# Fusion of Region and Point-Feature Detections for Measurement Reconstruction in Multi-Target Kalman Tracking

Michael Teutsch, Wolfgang Krüger, and Jürgen Beyerer
Fraunhofer IOSB
Fraunhoferstr. 1, 76131 Karlsruhe, Germany
Email: {michael.teutsch, wolfgang.krueger, juergen.beyerer}@iosb.fraunhofer.de

***Abstract*—Object tracking in 2D video surveillance image data is one of the key needs for many follow-up operations such as object classification or activity recognition. In scenes with multiple objects crossing each other's way, there is a high potential for split and merge detections disturbing the tracking process. In these situations, it is helpful or even necessary to reconstruct the object-related measurements to support tracking approaches such as Kalman or Particle Filter. We present a way of fusing three different detection approaches taking benefit from their specific advantages to reconstruct measurements, if a split or merge situation is recognized. The resulting split and merge handling shows better results than using each detection approach individually without fusion. Furthermore, the tracking process is fast with a computation time less than one millisecond per image. Experimental results are given in example video scenes of an infrared camera located on a buoy for maritime surveillance.**

**Keywords: multitarget tracking (MTT), Kalman filter, split and merge handling, multiple features, maritime object tracking, vessel surveillance, infrared video.**

## I. Introduction

In many applications for video surveillance, detected objects are to be classified or their activities are to be analyzed and recognized. The temporal dimension is very important to guarantee detection stability and robustness as well as the opportunity to track objects, re-identify them in case of multiple object occurrence and observe them for a period of time, not only in single frames.

In this work, we present a fast and robust multi-target tracking approach using the well-known Kalman filter [1]. Three object detection algorithms have been implemented with two of them delivering regions (bounding-boxes) and the third one detecting and tracking salient image points (point-features). The detection results are used to create measurements for the filter. Based on previous measurements, a Kalman prediction is performed for each Kalman track which represents one object in the scene. Together with the current measurement and the Kalman-Gain as weighting factor, the Kalman update is generated, which is the result of the filtering process. We focus on scenes with the appearance of multiple objects which cross each other and, thus, produce split and merged measurements. These split- and merge-situations harm tracking, so we use the three different detection approaches with their specific characteristics and advantages to reconstruct the expected measurements. The motivation for measurement reconstruction is given by the work of Grinberg et al. [2]: 3D point clouds and optical flow vectors coming from a stereo camera system are used successfully for multi object Kalman tracking. Based on an observability analysis together with the spatial and temporal relationship between stably tracked points and tracked objects, measurements are reconstructed in case of occlusion, split and merge situations.

The idea of fusing the three detection approaches considered in this paper was already proposed in [3], where basic multi-target tracking was introduced. In the following, we demonstrate how measurement reconstruction can be used to solve specific problems in the context of multi-target tracking such as occurring occlusion, split and merge situations.

Our application is maritime surveillance and the detection of criminal activities on the open sea such as drug trafficking, piracy or illegal immigration. A thermal infrared camera located on a buoy is used to identify suspicious objects around-the-clock [4]. To guarantee high robustness and stability of the detection results despite of strong variations in object appearance, image quality and weather conditions, it was decided to implement three different detection algorithms complementing each other [3]. For tracking, we aim to find a good tradeoff between robustness and speed: being as robust as possible towards multi-objects occurrences with split- and merge-situations on the one hand and being fast due to the already quite high computation time for the detection on the other. Standard Kalman filter is chosen because ship and boat motion is highly linear and well predictable.

### *Related work*

The fusion of different detection approaches for tracking is widely spread. Often, detection results coming from multiple sensors are fused to utilize the sensors' complementary characteristic. Some good examples can be found in driver assistance with radar or lidar delivering 2D-points in lateral and longitudinal dimension combined with a visual-optical or IR camera applying region-based detection considering e.g., optical flow vector clusters, symmetry, shadows, or intensity blobs [5], [6].

Since our detection results are coming from only one camera, we now focus on previous work where also only one visual sensor is used. Usually, Kalman tracking considering only region-detection suffers from merges, splits and occlusions [7], while approaches based on local features can handle them due to motion-information but sometimes generate partial or split detections, as local features have to be clustered [8]. General feature categories for fusion are shape (contour, edges), texture (intensities, wavelets), color, and motion. In [9], color, edge and texture features are used. A mean-shift algorithm is run for each features class separately and fused by weighting the results to handle ambiguities in one class (e.g., color) correctly. A combination of Mean Shift and Kalman filter is applied with the fusion of edge and color features in [10]. Wang et al. [11] combine edge and color features as input for an adaptive Particle Filter to handle occlusions in vehicle tracking successfully. In [12], Kalman filter-based motion and shape tracking are used to introduce the *geometric shape matching algorithm*, which is able to efficiently handle split and merge situations in color image sequences, if no significant motion change happens during occlusion. The shape is extracted by foreground/background segmentation. In [13], *multiple feature pseudo-color images (MFPCIs)* are generated from IR images. The first color-channel is used for the IR intensities, the second for Gabor features, and the third for Entropy features. This way, the object blob is emphasized as it is dominant in each channel compared to the background. Finally, since MFPCIs are color images, they can be used as input for a standard Mean Shift algorithm. In [14], *fragments* (local features) are determined in a manually set template. They are used to reconstruct the track in case of occlusion. Finally, in [15], a variety of features is offered to the tracking procedure. In an online feature selection, features are continuously ranked to find out, which features currently are the best to separate between object and background. Only the best ones are considered for tracking. The approach appears to be highly adaptive and dynamic, but some open questions remain such as the number of features to select, the number and kind of features to calculate, and how to reduce the computational effort.

This paper provides the following organization: The three used detection approaches are presented in section II, Kalman tracking with the fusion of only the region-based detection results is introduced in section III and the fusion of all three approaches in section IV. Some experimental results are given in section V and conclusions in section VI.

## II. The three detection approaches

In the existing literature, a variety of approaches from machine vision has been applied to detect ships and boats in image sequences. Examples cover template matching of ship silhouettes [16], mean-shift segmentation with minimum spanning tree clustering [17], multi-scale blob detection [18], Canny edge detection [19], and background models [20]. A more detailed description can be found in [3].

In this and previous work [3] a combination of algorithms relying on complementary image cues has been used to generate detections that are robust with respect to variations of boat appearance, image quality, and environmental conditions. The common idea of these boat detection algorithms is to search for temporally stable image features to separate detections at boats from those at sea clutter. In particular, three algorithms have been developed and implemented:

1) Extraction of stable image regions [21], adaptive thresholding, and tracking.
2) A track-before detect algorithm [22] which uses spatio-temporal integrated blob strength.
3) Detection and tracking of salient image points [23].

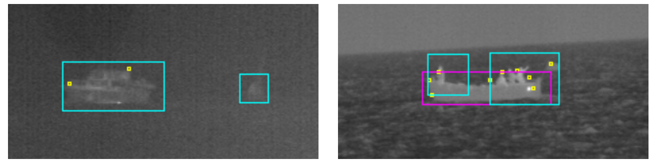
(a) (b)

Figure 1. Detection results from the three different algorithms: Stable regions (magenta), track-before detect (cyan), salient image points (yellow).

All three detection algorithms use their own short-term tracking to suppress spurious detections.

The idea of the stable image regions is to detect the body of boats or small ships as a single region (magenta bounding-box in Fig. 1b). Only a single image frame is needed to perform detection, but an additional short-term tracking is available to improve suppression of spurious detections. The drawback is that extraction of stable regions tends to get unreliable when image contrast is low (Fig. 1a). In such situations the track-before detect algorithm is more reliable (cyan bounding-boxes in Fig. 1a) but has a tendency to over-segment (isolated bounding-boxes in Fig. 1b).

Compared to these two approaches, detection and tracking of salient image points is able to yield much more accurate motion information which may be exploited in clustering and higher-level object tracking. Salient image points also have the advantage to be sensitive to people in small boats. The drawback is that points have to be tracked reliably over a larger number of frames to eliminate tracks arising from sea clutter, which becomes more and more difficult when image quality degrades (right object in Fig. 1a which is detected by track-before detect algorithm only).

The following sections demonstrate how fusion can be used to combine the respective strengths of these detection algorithms.

## III. Tracking using the regions only

Since the point-features are prone to get lost in case of weak contrast or occurring merge and occlusion situations, the measurement reconstruction and tracking have to be able to run independently of them.

### A. Concept

An overview of the tracking concept using all three detection approaches is given in Fig. 2. In this section, we

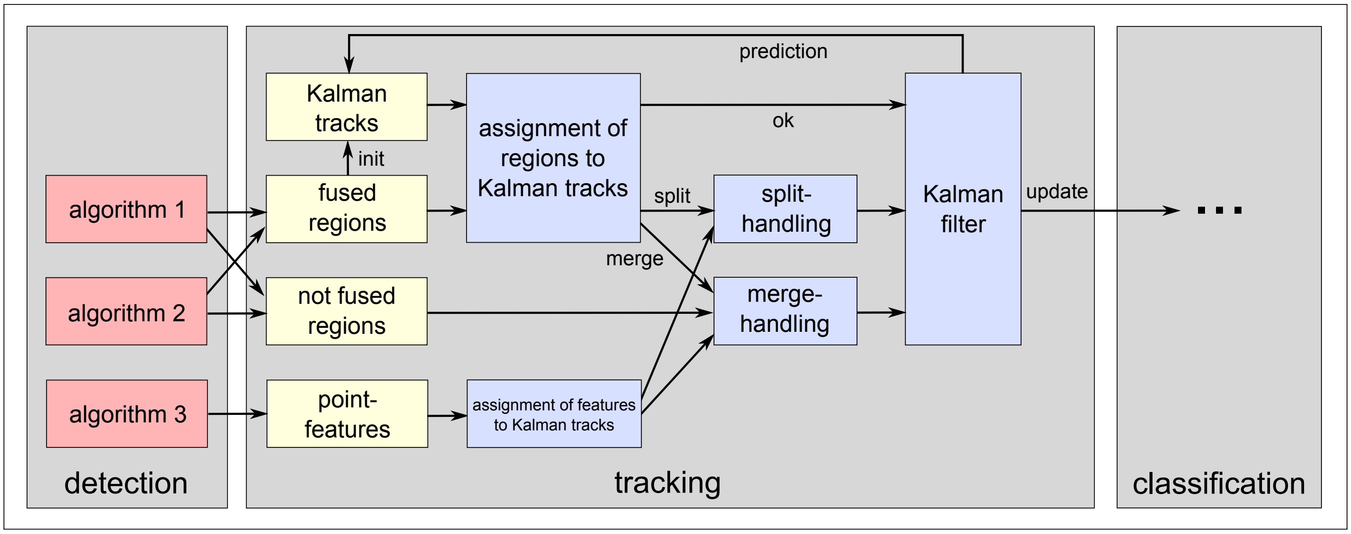


Figure 2. Tracking concept using the fusion of region- and point-feature-based detection algorithms.

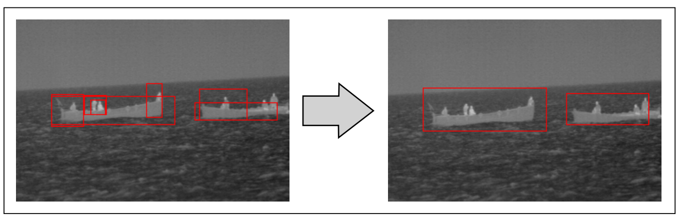

Figure 3. Trivial fusion of regions to create measurements.

consider only the two region-based detection approaches to create measurements and handle split and merge situations taking benefit of their approach-specific properties. The point-features, their assignment to Kalman tracks, and their influence to split- and merge-handling are not used, yet. The detected regions are delivered as bounding-boxes in two different sets: not fused and trivially fused, where all bounding-boxes are unified, which intersect each other. In cases of no merge, this trivial fusion is absolutely sufficient to create measurements as shown in Fig. 3. The not fused regions are kept to reconstruct the measurements, if a merge is detected. The measurements are assigned to related Kalman tracks or used to initialize new Kalman tracks. After successful assignment or split/merge-handling, the Kalman filter passes the updated Kalman tracks to the classification-module and performs a prediction for the next time step. Tracked parameters are center and size of the bounding-box.

### *B. Measurement assignment*

As already mentioned, measurements are created by trivial fusion of the detected regions. Normally, the regions spatially complement each other very well and generate a fused measurement covering the whole object. Since spatio-temporal stability of the regions is guaranteed by short-term tracking, severe clutter detections, which may disturb this trivial fusion, are not expected.

For fast assignment of measurements to Kalman tracks, we introduce a kind of *validation gating* [24]: the *intersection assignment matrix (IAM)* with one column for each existing Kalman track and one row for each measurement. The bounding-boxes of all current measurements and Kalman tracks are checked for intersecting each other. If Kalman track $i$ intersects measurement $j$, *1* is entered on matrix position $(i, j)$, or *0* otherwise. In standard case without split, merge or track-initialization/deletion, there will be exactly one *1* per row and column. So, the IAM for the situation in Fig. 3 is looking like

$$\left(\begin{array}{c|cc} Meas \setminus Track & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right)$$

and a good assignment is found immediately. For all other matrix value constellations, a special tracking case occurs, which needs specific handling. Split and merge situations are discussed in its own subsection. If a Kalman track gets no measurement at all (only *0* in its column), the track is kept alive using its Kalman prediction and a miss-counter is incremented. As soon as this counter exceeds a threshold, the Kalman track is deleted.

If there is no intersection between a measurement and any Kalman track (only *0* in measurement row), a new Kalman track is initialized and an initialization-counter is incremented. As soon as this counter exceeds a threshold, the Kalman track is trusted to represent an object.

### *C. Split-handling*

A split occurs, if a Kalman track is assigned to more than one measurement. This can be detected easily with a view to the IAM:

$$\left(\begin{array}{c|cc} Meas \setminus Track & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{array}\right)$$

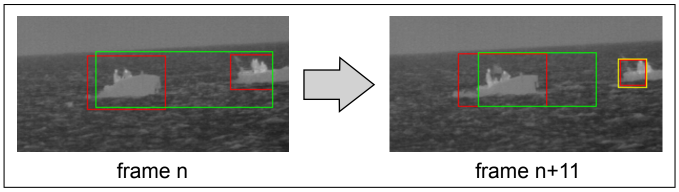


Figure 4. Split handling using *apostate* measurement: Regions in red, stable Kalman tracks in green, just initialized Kalman track in yellow.

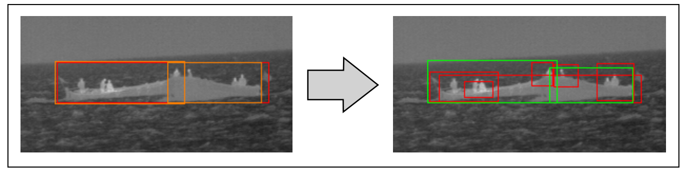

Figure 5. Merge handling using *trial-and-error* fusion.

Generally, a split happens either due to wrongly split detection (false split) or when a track is initialized while a merge already takes place (true split). Split-handling without additional features is done quite simple: in case of a false split, we assume the partial detections to be spatially near to each other, and vice versa for true splits. An *apostate* measurement is found, if the calculated distance between the partial measurements is bigger than a distance threshold derived from the measurements' size. Then, the coherent measurements are assigned to the related Kalman track and if the apostate measurement shows spatio-temporal stability, a new Kalman track will be initialized for it. Such a situation is shown in Fig. 4, where the stable Kalman track (green) determines the right measurement (red) to be an apostate and doesn't consider it anymore. Thus, a new Kalman track is initialized (yellow) for the apostate measurement 11 frames later.

### D. Merge-handling

A merge is detected for a measurement, which is assigned to more than one Kalman track according to the IAM:

$$\left(\begin{array}{c|ccc} Meas \setminus Track & 0 & 1 & 2 \\ \hline 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right)$$

We also tested a more sophisticated, probabilistic approach to detect a potential merge, which proved to be useful for multi-person tracking [25], but since boat motion is much more linear and predictable compared to a person's movement, there wasn't enough benefit for merge-handling considering the negative aspect of additional computation time.

Measurement reconstruction is now tried individually for each participating Kalman track by using the set of not fused regions. As seen in Fig. 1b, the track-before detect algorithm tends to generate multiple bounding-boxes per object, which sometimes only partially cover the object but withstand the merge-process. In the first step, all measurements intersecting the bounding-box of Kalman prediction are collected. To limit the computation time, the measurements are ordered by their Mahalanobis distance to the Kalman prediction and only the $n$ best are considered, if a maximum number of measurements is exceeded. For the chosen collection, a *trial-and-error* fusion is performed, where all possible measurement combinations are tried and evaluated. Therefore, we calculate the Mahalanobis distance of the fused bounding-box for each measurement combination to the Kalman prediction. If the Mahalanobis distance of the best fusion result goes below a certain distance threshold, measurement reconstruction was successful. An example of such a successful reconstruction for objects crossing each other is illustrated in Fig. 5. In the left image, the fused regions causing a merged measurement (red) and the Kalman tracks in trouble (orange) are shown and in the right image, the not-fused regions (red) and the Kalman tracks (green) supported by reconstructed measurements.

In case of no success, the Kalman track is kept alive using the Kalman prediction with two restrictions:

1) **Size constraint:** The spatial dimensions of the track's bounding-box are fixed to avoid over-sizing.
2) **Motion constraint:** Spatial shifting is limited by the bounding-box of the merged measurement.

## IV. Tracking using regions and point-features

Considering the point-features for tracking is optional as their existence can't be guaranteed. However, we found out, that they are a powerful support for the tracking process, if they exist.

### A. Concept

The concept itself remains the same as seen in Fig. 2. Now, it is assumed that point-features are available and can be used for measurement reconstruction in situations of split or merge. Therefore, they first need to be assigned to already existing Kalman tracks. Furthermore, point-features with no related Kalman track are not used for track initialization and assigned point-features, which get lost, are not used for track deletion.

### B. Measurement and point-feature assignment

Measurement assignment and initialization of Kalman tracks remains the same as in section III-B. In the follow-up, the assignment of point-features to existing Kalman tracks is presented. As the features are treated as optional measurement, they should not be used to initialize or delete Kalman tracks.

Each point-feature is described by 2D position, 2D velocity, existence-counter, and optional information about related Kalman track, if available. The assignment to an existing Kalman track is done using four criteria:

1) The feature is not assigned to another Kalman track.
2) The feature is located inside the bounding-box of the Kalman prediction.
3) The feature is not located inside the bounding-box of another Kalman prediction (merge area).
4) The feature roughly has the same motion as the Kalman track.

With the third criterion, ambiguous assignments are avoided. The fourth criterion can be improved as soon as a Kalman

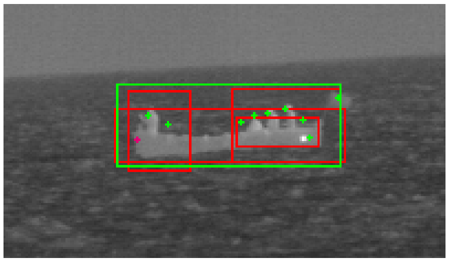

Figure 6. Example for point-features and their assignment.

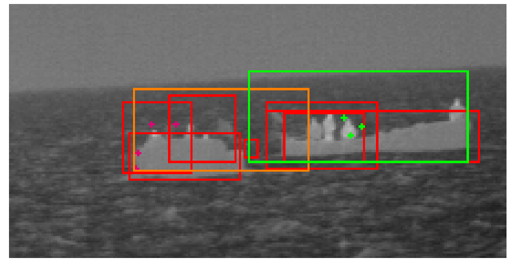

Figure 7. Split handling with not-fused regions and point-features.

track already owns some features. They offer much better motion information as the region-based Kalman tracking and, thus, can be used to estimate track movement much more precisely. This is helpful when further features are to be assigned to a Kalman track. An example for region and feature assignment is shown in Fig. 6. The point-features appear as small crosses. Green features are already assigned to the Kalman track, which is also green. The red feature is not assigned, yet, and the related, not fused regions are displayed in red, too.

Each Kalman track has a list of all currently assigned point-features. The track-velocity is calculated by the median of all feature-velocities to avoid the influence of outliers. Furthermore, outliers are detected by analyzing and comparing the velocities of all assigned features. If a feature's velocity deviates too much from the median, the feature will be erased from the list. Furthermore, each feature remembers its position in the Kalman track's bounding-box (after Kalman update) by storing the relative position of the bounding-box corners.

The last information about spatial dimensions is used, if region-based detections are missing. Then, the measurement is reconstructed by the assigned features only. Since the features are an optional measurement, a miss-counter is incremented anyway, but with a much higher threshold than the one mentioned in section III-B, because with the feature there still is some kind of detection.

### C. Split-handling

A split is detected like in section III with the IAM. It can be a false split, if the region-based approach delivered split detections, or a true split, if the Kalman track was initialized during a merge and the objects are splitting after some time. The point-features support the split-handling by analyzing the motion of the partial measurements (fused regions). Each region gets a list of related point-features. For assignment, a feature here only has to satisfy criterion 2 of section IV-B. The velocity median of the assigned features is used to estimate a region's motion. Two regions with obviously different motion are likely to be a true split while nearly the same motion indicates a false split. This way, it is possible to handle splits more accurate and earlier than in section III-C. An example for this case is given in Fig. 7. During the merge, one object has lost all its assigned features due to occlusion. The related Kalman track is kept alive by using only the Kalman prediction. This is visualized by the orange bounding-box. The other object's measurement has correctly been reconstructed by the features, so its bounding-box remains green. When a split was detected, the available features (green and red crosses) were assigned to the regions (red bounding-boxes) and different motion of the left and right regions was recognized. As the right regions were already related to the green Kalman track, it is clear that the left regions belong to the orange Kalman track. So, the true split was handled correctly.

Unfortunately, it can't be assumed that each region gets assigned point-features. Thus, the split-handling for all regions without features is performed as described in section III-C.

### D. Merge-handling

In case of a merge, two competing ways of measurement reconstruction are tried. On the one hand the idea of trial-and-error fusion as proposed in section III-D and on the other hand the analysis of the assigned point-features. In the latter one, each feature remembers about its relative position in the Kalman track's bounding-box of the previous time-step and reconstructs this bounding-box at the feature's current position. If multiple assigned features are available, the median for each bounding-box corner is taken to suppress outliers. For a longer merge period, this leads to a nearly fixed bounding-box size, because updates of relative feature positions are not done during a merge. However, the fixed spatial dimensions are admissible for most merge situations.

As soon as both ways of measurement reconstruction – the trial-and-error fusion and the point-feature analysis – have been applied, the resulting bounding-boxes are evaluated considering the Mahalanobis distance to the Kalman prediction. The reconstructed measurement with the lower distance is the better one. If this lower distance is also below a certain distance threshold, the measurement has been successfully reconstructed. If not, again the Kalman prediction is used to keep the Kalman track alive like in section III-D. Fig. 8 shows an example, where a merge happens and the point-feature-based measurement reconstruction performed better than the trial-and-error fusion.

### E. Algorithm overview

The proposed concept and methods can be implemented as shown in the algorithm flowchart in Fig. 9. After trivial fusion of the regions (see Fig. 3), Kalman prediction for all existing Kalman tracks is performed. The fused regions are assigned to the Kalman tracks as described in section III-B. If assignment

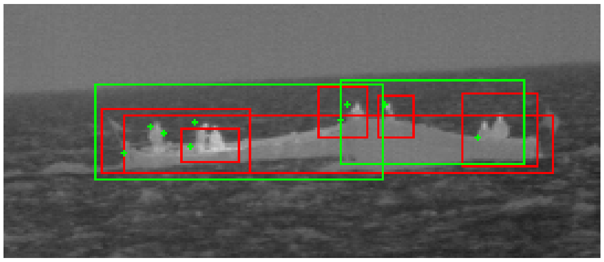

Figure 8. Merge handling with not-fused regions and point-features.

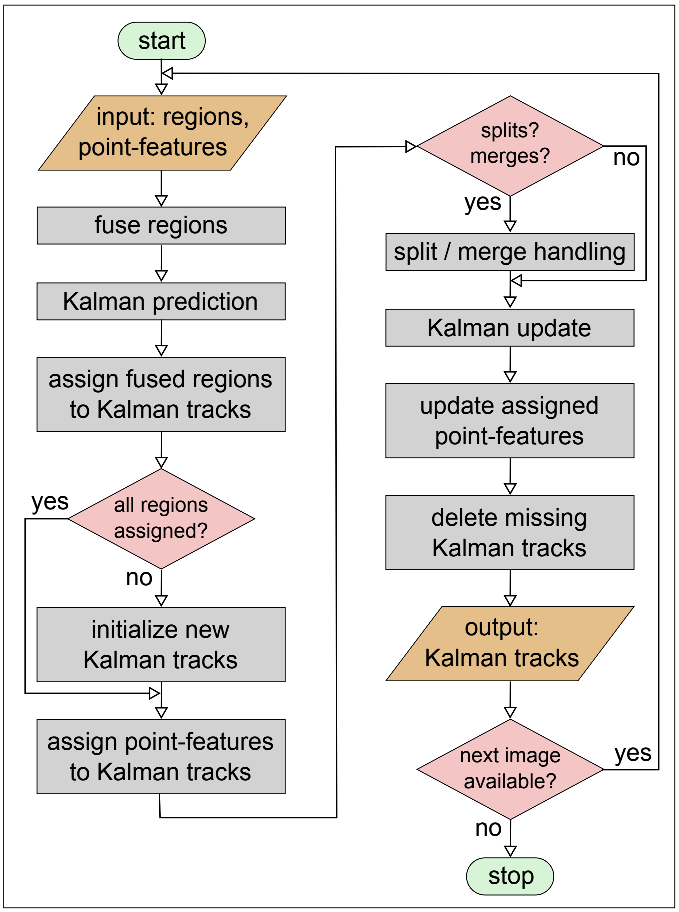


Figure 9. Tracking algorithm overview as flowchart.

wasn't possible for a region, this region is used to initialize a new Kalman track. Now all available point-features are assigned to the Kalman tracks either by using their index for previously assigned features or the four assignment criteria in section IV-B. If split or merge situations occur according to the IAM, specific split/merge handling is applied as presented in section IV-C and IV-D. Then, Kalman update is performed and the assigned point-features are updated regarding their relative position (see IV-B) in the related Kalman track for potential measurement reconstruction in the next time step. Kalman tracks, which didn't get any measurement for a certain time, are deleted. Finally, the updated Kalman tracks are passed to the next module (e.g., classification) for further processing.

## V. Experimental results

For testing and evaluation we chose four example scenes coming from a thermal infrared camera located on a buoy for maritime surveillance. Each scene has a length between one and two minutes and consists of single frames with a resolution of $576 \times 472$ per frame and a rate of 25 frames per second. In the four scenes different kinds of split- and merge-situations occur:

- two moving objects crossing each other
- two moving objects overtaking each other
- one moving object crossing a stationary object (buoy)
- multi-merge with one object crossing two stationary objects

These examples cover most split- and merge-situations to be expected in a maritime environment, as only few objects are likely to be at the same place at the same time. Nevertheless, the tracking concept as implemented is also able to handle more objects splitting or merging, but such situations didn't occur in the test data. However, while boat tracking is on the one hand easy because boat motion is quite slow, linear and predictable, merge-situations on the other hand are difficult to handle because they can last for 10 or more seconds. So, we also focused on such situations during testing.

Some results are shown in Fig. 10. In two example scenes we compare both fusion approaches from section III and IV with each other. Measurements are displayed as red bounding-boxes, unassigned point-features as red crosses, assigned point-features as green crosses, stable Kalman tracks as green bounding-boxes and *lost* Kalman tracks, where no measurement reconstruction was possible and only Kalman prediction is considered, as orange bounding-boxes.

In scene 1, two moving objects cross each other. When the merge situation occurs, the point-features help to reconstruct the measurements for both Kalman tracks correctly (frame $m+39$ (b)), while both tracks get lost without features (a). The Kalman prediction with two restrictions (size and motion constraint) keeps the tracks alive but leads to typical drifting problem due to velocity change of the related object as seen in frame $m+105$. On the contrary, for the Kalman track of scene 1 (b), which loses all point-features due to occlusion, Kalman prediction has to be used only for a quite short time, so no drifting problem occurs. Split handling with point-features as presented in section IV-C helps to resolve the split problem in frame $m+130$ earlier than without features.

In scene 2, a much more complex scenario is presented. The boat is slowly crossing a stationary buoy, which takes about 40 seconds and finally leads to a multi-merge of all objects in frame $n+1018$. Without the point-features, not only drifting occurs, but also track confusion takes place as seen in frame $n+1018$, where the buoy track is taken away from the buoy. The track not only had to be re-initialized (frame $n+1300$) but also is harming the tracking until the end as seen with two lost Kalman tracks for the boat (frame $n+1632$). With the point-features, correct tracking is possible even in difficult situations like in frame $n+1300$.


Figure 10. Comparison of the both fusion approaches from section III and section IV in two example scenes. Measurements are displayed as red bounding-boxes, unassigned point-features as red crosses, assigned point-features as green crosses, stable Kalman tracks as green bounding-boxes and Kalman tracks, where no measurement reconstruction was possible and only Kalman prediction is considered, as orange bounding-boxes.

In our tests, the computation time is less than $1\,ms$ per frame when using standard hardware with $2.66\,GHz$ processor and $3\,GB$ of RAM. In case of three objects without split and merged measurements, it was about $0.2\,ms$, and with all three objects merged, it was around $0.8\,ms$. In general, the computational costs grow linearly in number of objects, but exponentially in number of splits/merges. A potential bottleneck is the trial-and-error fusion, if the set of not fused regions per Kalman track is big. So, we limited the maximum number of considered regions and take only the ones with smallest Mahalanobis distance to Kalman prediction.

## VI. Conclusions

In this paper, we presented a way of fusing three different detection approaches for measurement generation and reconstruction in Kalman tracking. One point-feature-based and two region-based algorithms are working complementary and offer different properties and advantages, which are used for fusion. Since point-feature existence can't be guaranteed in situations of object occlusion or weak gray-value intensity contrast, we first discussed a way of combining only the region detection results. However, we assumed that point-features are a powerful support for the tracking process. So, specific split- and merge-handling using the features was implemented and evaluated. In two example scenes, we demonstrated the benefit of considering the point-features for fusion. Split, merge and occlusion situations were handled better than with fusion of the regions only. Compared to many other approaches of feature fusion for tracking, our concept appeared to be very fast with a runtime less than $1\,ms$ per frame.

The fusion and measurement reconstruction concept is independent of specific detection algorithms and just needs any kind of region and point-feature detections. Thus, it is possible to use it in other applications, too, where moving objects are to be tracked such as driver assistance, people tracking or surveillance in general.

## Acknowledgement

This work has been done in cooperation with Carl Zeiss Optronics GmbH and was supported with funds from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. SP1-Cooperation-218290.

## References


[1] E. R. Kalman, "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering*, vol. 82, no. 1, 1960.

[2] M. Grinberg, F. Ohr, J. Beyerer, "Feature-Based Probabilistic Data Association (FBPDA) for Visual Multi-Target Detection and Tracking under Occlusions and Split and Merge Effects", in *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, USA, October 3-7, 2009.

[3] W. Krüger, Z. Orlov, "Robust Layer-based Boat Detection and Multitarget-tracking in Maritime Environments", in *Proceedings of the 2nd NURC International Conference on WaterSide Security (WSS)*, Marina di Carrara, Italy, November 3-5, 2010.

[4] M. Teutsch, W. Krüger, "Classification of small Boats in Infrared Images for maritime Surveillance", in *Proceedings of the 2nd NURC International Conference on WaterSide Security (WSS)*, Marina di Carrara, Italy, November 3-5, 2010.

[5] F. Liu, J. Sparbert, C. Stiller, "IMMPDA Vehicle Tracking System using Asynchronous Sensor Fusion of Radar and Vision", in *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium*, Eindhoven, The Netherlands, June 4-6, 2008.

[6] U. Scheunert, H. Cramer, B. Fardi, G. Wanielik, "Multi Sensor based Tracking of Pedestrians: A Survey of Suitable Movement Models.", in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 14-17, 2004.

[7] X. Li, K. Wang, W. Wang, Y. Li, "A Multiple Object Tracking Method Using Kalman Filter", in *Proceedings of the 2010 IEEE International Conference on Information and Automation*, Harbin, China, June 20-23, 2010.

[8] N. Li, L. Liu, D. Xu, "Corner Feature Based Object Tracking Using Adaptive Kalman Filter", in *Proceedings of the 9th International Conference on Signal Processing (ICSP)*, Beijing, China, October 26-29, 2008.

[9] A. Babaeian, S. Rastegar, M. Bandarabadi, M. Rezaei, "Mean Shift-Based Object Tracking With Multiple Features", in *Proceedings of the 41st Southeastern Symposium on System Theory*, Tullahoma, TN, USA, March 15-17, 2009.

[10] X. Yang, H. Yang, J. Zeng, H. Yu, "Multi-feature Target Tracking Based on Lifting Scheme", in *Proceedings of the 9th International Conference on Electronic Measurement & Instruments (ICEMI)*, Beijing, China, August 16-19, 2009.

[11] H. Wang, C. Liu, L. Xu, M. Tang, X. Wu, "Multiple Feature Fusion for Tracking of Moving Objects in Video Surveillance", in *Proceedings of the 2008 International Conference on Computational Intelligence and Security (CIS)*, Suzhou, China, December 13-17, 2008.

[12] P. Kumar, S. Ranganath, K. Sengupta, H. Weimin, "Cooperative Multitarget Tracking With Efficient Split and Merge Handling", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 12, pp. 1477-1490, December 2006.

[13] R. Liu, Q. Liu, "Tracking Infrared Target with Constructing Multiple Feature Pseudo-color Image", in *Proceedings of the 2010 Symposium on Photonics and Optoelectronic (SOPO)*, Chengdu, China, June 19-21, 2010.

[14] C. R. Jung, A. Said, "Object Tracking Using Multiple Fragments", in *Proceedings of the 16th IEEE International Conference on Image Processing (ICIP)*, Cairo, Egypt, November 7-10, 2009.

[15] R. T. Collins, Y. Liu, M. Leordeanu, "Online Selection of Discriminative Tracking Features", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1631-1643, October 2005.

[16] A. Ondini, G. Marsiglia, L. Fortunato, G. Balzarotti, "Techniques for detection of multiple, extended, and low contrast targets in infrared maritime scenarios", *Optical Engineering*, vol. 45, 2006.

[17] W. Tao, H. Jin, J. Liu, "Unified mean shift segmentation and graph region merging algorithm for infrared ship target segmentation", *Optical Engineering*, vol. 46, 2007.

[18] C. Bibby, I. Reid, "Fast Feature Detection with a Graphics Processing Unit Implementation", in *Proceedings of the International Workshop on Mobile Vision 2006*, http://www.robots.ox.ac.uk/∼cbibby/pubs/paper994.pdf.

[19] S. Fefilatyev, D. B. Goldgof, L. Langebrake, "Towards detection of marine vehicles on horizon from buoy camera", in *Proceedings of SPIE Vol. 6736, Unmanned/Unattended Sensors and Sensor Networks IV*, 2007.

[20] S. P. van den Broek, E. J. Bakker, D.-J. de Lange, A. Theil, "Detection and classification of infrared decoys and small targets in a sea background", in *Proceedings of SPIE Vol. 4029, Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process*, 2000.

[21] J. Matas, O. Chum, M. Urban, T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions", in *Proceedings of the British Machine Vision Conference (BMVC)*, pp. 384-393, Oxford, England, September 24-27, 1990.

[22] L. D. Stone, C. A. Barlow, T. L. Corwin, *Bayesian multiple target tracking*. Artech House, 1999.

[23] J. Shi, C. Tomasi, "Good Features to Track", in *Proceedings of the 1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 593-600, Seattle, Washington, June 21-23, 1994.

[24] Y. Bar-Shalom, E. Tse, "Tracking in a Cluttered Environment with Probabilistic Data Association", *Automatica*, vol. 11, pp. 451-460, 1975.

[25] T. Stephan, M. Grinberg, "Probabilistic Handling of Merged Detections in Multi Target Tracking", *Submitted to 13th International Conference on Computer Vision (ICCV)*, Barcelona, Spain, November 6-13, 2011.