the gate will be evaluated and used whenever that sensor receives observations, therefore the superscript t_k is needed. We cumulate the probability $P(t_{now})$ of not finding anything in the gate between the time of the last observation that matched to the event (t_{last}) and the time at which we make the decision whether or not to terminate the event (t_{now}):

$$P(t_{now}) = \prod_{t_k=t_{last}+1}^{t_{now}} [1 - (\rho_{FA} + \rho_{NE})A_G^{t_k}]$$

If this probability is less than a threshold, the event is terminated. In the work described in Section 6 a value of $P(t_{now}) = 0.5$ was used.

5 A Strategy for Blind Zones

A blind zone is a part of the environment within the normal detection range of a multi-sensor system that no sensor can see directly and in which interesting events could be. It is important to deal properly with events falling into these zones since the performance of the tracking process may be affected by the entrances and exits to and from the blind zones. This will be true particularly when the number of entrances and exits is high.

We now briefly present a strategy for dealing with blind zones which uses the idea of *phantom events*.

5.1 Phantom Events

Informally, a phantom event is just the prediction of an event that is moving between sensors through a blind zone and hasn't been matched to any observations recently. More precisely we define a phantom event as a event that has entered (or re-entered) active or blind zones having previously been in a blind zone but has

not yet been matched to any observations originating from those zones' sensors since its entry. All events in blind zones are hence phantom.

Phantom events have other properties, which are:

- When a new event has passed the track initiation stage and is being compared with all existing events [5], it is compared with the phantom events as well.
- We do not attempt to match phantom events with observations.
- We use a different termination criterion for phantom events, based on a minimum speed and the maximum allowed distance to be travelled in a blind zone.

This is described in more detail in [4].

Having now presented a modified Mahalanobis distance, and strategies for terminating events and for events passing through blind zones, we can present a series of corresponding actions to the list of reasons for event disappearance presented earlier.

Reason	Action
Out of Range	terminate now (no danger)
In Blind Zone	blind zone strategy
Occlusion by a Known Event	blind zone strategy
Out-Manoeuvred Gate	wait then termination strategy
Not Detected	termination strategy

The first three reasons are detectable, and in the absence of any of the first three categories being detected, the termination strategy is employed.

6 Application to a Driver's Assistant System

The driver's assistant system ¹ uses several types of visual sensors, including telemetry, linear stereo and cameras to provide information about the environment of the vehicle. The perception information is passed to the tracking module which constructs a map of the environment, giving the relative position, velocity and size of all the perceived obstacles. The map is then analysed by a copilot module in order to determine which obstacles, if any, pose a danger to our vehicle.

We present three sets of results, two using a simulator program that generates realistic perception data with missed targets and false alarms. These results will illustrate the effect of switching on and off the various strategies. The third set of results is real data from the ProArt vehicle in a crossroads scenario. All the sets of results show a plan view of the environment around the ProArt vehicle, with the lines indicating the fields of view of the active sensors. Up on the figures is always ahead of the ProArt vehicle.

The first set of results is shown in Figure 5. Only one forward-facing sensor is active, our vehicle is just off-screen below the figure, in each case.

¹This work is a development of work done on the Eureka PROMETHEUS ProArt project

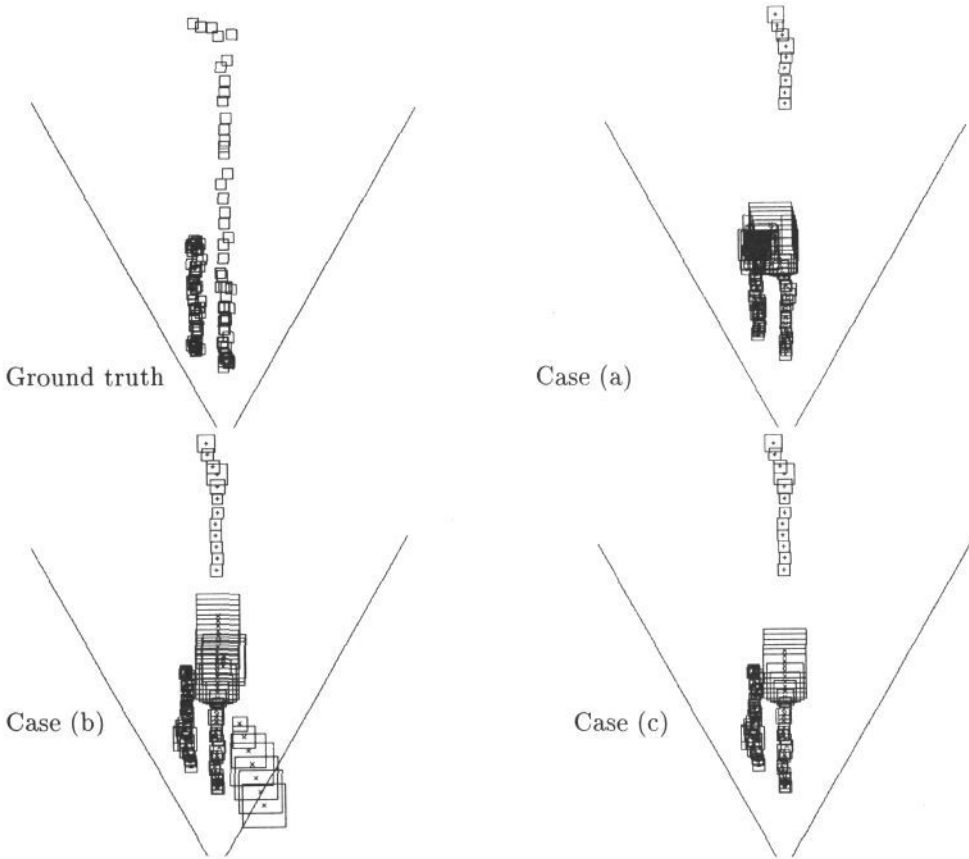


Figure 5: Rapidly Accelerating Vehicle Scenario

The top left case shows the ground truth of the scenario. The left-hand collection of squares corresponds to observations of one vehicle which is gradually pulling away from us, the right-hand collection corresponds to another vehicle that starts moving away gradually, then halfway through the sequence accelerates suddenly. The acceleration has deliberately been chosen to exceed the expected maximum acceleration for the tracker. A Poisson false alarm rate of mean $\mu = 3.0$ per time frame, uniformly distributed throughout the active tracking area has been added. Throughout, it will be seen that the track initiation strategy provides good immunity to false alarms.

In the other three cases, the size of the squares is related to the gate size or position uncertainty of the events, and the symbol in the centre represents the perceived identity of the event (check, triangle or cross).²

Case (a) shows what happens with the normal Mahalanobis distance. The second event loses lock with the observations, and then exhibits observation oscillation with the first event (cf. Figure 2). Meanwhile the tracker initiates another (different) event from the accelerating observations.

Case (b) shows the modified Mahalanobis distance, but with the track termination strategy switched off. Observation oscillation is prevented, but false alarms

²The size of the squares should not be interpreted as being identical to gate size, normally the 1σ gate is shown but the gate size saturates for clarity.

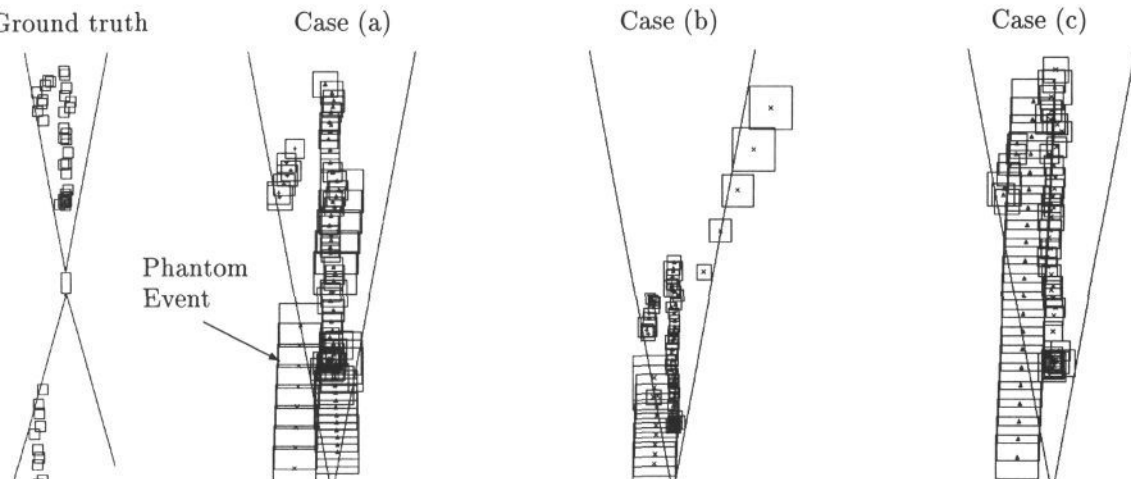


Figure 6: Overtaking Vehicle Scenario

in the gate of the right-hand event (check) lead to its matching with two successive false alarms and it exits to the bottom right of the figure.

Case (c) shows the modified distance and termination strategies switched on. In this case the right-hand event does not match with the left-hand event observations due to the modified Mahalanobis distance, and is terminated before it matches with any false alarms.

The second set of results is shown in Figure 6. The ProArt vehicle now has active forward- and backward-facing sensors. The first case again shows the ground truth of the scenario. A vehicle is observed behind our vehicle, and then it passes into the blind zone as it overtakes. We are gradually approaching another vehicle directly in front of us. The overtaking vehicle then re-appears in the front sensor.

Case (a) shows the normal Mahalanobis distance without the blind zone strategy. Here observation stealing by the phantom event occurs and the event in front (triangle) then matches with false alarms that send it towards our vehicle.

Case (b) shows the modified Mahalanobis distance, but without the blind zone strategy. The overtaking event is prevented from matching with the observations arising from the event in front by the modified distance, but false alarms again match with the overtaking event, sending it to the top right of the figure.

Case (c) shows both the modified Mahalanobis distance and the blind zone strategy. The overtaking event (triangle) matches correctly with the observations as they re-appear in the front sensor.

The third set of results in Figure 7 show real data taken in a crossroads scenario using a perception module that detects motion of the scene by image differencing from a combination of a background image and successive images. This is done from a static viewpoint [2] as a cue to detect obstacles. In the two cases shown, a vehicle is approaching from the right at speeds of 30km/h and 50km/h respectively. The modified Mahalanobis distance and the two strategies are switched on.

In both cases the vehicle is successfully tracked through the blind zone and re-matched with observations from the left camera. In the first case the prediction of the re-entering vehicle has a slightly slower velocity than the real vehicle, in the

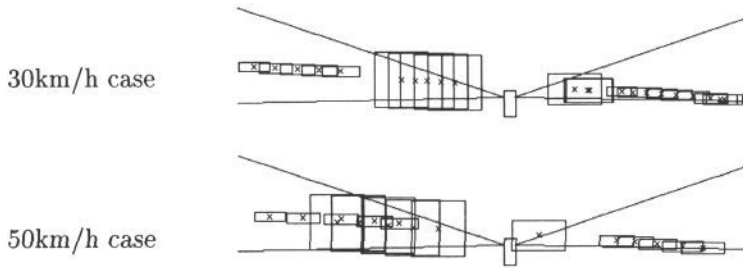


Figure 7: Crossroads Scenario

second case the prediction is slightly faster, but the velocity difference is within the predicted tolerance. The track initiation criterion [5, 4] requires 5 observations to be made before initiating an event, with the result that several false alarms in the original data are suppressed.

7 Conclusions

A modification to the Mahalanobis distance, and two strategies for track termination and dealing with blind zones have been presented. Together with the algorithms presented in [5], these enable the developed architecture for multi-sensor, multi-target tracking to cope with track initiation, maintenance and termination in a principled way related to the physics of the events being tracked.

References

- [1] Y. Bar Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, Boston, 1988.
- [2] A. Bellon, J.-P. Derutin, F. Heitz, and Y. Riquebourg. Real-time collision avoidance at road crossings on board the prometheus-prolab2 vehicle. In *Intelligent Vehicles*, pages 56–61, October 1994.
- [3] Samuel S. Blackman. *Multiple-Target Tracking with Radar Applications*. Artech House, 1986.
- [4] D. Hutber. *La Reduction de l'Incertitude dans le Contexte de Vision par Ordinateur Applique a un Vehicule dans un Environnement Routier*. PhD thesis, Universite de Nice - Sophia Antipolis, prevue en 1995.
- [5] D. Hutber and Z. Zhang. A two-stage approach to multi-sensor temporal data fusion. In *British Machine Vision Conference*, pages 721–730, September 1994.
- [6] D. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854, December 1979.
- [7] Z. Zhang. Token tracking in a cluttered scene. *Image and Vision Computing*, 12(2):110–120, March 1994.