Measuring the Quality of Figure/Ground Segmentations

Michael Arens, Claus Anderer
Fraunhofer Institute of
Optronics, System Technologies and Image Exploitation
Gutleuthausstr. 1
76275 Ettlingen
Germany
{michael.arens|claus.anderer}@iosb.fraunhofer.de

Abstract

Figure/Ground segmentation is of great interest within the analysis of video streams. We propose a new, continuous evaluation measure for figure/ground segmentation algorithms which allows for assessing the quality of a segmentation. The evaluation approach is based on set similarities and logic-motivated considerations. Results obtained with the new measure are shown for several well-known algorithms. Compared to existing detection rate/false alarm considerations, the proposed measure reveals advantages and shortcommings of different algorithms much better and can therefore be used as a selection criterion.

1. Introduction

Figure/Ground segmentation, i. e. the partition of an image in background and presently interesting foreground objects, is of great interest within the video analysis community. Obviously, separating a *figure* from some *ground* necessitates any algorithm for that problem to know of some property the one exhibits while the other does not. Examples of such properties are (color) intensity itself, texture (or the lack of it), and motion in the case of video streams.

The segmentation of single natural images into salient regions or "distinguished things" [1] can be regarded as a similar problem as figure/ground segmentation. A vast number of segmentation algorithms have been proposed, as well as several different measures to assess the quality of resulting image segmentations (see [1, 2] and especially [3, 4] for a survey of evaluation metrics). As has been discussed in [4] the segmentation of natural images is an ill-defined problem in the sense that there are several possible partitions of the image depending of the level of interest. Thus, in general there is no such thing as a ground truth segmentation to measure a given segmentation against. In the best case each pixel in an image belongs to not only one segment, but is assigned a sequence of labels associated with a ordered sequence of segments – the ordering relation being the subset relation on the pixel sets of segments.

Figure/ground segmentation as it is treated here approaches the segmentation problem from a different perspective: we assume to evaluate a video stream where certain objects in the scene are presently interesting and appear in front of a static, non-interesting background. Thus, each pixel in an image can be assigned one label of either background or one foreground object. An important property of figures in contrast to the background is their movement in the scene in front of a static background. However, algorithms for the problem of figure/ground segmentation based on the detection of motion in the observed scene tend to detect either too great or too small objects, because these algorithms actually detect changes of local image intensities rather than motion itself. It is our belief that this fact led researchers often to assess the quality of their algorithm on the basis of true detection/false alarm considerations, meaning that if an algorithm merely detected an object to be moving by marking some portion of it, it has successfully detected that object, and – if the algorithm marks some pixel with no moving object even near – that was simply a false alarm. To assess the quality of an algorithm supposed to create an alarm whenever there is something moving in the observed scene, this approach might suffice. It will certainly not, however, if one is interested in exact object masks in the sense of a figure/ground segmentation and, moreover, if different algorithms have to be evaluated regarding their ability to create such exact object masks. Pixel-exact object masks are of great interest if one wants to decide for a certain object to generate an alarm or not (e. g., by size) or if one wants to compute additional features based on the image of an object (such as mean and variance of apparent temperature in case of infrared data). In such cases a segmentation evaluation measure has to be introduced which

978-1-4244-7028-0/10/$26.00 ©2010 IEEE

allows for punishing over- and under-segmentation. This has been demonstrated for single images by [5]. The authors extended a segmentation measure proposed by Martin et al. in [1] for the purpose of punishing over- and under-segmentation.

One might additionally argue that motion is not the right property to discriminate figures from the background. Rather one should take intensity into account – leading to algorithms concentrating on contrasts in an image. Especially in infrared data, on which we will focus in the following sections, interesting objects, i. e., *figures*, tend to exhibit higher intensities than their surrounding, i. e., the background. Or even better, one should combine several image properties – such as motion hints and intensity borders – for figure/ground segmentation. However, the question still arises whether and to which extend the inclusion of further properties improves the quality of the resulting segmentation or not. We propose an evaluation measure for figure/ground segmentations created, e. g., by motion detection algorithms. We show evaluation results obtained with this new measure, too, on intensity-based algorithms and a combination of those two properties. As one result, these three groups of algorithms can clearly be discriminated in the resulting quality measures.

The approach quantitatively assesses the *quality* of a segmentation with respect to an a-priori known ground truth. How such a ground truth can be obtained with astonishingly little effort in case of infrared images is therefore shortly described in Sec. 2. The evaluation measure itself is based on a correspondence matrix, which can be formulated measuring the similarity of each ground truth segment to each result segment. The logic-motivated measure is described in detail in Sec. 3. We demonstrate the results of the proposed evaluation measure on several well-known detection algorithms in Sec. 4 and compare them to corresponding detection/false alarm rates, before we discuss these results in Sec. 5.

2. Ground truth generation

For the case of infrared imagery with moving people we propose a simple but quite successful way to generate *exact* ground truth pixel masks. Given that we already have an approximate ground truth in form of manually placed bounding boxes and we search for *hot*, i. e., high intensity objects, our approach refines the manually placed bounding box around an object. This is achieved by applying an intensity based segmentation of the given infrared image within the bounding box.

We cut the given image at each grayscale value, leading to a binary image with ones and zeros at pixels with a higher (or equal) or lower grayscale than the present cut value, respectively. From these binary images 8-connected components are created. Thus, we generate a set of segments for each grayscale value we cut the image. The set of segments obtained at all grayscales are processed together further on.

We then look at each bounding box in the approximate ground truth individually, intersecting each segment from the intensity segmentation process with that box. Empty results, i. e., intersection results from segments which were disjoint with the bounding box right from the start, are discarded. This leads to a set of segments which either coincide with the box or completely lie within that box. These segments are then sorted by size from greatest to smallest. We search for the biggest segment within the bounding box which at the same time exhibits the greatest contrast along its border. Thus, we search our list of size-sorted intersected segments for those with greatest (pixel-)size and highest contrast at its boundaries. In our experiments, using the product of squared contrast and pixelsize gave the best results, but the simple product of contrast and size might work as well. If two or more segments achive the same result, then we take the first one of these in the sorted list, i. e., we decide for the greater one.

An example result of this approach can be seen in Fig. 1. Here, six bounding boxes had been manually placed surrounding the six people visible in the image. The algorithm for generating exact ground truth segments created the six segments shown in the right image.

It should be noted at this point that the proposed ground truth generation algorithm works well for most isolated objects within infrared imagery we tested it on, but not quite so well for low contrast objects or objects where one occludes the other and the objects in question possess no significant mutual contrast (compare, too, Fig. 1). Still, the algorithm creates strictly more useful ground truth segments compared to the simple bounding boxes it starts with. And furthermore, the segments resulting from the algorithm can still be manually post-processed if needed – nevertheless resulting in less manual effort.

3. Measuring the accuracy of a segmentation

Wishing to quantitatively assess the accuracy of any figure/ground segmentation, several objectives to be met by the measure can be formulated:

1. That measure should reward truly recognized objects and penalize both detections where no object is present as well as not detected objects.

2. The measure should penalize deviations from true figure shapes.

3. The measure should penalize deviations in the number of detections compared to the number of true objects.

4. The measure should be continuous and normalized, allowing for direct comparison of results from different algorithms.

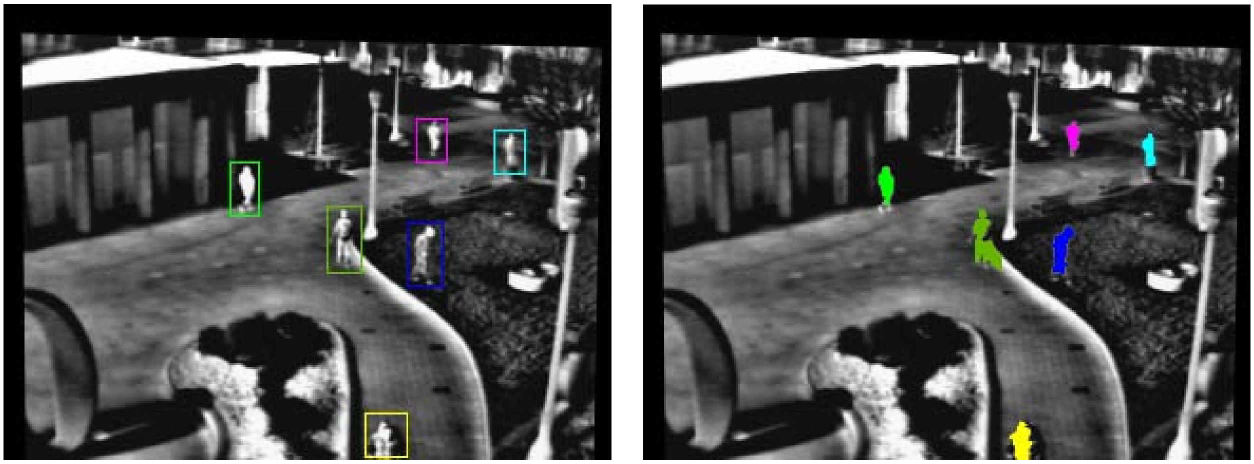


Figure 1. Example for the manually inserted ground truth rectangles (left) and the herefrom automatically generated pixel masks (right). These pixel masks build the basis for our evaluation measure. Examples shown for frame 120 of the OTCBVS-benchmark dataset 03, sequence 2 (see `http://www.cse.ohio-state.edu/otcbvs-bench/`).

We base our measure on two basic concepts: one is the *similarity* of ground truth segments and result segments. The other is a continuous version of the *correspondence matrix* introduced in [6] and logic-motivated observations which can be drawn from it.

As for the first foundation of our measure, we compute the similarity of a ground truth segment and a result segment obtained from some figure ground segmentation in terms of set similarity. The two segments are interpreted as pixel sets. Then, the similarity of these sets can be computed by means of the Jaccard-Index (see [7]). Alternatively, other set similarities are possible such as the Dice-Index (compare, too, [5]).

Using the Jaccard-Index, the similarity of the ith ground truth segment G_i ($i \in 1, \ldots, N$) and the jth result segment S_j ($j \in 1, \ldots, M$) results to:

$$s(i,j) = \frac{\#\,(G_i \cap S_j)}{\#\,(G_i \cup S_j)}. \tag{1}$$

Basing the assessment of figure/ground segmentation on such a set similarity allows for meeting the objective 2, as the similarity in itself constitutes a measure being one, if the two sets are identical, zero, if the two sets are disjoint, and a continuous value inbetween if the two sets deviate from these two extrema.

Based on this similarity, we reformulate the correspondence matrix C proposed, e. g., in [6] where each entry $C_{i,j}$ holds the similarity of the ith ground truth segment G_i and the jth result segment S_j, yielding simply $C_{i,j} = s(i,j)$. The authors of [6] had simply entered a one into their correspondence matrix whenever the two segments had overlapped and a zero, if the two segments were disjoint. By using the set similarity within the correspondence matrix, we allow for continuous measures based on this matrix keeping objective 4 in mind.

From the correspondence matrix, some qualitative observations can be drawn right away: with a perfect figure/ground segmentation we would get as many result segments as ground truth segments – resulting in a square matrix (compare objective 3). Furthermore, we would expect most if not all similarity aggregated in one element per row or column – meaning that each result segment perfectly matches the shape of exactly one ground truth segment and is disjoint with all others and vice versa. The deviation from this behavior can be measured as follows: we understand a single column or row of the correspondence matrix as a profile with one height for each matrix entry in that column or row, respectively. So we formulate the deviation of the optimal profile for a single row as

$$\operatorname{split}(i) = \min_j \{ \sum_{k=1}^{M} \left(\begin{cases} (1 - C_{i,k}) & k = j \\ \frac{1}{M-1} C_{i,k} & k \neq j \end{cases} \right) \} \tag{2}$$

such that the measure $\operatorname{split}(i)$ denotes the minimum deviation from the optimal profile, where the $\min$-function quasi searches for the peak in the profile. The minimum in Eq. (2) is reached if the index j is chosen as the index of the maximum within the row. Then, $split$ will add the deviation at that position from the optimal value of one and the average similarity present at other positions within the row. Thus by denoting m_i as the index of the maximum value in row i, we can reformulate Eq. (2) to

$$\operatorname{split}(i) = 1 - C_{m_i,j} + \frac{1}{M-1} \sum_{k \neq m_i} C_{i,k}. \tag{3}$$

The measure is named split because it describes how much a ground truth segment G_i is *split* into several result segments. Obviously $\mathrm{split}(i)$ yields values from the interval $[0..1]$, because the average similarity besides the peak is smaller (or equal) than the peak itself. Analogously, the measure

$$\mathrm{merge}(j) = \min_i \{ \sum_{k=1}^{N} \left(\begin{cases} (1 - C_{k,j}) & k = i \\ \frac{1}{N-1} C_{k,j} & k \neq i \end{cases} \right) \} \quad (4)$$

or shorter

$$\mathrm{merge}(j) = 1 - C_{m_j,j} + \frac{1}{N-1} \sum_{k \neq m_j} C_{k,j}. \quad (5)$$

formulates the same considerations for a single column, describing how much a result segment S_j *merges* several ground truth segments. Again, m_j denotes the index of the maximum element within the jth column.

Unfortunately, we do not exactly want to know how much a ground truth segment is split or how much a result segment is merging several ground truth segments, but rather would like to know whether – or to which degree – a ground truth segment has been detected correctly and whether or to which degree a result segment constitutes a false alarm. Nevertheless, both assertions can be derived from the two measures introduced so far in a continuous way. Speaking about correctness of the detection of a ground truth segment, we argue that a ground truth segment has been detected the better, the more it is similar to one result segment and the less it has been split into several result segments. A logic formulation of these relations is given by:

$$\mathrm{correct}(G_i) \leftarrow (\mathrm{similar}(\mathrm{G_i}, \mathrm{S_j}) \wedge \neg\, \mathrm{split}(G_i)), \quad (6)$$

meaning that the (degree of) correctness of the detection of some ground truth segment G_i follows from that segment being (maximally) similar to some result segment S_j and not being splitted. The mathematical formulation of this relation is given by

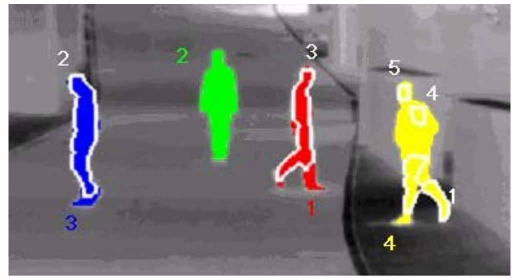
$$\mathrm{corr}(i) = \max_j (C_{i,j}) \cdot (1 - \mathrm{split}(i)) \quad (7)$$

where the maximum copes with the similarity part, whereas $(1 - \mathrm{split}(i))$ denotes the negation and the product stands for the logic conjunction. Analogously, a logic formulation of a result segment being a false detection would be:

$$\mathrm{false}(S_j) \leftarrow (\neg\, \mathrm{similar}(\mathrm{G_i}, \mathrm{S_j}) \wedge \mathrm{merge}(S_j)), \quad (8)$$

meaning that the result segment S_j is less acceptable the less it is similar to some ground truth segment G_i and the more it has been merging several ground truth segments. Mathematically, using the same transcriptions of logic operations as above, we obtain

$$\mathrm{false}(j) = (1 - \max_i (C_{i,j})) \cdot \mathrm{merge}(j). \quad (9)$$



	S1	S2	S3	S4	S5	split(i)	corr(i)
G1	0.00	0.00	0.81	0.00	0.00	0.19	0.66
G2	0.00	0.00	0.00	0.00	0.00	1.00	0.00
G3	0.00	0.87	0.00	0.00	0.00	0.13	0.76
G4	0.29	0.00	0.00	0.08	0.06	0.75	0.07
merge(j)	0.21	0.13	0.19	0.92	0.94		
false(j)	0.71	0.02	0.04	0.85	0.88		

Figure 2. Example for the proposed measures. Colored regions depict ground truth segments G1-G4, whereas white outlines show example result segments of a motion detection algorithm (S1-S5).

One can easily see that both measures reside within the interval $[0..1]$. $\mathrm{corr}(i)$ denotes the measure of one ground truth segment being correctly detected. $\mathrm{false}(j)$ describes to which degree a result segment constitutes a false detection. One should note here that the two measures are not converse in the sense that one is the negation of the other. Moreover, though not being completely independent, both measures express different aspects of a segmentation result. Additionally, both measures only state features of single (ground truth or result) segments. Example computations for the proposed measures can be seen in Fig. 2. To come to an overall measure for a single image frame, we compute the mean value for all $\mathrm{corr}(i)$ and the mean value of all $\mathrm{false}(j)$, yielding two values $\mathrm{correct}(t)$ and $\mathrm{false}(t)$ for each image time t. Finally, to come to an assessment of the results of a segmentation algorithm for a whole image sequence, we compute the mean of all frame results for $correct$ and $false$, yielding two values Correct and False for a single algorithm. Thus, as the result of a quantitative assessment of a segmentation algorithm, we obtain a point within the interval $[0..1] \times [0..1]$.

4. Experiments & Results

4.1. Data preparation

We tested the proposed measure on the OTCBVS-benchmark dataset 03, sequence 2 (see http://www.cse.ohio-state.edu/otcbvs-bench/) with a total of 601 frames. For the experiments, we manually placed a bounding box around each person in each frame. Then, we refined these bounding boxes with the approach described in Sec. 2, yielding a pixel mask for each person in each image frame. These define the ground truth segments to be obtained by the figure/ground segmentations.

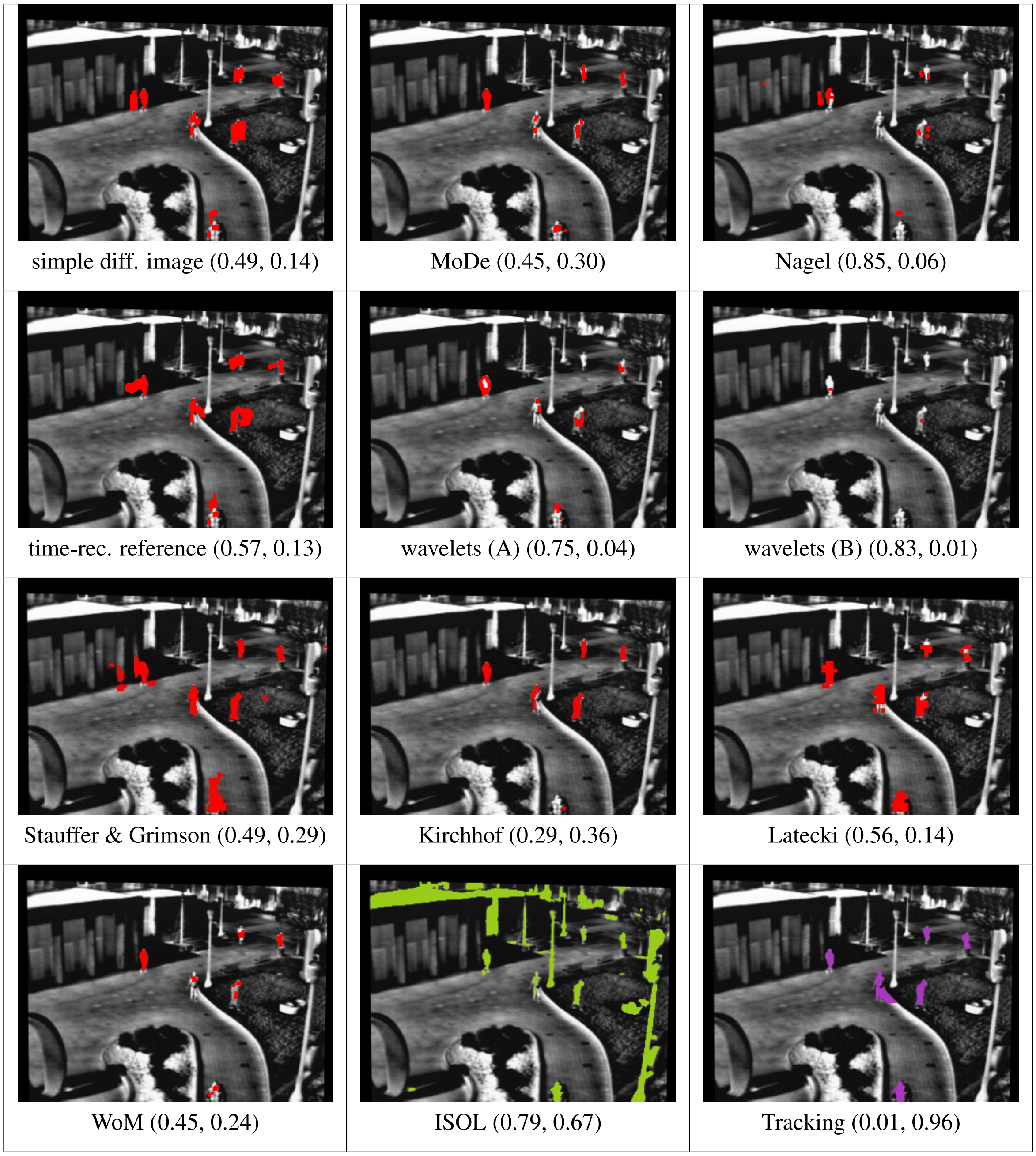


Figure 3. Example results for frame 120 obtained with different algorithms, each with apparently best parametrization. The numbers in brackets denote the falseness and correctness of the respective algorithm for this frame. Results of purely motion based algorithms are depicted in red. The results of *ISOL* – overlayed in green – are based on intensity contrasts only. A combination of motion and contrast based figure ground segmentation is shown in magenta.

4.2. Different algorithms

As proposed in the introduction, we look at motion segmentation as being one example for figure/ground segmentation, where *motion* is the one property distinguishing foreground objects (figures) from the (back-)ground. Besides that, we will look at intensity based segmentation for the special case of infrared data, where we assume that figures exhibit high intensities. And, finally we will combine these two properties for a both motion and intensity based segmentation approach.

We tested the proposed measures on several different algorithms for motion based figure/ground segmentation. All algorithms result first in a binary classifaction image (motion/no motion), which then was segmented using simple 8-connected component search. Among the algorithms were

- a simple image differencing (with different temporal distance between the two images and different thresholds for binarization),
- a method named *MoDe* described in [8], which works with two temporally symmetric image differences (again, several parametrizations),
- an image differencing scheme based on the time-recursive computation of a reference image or background image (see [9], several parametrizations),
- Several parametrizations of two variants of the Haar-Wavelet based motion detection (compare [10, 11]).
- An own development based on the choice of optimal parameters and sub-procedures within our experimental system for testing all these algorithms and variants (called WoM, which, too, allows for several parameters to tune).
- The very well known method of Stauffer and Grimson based on modelling the background with a field of Gaussian Mixtures (see, e. g., [12]), where we used different adaption rates and different background thresholds,
- A modification of the Stauffer & Grimson approach proposed in [13] which allows for fast background adaption and thus for (moderately) moving cameras,
- the statistically motivated change detection algorithm introduced by Nagel et al. in [14] with different thresholds for classification,
- and the method proposed by Latecki et al. in [15] based on the observation of spatio-temporal patches called Motion Orbits. These authors actually tested their approach on infrared images.

Overall, we tested ten different motion based algorithms with a total of 80 parametrizations. In addition to these motion based approaches we reimplemented an intensity based segmentation approach described in [16] called *ISOL*. That approach is based on cutting a grayvalue image at several grayvalues and deciding for resulting segments which maximize the contrast to their surrounding. For this algorithm, we again examined different parametrizations.

Finally we combined one of the motion based approaches (namely *MoDe*, because it seemed to gain quite average results) with the intensity based approach. For the combination, we first compute the motion based segments obtained by *MoDe* as a kind of interest operator. Then we build a bounding box arround each segment. Overlapping bounding boxes are merged, resulting in an encompassing bounding box. In each such bounding box, the refinement algorithms described in Sec. 2 is applied which produces the final figure mask. We also convey information from the last frame, as we include slightly enlarged bounding boxes around each pixel mask from the last frame into the present frame before the merging of bounding boxes. In such way, we allow for a region – or figure – growing even in cases where the mere motion detection can only mark some pixels at the edges of a moving object. The propagation of information from the last frame is the reason why we refer to this approach as *tracking* in the sequell. Again we tested several parametrizations of the approach.

Example results for each of the described algorithms can be seen in Fig. 3. As can be seen, the various motion based algorithm widely vary in the pixel masks they produce.

4.3. Results

The assessment results, i. e., the two-dimensional coordinates obtained by each parametrization of the algorithms examined can be seen in the lower diagram of Fig. 4, which shows the achieved correctness on the ordinate and the falseness on the abscissa. Thus, the optimal position within this diagram lies at $(0, 1)$, i.e., minimum falseness and maximum correctness. The upper diagram of Fig. 4 shows common evaluation in terms of detection/false alarm rates, where the latter are depicted in logarithmic scale. For creating this diagram, we counted each result segment of some algorithm as a true detection if it overlapped with some ground truth segment. As false alarms, we counted each result segment which overlapped with the ground truth background, i.e., the image without all pixels which are marked as figure in the ground truth.

An initial impression from looking at the result segments (compare Fig. 3) is that the motion based algorithms tend to produce only few false alarms, but also fail to produce fair pixel masks in several cases. On the contrary, the intensity based algorithm extracts the pixel masks for true objects quite well, but produces many false positives in the background. And last but not least the *tracking* approach combining motion and intensity based features seems to combine the good properties of both approaches, too, resulting in few false alarms and fairly exact pixel masks. This observations have also been indicated in the lower diagram of Fig. 4, where colored ellipses surround the three groups of algorithms/parametrizations.

5. Discussion and Conclusion

The argument of this paper is not to bring forward a certain algorithm for figure/ground segmentation, nor is it to declassify some other. Rather it is our belief that algo-

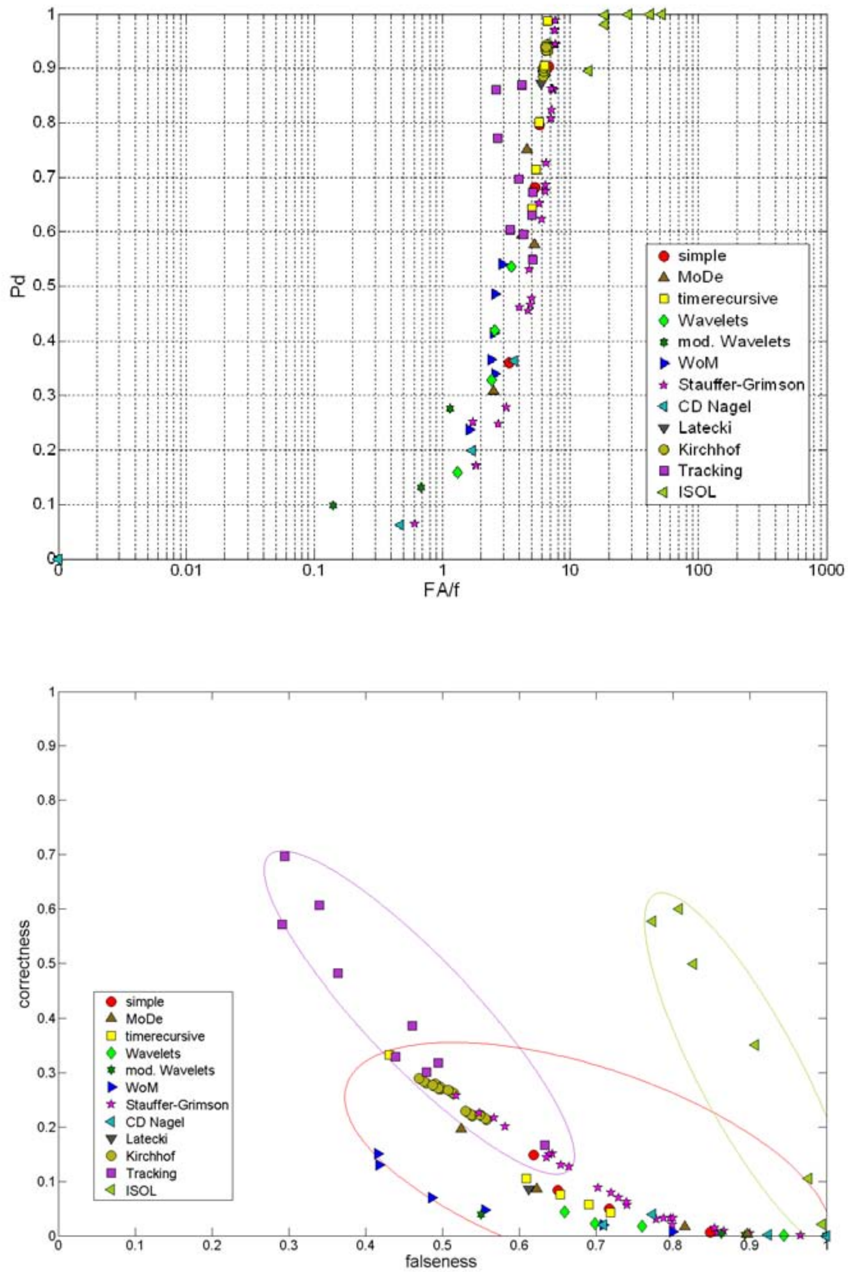


Figure 4. Above: evaluation results in terms of detection/false alarm rates. Below: evaluation results with the proposed measure. Each approach obtains a coordinate in the interval $[0, 1] \times [0, 1]$. Manually inserted ellipses surround only motion based (red ellipse), only contrast based (green ellipse), and combined motion and contrast based approaches/parametrizations (magenta ellipse).

rithms for figure/ground segmentation in general should be assessed strictly with regard to their ability to extract exact pixel masks for each figure present in the observed scene. One possible measure to do so has been proposed.

The proposed measure is based on a continuously measured similarity of ground truth segments and segments resulting from an algorithm in question. By means of a correspondence matrix with continuously valued entries, we sucsessively derive two values indicating the *correctness* of the detection of ground truth segments and the *falseness* of result segments. Together, these measures result in a two dimensional coordinate within the intervall $[0..1] \times [0..1]$

indicating the quality of the results of a given algorithm for figure/ground segmentation.

Looking at the results obtained, one can draw some conclusions. First of all it might strike how bad most of the motion based approaches perform in terms of falseness/correctness, given that the optimal result would be (0,1), and the best result achieved resides around (0.45, 0.3). The reasons for that certainly include the following:

- we measured the results against the ground truth half-automatically generated with the algorithm introduced in Sec. 2. No post-processing had been performed on these results, despite the drawbacks of that ground truth already discussed,
- in some cases we might not have found the best parameter settings, yet. An indication for this reason lies within the fact that some algorithms achieve quite scattered coordinates in Fig. 4,
- and, probably most important of all, most motion based algorithms examined cannot achieve much better results with regards to our measure, in principle, as they are based on image differencing and thus on the detection of intensity changes in the image plane. Though an object is moving as such, still there are parts of the object in the image which do not change at all, at least within the temporal window of some algorithm in question. Thus, the algorithm cannot detect the object as a whole pixel mask, but only parts in and in opposite to the direction of motion.

Apart from that, the only intensity based approach, too, fails to achieve good results in terms of falseness/correctness. The best result here resides about (0.8,0.55). The reason for that is quite obvious: high contrast regions are deemed to appear in the background, too, and thus lead to high falseness values in the end.

Third, the combined approach of motion and contrast based figure/ground segmentation performs best, as indicated by the example images shown in Fig. 3. Best values for falseness/correctness reach up the point of (0.3, 0.7). This might not be surprising to most readers, but it is much more emphasized in the falseness/correctness diagram (see Fig. 4, lower part) than in the detection/false alarm rate plot. The reason for this better indication of the segmentation quality lies within the way both measuring schemes work: where detection/false alarm rates simply count result segments as one or the other, the proposed measures take into account the shape of result and ground truth segments on a pixel basis. This, in fact, is the main contribution of this article.

References

[1] D. Martin, C. Fowlkes, D. Tal, J. Malik: *A database of human segmented natural images and its application to evaluating algorithms and measuring ecological statistics.* In: Proc. of the 8th Int. Conf. on Computer Vision (ICCV-2001), 7-14 July 2001, Vancouver, BC, Canada. IEEE, pp. 416-425. 1, 2

[2] A. Hanburry and J. Stöttinger: *On Segmentation Evaluation Metrics and Region Counts.* In: Proc. of the 19th Int. Conf. on Pattern Recognition (ICPR-2008), 8-11 Dec. 2008, Tampa, FL, USA, IEEE 2008, pp. 1-4. 1

[3] J. S. Cardoso and L. Corte-Real: *Toward a Generic Evaluation of Image Segmentation.* IEEE Transactions on Image Processing **14**:11(2005) 1773–1782. 1

[4] R. Unnikrishnan, C. Pantofaru, and M. Hebert: *Toward Objective Evaluation of Image Segmentation Algorithms.* IEEE Transactions on Pattern Analysis and Machine Intelligence **29**:6(2007) 929-944. 1

[5] M. Polak, H. Zhang, and M. Pi: *An evaluation metric for image segmentation of multiple objects.* Image and Vision Computing **27**:8(2009) 1223-1227. 2, 3

[6] J. Nascimento and J. S. Marques: *New Performance Evaluation Metrics for Object Detection Algorithms.* Proc. of the 6th IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2004), May 10 2004, Prague, Czech Republic; pp. 7-14. 3

[7] P. Jaccard: *Nouvelles recherches sur la distribution florale.* Bulletin de la Societe Vaudoise de Sciences Naturelles 44 (1908) pp. 223-370. 3

[8] D. Borghys, L. Svigny, R. Gabler, J. Blanc-Talon, P. B. W. Schwering, H. C. Palm, J. F. Haddon, and N. Nasrabadi: *Multi-sensor Image Exploitation.* NATO RTO Technical Report, AC323(SET)TP/10, December 2001. 6

[9] C. Wren, A. Azabayejani, T. Darrell, and A. Pentland: Pfinder: *Real-time tracking of the human body.* IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 1997. 6

[10] J.–M. Letang, P. Bothemy, and V. Rrebuffel: *Robust motion detection with temporal decomposition and statistical regularization.* Institut national de recherche en informatique et en automatique, 1995. 6

[11] C. Blatter: *Wavelet – Eine Einführung* Vieweg, 1998 (in German). 6

[12] C. Stauffer and W. E. L. Grimson: *Learning Patterns of Activity using Real-Time Tracking.* IEEE Trans. on Pattern Analysis and Machine Intelligence 22:8(2000) 747-757. 6

[13] M. Kirchhof: *Statistical Motivated Background Modelling.* Proc. of the Int. Symposium on Spectral Sensing Research (ISSSR), New York, 2008. 6

[14] Y. Z. Hsu, H.–H. Nagel, G. Rekers: *New Likelihood Test Methods for Change Detection in Image Sequences.* Computer Vision, Graphics, and Image Processing **26**:1(1984) 73-106. 6

[15] L. J. Latecki, R. Miezianko, and D. Pokrajac: *Tracking Motion Objects in Infrared Videos.* Proc. IEEE Int. Conf. on Advanced Video and Signal Surveillance, Como, Italy, Sep. 2005. 6

[16] C. Anderer and R. Gabler: *ISOL – Ein Verfahren zur Segmentation von Grauwertbildern.* Internal Report, FGAN-FOM, Ettlingen, Germany, SZA 2002/2 And/Gab, December 2002 (in German). 6