Integration of Automatic Detection Algorithms into Interactive Image Interpretation

A. Berger, R. Eck, N. Heinze, E. Peinsipp-Byma
Fraunhofer – IITB
Fraunhofer Str. 1
76131 Karlsruhe - Germany

1. ABSTRACT

An automatic Target Detection Algorithm (ATD) is integrated in the interactive image interpretation process. The detection algorithm works on large infrared line scanner images and detects military vehicles. A concept for the integration into an image interpretation system was developed and discussed with military image interpreters. After implementation test experiments were carried out with image interpreters, which showed improved performance by using ATD in comparison to "human interpreter" alone.

Keywords: interactive image interpretation, ATR, human-machine interface, screening, detection

2. INTRODUCTION

In the field of airborne and satellite image interpretation the performance of digital sensors is improving continually. Image interpreters have to exploit this growing amount of data as fast as possible. A clipped part of a typical image is shown in Figure 1.

Figure 1: Section of an infrared line scanner image.

To be able to manage such a heap of data the image interpreters need assistance. One possibility for assisting the image interpreter is the use of automatic image exploitation algorithms. In the field of remote sensing conditions as day time, weather and season may vary with each image. But automatic algorithms need defined conditions to produce perfect results. Anyway they offer advantages like high processing speed and objectivity and they don't get tired.

Goal of this project was to integrate an automatic target detection algorithm (ATD) into the interactive image exploitation process to improve the performance of image interpreters and to reduce work load. A workflow analysis allowed finding out how to optimize the integration of the algorithm. In addition the human-machine interface (HMI) was optimized.

3. STATE OF THE ART

In general it is assumed that work load of humans is reduced if they get support of automatic algorithms [1]. With regard to support the image interpreters by automatic image exploitation algorithms different aspects are considered. [2] investigated the perceived utility of human and automated aids. [3] reviewed the correlation between the ATR (target recognition algorithm) result visualization and the performance of the image interpreter using ATR. [4] investigated the correlation between ATR performance and image interpreters' performance using ATR. [5] carried out experiments with few small images and unskilled image interpreters. The result was that in simple optical images the performance with aid of ATD did not improve. With more demanding IR images the results improved. Another observation was that image interpreters relied too strongly on ATD-results.

In most of these investigations simulated ATD or simulated ATR results were used. There are two reasons to do so. On one hand investigations are easier to interpret when defined boundary conditions are used. But, as mentioned above, available automatic image exploitation algorithms can't deliver results with absolutely stable performance. On the other hand automatic algorithms with high performance are not easily available. In 1999

Fraunhofer IITB carried out experiments to find out how the image interpreters performance behaves when using a real ATD [6]. The results of this investigation showed that the performance of a well trained interpreter will not be improved using ATD if the image interpreter has to exploit small images without time pressure.

4. AUTOMATIC TARGET DETECTION ALGORITHM

In our present research with the objective to develop an optimized system to assist the image interpreter effectively we used and improved an ATD which was developed at Fraunhofer IITB to detect man made objects [7], [8], [9]. This algorithm is especially used to assist human image interpreters by detecting vehicles, single or in groups, in infrared line scanner images (IRLS) under high time pressure. IRLS images have a typical size of up to 2 GB. The large image size makes it difficult for human image interpreters to perform their exploitation task in the demanded short time.

ATD uses a training approach for detecting suspicious regions in IRLS. The tuning phases of the ATD are shown in upper part of Figure 2. There is a tuning stage in the laboratory, in which the classificatory is trained to detect the relevant types of objects and a screening phase, in which the ATD- process screens large images in near real time to detect those objects (lower Figure 2).

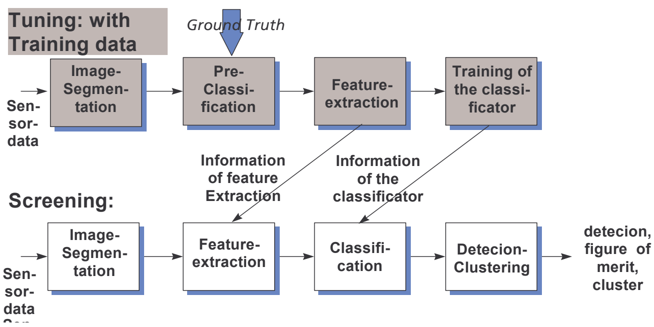

Figure 2: Processing diagram of the ATD.

The automatic image exploitation starts with the elimination of specific distortions like jitter before further processing steps can be done. A segmentation step follows to extract preliminary suspicious areas. Only these areas are processed in further steps (Figure 3, Figure 4).

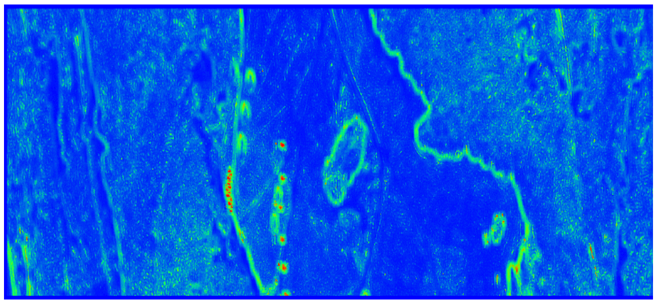

Figure 3: Intermediate result of ATD-algorithm, suspicious areas are bright.

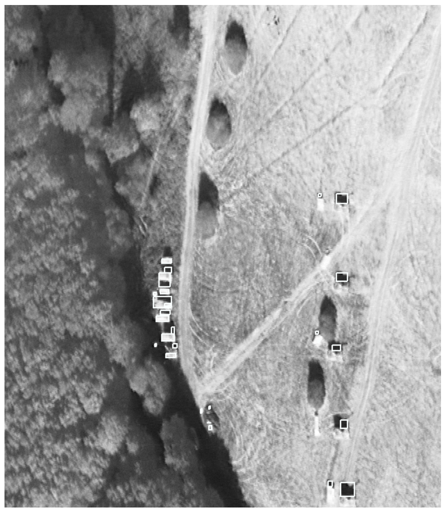

Figure 4: Result of first detection step. White boxes mark suspicious areas.

The task of classification is now to distinguish between signatures of vehicles (Figure 5) and of clutter (Figure 6).

The identification of features which are significant for the objects of interest and also their confusion objects is task of the tuning phase. During this phase ground truths for the images are available. So for every signature the correct class is known. Features with good separating capacity are selected by using separate measures [10]. In the suspicious regions the selected features are extracted, classified and result in single target detection hypotheses. Every detection hypotheses is combined with a figure of merit, so selection of the best hypothesis can be done afterwards in the interactive exploitation interface. This figure of merit is a measure of the similarity of the extracted signatures compared to signatures presented during the tuning phase.

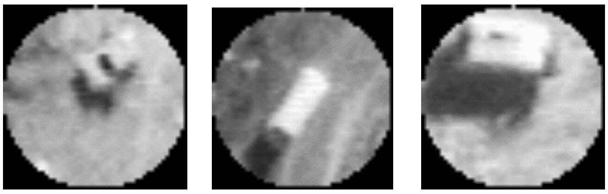

Figure 5: Examples of vehicles.

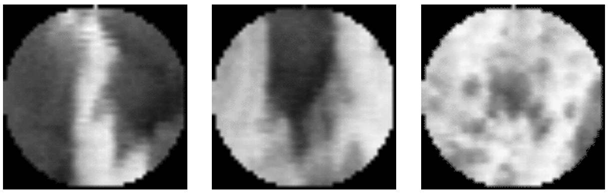

Figure 6: Examples of clutter.

Based on these single target detections the next step generates clusters of detection hypotheses. Based on the figure of merit of the single targets a figure of merit is calculated for these clusters which are annotated by a convex polygon – see Figure 8. This step is simple but most important for the acceptance of the algorithm by the image interpreters. Otherwise a large amount of single object hypotheses stresses their visual system. These Regions of Interest (ROI) can be selected due to the scoring of their figure of merit. Because of the clustering step ROIs generate fewer cues - see Figure 9. An evaluation of the algorithm can be done by computing the Receiver Operation Characteristic (ROC) with the detection-rate and the false alarms [11]. The ATD works in real time and is already integrated in a demonstrator for image exploitation [12].

The used ATD gains high detection rates but also has high false alarm rates too. Due to the high false alarm rate in this non-cooperative environment it is not possible to use ATD results without human supervision. So it is important to integrate the ATD into the image exploitation process in a manner that the assistance of ATD shows improved performance of the human interpreter.

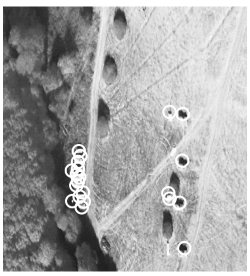

Figure 7: Detection result (white circles) of single vehicles.

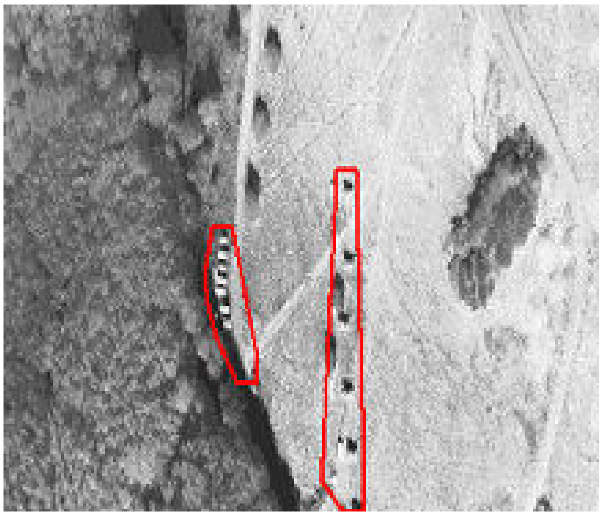

Figure 8: Clustering step results in groups of vehicles.

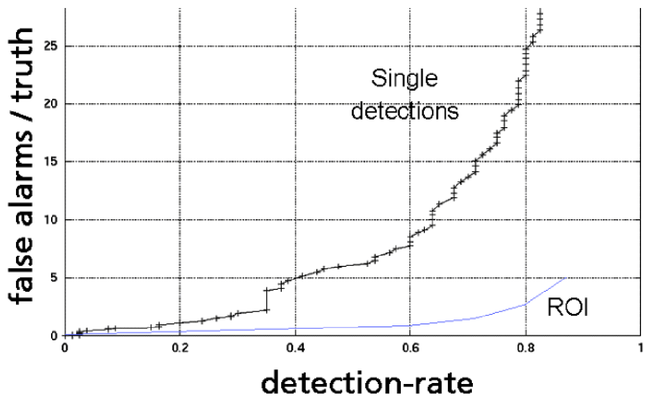

Figure 9: Typical ROC of single detections and ROI.

5. SCENARIO

The reconnaissance cycle in this examination is based on the UAV (Unmanned Aerial Vehicle) CL-289 scenario of the German Army. The UAV CL-289 (Figure 9) is used in close-up range up to 40 km for combat field reconnaissance. For situation reconnaissance the UAV has a range up to 170 km. Up to maximum 75 km a radio link is available for live downloading infrared line scanner data to the ground station. At the same time a wet film is exposed by optical image. The downloaded IR image only is currently used for image interpretation and used for ATD in our examination. Up to 10 images are produced each flight. The images show areas of up to 10 km length. The cutter distributes the image data to the interpreter groups.

The time budget to send the first report depends on the kind of report which has to be done. A first report in

ADLER mission must be sent within 5 minutes after interpretation start, time counting starts with the beginning of the image exploitation. Reporting in RECCEXREP gives the interpreter up to 45 minutes.

The interpreters prepare themselves by flight data plan, initiator demand, and digital maps. According to the demands of German Army artillery communication system ADLER or of German Air Force RECCEXREP reports were compiled by the interpreter.

6. Integration of the ATD in the Exploitation Work-Flow

One of the main tasks of a CL-289 image interpreter is to detect military vehicle groups and single military vehicles as fast as possible. It is to assume that the performance of the image interpreter will grow if he is assisted by the ATD as described in chapter 3. In order to reach the best assistance by ATD one main focus of this work was the integration of the ATD into interactive image exploitation workflow. HMI specialists developed different HMI versions, working together closely with both, image interpreters of the CL-289 units and specialists for the ATD. The favorite versions developed were used in the experiments.

The main task of the ATD is to reduce the image regions which have to be inspected by the image interpreter. To recognize these regions at a glance an overview image is necessary, including the ATD results. Figure 10 shows the HMI with the overview image on the left side. A rectangle within the overview image shows the region which is presented in detail on the right side.

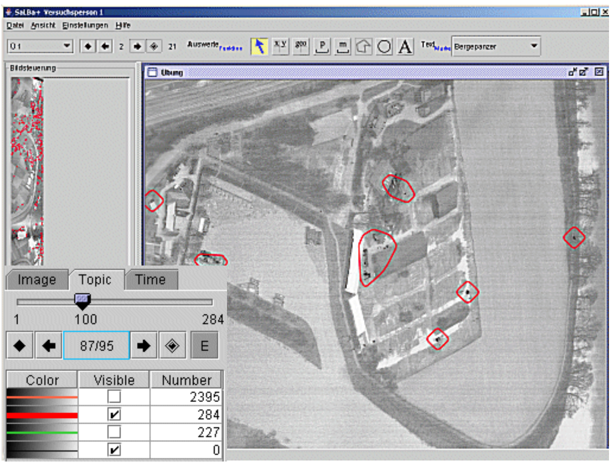

Figure 10: HMI with overview image on the left side, detail image on the right side, and ATD results.

To manipulate the image a menu is available by clicking into the image with the right mouse button. This context menu offers the usual image manipulation possibilities as contrast and brightness adjustment, image rotation, zoom, and grey value inversion.

As mentioned in chapter 3 each ATD result has a figure of merit. This figure of merit is used to offer the image interpreter the possibility to control the number ATD results. This can be accomplished in two ways: first the best 10, 20, or 50 ATD results can be selected by radio buttons or second an arbitrary number of ATD results is chosen by moving a slider until the appropriate number is displayed. Moving the slider changes the visualization of the chosen ATD results in real time.

To pass in a very short time through the ATD results the image interpreter can use keyboard buttons or a menu to jump from one ATD result to the next. The ATD results are sorted by their image coordinates. This allows the image interpreter to move from the bottom of the image up to the top or vice versa. Of course any path through the image is possible too.

7. Experiments

Goal of the experiments was to find out whether the workload of the interpreter will be reduced and therefore performance will rise by using ATD assistance under time pressure. The experiments were executed with six trained CL-289 interpreters of the German Army using image data from military training scenarios. The experimental image exploitation system SaLBa+ shown in figure 10 was used. SaLBa+ offers the possibility to log user interactions with time stamp. These log files are the basis for a later statistical analysis.

Figure 11: Military interpreters in mobile laboratory.

Interpreters had the task to detect all military vehicles in several images under time pressure. The available time varied between 3 and 10 minutes for each image.

Interpreters used three different user interfaces (HMI). HMI-1 had no ATD implemented. HMI-2 showed ATD results as polygons. The number of visualized polygons could be controlled in two manners. The number of polygons with the best figure of merit could be selected (20, 50, or 100). A threshold of the figure of merit could alternatively be used to select the best polygons. HMI-3 had additional controls to select the polygons step by step. The polygons were sorted by location from upper left to bottom right of the image. The image was automatically centered to the position of the polygon selected when using the jumping mode. Interpreters placed an indicator over the objects they find. Performance was measured by detection- and false alarm rates.

First of all interpreters were introduced to and trained in ADT and the experimental image exploitation system SaLBa+. Introductory exercises were performed to verify the image interpretation skills of each participant. Then three interpreter groups were constituted. All interpreters got the same images and tasks but on different combinations of HMI and image. To each interpretation task a map with corresponding flight route was provided. A decrementing counter indicates the remaining time for the interpretation task.

To measure the workload of the interpreters NASA-TLX test was used. With this standard test, developed by NASA, a subjective measurement of human demand and performance can be done. Image interpreters' satisfaction was acquired by a questionnaire. A final discussion allowed the interpreters to comment the introduced ideas free from formal experimental restrictions.

NASA-TLX measures mental demand, physical demand, temporal demand, effort, and frustration level separately. The results of NASA-TLX test results showed that the highest human demand appeared without assistance of the ATD (HMI-1). Using the ATD with HMI-2 decreases all types of demand obviously. All types of demand are lowest by using the HMI-3 (see figure 12).

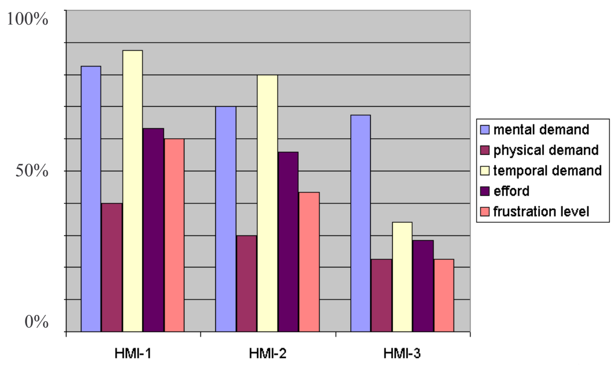

Figure 12: NASA-TLX weighted demand rating.

NASA-TLX asks the interpreter to weight the demand by their importance for their job. So it is possible to calculate a weighted summary of all types of demand, which is shown in figure 13.

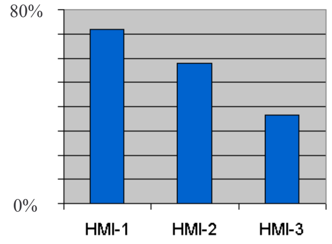

Figure 13: NASA-TLX demand ratings.

Figure 14 gives an overview about the subjective evaluation of the interpreters' performance. HMI-1 got the worst result, the subjective evaluation of HMI-2 is better. At all, HMI-3 also got again the best result.

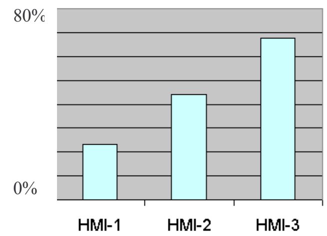

Figure 14: NASA-TLX performance rating.

NASA-TLX confirmed what the image interpreters reported: The introduced ATD assistance reduces their demand and allows them to reach higher performance, especially using HMI-3. All involved interpreters accepted ATD assistance in our experiments and argued for ATD assistance in their daily tasks. One participant reported that "ATD is the third eye of the image interpreter".

8. Results and Outlook

With the work described above it was shown how to integrate an available ATD into a nowadays existing workflow, so that the workload of the image interpreter is reduced.

Further it was shown that a detailed analysis of the existing image exploitation workflow and the needs of the image interpreters are absolutely necessary and the development of the HMI has to be done by experts in closed cooperation with the ATD developers and the image interpreters. Especially the design of the HMI affects the performance which is reached by the image interpreter using ATD.

In a next step experiments will be carried out to examine the performance of the image interpreter using the ATD in an objective way.

9. ACKNOWLEDGMENT

The authors would like to acknowledge the support of the German Ministry of Defense, the German Federal Office of Defense Technology and Procurement (BWB) and the Wehrtechnische Dienststelle 81 (WTD 81). The authors also want to thank the personnel in several reconnaissance units of the German Army and German Air Force for their assistance and the valuable discussions.

10. REFERENCES

[1] K.-F. Kraiss, “Benutzergerechte Automatisierung – Grundlagen und Realisierungskonzepte“, at – Automatisierungstechnik 46, R. Oldenbourg Verlag, 1998, pp. 457-467.

[2] M .T. Dzindolet, L .G. Pierce, H.P. Beck, L.A. Dawe, “The Perceived Utility of Human and Automated Aids in a Visual Detection Task”, Human Factors, Volume 44, Number 1, Spring 2002, The Journal of Human Factors and Ergonomic Society – HFES, pp. 79-94.

[3] E.B. Entin, E.E. Entin, D. Servaty, Optimising “Aided Target-Recognition Performance”, Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting 1996, pp. 233-237.

[4] B.A. Eckstein, J. Irivine, “Evaluating the Benefits of Assisted Target Recognition”, Proceedings 30th Applied Imagery Pattern Recognition Workshop (AIPR) 2001, pp. 39-45

[5] M. Maltz, D. Shinar “New Alternative Methods of Analyzing Human Behavior in Cued Target Acquisition” in Human Factors, Volume 45, Number 2, summer 2003, pp. 281-295.

[6] P. Klausmann, E. Peinsipp-Byma, W. Roller, G. Saur, D. Willersinn, "Assessment of machine assisted target detection", F.A. Sadjadi, Society of Photo-Optical Instrumentation Engineers - SPIE-, Bellingham/Wash. Automatic target recognition IX : 7-9 April 1999, Orlando, Florida Bellingham/Wash.: SPIE, 1999 SPIE Proceedings Series 3718, pp. 430-436.

[7] N. Heinze, A. Korn, C.-K. Sung, L. Berger, “Vehicle Classification within the System SAFIR”, in Proceedings of the IAPR TC-7 Workshop “Remote Sensing and Mapping” Technical University Graz, Austria 2.-3. September, 1996, Editors: F. Leberl, R. Kalliany, M. Gruber, R. Oldenbourg, Wien München 1996, pp. 309-317.

[8] M. Mueller, W. Krueger, N. Heinze, “Model-based target and background characterization“ in Proc. SPIE Vol. 4029, Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process; W. R. Watkins, D. Clement, W. R. Reynolds; Eds. Publication July 2000, pp. 88-92.

[9] M. Mueller, W. Krueger, N. Heinze, “Multisensor data fusion and GIS utilization for ATR“ in Proc. SPIE Vol. 4370, Targets and Backgrounds VII: Characterization and Representation; Wendell R. Watkins, Dieter Clement, William R. Reynolds; Eds. Publication September 2003. pp. 36-41.

[10] A. König, A. Korn, F. Quint, M. Glesner. “Two Neural Models for fast Category Learning – Neural Associative Memories and the Restricted Coulomb Energy Mold“, in Proceedings of the International Conference on Artificial Neural Networks ICANN 93, Amsterdam, The Netherlands, September 1993.

[11] U. E. Jaeger, H. Maier-Herburger, C. Stahl, N. Heinze, D. Willersinn "IR and SAR automatic target detection benchmarks" in Proceedings of SPIE Vol. 5426 AUTOMATIC TARGET RECOGNITION XIV 13-15 April 2004, to be published.

[12] H. Maier-Herburger, W. Drescher, D. Menges, J. Schroeder, J. Werner, “Automatic Target Screening and Recognition in IR and SAR Imagery“ in Fifth International Military Sensing Symposium, Vol. 47, No. 1, 9-12 December 2002 NIST, Gaithersburg

[13] Wolfgang Roller, Elisabeth Peinsipp-Byma, Anton Berger, Thomas Partmann, „Experimentalsystem für die interaktive Bildauswertung“ in IITB Mitteilungen 1999