Comparison of Edge-Driven Algorithms for Model-Based Motion Estimation

Hendrik Dahlkamp, Artur Ottlik, and Hans-Hellmut Nagel

Institut für Algorithmen und Kognitive Systeme (IAKS),
Universität Karlsruhe (TH), 76128 Karlsruhe, Germany
`nagel@iaks.uni-karlsruhe.de`

Abstract. 3D-model-based tracking offers one possibility to explicate the manner in which spatial coherence can be exploited for the analysis of image sequences. Two seemingly different approaches towards 3D-model-based tracking are compared using the same digitized video sequences of road traffic scenes. Both approaches rely on the evaluation of grey-value discontinuities, one based on a hypothesized probability distribution function for step-discontinuities in the vicinity of model-segments, the other one based on extraction of Edge Elements (EEs) and their association to model-segments. The former approach could be considered to reflect a stronger spatial coherence assumption because the figure-of-merit function to be optimized collects evidence from all greyvalue discontinuities within a tolerance region around visible model segments. The individual association of EEs to model-segments by the alternative approach is based on a distance function which combines differences in position and orientation, thereby taking into account the gradient direction as well as the location of a local gradient maximum in gradient direction.

A detailed analysis of numerous vehicles leads to the preliminary conclusion that both approaches have different strengths and weaknesses. It turns out that the effects of how greyvalue discontinuities are taken into account are in general less important than the inclusion of Optical Flow (OF) estimates during the update-step of the current state vector for a body to be tracked. OF estimates are evaluated only within the area of the body to be tracked when projected into the image plane according to the current state estimate. Subtle effects related to simplifications and approximations during the implementation of an approach thus may influence the aggregated result of tracking numerous vehicles even in the case where spatial coherence appears to be rigorously exploited.

1 Introduction

Obviously, spatiotemporal coherence plays an important role in human visual perception. Exploitation of spatiotemporal coherence thus offers itself for Computer Vision (CV) although it is by no means evident how to do this, in particular if an 'optimal' use is postulated as an additional requirement. The latter postulate implies that alternative means are compared regarding how to exploit

spatiotemporal coherence in CV. With the advent of solid-state video cameras, sufficient bandwidth into and out of memory, large cheap memories, and powerful processors, recording and processing of entire image sequences constitutes today's challenge rather than that of single image frames. The concept of an image sequence implies already some kind of temporal coherence. Unless explicitly mentioned, its exploitation will not be discussed separately in what follows.

One of the alternatives with respect to exploitation of spatial coherence involves the question whether this aspect enters explicitly into the approach to be studied or only implicitly. We selected the first alternative. In a second step, one then has to decide whether the approach comprises only representations in the 2D image plane or relies on 3D models for objects of interest in the 3D scene. Examples for explicitly taking into account spatial coherence in a 2D image plane approach are provided by [1] and [4].

The use of 3D models for bodies and their movements in the depicted scene usually involves more complicated algorithms and thus more computing power and/or time, but has the advantage that size variations and occlusion effects can be taken into account in a more 'natural' manner. Spatial coherence in the image plane then is the consequence of models for the camera, the objects, and their movements rather than a postulate for image regions.

Unless considerable prior knowledge about the depicted scene and its temporal development is available, however, a 3D model-based approach necessitates an initialisation phase which either is based on a – potentially quite involved – search or substitutes assumptions for search. Such assumptions usually comprise some kind of spatial coherence. Once the image of a body has been detected, its tracking becomes much simpler based on the postulate of temporal coherence. Experience has shown that object tracking can depend critically on the initialisation phase. Tracking results are thus used frequently in order to assess the initialisation phase, its underlying assumptions and parameterisations. Such an approach implies, however, that the tracking process itself is fairly robust. The desire to exploit tracking results in order to study the complex initialisation phase and its implicit assumptions about spatial coherence thus leads to efforts to clarify the effects of various factors on tracking itself. As a consequence of these considerations, this contribution compares two different model-based approaches for tracking road vehicles, using different traffic videos.

In the next section, we sketch the differences between the two approaches and discuss expected relations between properties of the alternatives and results obtained by each alternative. Section 3 discusses experimental results. Section 4 includes results obtained by an extended approach which comprises Optical Flow (OF) estimates within the image of a body to be tracked, exploiting a different manifestation of spatial coherence observable in the image plane.

2   Discussion of the Algorithms to Be Compared

Both approaches to be compared rely on the evaluation of greyvalue discontinuities. The one reported in [5,6] is based on a hypothesized probability

distribution function for step-discontinuities in the vicinity of model-segments. This approach will be referred to in the sequel as 'Expectation Maximisation Contour Algorithm (EMCA)'. The other one is based on extraction of Edge Elements (EEs) and their association to model-segments (see, e. g., [3]). It will be referred to by 'Edge Element Association Algorithm (EEAA)'. In order to facilitate a self-contained discussion, essential aspects of both approaches have been outlined in Appendix A.

Several components are common during edge-based tracking. Their use, however, differs in the two analyzed algorithms:

- Use of *gaussian convolution*: While EMCA convolves the probability distribution for the greyvalue discontinuity position by a gaussian, EEAA convolves the image greyvalues directly. Therefore, EMCA use the gaussian to smooth its figure-of-merit function while EEAA increases the quality of its features, i. e. edge elements.

- Use of *edge intensity*: EMCA includes edge intensity in its figure-of-merit function. To be more specific, the figure-of-merit function is an exponential of the square root of the edge intensity, which can be seen by inserting equ. (2) into equ. (5) and equ. (5) into equ. (10). As the algorithm maximizes equ. (10), the gain of minimizing $\|\boldsymbol{\nu}_{k,j} - \boldsymbol{\mu}_k\|$ is larger if the edge intensity of measure point $\boldsymbol{\nu}_{k,j}$ is larger. In equation (11), however, every sample point is weighted the same, independent of its largest orthogonal greyvalue difference. Therefore, it can be said that the EM Contour Algorithm does not weight the influence of the sample points with their edge intensity. The figure-of-merit function contains such a weighting but the intermediate maximization goal given by equ. (10) does not. The estimation accuracy of the discontinuity location $\hat{\boldsymbol{\nu}}_k$ definitely depends on the edge intensity, as shown in [2, appendix B.1]. EEAA weights its distance measure between edge element and closest projected model segment by the edge element intensity, normalized to the strongest intensity of the vehicle.

- Use of *edge direction*: While EEAA includes the orientation difference between EE and projected model segment, EMCA computes the greyvalue difference in the direction obtained from the model.

- *locality of decisions*: Obviously, it is desirable for an algorithm to avoid making decisions based on local greyvalue structure, because smaller structures are more susceptible to noise than the entire vehicle image. EMCA contains no local decisions in the figure-of-merit function. EEAA decides locally where to put edge elements and where to map them onto the model. Given large filter masks, local EE computation is stable enough but mapping can produce significant jumps in the figure-of-merit function even if the model moves only very little.

In summary, EEAA extracts few, reliable features but can be expected to have trouble associating them to model-segments while EMCA uses many simple features and can be expected to have trouble with their accuracy and fusion.

3   Experiments with Edge-Based Approaches

First, we applied EMCA to the PETS2000 sequence [8] which was used to intro-
duce the algorithm. The results, as displayed in Fig. 1, show successful tracking
of all 3 vehicles if vehicles are initialized interactively, vehicle models are assigned
individually, and the same parameters are used as introduced in [5]. Using a sin-
gle model without wheelarches for all three vehicles yields similar results except
for wrong directional estimates of the second vehicle. The results are essentially
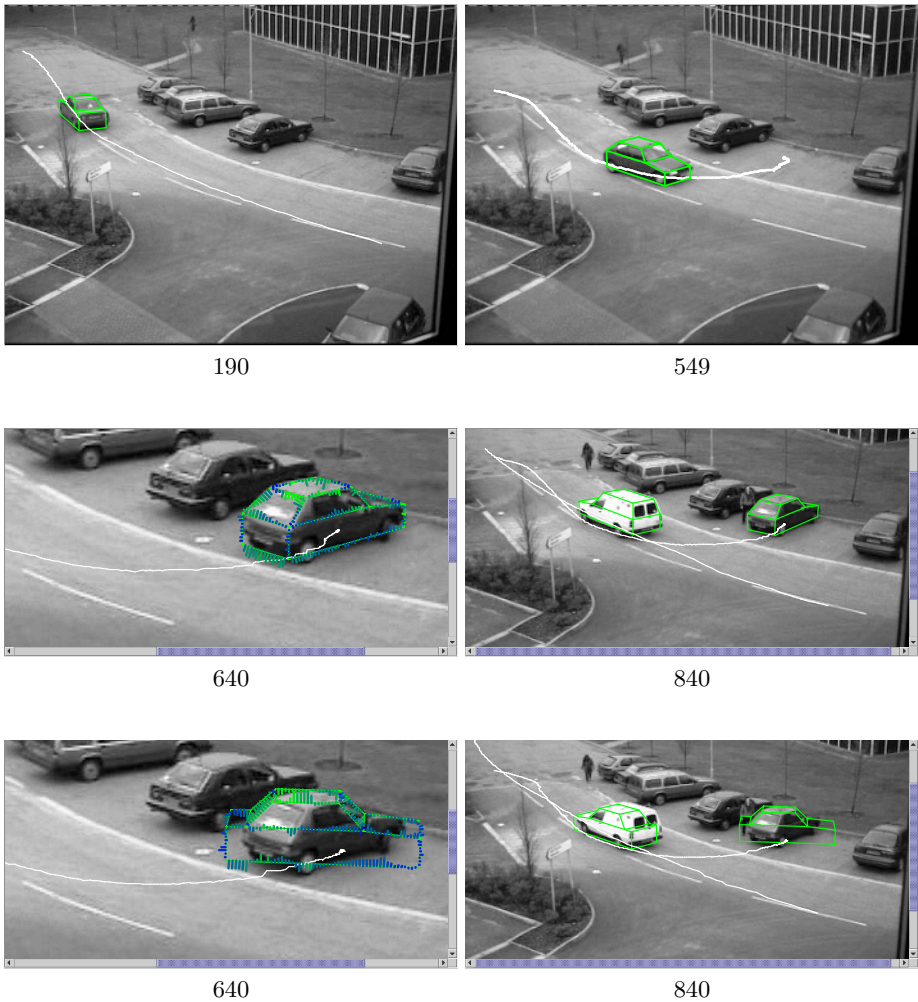compatible with those reported in [5].

<div align="center">190   549</div>

<div align="center">640   840</div>

<div align="center">640   840</div>

Fig. 1. Tracking results on the PETS2000 image sequence: The first four images show
tracking for individually assigned vehicle models at the given frame times. The lower
two images show results obtained with the *same* model for all vehicles.

We then compared the performance of both algorithms on the more challenging `dt_passat03` sequence [7] displayed in Fig. 2. This sequence contains more vehicles of significantly smaller size than in the `PETS2000` sequence. The results obtained with EEAA are essentially compatible with those reported in [3].

Fig. 2. Overview of the `dt_passat03` image sequence at frame 953. The vehicles at the traffic light are standing and, therefore, have not been tracked yet.

Figure 3 contains the tracking results showing visibility duration for each vehicle and the period during which the algorithms managed to track it. It can be seen that the EEA-approach scores better than the EMC-approach in a majority of cases, as represented by longer tracking duration until the vehicle is lost.

For a closer examination of the results, Figure 4 contains an exemplary look at two vehicles. The first one (upper left) is partially occluded by a leafless tree which darkens about a third of the vehicle pixels. For EMCA, this results in some mispredicted border positions but the errors cancel out and we are left with a successful prediction. The center left part of Fig. 4 demonstrates this by plotting the figure-of-merit function given by equ. (10) against vehicle displacement in the scene parallel and orthogonal to the vehicle orientation. The lower left part of Fig. 4 depicts the edge elements of the same image part. Because of the random placement of tree branches, these edge elements are very disturbed and mapping to model segments becomes practically arbitrary so the vehicle is lost. EMCA, therefore, performs systematically better than EEAA in cases where the tracked vehicle is occluded by many small objects. The second image (on the right) shows a dark grey vehicle in front of only slightly brighter background. Due to the dependence on the clear distinction between on-border and off-border
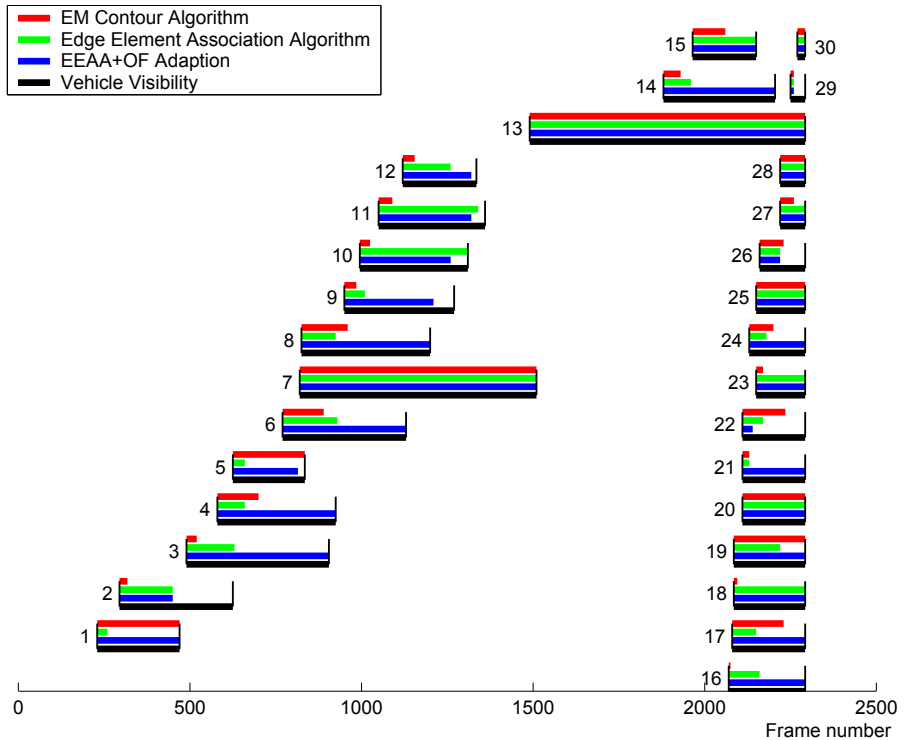
Fig. 3. Tracking results for the `dt_passat03` sequence. For each vehicle, four bars show its duration of visibility and duration of succesful tracking using different approaches. From top bar to bottom bar: duration of tracking with EMCA (red), tracking with EEAA (green), tracking with a combination (see Section 4) of EEAA and OF adaption (blue), vehicle's visibility (black) in the image sequence.

greyvalue differences, EMCA fails during the first few tracking frames. As shown in the center right part of the image, this tracking failure can be attributed to its figure-of-merit function plot having a maximum off the center. EEAA performs better in this case as the global edge intensity is not important. The lack of visual appearance of some model edges is irrelevant for EEAA as long as there are at least a few features present.

4   Supplemental Experiments Including OF Estimates

So far, only the exploitation of greyvalue transitions has been discussed. One of the two approaches to be compared offers an additional option, namely the inclusion of OF estimates into the state vector update step. Results presented, for example, in [3] support the expectation that the inclusion of OF estimates considerably stabilizes the tracking process. OF estimates directly influence the velocity estimate in the state vector. In addition, the number of OF vectors which
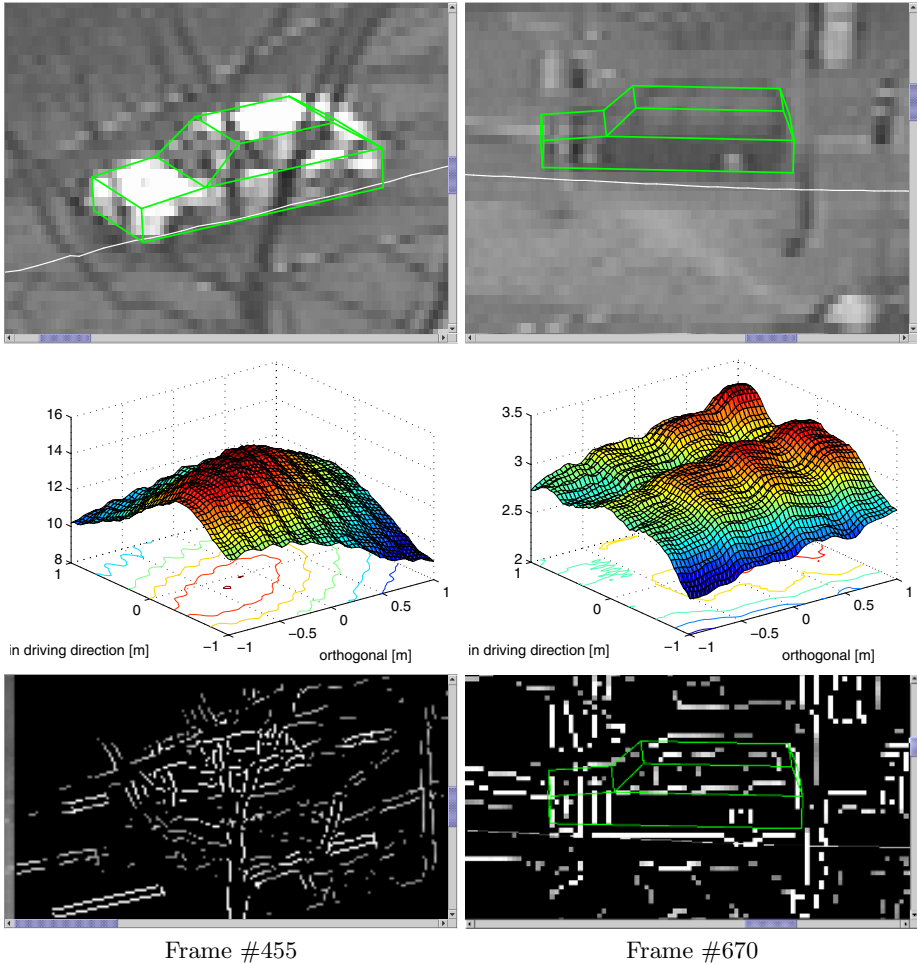
Frame #455   Frame #670

Fig. 4. Image zoom, EMCA figure-of-merit function plot and edge elements for two vehicles of the `dt_passat03` sequence

can be taken into account during the state update is proportional to the area of the projected model whereas the number of EEs is more closely proportional to its boundary length and thus smaller.

As can be seen in Figure 3, the combined EEAA and OF adaption outperforms both other algorithms, especially for the vehicles which drive from the right hand side of the image to the upper side (i. e. vehicle number 3, 4, 6, 8, and 14; with the exception of vehicles 2 and 15 which are tracked equally well by the EEA-approach). Their trajectory includes a right-turn within an area with both lane markers and partial occlusions by traffic lights. Such situations constitute a significant challenge for edge-based algorithms. This structured background, however, helps to estimate Optical Flow correctly, so the combined algorithm can

at least estimate speed and direction of the vehicles in this area which is enough to bridge tracking until edge estimation works reliably again. It can be assumed that a combination of EMCA and OF adaption would perform similarly.

5   Discussion and Conclusions

Both the EMC and the EEA approach rely solely on greyvalue discontinuities for 3D model-based tracking. The EMC approach hypothesizes a probability distribution function (pdf) for greyvalue discontinuities in the vicinity of model-segments resulting from the projection of a 3D body model into the image plane according to the current pose estimate. This pdf determines the probability to find greyvalue discontinuities in the vicinity of a model projection. An expectation maximisation step (re-)estimates corrections to the pose parameters in order to maximize this figure-of-merit.

A weakness of the EEA approach becomes particularly discernible in cases where the number of EEs is small such that any incorrect associations between EEs and model-segments begin to distort the pose parameter update step. Another difficulty for the EEA approach is presented by heavily textured image regions with many EEs whose orientation accidentally matches that of model-segments within a given tolerance. It is here where the *local* restriction for EE association with a given model-segment appears to work against this approach, despite the fact that the orientation difference is taken into account explicitly. This observation suggests to reconsider the manner in which position and orientation information of each EE is combined into a compound distance function between an EE and a model-segment. It should be noted, however, that the number of cases where this effect begins to endanger the tracking performance appears to be small with the result that the EEA approach performs still better than the EMC approach. The statistics are not yet large enough, however, to proclaim a clear superiority of one approach over the other. In addition, attention has to be drawn to the fact that no efforts have been made to optimize *either* parameter-set for the implementation used in these experiments beyond copying values available from original publications.

In general, our experiments support the hypothesis that the incorporation of OF-estimates into the tracking phase is more important than the particularities of how EEs are taken into account. This finding is compatible with statements which can be found, e. g., in [3]. One may now argue '... the integration of optical flow information significantly improves the tracking result, which is logic and expected. Such a study does not result in new understanding of the algorithms and most of the results can be predicted'. An analogous critique could be formulated like '... the difference between the two approaches being compared is small compared to the benefits of adding a new source of information from the optical flow. This weakens the importance of the original comparison.'

It is admitted freely that the comparison may be statistically weak, despite the fact that the second image sequence comprises about ten times the number of vehicles than the PETS-2000 sequence, not to mention the differences in

vehicle movements and recording conditions. The *experimental* evidence is *significant* that the difference between the two approaches with respect to handling greyvalue transition information appears to be small compared to the effect of incorporating optical flow estimates into the state-vector update. There does not yet exist an accepted theory in this field, however, which would allow *to derive* such a conclusion. 'Hunches' may vary from scientist to scientist or from group to group. Such an observation should not distract, though, from the fact that more attention has to be paid to the manner in which information from greyvalue discontinuities is exploited by the tracking process. Even if such effects may become relevant only rarely during the tracking phase, a negative effect caused by an inappropriate handling of greyvalue transition information can disrupt the tracking process and thereby endanger efforts to maintain the identity of a body moving within the scene.

It is unclear to which extent the consequences of unavoidable simplifications counteract theoretically appealing advantages of alternative approaches. Such a question can only be answered by careful comparative experiments of the kind reported here which – according to the authors' knowledge – constitute by no means the standard manner of research in the field of model-based tracking. A re-appraisal of our original assessment suggests the following conclusions:

1. Comparison experiments, which do not simply reproduce earlier ones under identical conditions, are necessary *even if they largely confirm prior expectations.* The field does not command a proven and accepted theory which allows to predict the outcome of an experiment beyond reasonable doubt.
2. Given the small margin in the currently reported experiments, we refrain from making any final decisions regarding superiority of one approach over the other.
3. Handling of greyvalue transition information during model-based tracking can become crucial in some cases, but need not determine the outcome in general. It thus is not easy to investigate which approach (in which implementation and with which parameterisation) is superior to which alternatives. This observation suggests that larger test samples need to be evaluated, implying a *substantial* increase of efforts to perform such experiments.
4. In addition to the preparation of experimental runs with a larger number of vehicles, other aspects have to be investigated which might potentially influence the outcome of such experiments, e. g., details of the vehicle and/or motion model.
5. If the difference of alternative ways to handle the greyvalue transition information appears to be small, *all* potentially relevant influences have to be uncovered and investigated, including theoretical simplifications and numerical effects.
6. True progress towards better understanding model-based tracking thus appears to be more expensive in terms of eperimental expertise and theoretical analysis than has been anticipated by many in the field.

Additional experiments have been prepared in order to answer at least some of the new questions raised.

Acknowledgments

References

1. T. Brox, M. Rousson, R. Deriche, and J. Weickert: *Unsupervised Segmentation Incorporating Color, Texture, and Motion.* In: Proc. 10th International Conference on Computer Analysis of Images and Patterns CAIP'03, 25-27 August 2003, Groningen, The Netherlands. LNCS 2756, pp. 353–360, Springer-Verlag: Berlin·Heidelberg·New York/NY 2003.
2. H. Dahlkamp: *Untersuchung eines Erwartungswert-Maximierung (EM)-Kontur-Algorithmus zur Fahrzeugverfolgung* (in German). 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. N. Paragios and R. Deriche: *Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision.* Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics, (March/June 2002) 249-268.
5. 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. Springer LNCS 2350, pp. 3-17.
6. 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.
7. http://i21www.ira.uka.de/image_sequences/
8. ftp://pets2001.cs.rdg.ac.uk/PETS2000/test_images/

A   Outline of the Two Algorithms

A.1   Edge Element Association Algorithm

During edge element adaption, a new state is searched which minimizes the distance between edge elements and projected model segments. An edge element $\mathbf{e} = (u_e, v_e, \phi_e)^T$ represents a local maximum of the gradient norm in gradient direction ϕ_e at the position (u_e, v_e). As illustrated in Figure 5, the difference measure between an edge element $\mathbf{e}$ and a projected model segment $\mathbf{m}$ considers both the Euclidean distance and the difference between the measured and predicted gradient direction:

$$
\begin{aligned}
d_{\mathbf{m}}(\mathbf{e}, \mathbf{x}) &= \frac{b}{\cos \Delta} \\
&= \frac{-(u_e - u_m)\sin\theta + (v_e - v_m)\cos\theta}{\cos\left(\phi_e - (\theta + \frac{\pi}{2})\right)} \quad .
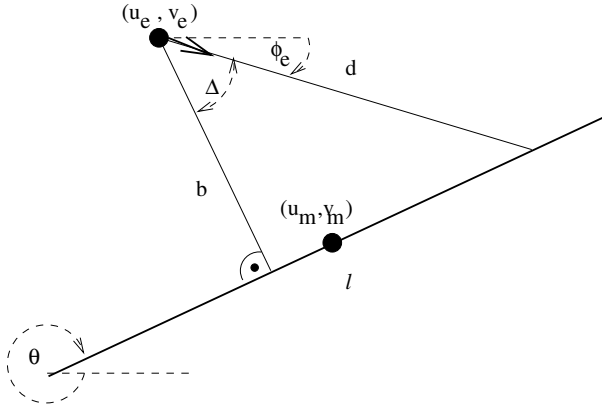\end{aligned}
\tag{1}
$$

Fig. 5. Distance metric between an edge element and a projected model segment
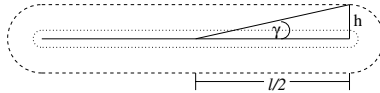

Fig. 6. Region of a model segment in which edge elements are accepted. This acceptance region is specified based on the length of (the visible part of) the model segment.

It is assumed that $d_{\mathbf{m}}$ is normally distributed with zero-mean and variance σ^2. The distance measure induces 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 with the smallest Mahalanobis distance. Furthermore, only those edge elements within an acceptance area around visible model segments are taken into account (see Figure 6).

A.2   Expectation Maximisation Contour Algorithm

According to the publication [5], the assumptions underlying this algorithm are

1. Greyvalue differences of adjacent pixels have a prior probability distribution

$$P(\Delta I = \Delta i) = \text{const} \cdot \exp\left(-\left|\frac{\Delta i}{\lambda}\right|^{\beta}\right) \qquad (2)$$

with $\beta = 0.5$ and $\lambda = \left(\frac{E\{|\Delta I|^{\beta}\}}{2}\right)^{1/\beta}$, where ΔI is a random variable for the difference between greyvalues (depending on the distance between the measured image points).

2. Greyvalue differences of pixels across an edge have a uniform distribution.
3. The visible shape of the model deviates from the projection of the model into the image plane by gaussian variability orthogonal to the projected model segments with a variance equal to the projection of σ into the image ($\sigma \in [0.1m, 0.5m]$).

While these assumptions sound reasonable, the algorithm contains in addition the following implicit assumptions:

4. All modelled edges are visible in the image as a a greyvalue *discontinuity*, because the model assumes that they affect only one greyvalue difference. Slow and gradual differences which commonly occur on aerodynamically shaped cars, however, affect more than one greyvalue difference.
5. Each line orthogonal to a projected model segment contains exactly one grey-value discontinuity. In practice, this assumption can be violated by having zero discontinuities (for example if there is no vehicle-background-contrast) or more than one (in tapered model-segment configurations where model segments are close to each other).

In order to combine these assumptions, the algorithm selects points on the projected model segments, called *sample points* $\boldsymbol{\mu}_k \in I\!\!R^2 (k = 1..n)$ in this document. Subsequently, the algorithm looks for an image discontinuity on a line of points $\boldsymbol{\nu}_{k,j} = \boldsymbol{\mu}_k + j\, d\, \mathbf{r}_k$ ($j = -8..8, |\mathbf{r}_k| = 1$) orthogonal to the projected model segment. At each of these measure points $\boldsymbol{\nu}_{k,j}$, it computes the greyvalue difference

$$\Delta i(\boldsymbol{\nu}_{k,j}) = I(\boldsymbol{\mu}_k + (j + 1/2)d\, \mathbf{r}_k) - I(\boldsymbol{\mu}_k + (j - 1/2)d\, \mathbf{r}_k) \tag{3}$$

using the image greyvalues $I(\cdot)$. The distance d between the measure points is chosen to be $\mathrm{proj}(\sigma)/4$.

Then, the algorithm computes the probability of all greyvalues on the measure point line to be $P(\Delta \mathbf{i}_k) = \prod_{m=-8}^{8} f_L(\Delta i(\boldsymbol{\nu}_{k,m}))$. The conditional probability of this given that the object border (specified by the random variable $\mathbf{N}_k$) crosses near measure point $\boldsymbol{\nu}_{k,j}$ is given by

$$P(\Delta \mathbf{i}_k | \mathbf{N}_k = \boldsymbol{\nu}_{k,j}) = \prod_{m=-8}^{j-1} f_L(\Delta i(\boldsymbol{\nu}_{k,m})) \frac{1}{256} \prod_{m=j+1}^{8} f_L(\Delta i(\boldsymbol{\nu}_{k,m})) \tag{4}$$

$$= \frac{P(\Delta \mathbf{i}_k)}{256\, f_L(\Delta i(\boldsymbol{\nu}_{k,j}))}. \tag{5}$$

Assumption 3 expands to a prior distributing for the object border location

$$P(\mathbf{N}_k = \boldsymbol{\nu}_{k,j} | \boldsymbol{\mu}_k) = \mathrm{const} \cdot \exp\left(-\frac{\|\boldsymbol{\nu}_{k,j} - \boldsymbol{\mu}_k\|^2}{2\mathrm{proj}(\sigma)^2}\right) = \mathrm{const} \cdot \exp\left(-\frac{|j|}{16}\right). \tag{6}$$

Using these two formulas, the following figure-of-merit function is defined which has to be optimized:

$$l(\boldsymbol{\mu}) = \ln \frac{P(\Delta \mathbf{i}|\boldsymbol{\mu})}{P(\Delta \mathbf{i})} = \ln \prod_{k=1}^{n} \frac{P(\Delta \mathbf{i}_k|\boldsymbol{\mu}_k)}{P(\Delta \mathbf{i}_k)} \tag{7}$$

$$= \ln \prod_{k=1}^{n} \sum_{j=-8}^{8} \frac{P(\Delta \mathbf{i}_k|\boldsymbol{\nu}_{k,j}) \, P(\boldsymbol{\nu}_{k,j}|\boldsymbol{\mu}_k)}{P(\Delta \mathbf{i}_k)} \quad . \tag{8}$$

To perform the maximisation of this function, the Bayes formula is applied to combine (5) and (6) to a posterior probability for the border location:

$$P(\boldsymbol{\nu}_{k,h}|\Delta \mathbf{i}_k, \boldsymbol{\mu}_k) = \frac{P(\Delta \mathbf{i}_k|\boldsymbol{\nu}_{k,h}, \boldsymbol{\mu}_k)P(\boldsymbol{\nu}_{k,h}|\boldsymbol{\mu}_k)}{P(\Delta \mathbf{i}_k|\boldsymbol{\mu}_k)} \tag{9}$$

$$= \frac{P(\Delta \mathbf{i}_k|\boldsymbol{\nu}_{k,h})P(\boldsymbol{\nu}_{k,h}|\boldsymbol{\mu}_k)}{\sum_{j=-l}^{l} P(\Delta \mathbf{i}_k|\boldsymbol{\nu}_{k,j})P(\boldsymbol{\nu}_{k,j}|\boldsymbol{\mu}_k)}. \tag{10}$$

The expectation value of this distribution gives an estimator for the border location: $\hat{\boldsymbol{\nu}}_k = E\{\boldsymbol{\nu}_k|\Delta \mathbf{i}_k, \boldsymbol{\mu}_k\}$.

Finally, it can be shown that setting $\boldsymbol{\mu}$ to

$$\boldsymbol{\mu} = \min_{\boldsymbol{\mu}} \sum_{k=1}^{n} \|\hat{\boldsymbol{\nu}}_k - \boldsymbol{\mu}_k\|^2 \tag{11}$$

increases the figure-of-merit function $l(\boldsymbol{\mu})$ and the iterated execution of this approach eventually leads to a (possibly local) maximum. Therefore, it is possible not only to design a single figure-of-merit function measuring the overall quality of model fitting, but also to derive a greedy algorithm for finding its optimum. Furthermore, it is also possible to take $\hat{\boldsymbol{\nu}}_k$ as the measurement of a Kalman filter and $\boldsymbol{\mu}_k$ as the predicted measurement, which also results in equ. (11) to be minimized, just with a second additive term minimizing the squared distance between estimated and corrected vehicle state.