Three-Dimensional Intra- and Inter-Frame Target Detection in IR/mmW Image Sequences

Klaus Jaeger[*], Karlheinz Bers, Klaus Jurkiewicz

FGAN-FOM Forschungsinstitut für Optronik und Mustererkennung
(Research Institute for Optronics and Pattern Recognition)

Gutleuthausstr. 1, D-76275 Ettlingen, Germany

ABSTRACT

This paper describes a model-based method for the automatic recognition of high value targets in multisensor data. A production net is used to represent the knowledge about target structures. The analysis is carried out by a blackboard-based production system with a database stored in an associative memory. The efficiency of the analysis system is illustrated by an example involving the detection of bridges. For this, sensor data has been interpreted which were recorded with an experimental dualmode sensor (IIR, mmW) mounted to an helicopter while flying over the scenario. Image sequences are taken in oblique view with high frequency. The analysis starts with an intra-frame process by extracting cues in the actual sensor data and by the combination of orthogonal information of IR (intensity, direction) and radar (RCS, range) data to estimate target location in space. To improve detection and reduce false alarms, inter-frame processing is applied to exploit intra-frame results of overlapping images of the scanning sensor system resulting in a higher confirmation of the real target and a better discrimination of false alarms.

Keywords: IR image processing, radar data processing, dualmode IR/mmW-seeker, multisensor fusion, model-based object recognition, autonomous systems

1. INTRODUCTION

At present the engagement of strategic High Value Targets (HVT) like airfields, bridges, depots etc. is mainly done from an airplane as platform.[1] In the future, autonomous detection and recognition of military targets will be done more and more by stand-off weapons with multisensor systems and the capability of automatic target detection and tracking from greater distances even under adverse conditions.[2,3] Consequently the successful mission of autonomous missile systems will strongly depend on the performance of intelligent techniques for data interpretation used for navigation, target acquisition and terminal homing.

Within the framework of former investigations,[4] we realized a model-based structure analysis method for target recognition and demonstrated its suitability for seekerhead applications in principle. The basic approach was based on a two-dimensional approximation of the target scenario located on a plane terrain.

In this paper, we describe the extension of the method to a three-dimensional multisensorial data analysis, also considering the requirements of an operational IR/mmW seekerhead system. For specific scenarios, like highly defended fixed installations (e.g. airfields, bridges, shelter etc.) a higher detection probability and a reduction of the false alarm rate will be expected by using model-based target recognition.

Conventional system design for autonomous target acquisition considers the requirements of high detection probability and low false alarm rate through an optimal selection of the detection algorithms and the associated parameters, with an optimization for each individual sensor channel. In comparison to this procedure, the proposed method for the analysis of

---

[*] Correspondence: Email: klaus.jaeger@fom.fgan.de; Phone: +49-7243-992-321; Fax: +49-7243-992-299

multispectral data meets the desired requirements by sensor fusion with an optimized combination of target features for a high probability of detection mainly based on standard image processing methods (low-level processing), whereas the reduction of false alarms is done by the structural model-based approach (high-level processing). In doing so, the algorithm combines structural objects, extracted in each single IR/mmW record (intra-frame fusion), resulting in a symbolic and geometrical description of the complete target scenario (inter-frame fusion).

The application of the developed method to multisensor data of a dualmode seeker system (IIR/mmW) for special HVT-scenarios demonstrates its suitability for autonomous applications.

2. MODEL-BASED SENSOR FUSION

2.1. System Architecture

For the task of target detection in multispectral data the Blackboard-based Production system for Image understanding (BPI) is used.[5,6,7] The BPI-System is a framework for model-based structure analysis of complex scenes and the architecture consists of three basic components: a global database (blackboard), a control unit and a set of processing modules (Fig. 1):

- The analysis starts with a low-level processing or preprocessing of the sensor raw data by extracting proper primitive objects from each individual sensor channel.
- Next, more and more complex parts (partial objects) of the searched target are created from less complex objects by applying production rules of the medium-level processing. Interpreting information of neighbouring and partially overlapping images, the fusion process is activated as soon as possible keeping up with the mission progress of the autonomous system.
- Finally, productions of the high-level process are used to build up the target from partial objects.

The productions are implemented by processing modules which communicate exclusively via the blackboard. The analysis is data-driven and is designed for running in parallel.

In addition to the modules for the dualmode target recognition, modules for background analysis (e.g. texture analysis[8]) and the automatic interpretation of auxiliary information (situation and map analysis) are integrated in the system.

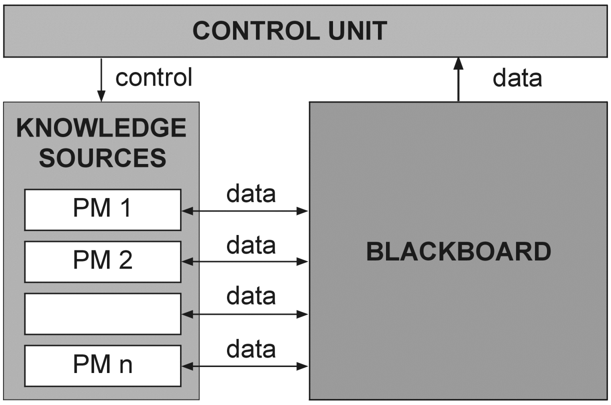


Fig. 1: Architecture of the analysis system

2.2. Target Modelling

The general interactions of productions and the stepwise transfer of objects into objects of higher complexity can be displayed by a production net.[9] It represents the target model or the knowledge about the target. For the description of different structures, different object types with a proper set of attributes are used, e.g. geometric attributes like coordinates, length and width of structural parts.

The model generation process presupposes a map, an aerial photo or comparable target information, given for example by military studies. By this, the model generation process needs surveillance in advance. The level of detail of model description is determined by the requirements of the actual mission phase also defining sensor scan mode and target resolution. For example, due to the system constraints, the target acquisition phase is usually characterized by scanning the scenario in high oblique view. As consequence, target resolution is low and 2D modelling is adequate. The principle procedure is given in the following by modelling a bridge guided e.g. by a map description. The partial objects of the bridge, their size and their geometric relations to each other are extracted from the map.

In our example, the object BRIDGE can be built by applying the productions (P1-P5) repeatedly (Fig. 2). Starting with the object primitives LINE (L), the partial objects LONGLINE (LL) are generated from collinear, prolongable objects LINE. Parallel objects LONGLINE produce the object STRIPE (S) and prolongable objects STRIPE with internal structure are used to build up the object ROAD (RD). Prolongable objects STRIPE without internal structure are combined to the object RIVER (RV) and the object BRIDGE (BG) is produced by an object ROAD crossing an object RIVER.

With this model and the use of known flight- and sensor-parameters the geometric representation of the bridge and its parametric description can be computed at the actual mission time.

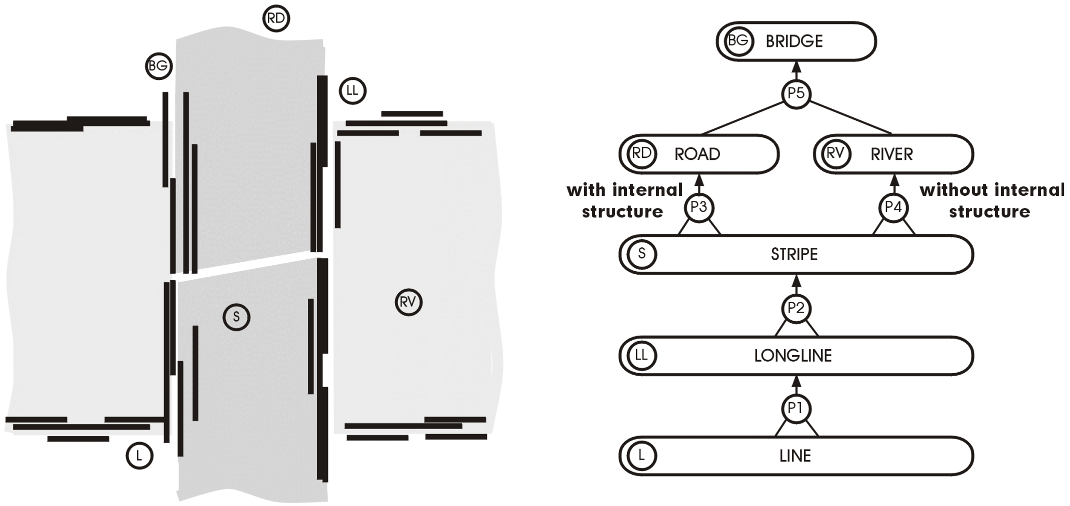


Fig. 2: Production net for object BRIDGE

A more detailed description of geometrical and structural features could be generated from a suitable 3D wire-frame model. This representation is especially useful for target recognition and localization during the terminal homing phase of an operational seekerhead system.

To demonstrate the method of 3D modelling, we applied – as an example - the *3D Builder* productivity tool,[10] which is designed to reconstruct buildings out of photos. It is a commercial desktop system used in photogrammetry, which does not require any special camera system. The bridge model is constructed by combining information from 3 photos taken with a commercial 35mm camera. Fig. 3 shows the input and the result based on the interpretation of about 170 reference points interactively identified by the user.

Fig. 3: Generation of the 3D target model

2.3. Fusion

The fusion of multisensor data is based on a spatial confirmation of geometrical target features in a world reference system. In the next chapter, it will be demonstrated, how the fusion of objects ROAD in infrared images crossing an object RIVER and objects STRIPE in radar images will yield the bridge detection.

For spatial fusion it is necessary to know the process of image generation, which enables the transformation of sensor data into the reference system (registration). In particular, the sensor parameters must be available which consist of the extrinsic (position, orientation) parameters - given by inertial measurement units (IMU) and/or (D)GPS systems - and intrinsic (focal length, range resolution...) parameters, calibrated by experiments or given by hardware specification.

3. INTRA- AND INTER-FRAME SENSOR FUSION

The main emphasis of this work is the realization of the proposed solution concept and the examination of the developed method applied to dualmode sensor data. In the following, the results of each module of the implemented HVT-recognition system will be explained on data taken by a multisensor seeker system in IR and mmW spectral range. Low-level processing is applied to each single image of the sequence resulting in intra-frame fusion of primitive objects (e.g. line elements) detected in both sensor channels. Based on known sensor orientation parameters intra-frame objects are projected into a reference coordinate system to enable the construction of more complex objects by inter-frame interpretation.

3.1. Dualmode Sensor System

The dualmode data were recorded during flight trials with an experimental multispectral seeker system.[11] The sensor consists of an imaging infrared sensor (8-12µm) and a frequency modulated continuous wave radar (94GHz) and is provided with a scanning antenna. The individual sensors are gimbal-mounted and built into a common boresighted aperture system. In addition both sensors cover the same field-of-view enabling fully synchronization of the recorded image sequence in time and space (Fig. 4a).

Each frame of the multisensor-system consists of an infrared image and two radar-spectra of reflected power (radar-ramp) corresponding to two channels of co- and cross-polarization (Fig. 4b). The infrared sensor supplies a two-dimensional angular resolution and the radar a one-dimensional range resolution inside the common field-of-view.

During the captive flight tests, the sensor frontend was fixed forward to the transport helicopter under a variable depression angle. The scene is scanned in a pushbroom way producing a sequence of overlapping images. Scan rate, angular and flight velocity leads to a 100% coverage of the scenario. Sensor- and flight-parameters were recorded during the measurements enabling inter-frame fusion of the structural objects extracted in the single frames by intra-frame processing.

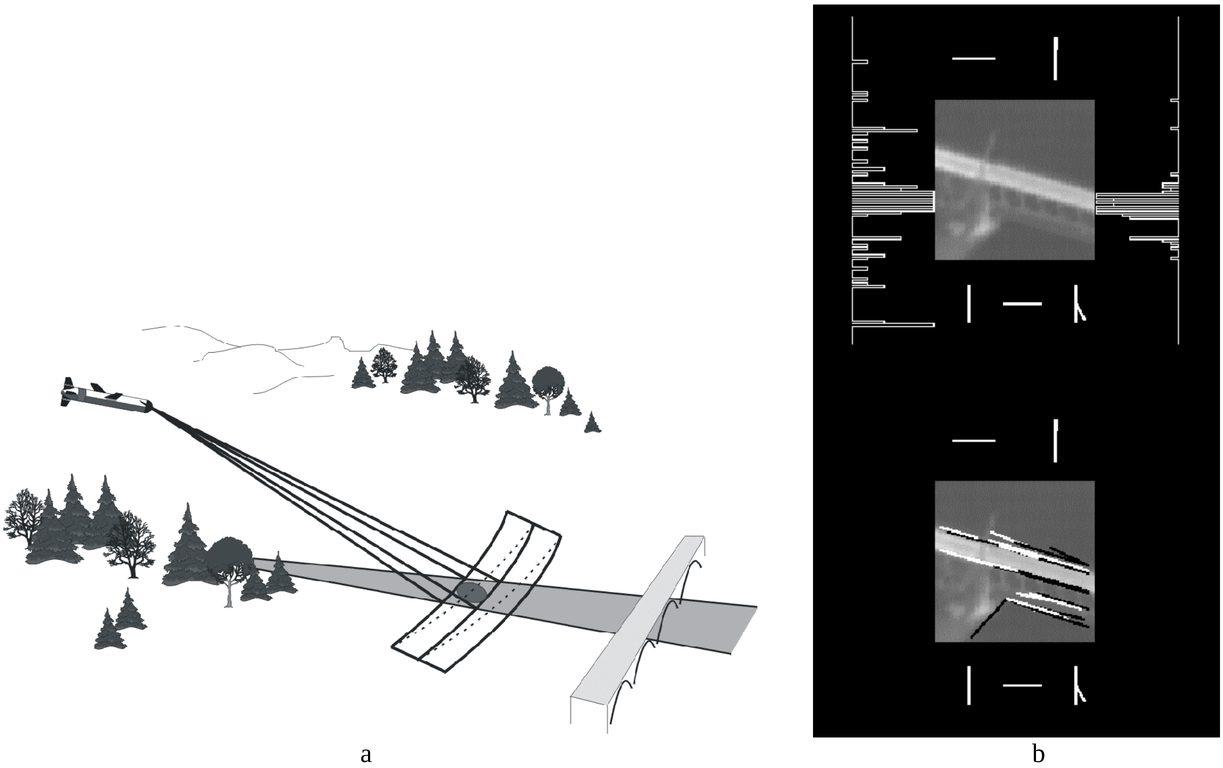


Fig. 4: a) Scan technique during flight, b) Single frame and IR-preprocessing

3.2. Bridge model

A general production net for an object DUALMODE-BRIDGE which reflects the benefits of a multisensor system for target detection is shown in Fig. 5: The bridge can be generated by the combination of structures extracted in the same or in different sensor channels. This takes into account that the bridge can be detected even if one sensor channel fails or certain structures are not found in an individual sensor channel.

The object DUALMODE-BRIDGE is produced by the overlapping region of objects RADAR-ROAD which are detected in the radar range map and objects IR-ROAD crossing an object IR-RIVER. The production net for the partial objects RADAR-ROAD, IR-ROAD and IR-RIVER can be deduced from the production net for the object BRIDGE crossing a river (Fig. 2).

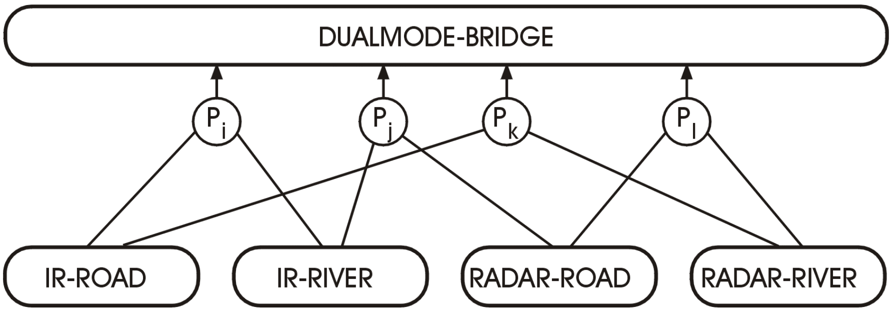
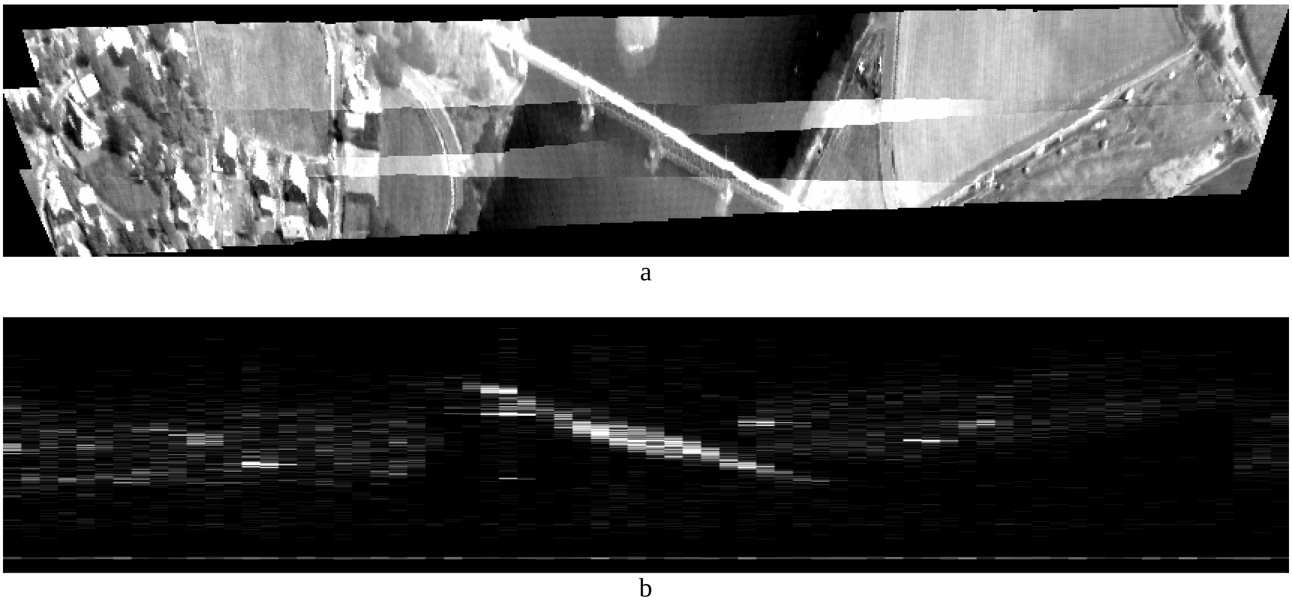


Fig. 5: Production net for object DUALMODE-BRIDGE

Because of the narrow field-of-view of the system, a single image contains only a small section of the scene (Fig. 4b). To get a complete covering of an extended HVT, like a bridge, it is necessary to assemble the partial results of the analysis, obtained in each IR-frame and each radar-ramp. This is done by registering the objects extracted from each snapshot in a common reference system of coordinates.

In Fig. 6a an infrared ground map is represented to visualize the acquisition of the bridge. The map shows the result of the registration of successive images belonging to horizontal scans. The projection of the pixels into a plane ground is computed on the basis of the sensor orientation measured by an INS. The corresponding sequence of radar spectra (e.g. co-channel) is shown in Fig. 6b. The range map is a distance-to-sensor (row) over scanning angle (column) representation of the reflected radar power. In the range map, the radar signature of the bridge appears as a bright stripe and in the infrared ground map the bridge can be interpreted as a road crossing a river.



a

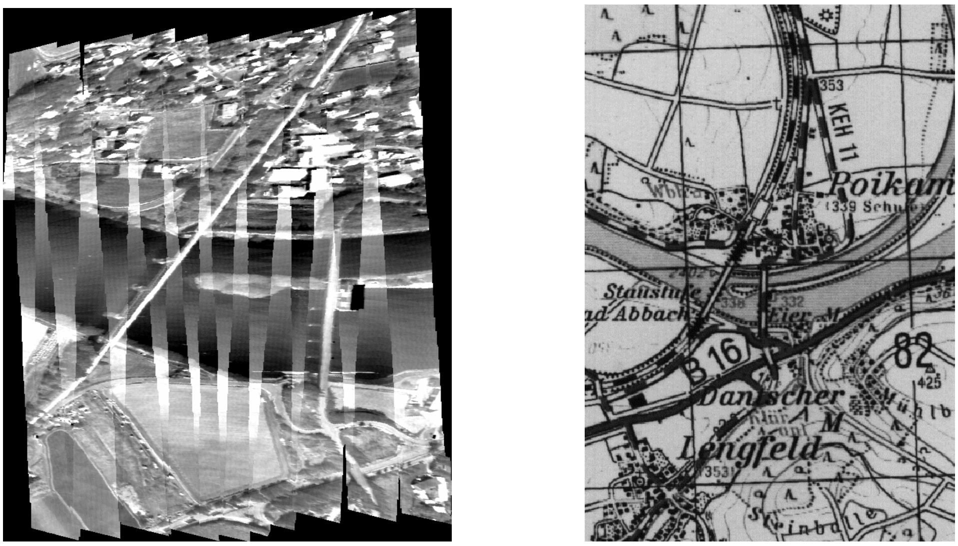
b

Fig. 6: a) Infrared ground map, b) Radar range map

3.3. Results

Although the results of the experiments with data of the dualmode system are visualized in the following figures in the infrared groundmap, the low-level processing of IR-primitive objects (LINE) is done directly on each single IR-image (Fig. 4b) and the objects RADAR-ROAD are detected in radar range maps (Fig. 6b).

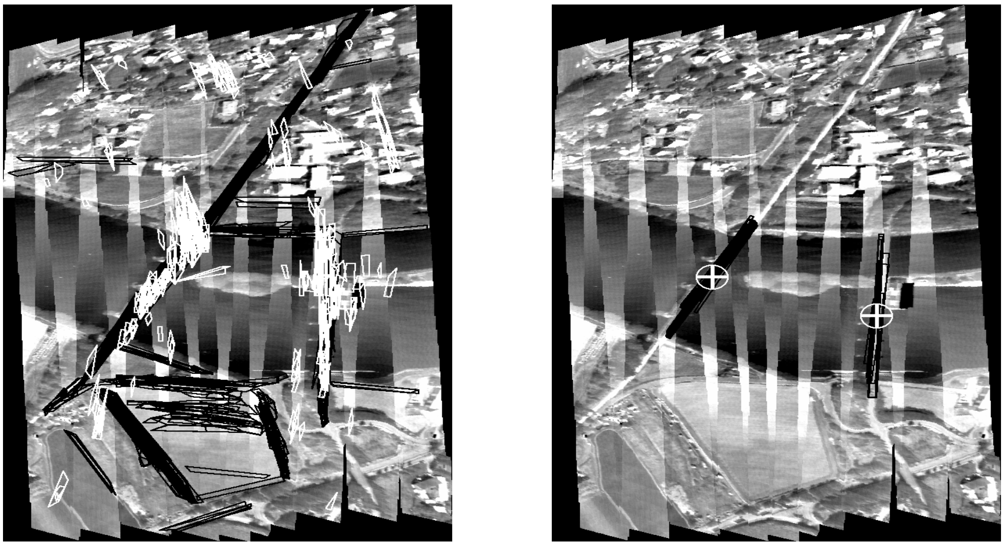
The interpreted sensor data belong to a scenario with two bridges crossing a river. Besides the targets, the scene contains a variety of natural and man made background objects like meadows, trees, houses and roads. Fig. 7a shows the computed infrared ground map which is generated by the registration of several thousands of IR-images. For the model generation the geometrical parameters of the bridges (e.g. length and width) were taken from a map (Fig. 7b).



a b

Fig. 7: a) Infrared ground map, b) Map information

The results of generating the objects IR-ROAD and RADAR-ROAD are represented in Fig. 8a. The objects IR-ROAD (dark) are built up by lines which were extracted in the single IR-images. Objects RADAR-ROAD (bright) are created by the ground projection of stripes which are analysed in radar range maps and again built up by lines. Due to the lack of lateral angular resolution of the radar sensor, the area of an radar object was interpreted as a rough spatial estimation of a target section. For the extraction of lines in the IR-image and in the range maps standard preprocessing methods were used.



a b

Fig. 8: 2D target recognition (acquisition phase)
a) Objects IR-ROAD (dark) and RADAR-ROAD (bright), b) Bridge detection

Fig. 8b shows the results of the fusion process: the objects BRIDGE (dark) are detected unambiguously and their centre is marked by a reticle (bright). The models of the two targets BRIDGE only differ in their parameters.

4. 3D INTER-FRAME ANALYSIS

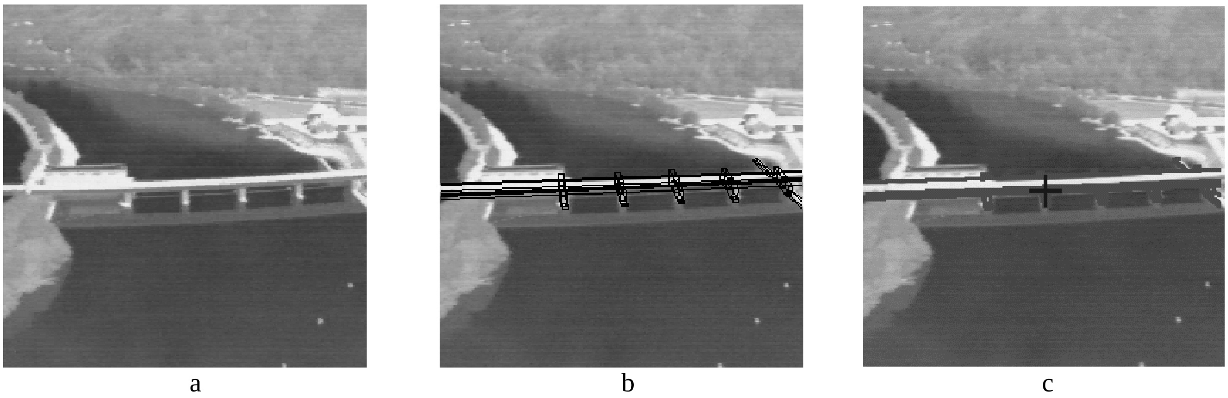
Terminal homing with imaging sensors usually starts at distances of one or two kilometers to the target. During this phase, high resolution sensors provide imagery used to classify military targets and improve the selection of missile impact aim points. However, due to the fact that inertial data are affected by noise or drift and GPS-data may be disturbed by countermeasures, accurate measurement of missile pose relative to the target may be difficult. To overcome this problem, model-based image processing is proposed combining automatic target recognition and sensor pose determination.[12] Moreover, because of the strongly distinctive 3D geometrical attributes of objects, it is assumed that this approach will yield a system with high target detection probability and precise 3D aim point localization.

The method is based on a geometrical match of target features - extracted by 2D structural image analysis - with the projected elements of the associated 3D model (Fig. 3). Knowledge about the approximated sensor pose, measured for example by IMU or GPS, enables to estimate an expected target projection used for solving the correspondence problem of target and model elements. These correspondences are presupposed for pose computation carried out by a nonlinear numerical optimization algorithm.

Because of tolerances in parameter values and inaccurate knowledge about the model, at each level of the inference process it is possible, that cues of target-like objects occur in some images of the sequence, which do not correspond to the target. To eliminate such false alarms and to reduce the number of hypotheses for the pose determination, an inter-frame tracking algorithm is implemented based on the verification of the previous target position and the expected target size as well as a filter considering the plausibility of the platform behavior.

For a qualitative assessment the estimated sensor pose was used to calculate the projection of the 3D wire-frame model into the image plane and to superimpose these model structures with the image data frame by frame. Fig. 9 shows for a single IR image (a) the results of pose determination (b) and aim point selection (c). The image is selected out of a sequence taken during target approach by an helicopter. The potential aim point between pillar and roadway of the object BRIDGE is marked by a dark cross hair (c).

Interpreting the complete sequence, good alignment of target image signature with model structures could be stated.



a b c

Fig. 9: 3D target recognition (homing phase)
a) IR image, b) Pose determination, c) Aim point selection

5. CONCLUSIONS

The examinations carried out with data of a multisensor system have shown that knowledge-based structure analysis combined with intra- and inter-frame fusion is suitable for autonomous detection, classification and identification of characteristic targets. Therefore the method is appropriate for seekerhead and drone applications, presupposed that the structures and spectral signatures characterizing the targets can be taken by the used sensor system and can be segmented from the background/clutter by image processing methods.

REFERENCES

1. I. Niemzig, R. Steffen, “Engagement of Hard Targets from the Air,” *Military Technology* **11/92**, pp. 48-55, 1992.
2. K.-H. Bers, K. Jurkiewicz, “Model-based Acquisition and Terminal Homing of High Value Targets for Precision Guided Weapons,” 34th DRG's Seminar on Precision Guided Weapons, *NATO AC/243-TP/6*, Brussels, 1993.
3. K.-H. Bers, K. Jaeger, ”Model-based Target Detection in Multi-Sensor Data,” In: Buser R. G., Wolff H. (ed) Symposium on Multisensors and Sensor Data Fusion. *NATO AC/243-TP/8*, Brussels, 1994.
4. K. Bers, K. Jaeger, K. Jurkiewicz, “Model-based Sensor Fusion for Target Detection,” *SPIE Proceedings Vol. 2756*, Orlando, 1996.
5. L.D. Erman, F. Hayes-Roth, V.R. Lesser, R. Reddy, “The HEARSAY-II speech-understanding system,” *Comp.Surveys* **12**, 1980.
6. P. Nii, “The blackboard model of problem solving,” *AI Magazine* **7/2**, 1986.
7. K. Lütjen, “BPI: Ein blackboardbasiertes Produktionssystem für die automatische Bildauswertung,” *Informatik-Fachbericht 125*, G.Hartmann (Hrsg), pp. 164-168, Springer Verlag, Berlin, Heidelberg, 1986.
8. B. Bargel, ”Automatische Klassifikation von Fernerkundungsdaten durch statistische und strukturelle Texturanalyse,” Dissertation, *Universität Karlsruhe*, 1983.
9. U. Stilla, “Map-aided structural analysis of aerial images,” ISPRS Journal of Photogrammetry and Remote Sensing, **50(4)**, pp. 3-10, 1995.
10. *3D Builder user’s guide 3.0*, 3D Construction Company, Elizabethton, Tennessee, USA, 1998.
11. F. Nivelle, W. Rudolph, F. Winter, “ASTRID: Autonomous Search and Track with Radar and Infrared Dual-Mode IR/mmW Dual-Mode Seeker,” In: Proc. to MSP 4th Symposium on Technologies for Precision Air Strike Operations in Rapid-Reaction and Localised-Conflict Scenarios, Sevilla, Spain, 1995.
12. K. Jaeger, K.-H. Bers, K. Jurkiewicz, “ATR-System for 3D-Target Approach,” In: Fourth Joint International Military Sensing Symposium (MSS), NATO RTO-SET PANEL, Paris, 2000 (in prep.).