Differential Analysis of Two Model-Based Vehicle Tracking Approaches

Hendrik Dahlkamp[1], Arthur E.C. Pece[2,3], Artur Ottlik[1], and
Hans-Hellmut Nagel[1]

[1] Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe (TH)
Postfach 6980, 76128 Karlsruhe, Germany
[2] Heimdall Vision, Bjørnsonsvej 29, DK-2500 Valby, Denmark
[3] Department of Computer Science, University of Copenhagen
Universitetsparken 1, DK-2100 Copenhagen, Denmark

Abstract. An experimental comparison of 'Edge-Element Association
(EEA)' and 'Marginalized Contour (MCo)' approaches for 3D model-
based vehicle tracking in traffic scenes is complicated by the different
shape and motion models with which they have been implemented orig-
inally. It is shown that the steering-angle motion model originally asso-
ciated with EEA allows more robust tracking than the angular-velocity
motion model originally associated with MCo. Details of the shape mod-
els can also make a difference, depending on the resolution of the images.
Performance differences due to the choice of motion and shape model
can outweigh the differences due to the choice of the tracking algorithm.
Tracking failures of the two approaches, however, usually do not happen
at the same frames, which can lead to insights into the relative strengths
and weaknesses of the two approaches.

1   Introduction

Detection and tracking of visible objects constitute standard challenges for the
evaluation of image sequences. If the camera is stationary, then tracking by
change detection (e.g. [11,7,5]) is feasible in real time, but can generate prob-
lems when the images of vehicles in a road traffic scene overlap significantly.
Switching from 2D tracking in the image plane to 3D tracking in the scene do-
main often results in a substantially reduced failure rate, because more prior
knowledge about the size and shape of objects to be tracked, their motion, and
their environment, can be brought to bear.

Very few 3D model-based trackers have been developed for road vehicles. In
order to understand the strengths and weaknesses of these trackers, they should
be compared under a variety of driving conditions, camera geometry, illumina-
tion, and occlusion conditions. Such comparisons are complicated by the fact
that differences in performance can be due to differences between initialisation
methods, shape models, motion (dynamical) models, or pose-refinement meth-
ods. Pose-refinement methods can be further decomposed into the criterion used
to evaluate the vehicle's pose and the method used to optimise the evaluation

criterion. Differences between pose-refinement methods are perhaps of greater scientific interest, yet differences between shape and/or motion models can be equally important for robust tracking. In the following, the term "approach" will be used as shorthand for pose-refinement method.

Here, two approaches to vehicle tracking with 3D shape and motion models are analysed: an approach based on optimising a marginalised likelihood ratio [8, 9] and an approach based on *edge-element association* [3]. A first comparison of these approaches has been reported in [1], but was carried out with two different implementations and different dynamical models. Experience with these initial experiments stimulated us to integrate modules corresponding to *each* approach *within the same system*. In this manner, the same shape and motion models can be used for both tracking approaches.

The next section outlines the test system (see also [2,10]) and the two tracking approaches. We then discuss the effects of different vehicle shape and motion models in Section 3, using the `rheinhafen` image sequence [12]. The insights gained thereby are exploited in Section 4 to compare the two approaches on a more complex image sequence (the 'dt_passat03' sequence, see [12]) with more vehicles being seen from a larger distance. Experiments have been carried out on the well-known `PETS-2000` [13] image sequence as well. Results are comparable to those obtained for the `rheinhafen` sequence and will be presented elsewhere.

2   The Approaches to Be Compared

The tracking experiments were performed within `MOTRIS` (**Mo**del-based **Tr**acking in **I**mage **S**equences), a framework implemented in Java (see, e. g., [10,2]), partially based on modules developed by other members of our laboratory and released under the GNU GPL [14]. In the experiments reported in this contribution, all vehicles were initialised interactively in order to avoid errors by automatic initialisation which might complicate the interpretation of the results.

2.1   Edge Element Association (EEA)

In the EEA approach, the boundaries between object and background in the image plane, as well as the internal boundaries of the object, are assumed to generate *edge elements* in the image. An edge element $\mathbf{e} = (u_e, v_e, \phi_e)^T$ represents a local maximum of the gradient norm in gradient direction ϕ at the position (u_e, v_e). As illustrated in Figure 1 (left), the distance measure $d_{\mathbf{m}}$ between an edge element $\mathbf{e}$ and a projected model segment $\mathbf{m}$ depends on (i) the Euclidean distance b between the edge element and the model segment and (ii) the difference Δ between the gradient direction of the edge element and the normal to the model segment:

$$d_{\mathbf{m}}(\mathbf{e}, \mathbf{x}) = \frac{b}{\cos \Delta} \quad . \tag{1}$$

It is assumed that $d_{\mathbf{m}}$ is normally distributed with zero mean and variance σ^2, implying a Mahalanobis distance which follows a $\chi^2(1)$ distribution. Edge

elements which exceed the $(1 - \alpha)$ quantile of the Mahalanobis distance are rejected. Edge elements which can be assigned to several model segments are assigned to the segment associated with the smallest Mahalanobis distance. The object pose is refined by iteratively minimising the sum of squared distances between edge elements and projected model segments. Thus, the optimisation of the model pose is reduced to an iterated least-squares problem.
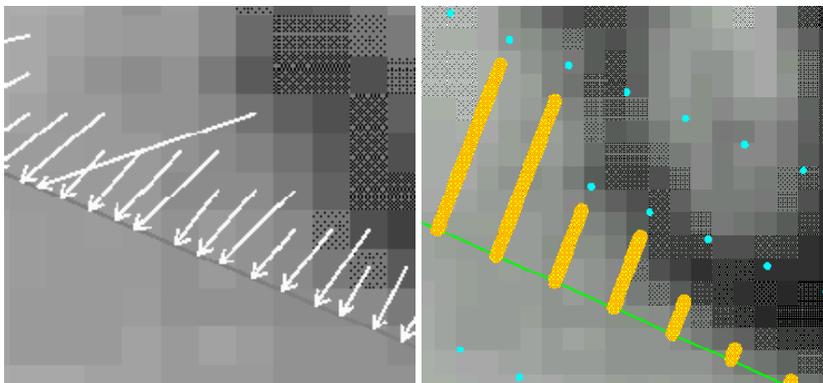

Fig. 1. Left: Edge elements and their distances from the closest model segment in the direction of the edge gradient. Right: Measurement points on the normals to the same model segment and estimated displacements for each normal, in the MCo approach.

2.2 Marginalised Contours (MCo)

The MCo approach is based on a statistical model of image generation and Maximum *a Posteriori* estimation of the vehicle pose. The principal assumptions underlying the MCo statistical model are

1. Grey value differences ΔI between adjacent pixels from the same region have a prior probability density $f_L(\Delta I)$ which is sharply peaked around zero, while grey value differences between pixels across an edge (object boundary) have a broad probability density $f_E(\Delta I)$ which may be considered to be uniform.
2. The visible object boundaries occur at distances, from the projection of the model into the image plane, which are randomly distributed with a Gaussian probability density f_D with zero mean and variance σ^2.

Assumption (1) implies that the likelihood ratio $f_E(\Delta I)/f_L(\Delta I)$ can be used as an estimator of whether a projected model edge falls between the pixels whose grey values are sampled. When assumption (2) is taken into account, it becomes necessary to integrate (marginalise) the likelihood ratios over all possible distances between projected model segments and object boundaries in the image

plane. By approximating the integration with a discrete summation, we obtain an equation for the marginalised likelihood ratio r_k at a sample point k:

$$r_k = \sum_j \frac{f_E(\Delta I_{k,j})}{f_L(\Delta I_{k,j})} f_D(j\Delta\nu) \qquad (2)$$

where the sum is over a number of measurement points indexed by j, equally spaced by a distance ν on the normal line to the sample point, see Figure 1 (right). By summing logarithms of marginalised likelihood ratios over all sample points k, we arrive at the evaluation function $E = \sum_k \log r_k$ which is to be maximised w.r.t. the vehicle's pose parameters. The Hessian of this evaluation function is not guaranteed to be negative definite and therefore the simple Newton method is not applicable for maximisation. However, a number of alternative gradient-based methods can be applied, for example the EM method used in [8].

3   Experiments Using the `rheinhafen` Image Sequence

In order to set the stage, we investigate how the two tracking approaches perform on the `rheinhafen` sequence with different shape and motion models.

In the original report [8], the MCo approach relied on an 'angular-velocity' motion model: the state of the vehicle to be tracked is described by position on the ground plane, orientation, tangential velocity, angular velocity, and tangential acceleration. An alternative 'steering-angle' motion model [4] was used in some of the experiments to be described in the following. This alternative includes the same state variables except for angular velocity and tangential acceleration, which are replaced by a steering angle and thus additional nonlinearities. Using the steering angle implies that the vehicle orientation change depends on the vehicle speed – this model does not allow an orientation change at zero tangential velocity. Since such a behaviour reflects driving physics more accurately, the steering angle model is expected to perform better on vehicles with changing speed.

3.1   Influence of Wheel Arches in the Vehicle Model

Initial tracking experiments indicated that the presence or absence of wheel arches constitutes an important feature of the shape model. This is confirmed by more systematic experiments.

Figure 2 illustrates the impact of the shape model for tracking of a notchback using different motion models and tracking approaches. The model has been manually optimised for this particular vehicle. Except in one case (using the EEA algorithm and the angular velocity motion model, discussed above and in Section 3.2), tracking results *improve* when using a model *with* wheel arches. The choice of the less accurate model without wheel arches leads to tracking failure for the combinations of EEA/steering angle and MCo/angular velocity.
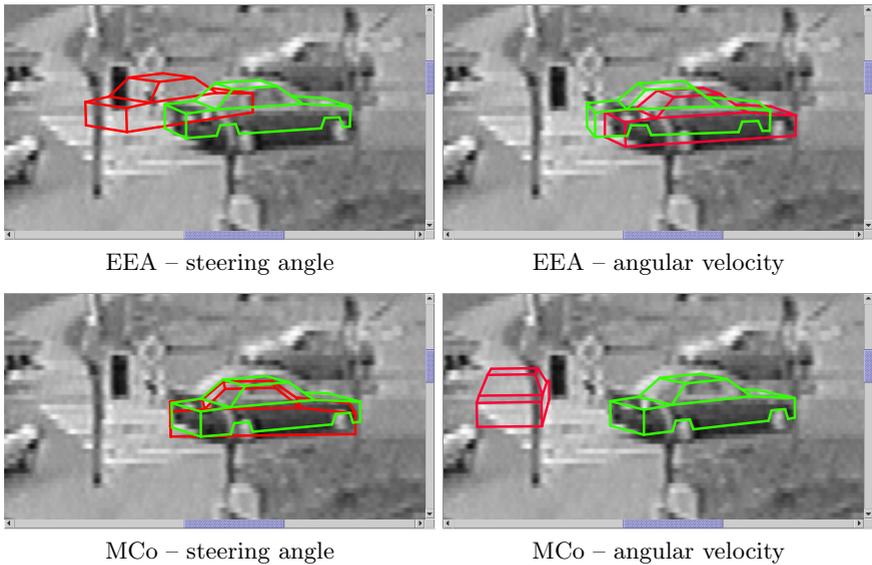
EEA – steering angle   EEA – angular velocity

MCo – steering angle   MCo – angular velocity

Fig. 2. Comparison of tracking results using vehicle model with (green) and without (red) wheel arches for a notchback at half-frame 2320 of the `rheinhafen` sequence for all combinations of tracking approaches and motion models.

3.2   Angular-Velocity Versus Steering-Angle Motion Model

Further experiments were carried out to compare the two motion models. Figure 3 shows tracking results on the same vehicle as in Figure 2 obtained for all possible combinations of tracking approaches and shape models.

As expected, introduction of the steering angle improves tracking performance. Only for the combination of the EEA tracking approach with a no-wheel-arch geometric model, the use of the steering-angle motion model leads to worse tracking results than the angular-velocity model.

Note that in all the experiments reported above, the same parameters were used for shape and motion models except, of course, for the shape parameters of the wheel arches and for the Kalman process noise associated with state variables present only in one of the two motion models.

Moreover, it can be seen that the choice of shape and motion model can outweigh the choice of the tracking approach (compare, e.g., top left panel (EEA / steering angle / wheel arches) and bottom right panel (MCo / angular velocity / no wheel arches) in Figure 2).

In conclusion, the experiments discussed so far suggest that the combination of a shape model with wheel arches and a motion model based on the steering angle - which has not been attempted before - is most efficient for model-based tracking in image sequences like those studied here.
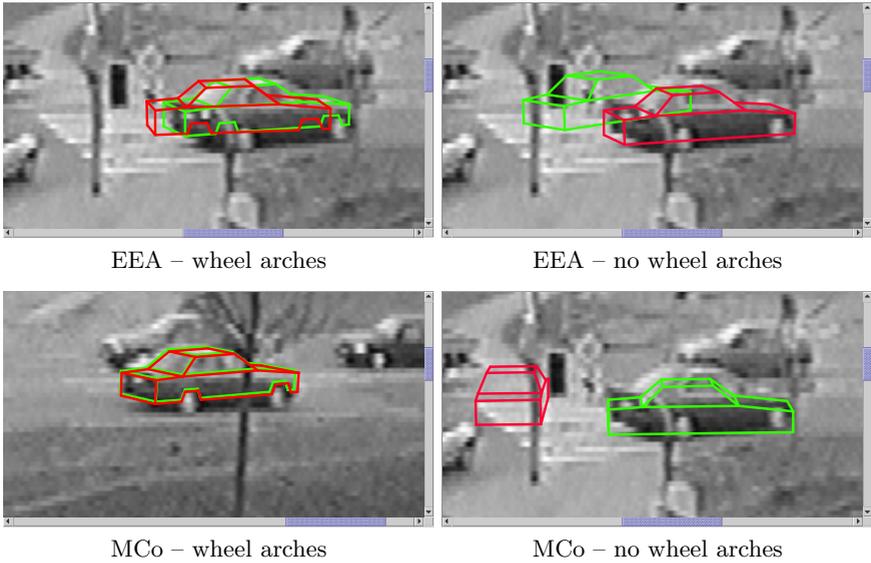
EEA – wheel arches   EEA – no wheel arches

MCo – wheel arches   MCo – no wheel arches

Fig. 3. Comparison of tracking results using steering angle (green) and angular velocity (red) motion model for a notchback at half-frame 2380 (lower left panel) and half-frame 2320 (other panels) of the `rheinhafen` sequence for all combinations of tracking approaches and shape models.

4   Experiments with a Traffic-Intersection Sequence

Figure 4 (left) shows a representative frame from the sequence 'dt_passat03' (see [12]). Here, vehicle images are smaller than in the `rheinhafen` sequence so that the presence or absence of wheel arches is no longer very important: in fact, in many cases the wheel arches can not be seen at all, due to the viewing distance and especially due to the viewing angles.

Figure 4 (right) summarises tracking results obtained by the 'optimal' combination of shape and motion models (wheel arches in the shape model and steering-angle motion model). The results obtained with either the EEA or the MCo approaches are not as good as those reported in [3]. However, this comparison is compromised by the fact that the latter results were obtained by a tracker including optical flow in addition to the EEA-approach: the integration of optical flow information reportedly leads to more robust tracking.

Vehicles 1, 6, 7, 13, and 15 are tracked successfully with the MCo approach while visible. Vehicles 1, 2, 6, 7, 9, 10, 12, 13, and 15 are tracked successfully with the EEA approach; vehicle 14 is tracked for much longer with the MCo approach, whereas vehicle 5, 9, and 10 can be tracked for much longer with the EEA approach. Tracking failures occur with both approaches for six other vehicles. Only five vehicles (1, 6, 7, 13, and 15) are tracked successfully by both approaches.
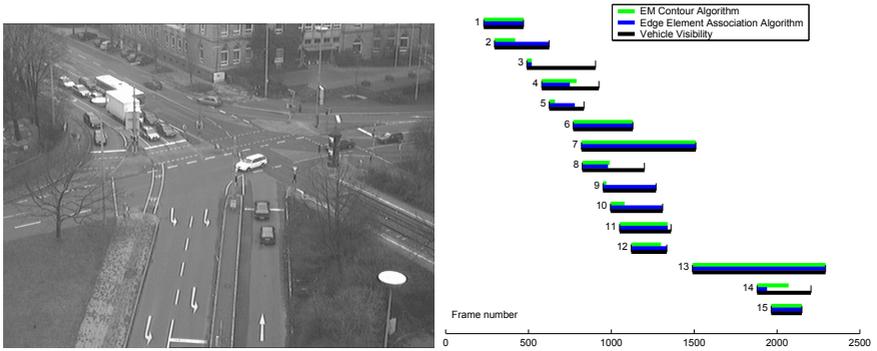
Fig. 4. Left: a typical frame of the dt_passat03 test sequence. Right: a graph illustrating the visibility of vehicles 1 to 15 as a function of frame number and the intervals during which the vehicles are successfully tracked by the EEA and MCo approaches, as determined by visual inspection.

5   Conclusions

Our experiments aim, first, at clarifying the effects of implementation alternatives and separate these effects from the effects of more fundamental assumptions underlying the two approaches under investigation; second, at gaining insight into the differential strengths and weaknesses of the two approaches when applied to challenging traffic sequences.

The experiments reported in this paper support the following conclusions:

1. Seemingly innocuous differences in the motion models used may, under certain circumstances (see Figures 2 and 3), be more important than the difference between the EEA and the MCo approaches.
2. Differences between the shape models can be more important than differences between motion models, again independently of the approach (EEA or MCo).
3. One must be careful, therefore, regarding tracking approaches carried out with different shape or motion models.
4. Similar to the results reported in [1], it is inconclusive which tracking approach performs best. The `rheinhafen` sequence seems to favor the MCo algorithm while the `dt_passat03` sequence gives an edge to the EEA approach.
5. Failures of the two approaches at different frames in the dt_passat03 sequence suggest that, even though the two approaches use the same information (grey value gradients in the image plane), a combination of their advantages might result in more robust tracking.

As a consequence, our experiments suggest to analyse the influence of other alternatives and parameter settings prior to the formulation of a 'final verdict' about a particular approach.

References

1. H. Dahlkamp, A. Ottlik, and H.-H. Nagel: *Comparison of Edge-driven Algorithms for Model-based Motion Estimation.* In Proc. Workshop on Spatial Coherence for Visual Motion Analysis (SCVMA'04), 15 May 2004, Prague, Czech Republic.

2. H. Dahlkamp: *Untersuchung eines Erwartungswert-Maximierung (EM)-Kontur-Algorithmus zur Fahrzeugverfolgung.* Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Januar 2004.

3. M. Haag and H.-H. Nagel: *Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Image Sequences.* International Journal of Computer Vision **35**:3 (1999) 295-319.

4. H. Leuck, and H.-H. Nagel: *Automatic Differentiation Facilitates OF-Integration into Steering-Angle-Based Road Vehicle Tracking.* In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99), 23-25 June 1999, Fort Collins, Colorado/USA; IEEE Computer Society Press, Los Alamitos/CA, Vol. 2, pp. 360-365.

5. D.R. Magee: *Tracking Multiple Vehicles Using Foreground, Background and Motion Models.* Image and Vision Computing **22**:2 (2004) 143-155.

6. H.-H. Nagel, M. Middendorf, H. Leuck und M. Haag: *Quantitativer Vergleich zweier Kinematik-Modelle für die Verfolgung von Straßenfahrzeugen in Video-Sequenzen.* In S. Posch and H. Ritter (Hrsg.), *Dynamische Perzeption*, Proceedings in Artificial Intelligence Vol. 8, Sankt Augustin: infix 1998, pp. 71-88 (in German).

7. A. E. C. Pece: *Generative-model-based Tracking by Cluster Analysis of Image Differences.* Robotics and Autonomous Systems **39**:3-4 (2002) 181-194.

8. A.E.C. Pece and A.D. Worrall: *Tracking with the EM Contour Algorithm.* Proceedings of the 7th European Conference on Computer Vision 2002 (ECCV2002), 28-30 May 2002, Copenhagen, Denmark; A. Heyden, G. Sparr, M. Nielsen, P. Johansen (Eds.), LNCS 2350, Springer-Verlag, Berlin·Heidelberg·New York (2002), pp. 3-17.

9. A.E.C. Pece: *The Kalman-EM Contour Tracker.* Proceedings of the 3rd Workshop on Statistical and Computational Theories of Vision (SCTV 2003), 12 October 2003, Nice, France;
   `http://www.stat.ucla.edu/~yuille/meetings/2003_workshop.php`.

10. P. Reuter: *Nutzung des Optischen Flusses bei der modellgestützten Verfolgung von Fußgängern in Videobildfolgen.* Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Oktober 2003.

11. C. Stauffer and W.E.L. Grimson: *Learning Patterns of Activity Using Real-Time Tracking.* IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) **22**:8 (2000) 747-757.

12. `http://i21www.ira.uka.de/image_sequences/`

13. `ftp://pets2001.cs.rdg.ac.uk/PETS2000/test_images/`

14. `http://kogs.iaks.uni-karlsruhe.de/motris/`