Evaluation of Motion Segmentation Quality for Aircraft Activity Surveillance

Josep Aguilera, Horst Wildenauer, Martin Kampel
Pattern Recognition and Image Processing Group
Vienna University of Technology
Favoritenstr.9, 183-2, A-1040 Vienna, Austria
{agu, wilde, kampel}@prip.tuwien.ac.at

Mark Borg, David Thirde, James Ferryman
Computational Vision Group
The University of Reading
Whiteknights, Reading, RG6 6AY, UK
{m.borg, d.j.thirde, j.m.ferryman}@reading.ac.uk

Abstract

Recent interest has been shown in performance evaluation of visual surveillance systems. The main purpose of performance evaluation on computer vision systems is the statistical testing and tuning in order to improve algorithm's reliability and robustness. In this paper we investigate the use of empirical discrepancy metrics for quantitative analysis of motion segmentation algorithms. We are concerned with the case of visual surveillance on an airport's apron, that is the area where aircrafts are parked and serviced by specialized ground support vehicles. Robust detection of individuals and vehicles is of major concern for the purpose of tracking objects and understanding the scene. In this paper, different discrepancy metrics for motion segmentation evaluation are illustrated and used to assess the performance of three motion segmentors on video sequences of an airport's apron.

1 Introduction

Over the last decade, increasing interest in the field of visual surveillance has led to the design of a plethora of systems for automated visual tracking of moving objects. In many of these systems, the detection and segmentation of moving objects represent the first step on which subsequent processing stages heavily depend. Deteriorations of segmentation quality can have a severe impact on the performance of a surveillance system, and thus, the ability to effectively segment moving objects under a wide range of disturbing conditions is a critical requirement.

Although considerable efforts have been spent on the development of robust motion segmentation algorithms, no comparable attention has been given to their evaluation. As a consequence, there is rising demand for quantitative evaluation of segmentation quality in order to assess the reliability of existing approaches and to facilitate their comparability [7]. Especially in the case of outdoor surveillance, where illumination changes, weather conditions, shadows, and occlusions strongly impact the segmentation quality.

Empirical methods developed for the assessment of motion segmentation quality can be characterised by their basis of evaluation: (1) *goodness methods* that operate without reference segmentations (ground truth), and (2) *discrepancy methods* based on the use of ground truth.

Recently, a set of performance metrics for motion segmentation evaluation have been proposed in the case of non available ground truth. In [5] Correia and Pereira present a methodology based on the idea of measuring intra-object homogeneity features and inter-object disparity features. Erdem et al. [9] propose two error metrics based on colour information and motion features.

Ellis [7] proposes error metrics based on correct and false matches between ground truth and observations. Erdem et al. [8] suggest the use of spatio-temporal segmentation measures for object-based motion segmentation. Both an evaluation methodology and metrics for video segmentation quality analysis separating individual object evaluation from overall evaluation have been introduced by Correia and Pereira [4]. Perceptually-weighted criteria, which take into account visually desirable properties of reference segmentations, have been designed by Villegas and Mariachal [15] and Cavallaro et al. [3].

To facilitate and accelerate the creation of ground truth, semi-automatic frameworks such as ViPER [6] and ODViS [11] have been designed. In [13] Schlögl et al. present a fully-automatic evaluation framework, which introduces ground truth by the use of synthetic objects.

In this paper, we investigate the use of discrepancy metrics for the quantitative analysis of spatial accuracy of segmentations provided by motion segmentors. Specifically, we are concerned with the case of visual surveillance on an airport's apron addressed by the European AVITRACK project [1]; that is, the robust detection of individuals and vehicles for the purpose of tracking and categorising objects in the scene. We evaluate three motion segmentation algorithms on airport's apron sequences across a range of conditions. The conditions studied are: Varying weather and illumination conditions, different camera viewpoints, and different scene complexity. The motion segmentors used in the evaluation were: Mixture of Gaussians [14], Colour and Edge fusion [10], and Colour Mean and Variance [16].

The remainder of this paper is organised as follows: Section 2 gives a short introduction into the European AVITRACK project. A methodology for object-based segmentation evaluation is presented in Section 3. Discrepancy metrics for the evaluation of segmentation quality are described in Section 4. Section 5 presents experimental results and discusses them. Finally, in Section 6 we draw the conclusions.

2 The AVITRACK Project

The European AVITRACK [1] project addresses the specific case of automatically supervising commercial aircraft servicing operations at a designated airport. It is intended to add an additional technological layer to a specific airport in order to improve efficiency and security of the handling services. Visual surveillance in the AVITRACK project includes tracking and categorisation of vehicles and individuals. The apron on which the prototype is installed is observed by eight cameras in order to see a maximum of operations around the aircraft. Figure 1 illustrates the complexity of the specific application scenario showing an apron with tracked vehicles and individuals.

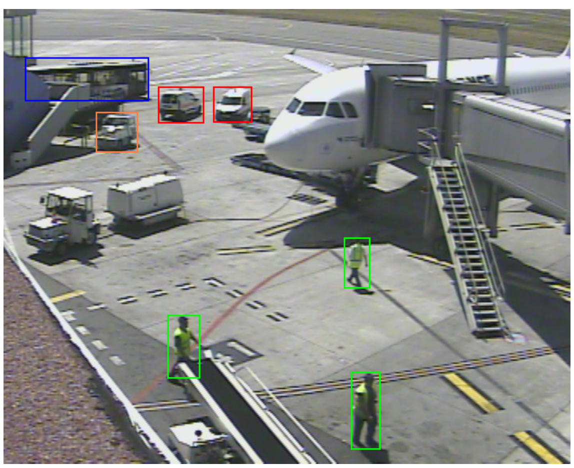


Figure 1: Airport's apron: tracking of vehicles and individuals

Object categorisation is necessary in order to perform a high level interpretation of the tracking results and to recognise the activities on the airport's apron. Such categorisation strongly depends on the motion segmentor's output. Hence, one aim of the project is to apply and analyse robust statistical techniques for motion segmentation. 16 motion segmentation algorithms were implemented and evaluated [2]. The performance criteria for the evaluation were 'Susceptibility to Noise', 'Robustness to Illumination Changes', 'Detection Sensitivity' and 'Speed'. Of these, the following three algorithms have shown acceptable susceptibility to noise and good detection sensitivity: Mixture of Gaussians [14], Colour and Edge fusion [10] and Colour Mean and Variance [16]. A quantitative evaluation of the selected algorithms will illustrate which motion segmentor provides the best segmentation quality results therefore helping to come to a decision about the approach to be chosen for our application.

3 Performance Evaluation of Motion Segmentors

One of the most popular approaches to performance evaluation of motion segmentors is to address it as a simple two-class (foreground/background) segmentation evaluation problem [7, 6]. That is, in the case of several foreground objects, these are aggregated into a compound foreground region, which is compared to the compound foreground estimated by the motion segmentor. While such an approach can provide a meaningful quantification of overall frame-wise segmentation quality, it fails to deliver insight into the segmentation algorithms behaviour regarding individual objects. Unfortunately, this shortcoming limits it's usefulness in the area of visual surveillance, where the stable segmentation of individual objects is usually favoured over a good frame-wise segmentation.

In [4, 8] methodologies for object-based quality evaluation have been proposed. These approaches combine the results of individual object evaluation into an overall measure, assuming that a unique correspondence between estimated and reference objects exists. In the context of video tracking, however, such one-to-one mappings are seldom available. Single objects may be missed or split into disconnected regions, and estimated segmentations often overlap with more than one reference object [15].

In the next section, we propose a methodology for segmentation evaluation, facilitating the assessment of individual objects segmentation quality while taking into account the segmentation problems mentioned above.

Evaluation Methodology

We loosely follow the methodology for object-based segmentation quality evaluation introduced by Correia and Pereira [4]. The methodology for quantitative analysis of motion segmentors is described in the following:

1. **Establishing of correspondence**:

   (a) All regions overlapping the reference object are assigned to the object. (Multiple assignments of estimated regions to reference objects are allowed.)

(b) Splitting of multiply assigned regions: When estimated regions overlap more than one reference object, regions are split by assigning individual pixels to the closest object.

(c) If objects are missed in the segmentation, the missing regions are processed with all reference pixels as false negatives.

(d) Estimated foreground regions not associated with any reference objects are mapped onto the reference background; thus treated as false positives with respect to the background mask.

2. **Individual object segmentation evaluation**:
Empirical discrepancy metrics are used to compute the segmentation quality for each reference object.

3. **Overall segmentation quality evaluation**:
Results from individual object evaluation are weighted by their relevance and combined into an overall quality metric.

4 Evaluation Metrics for Motion Segmentation

The quality of motion segmentation can in principle be described by two characteristics. Namely, the spatial deviation from the reference segmentation, and the fluctuation of spatial deviation over time. In this work, however, we concentrate on the evaluation of spatial segmentation characteristics. That is, we will investigate the capability of the error metrics listed below, to describe the spatial accuracy of motion segmentations.

- False negative rate (fnr) and false positive rate (fpr)
These normalised metrics are based on pixel-wise mismatches between ground truth and observations in a frame [7].

$$fnr = \frac{N_{fn}}{N_{tp} + N_{fn}} \tag{1}$$

$$fpr = \frac{N_{fp}}{N_{fp} + N_{tn}} \tag{2}$$

where N_{fn} and N_{fp} denote the number of false negative and false positive pixels respectively. N_{tn} and N_{tp} are the number of true negatives and true positives.

- Misclassification penalty (*MP*)
The obtained segmentation is compared to the reference mask on an object-by-object basis; misclassified pixels are penalized by their distances from the reference objects border [8].

$$MP = MP_{fn} + MP_{fp} \tag{3}$$

with

$$MP_{fn} = \frac{\sum_{j=1}^{N_{fn}} d_{fn}^{j}}{D} \tag{4}$$

$$MP_{fp} = \frac{\sum_{k=1}^{N_{fp}} d_{fp}^{k}}{D} \tag{5}$$

Here, d_{fn}^{j} and d_{fp}^{k} stand for the distances of the j^{th} false negative and k^{th} false positive pixel from the contour of the reference segmentation. The normalised factor D is the sum of all pixel-to-contour distances in a frame.

- Rate of misclassifications (*RM*)
The average normalised distance of detection errors from the contour of a reference object is calculated using [13]:

$$RM = RM_{fn} + RM_{fp} \tag{6}$$

with

$$RM_{fn} = \frac{1}{N_{fn}} \sum_{j=1}^{N_{fn}} \frac{d_{fn}^{j}}{D_{diag}} \tag{7}$$

$$RM_{fp} = \frac{1}{N_{fp}} \sum_{k=1}^{N_{fp}} \frac{d_{fp}^{k}}{D_{diag}} \tag{8}$$

N_{fn} and N_{fp} denote the number of false negative and false positive pixels respectively. D_{diag} is the diagonal distance within the frame.

- Weighted quality measure (*QMS*)
This measure quantifies the spatial discrepancy between estimated and reference segmentation as the sum of weighted effects of false positive and false negative pixels [15].

$$QMS = QMS_{fn} + QMS_{fp} \tag{9}$$

with

$$QMS_{fn} = \frac{1}{N} \sum_{j=1}^{N_{fn}} w_{fn}(d_{fn}^{j}) d_{fn}^{j} \tag{10}$$

$$QMS_{fp} = \frac{1}{N} \sum_{k=1}^{N_{fp}} w_{fp}(d_{fp}^{k}) d_{fp}^{k} \tag{11}$$

N is the area of the reference object in pixels. Following the argument that the visual importance of false positives and false negatives is not the same, and thus they should be treated differently, the weighting functions w_{fp} and w_{fn} were introduced:

$$w_{fp}(d_{fp}) = B_1 + \frac{B_2}{d_{fp} + B_3} \tag{12}$$

$$w_{fn}(d_{fn}) = C \cdot d_{fn} \tag{13}$$

In our work, we used the parameters $B_1 = 19$, $B_2 = -178.125$, $B_3 = 9.375$, and $C = 2$, resulting in the weighting functions depicted in Figure 2. One can see, that missing (false negative) pixels gain more importance with increasing distance than added foreground pixels. Thus, our weighting favours algorithms which provide larger foreground estimates, over more conservative ones. Naturally, the choice of weighting functions depends on the targeted application; see [12, 3] for examples.

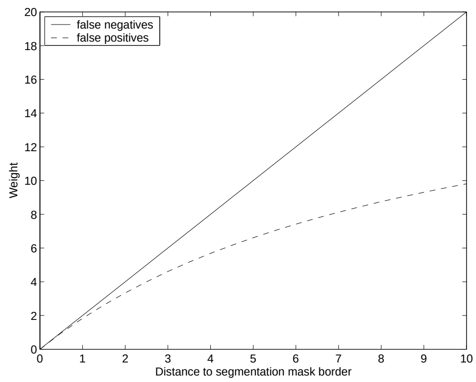


Figure 2: Weighting functions for false positives and false negatives.

5 Experiments

We evaluated three different motion segmentation algorithms on airport's apron datasets using the methodology and the error metrics introduced in Section 3 and 4 respectively. These algorithms were Colour and Edge fusion (CEF), Mixture of Gaussians (MoG), and Colour Mean and Variance (CMV).

5.1 Data sets

In outdoor surveillance applications such as AVITRACK a wide range of disturbing conditions are influencing the scenery. Varying Illumination, different weather conditions, and scene complexity are issues to be considered during the performance evaluation of motion segmentation algorithms. According to these points, we have chosen the following representative apron sequences:

S3 - Camera 2: Shows an aircraft parking on the apron. Moreover, it contains individuals and vehicles such as conveyor belts, transporters with dollies and a stair vehicle working on maintenance tasks.

S4 - Camera 5: A tanker and a service vehicle move across the apron. A ground power unit (GPU) parks in the maintenance area and a person leaves the GPU.

Figure 3: Sample image (frame 5814) from the S21-Camera 7 sequence (foggy conditions).

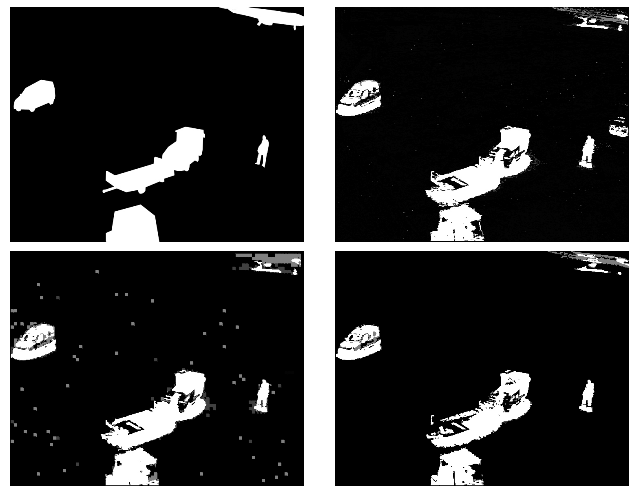


Figure 4: The manually created ground truth and CEF, MoG, and CMV detection results for frame 5814 taken from sequence S21-Camera 7 (clockwise from top-left).

S5 - Camera 5: Three individuals walk around the apron while a transporter and a GPU park in the maintenance area. The sequence contains reflections either caused by liquid on the ground or by the paint on the ground.

S8 - Camera 6: A GPU enters in the scene and two individuals walk on the apron. The sequence presents in close-up a transporter with dollies in movement. As a night sequence, the vehicle lamps produce large reflections on the ground.

S21 - Camera 7: The sequence contains individuals walking on the apron. Vehicles in movement such as a GPU, a tanker, a catering vehicle and service vehicles are shown.

S26 - Camera 6: A group of individuals walking and a conveyor belt in movement are shown. An aircraft starts its departure.

Sequences S3-Camera 2, S4-Camera 5, and S5-Camera 5 were acquired under bright daylight. S21-Camera 7 and

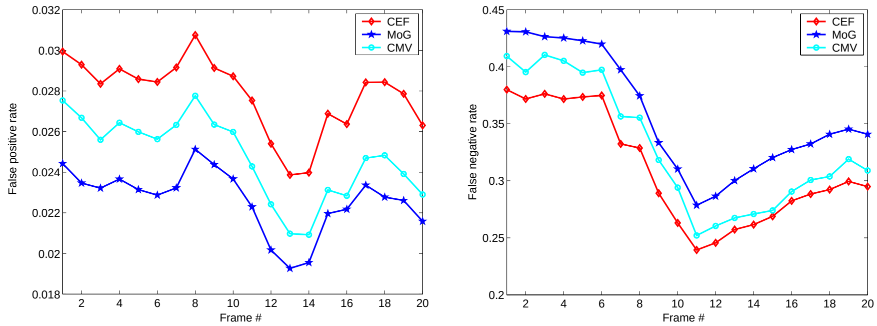


Figure 5: False positive (left) and false negative (right) error rates versus number of frames for the S21-Camera 7 sequence.

S26-Camera 6 were taken under foggy conditions. S8-Camera 6 is a night sequence. All sequences were stored at a size of 720x576 pixels, and at a frame rate of 12.5 fps. 20 reference frames were extracted from each of the six sequences. Additionally, 25 frames from S3-Camera 2 and 10 frames from S21-Camera 7 were selected, resulting in a total of 155 reference images. For these, ground truth segmentation masks have been created manually.

5.2 Results

To illustrate the application of object-based discrepancy metrics, we conducted an experiment using 20 frames from the particularly challenging sequence S21-Camera 7. This sequence was acquired under foggy conditions and features five objects (aircraft, auto, transporter, GPU and a pedestrian) moving at different distances from the camera (see Figure 3) . Results of the segmentation process are shown in Figure 4. Objects such as the aircraft are only partially detected due to the achromaticity of the scene. Shadows are segmented as part of the mobile objects, and holes and fragmentations appear where objects have the same colour as the background.

At first, *fpr* and *fnr* were calculated for whole frames. The results of this evaluation are depicted in Figure 5. The algorithms achieve relatively low false positive rates (1.9% - 3.1%) at the cost of high false negative rates between 24% and 44%, caused by the achromatric nature of the scene. Note, that the CEF algorithm generates a ghost (close to the pedestrian), resulting in an increased false positive rate.

In addition, the weighted quality measure *QMS*, the misclassification penalty *MP*, and the rate of misclassifications *RM* were computed separately for each object (see Figure 6). We obtained the overall object-based segmentation quality as an average of the individual object's segmentation errors. At frame one, three moving objects are present in the scene (aircraft, auto and transporter). A GPU and a pedestrian enter the scene at frame five and eight respectively. These objects produce a lower individual *QMS* error than the aircraft, auto or transporter. (Figure 6 (d, e, f)). This is also reflected in Figure 6 (a) which shows the decrease in overall *QMS* after frame five and eight. Notice, that the *MP* metric is dominated by the segmentation error of the transporter (Figure 6 (g, h, i)). This can be explained by the fact that the transporter generates a larger segmentation error due to it's size, which, in contrast to *QMS* is not normalized by the area of the reference object mask (for details see Section 4). The rate of misclassifications *RM* (Figure 6 (c)) provides unstable results for the selected data set.

To further analyse the usefulness of the object-based discrepancy metrics, we conducted an experiment on 10 images of the S21-Camera 7 sequence using the MoG algorithm. For this purpose, we varied the factor *k* for the matching threshold of the Gaussians (pixel values outside *k* standard deviations of a distribution are classified as foreground) between 0.5 and 3.25 in steps of 0.25. In Figure 7 a sample image, showing three moving pedestrians is depicted. The corresponding ground truth image and the segmentation results at *k*=1, *k*=2.25, and *k*=3.25 are depicted in Figure 8. The misclassification penalty MP, the weighted quality measure QMS, and the rate of misclassifications RM were computed for each pedestrian. The overall frame segmentation errors were obtained by averaging over the individual objects (see Figure 9).

One can see that there is a high amount of perceptual agreement between the segmentation results depicted in Figure 8 and the objective values of the overall (per frame) MP and QMS. Again, less stable results are obtained using the rate of misclassifications RM. This can be explained by it's sensitivity with respect to certain types of segmentation errors. The RM computes the average distance of misclassified pixels from the reference object's contour. Therefore already a small number of erroneous pixels can produce a relatively high error rate.

The performance results of the motion segmentation algorithms for all 155 test images are depicted in Table 1. False positives and false negatives of *QMS*, *MP*, and *RM* error metrics were computed for each object and averaged per frame. Furthermore, *fpr* and *fnr* (ER) were calculated for whole frames. *CMV* produces the best false negative

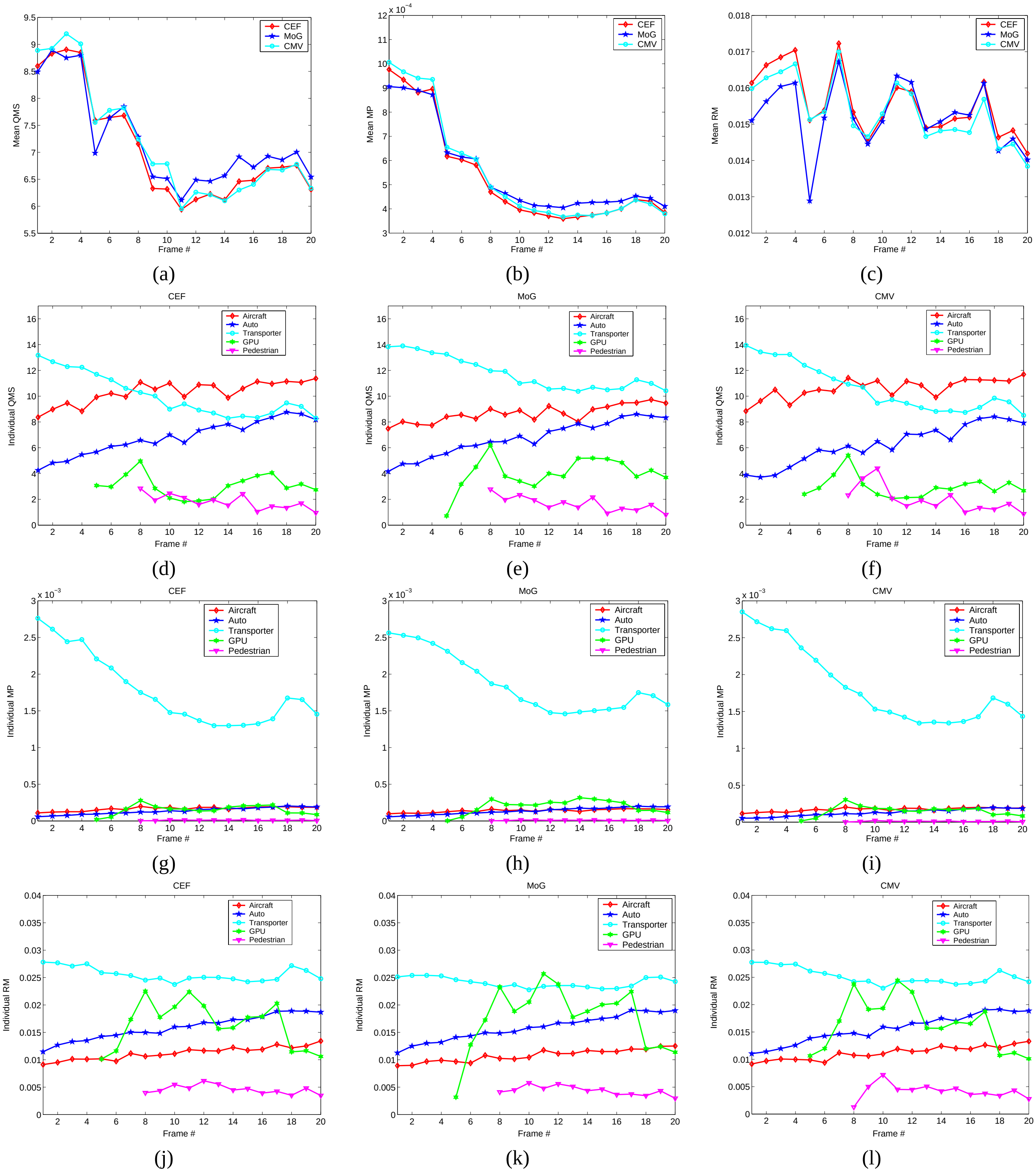
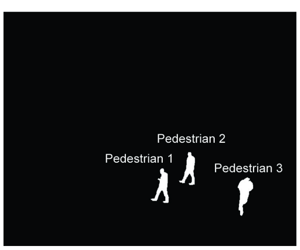


Figure 6: Average and individual object segmentation errors. (a, b, c) Average *QMS*, *MP* and *RM* object-based segmentation errors of the motion segmentors, (d, e, f) *QMS* individual object segmentation error of CEF, MoG and CMV, (g, h, i) *MP* individual object segmentation error of CEF, MoG and CMV, (j, k, l) *RM* individual object segmentation error of CEF, MoG and CMV.

	S3-Camera 2		S4-Camera 5		S5-Camera 5		S8-Camera 6		S21-Camera 7		S26-Camera 6	
	FP	**FN**	**FP**	**FN**	**FP**	**FN**	**FP**	**FN**	**FP**	**FN**	**FP**	**FN**
QMS-CEF	5.517	1.576	4.674	1.324	4.636	0.445	1.898	10.74	12.37	4.501	1.45	0.849
QMS-MoG	4.653	2.888	4.621	1.407	4.534	0.551	1.901	16.49	2.454	6.237	0.716	2.089
QMS-CMV	4.843	2.026	4.632	1.199	4.905	0.392	3.28	6.61	2.902	5.727	0.527	1.182
MP-CEF	6.3e-4	5.5e-5	4.44e-4	4.5e-5	4.4e-5	2.3e-6	1.1e-3	2.1e-3	3.7e-4	3.6e-4	4.5e-5	1.4e-5
MP-MoG	5.93e-4	9.7e-5	4.41e-4	4.8e-5	5.3e-5	2.9e-6	4.5e-4	3.7e-3	2.6e-4	3.7e-4	2.8e-5	3.3e-5
MP-CMV	5.95e-4	6.8e-5	4.43e-4	3.9e-5	5.9e-5	2.1e-6	2.1e-3	1.4e-3	3.1e-4	3.3e-4	1.6e-5	1.3e-5
RM-CEF	4.53e-3	1.06e-2	5.4e-3	1.13e-2	2.44e-3	8.1e-3	1.63e-2	0.013	8.8e-3	8.1e-3	3.7e-3	3.8e-3
RM-MoG	4.94e-3	1.11e-2	5.3e-3	1.15e-2	2.43e-3	8.2e-3	1.75e-2	0.011	8.5e-3	6.6e-3	3.6e-3	3.2e-3
RM-CMV	4.51e-3	1.09e-2	5.1e-3	1e-2	2.16e-3	8e-3	1.62e-2	0.017	8.2e-3	0.011	1.8e-3	2.7e-3
ER-CEF	1.72e-2	0.171	1.98e-2	0.131	7.2e-3	0.076	0.031	0.375	0.032	0.232	0.01	0.09
ER-MoG	1.57e-2	0.276	1.91e-2	0.139	5.7e-3	0.092	0.03	0.564	0.018	0.375	5.2e-3	0.228
ER-CMV	1.58e-2	0.213	1.93e-2	0.121	6.5e-3	0.075	0.046	0.243	0.021	0.325	2.5e-3	0.184

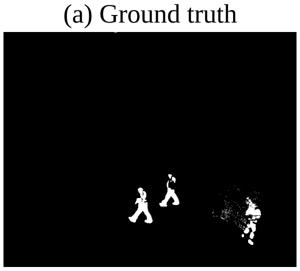
Table 1: Performance results of the motion segmentation algorithms on the apron data sets.

Figure 7: Sample image (frame 1659) from the S21-Camera 7 sequence.

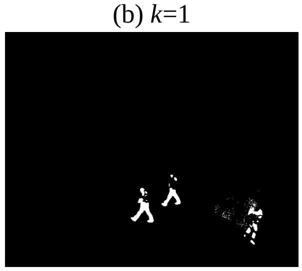


(a) Ground truth

(b) k=1



(c) k=2.25



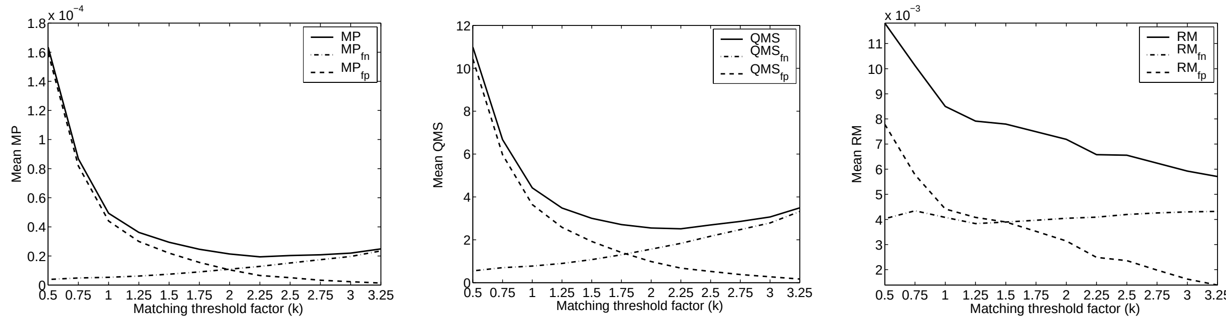
(d) k=3.25

Figure 8: The ground truth and the MoG segmentation results for different matching threshold factors k.

error results (for all metrics) on sequences S4-Camera 5, S5-Camera 5, S8-Camera 6 (not for *RM*) and S26-Camera 6 (not for *fnr*). All segmentors provide similar results w. r. t. effects caused by strong illumination (such as cast shadows) (see S3-Camera 2 and S4-Camera 5 results). A high amount of false negatives is produced by the motion segmentors on the night sequence S8-Camera 6. *CMV* and *MoG* produce the best false positive error results on the sequences with the presence of fog (see S21-Camera 7 and S26-Camera 6 results). The *CMV* exhibited the most promising results for the selected data sets. Therefore, for our application scenario we choose the *CMV* as reference motion segmentation algorithm.

6 Conclusions

In this paper we proposed an object-based methodology and four discrepancy metrics for motion segmentation evaluation. The metrics usability was tested on video data acquired at an airports apron using three selected motion segmentation algorithms. Representative scenes according to variability in illumination condition, weather conditions, and complexity of scene activity were selected. Two of the tested metrics, namely the misclassification penalty *MP* and the weighted quality measure *QMS*, have shown their usefulness in providing meaningful quantifications of segmentation quality. An advantage of these metrics is that they can be applied on an object-by-object basis. Due to *RM*'s sensitivity to segmentation errors it is a less robust error metric then *MP* and *QMS*. *fnr* and *fpr* provide a valuable insight into a segmentation algorithms behaviour regarding an overall scene. Our future research direction is to determine how factors like the defragmentation of estimated objects, as well as the overlap with more than one reference object,



(a) (b) (c)

Figure 9: Discrepancy measures MP (a), QMS (b), and RM (c) versus matching threshold factors k for the pedestrian sequence S21-Camera 7.

can be incorporated directly into the performance metrics.

Acknowledgments

This work is supported by the EU, grant AVITRACK (AST3-CT-2003-502818). The AVITRACK consortium is composed of the following partners: SILOGIC (France), University of Reading (UK), INRIA (France), Toulouse-Blagnac Airport (France), Fedespace (France), Euro Inter (France), IKT (Norway), Tekever (Portugal), ARC (Austria) and PRIP (Austria).

References

[1] P. Blauensteiner and M. Kampel. Visual Surveillance of an Airport's Apron - An Overview of the AVITRACK Project. In *Workshop of the Austrian Association for Pattern Recognition, Hagenberg, Austria*, volume 179, pages 213–220, June 2004.

[2] M. Borg, D. Thirde, J. Ferryman, F. Fusier, F. Brémond, and M. Thonnat. An Integrated Vision System for Aircraft Activity Monitoring. In *Proceedings of the IEEE International Workshop on WAMOP-PETS*, pages 40–47, Jan 2005.

[3] A. Cavallaro, E.D. Gelasce, and T. Ebrahimi. Objective Evaluation of Segmentation Quality using Spatio-Temporal Context. In *Proceedings of the IEEE Conference on Image Processing*, pages 301–304, Sept 2002.

[4] P. Correia and F. Pereira. Objective Evaluation of Relative Segmentation Quality. In *Proceedings of International Conference on Image Processing,* , volume 1, pages 308–311, Sept 2000.

[5] P. Correia and F. Pereira. Stand-Alone Objective Segmentation Quality Evaluation. *Eurasip Journal on Applied Signal Processing*, 2002(4):389–400, 2002.

[6] D. Doermann and D. Mihalcik. Tools and Techniques for Video Performance Evaluation. In *Proceedings of the International Conference on Pattern Recognition*, volume 4, pages 167–170, Sept 2000.

[7] Tim Ellis. Performance Metrics and Methods for Tracking in Surveillance. In *IEEE Workshop on Performance Evaluation of Tracking and Surveillance*, pages 26–31, June 2002.

[8] C.E. Erdem and B. Sankur. Performance Evaluation Metrics for Object-Based Video Segmentation. In *10th European Signal Processing Conference*, pages 917–920, Sept 2000.

[9] C.E. Erdem, B. Sankur, and A.M. Tekalp. Performance Measures for Video Object Segmentation and Tracking. *Journal on Image Processing*, 13(7):937–951, July 2004.

[10] S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld. Detection and Location of People in Video Images Using Adaptive Fusion of Color and Edge Information. In *Proceedings of the International Conference on Pattern Recognition*, pages 4627–4631, 2000.

[11] C. Jaynes, S. Webb, R.M. Steele, and Q. Xiong. An Open Development Environment for Evaluation of Video Surveillance Systems. In *Proceedings of the International Workshop on Performance Evaluation of Tracking and Surveillance*, pages 32–39, June 2002.

[12] V. Mezaris, I. Kompatsiaris, and M.G. Strintzis. Still Image Objective Segmentation Evaluation Using Ground Truth. In *Proceedings of the Fifth COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication*, pages 99–14, Oct 2003.

[13] T. Schlögl, C. Beleznai, M. Winter, and H. Bischof. Performance Evaluation Metrics for Motion Detection and Tracking. In *Proceedings of the International Conference on Pattern Recognition*, volume 4, pages 519–522, August 2004.

[14] C. Stauffer and W.E.L Grimson. Adaptive Background Mixture Models for Real-Time Tracking. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 246–252, 1999.

[15] P. Villegas and X. Marichal. Perceptually-Weighted Evaluation Criteria for Segmentation Masks in Video Sequences. *IEEE Transactions on Image Processing*, (8):1092–1103, August 2004.

[16] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-Time Tracking of the Human Body. In *In IEEE Transactions on PAMI*, volume 19 num 7, pages 780–785, 1997.