FEATURE EVALUATION FOR TARGET/BACKGROUND DISCRIMINATION IN IMAGE SEQUENCES TAKEN BY APPROACHING SENSORS

Rene Schoene*, Jochen Meidow, Edmond Mauer
Fraunhofer-Institute of Optronics, System Technologies and Image Exploitation,
Gutleuthausstraße 1, 76275 Ettlingen, Germany

ABSTRACT

The conspicuity of different targets in image sequences taken by approaching sensors is addressed in applications such as the assessment of camouflage effectiveness or the performance evaluation of autonomous systems. In such evaluation processes the consideration of background characteristics is essential due to the propensity to confuse target and background signatures. Several discriminating features of target and background signature can be derived. Furthermore, the changing aspect and spatial resolution during an approach on a target have to be taken into account.

Considering salient points in image sequences, we perform a nominal/actual value comparison by evaluating the receiver operating characteristic (ROC) curve for the detections in each image. Hence, reference regions for targets and backgrounds are provided for the entire image sequence by means of robust image registration. The consideration of the uncertainty for the temporal progression of the ROC curve enables hypothesis testing for well-founded statements about the significance of the determined distinctiveness of targets with respect to their background. The approach is neither restricted to images taken by IR sensors nor applicable to low level image analysis steps only, but can be considered as a general method for the assessment of feature evaluation and target distinctiveness.

The analysis method proposed facilitates an objective comparison of object appearance with both, its relevant background and other targets, using different image analysis features. The feasibility and the usefulness of the approach are demonstrated with real data recorded with a FLIR sensor during a field trial on a bare and mock-up target.

Keywords: automatic registration, feature extraction, ROC curves analysis, Mann Whitney U-Test

INTRODUCTION

The given sequences of IR images describe scenarios with an approaching sensor to targets in the field, with typical scale increasing during the approach as illustrated in the upper part of Figure 1. Obviously, a moving airborne platform, e.g. a helicopter, cannot fly along a straight trajectory towards the targets. Platform motions due to air currents and

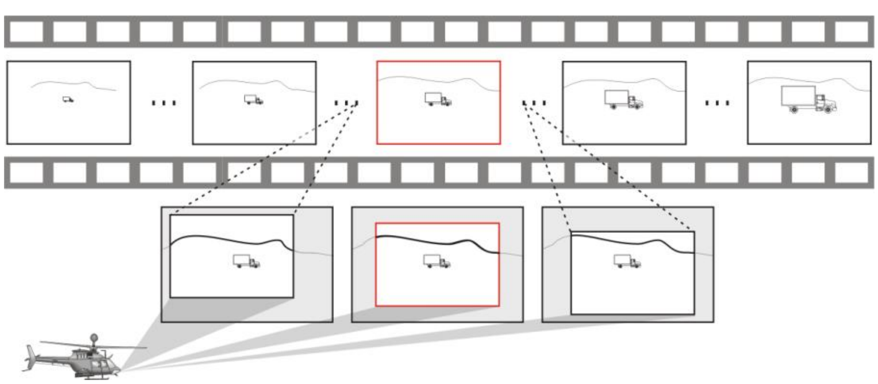

Figure 1: Demonstration of an approaching sensor scenario.

*rene.schoene@iosb.fraunhofer.de

camera movements due to operator interactions result in random displacements of the target position relative to the image coordinates. The target location is roaming through the images in the video sequence, as demonstrated in lower part of Figure 1. Any assessment has to account for these typical properties of a scenario captured by approaching sensors.

Because of continuous changes of the background scenarios through the approaching sensor different features are distinctive in the complete sequence. In far distances the silhouettes of the targets are so small, that they are hardly visible in the frames. The regions of target are comprised only a few two dimensional points in the sequence images. To discriminate between these areas and the background, only point-like feature extraction methods are expedient. Least references (e.g. hot spots from warmer targets in the directly neighborhood) are detectable. But with increasing target silhouettes more and more details are visualized in the IR image. So, entire regions in the target silhouettes are assignable to the real target objects (e.g. wheels or engine hoods) and the uses of complex feature methods are conceivable.

This paper deals with a diversity of feature extraction methods regarding to the evaluations for target/background discrimination. To assess the comparison of targets and backgrounds in the complete sequence we restricted only point-like methods to extract characteristics in the images. To assess the chosen features in each image two evaluation measures are performed on the ROC curve. The considering effect sizes are the Area under the ROC curve and the statistical Mann Whitney U-test for the samples of targets and background.

GENERAL APPROACH

Every evaluation task is based on a comparison between a true reference and the actual samples of the evaluation item. Depending on the evaluation criteria representative measurements of the discrepancy between the expected ideal (e.g. effective target position, characteristic background properties) and the achieved truth (e.g. characteristic target properties) are to be derived. This first evaluation step leads to a set of evaluation values, as shown in Figure 2. Considering the costs and benefits of the evaluation task's aim, these values are summarized to a final figure of merit.

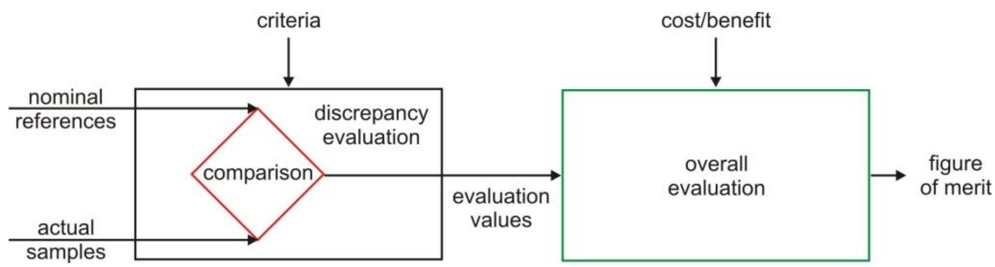

Figure 2: Scheme of the evaluation task.

For evaluation in image sequences using an approaching sensor, the references for the comparison step are basically the truth target and background silhouette in each image of the video sequence, as illustrated in Figure 3. Because of the sensor movement and changing relative camera perspectives, the position of the target and background silhouette is changing from frame to frame as well as the silhouette's appearance. It would be a hard work to interactively define the related image areas in each frame accurately, especially if the background is difficult to distinguish, if the target is disappearing due to effective camouflage or due to the reduced object size at long distances.

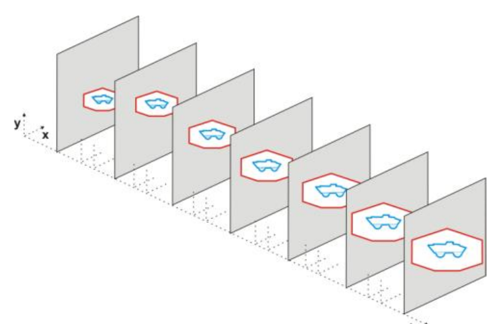

Figure 3: Example of regions of interest (ROI) definition over video frames.

Therefore, we apply an automatic image to image registration, prior to any further analysis, in order to determine motion-corrected references at any time and any position of the image. With the resulting knowledge of geometric transformation behavior from one frame to another it is possible to define the ground truth reference areas once at the best suited distance and to calculate the related silhouettes and positions throughout the entire video sequence. Once this step of defining all regions of interests (ROI) is done, the comparison between nominal references and actual sample values can be performed.

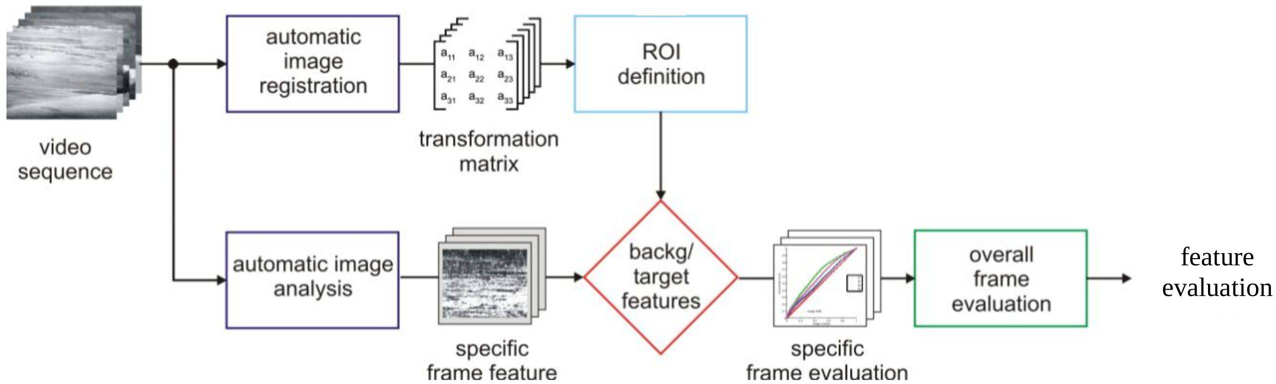

Figure 4: General analysis scheme for feature evaluation by approaching sensor.

Figure 4 outlines the main steps of the general approach. The central part is the automatic registration application to the original video sequence for the definitions of the regions of interests. They are used the comparison of the extracted background/target features by automatic image analysis. The comparative specific results are summarized in a final figure of merit value. We can identify the main steps for evaluation, as shown in Figure 2. In the following we describe more details to the steps of automatic image registration, feature extraction and feature evaluation.

AUTOMATIC IMAGE REGISTRATION

For an automatic registration process, it is generally necessary to have a priori knowledge of the detailed 3D scene structure. In our application the observed sceneries are considered to be approximately planar and so the 3D information is not essential. By using planar homographies the transformations between two consecutive images can be performed [1]. Homography is a straight-line-preserving mapping, which can be described by a 3x3 transformation matrix H_k for two images.

To determine these transformations H_k, firstly conspicuous points (tie points) are automatically extracted in each sequence image through using of feature extraction methods [2]. Thereafter, the statistical method RANSAC is performed for an automatic determination of the correspondences between the tie points in subsequent images of the video [3]. This is a non-deterministic algorithm and is attempted to avoid outliers in the set of correspondences. The outliers can result, e.g., from tie point extraction only in one image or from erroneous measurements through the recording of the camera sensors (defective pixels). Let x_k and x_{k+1} be a representation in homogeneous coordinates of corresponding tie points in the images k and $k+1$ of the sequence. Then, the following relation holds for the transformations H_k

$$x_{k+1} \sim H_k x_k , \tag{1}$$

where $\sim$ means identity between elements in projective spaces. Figure 5a shows those tie points in an IR image, colored in yellow, which indicate correspondences between two consecutive images. The blue arrows visualize the angular direction and the amount of the image to image motion at each tie points.

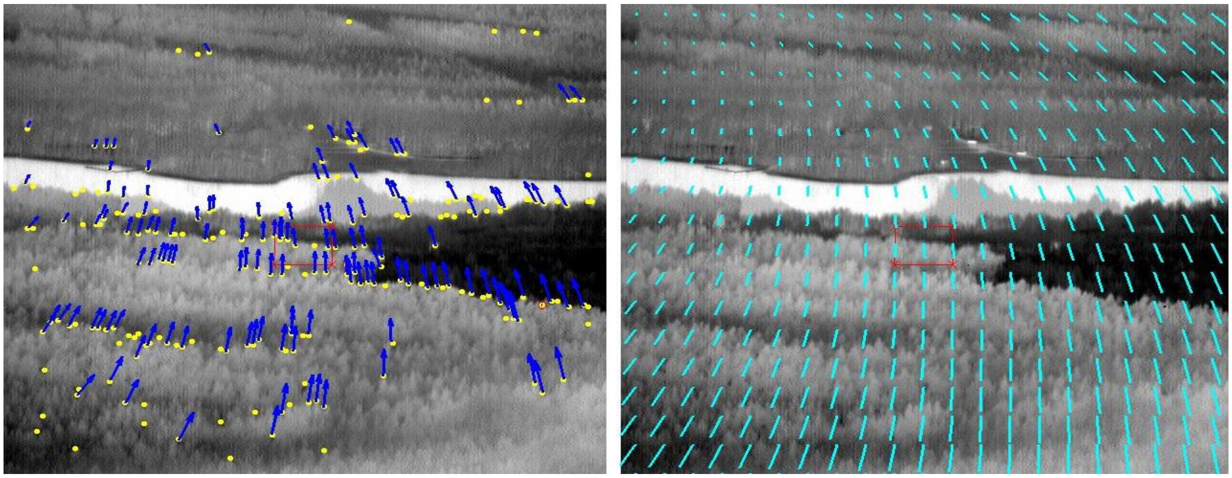

Figure 5: Image to image displacement vector fields for
a) tie points (yellow) with blue arrows (length of arrows is magnified for better visualization) and
b) visualization of the overall vector field represented by fixed grid displacement vector bars.

Determining the image to image transformation from a large number of correspondences leads for the matrix entries in eq. (1) to an over-determined set of equations which may be solved by least-squares techniques. Through the geometric transformation, the displacement in every real-valued image point is calculable. For demonstration, a regular grid was defined in Figure 5b. The cyan lines show the direction and amount of the displacement in each grid point. This registration allows us to determine the motion of the target relative to the sensor.

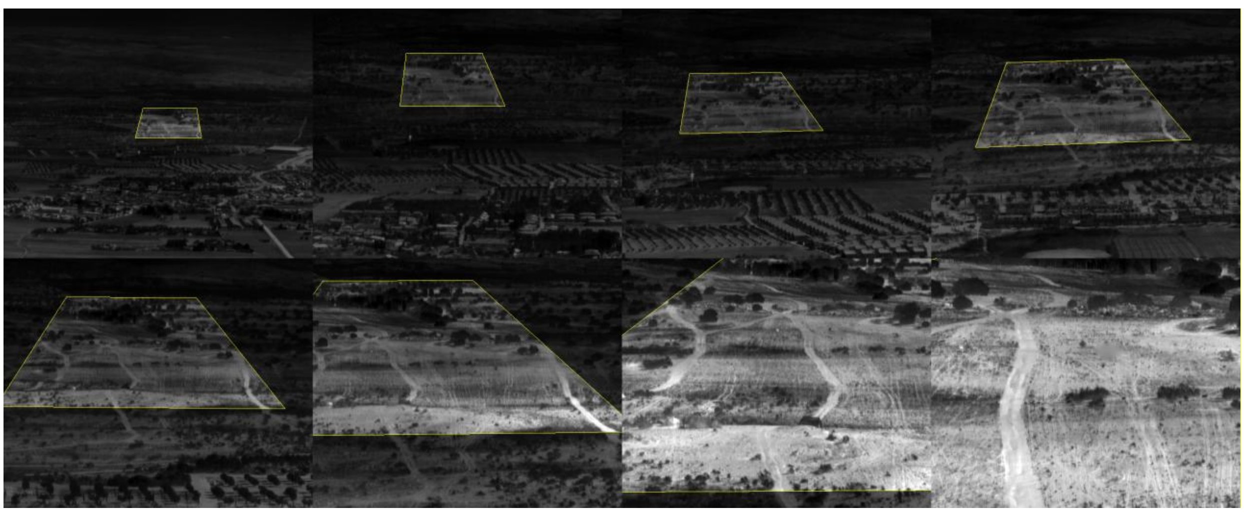

Figure 6: Demonstration of transformations for a background area of interest through the registration.

For any given ground truth position in one image of the video sequence, the nominal target positions in every frame of the video sequence can now be determined by applying the transformations subsequently. The following Figure 6 gives an impression of the transformations of a ROI for the background area in an approaching scene through the automatic image registration. From top left to bottom right it can be seen that the defined background region is increasing, but the shape of the polygonal ROI is changing and more and more details are visible in scene (e.g. scrubs and paths).

FEATURE EXTRACTION METHODS

To analyze different features in the ROIs of the targets and backgrounds, the application of well known image feature extraction methods is necessary. A large diversity of feature extraction methods are already established in the pattern recognition community. Here, only a few features are applied to demonstrate the evaluation process. If we consider, as illustrated in Figure 6, that different details in varied sizes are distinctive at different distances of an approaching scene, the feature extraction step over the complete image sequence is not obvious. The more complex the related feature properties, the more unlikely that a good performance at any resolutions of the approach can be achieved. Basic feature properties at pixel level are e.g. hot spots, conspicuous corners, edges, or small homogeneous areas in an image can be calculate in a large number of the complete image sequence. In the focus of this paper we selected the following feature extraction methods:

- **Intensity (Thermal value):** This is original feature of the thermal emission captured by a thermal camera, the original image. It is the thermal signature of the actual scene. A higher intensity value of an IR image is directly associated to a higher temperature value of the objects in the scene.
- **Foerstner operator:** This operator, a so called interest operator, is used to extract and assess interest points in images [2]. These landmarks decide on the inspected image position, if it is a distinctive point, a point on a straight-lined edge or a point in an unstructured homogeneous area. The operator uses a local auto correlation matrix for the determination of distinctive points in an image. The method calculates two parameters by analyzing the eigenvalues of the matrix.
- **Hot spot operator**: The hotspot filter extracts all regions in an image, which have higher intensity than their neighborhood. This is a point-like method and checks the inspected positions, if these points are maxima in a local subregion of the image. The filter is especially suited for the detection of thermal sources (hot spots), e.g. from in IR images [4].

Figure 7 shows typical results of these three features at the same frame number of a video sequence recorded during the field trial THEO 2009 in Storkow (Germany). Both targets are bounded by an individual polygon. In the left intensity image the left object seems to be warmer as the right vehicle, but e.g. due to sun reflection effects this must not be necessarily the case. Also, the heat path with two bollards and the left electricity shaft are clearly visible. The Foerstner operator in the middle image separates concisely both, the object areas and the background area. In this case, the ROI of the right target is more intensive than the other. Additionally, the two bollards are easily detectable. The right image shows the result of the hot spot operator. Here, the wheels of the vehicles have the highest intensity values in the image, but the path and some trees are also easily visible. These three images give only a little impression of diversified effects by using of feature extraction methods. It cannot be expected that these low level operators are comparable with the complex human detection capacity.

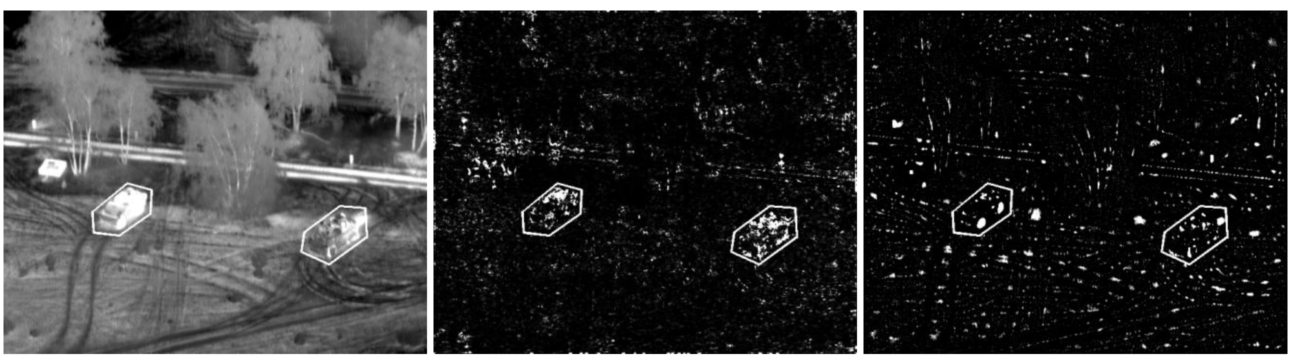

Figure 7: Results of the intensity, Foerstner and hot spot operator with polygonal silhouettes around two targets.

FEATURE EVALUATION

To evaluate the feature characteristics of the targets and backgrounds in image sequences, the automatic registration allows a precise automatic localization of the ROIs at any time of the approach. Considering the approaching scene, as shown in Figure 6, the target silhouettes in earlier frames of the sequence are very small. Additionally, background areas must be chosen in adequate neighborhood to the targets, because other regions located far away from the targets are not relevant for a meaningful evaluation. So, for a target/background comparison over the complete sequence a compromise between a full target description and the less possible background overlapping is necessary. Figure 8 shows three potential interactively designed ROIs, one green and one blue polygon for the targets, and a red polygon for the related background.

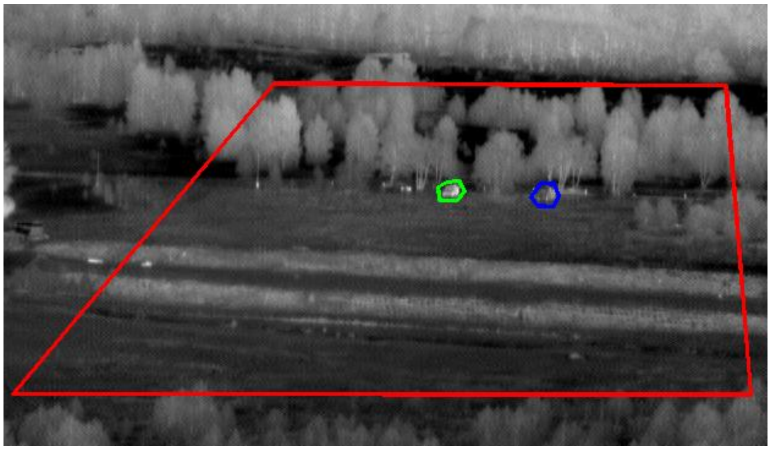

Figure 8: ROI-definitions of bare (left), camouflaged (right) target and common background.

To analyze the ROIs of targets and backgrounds, characteristic properties in these regions must be determined. By applying of the above described feature extraction methods, we get specific distributions for all targets and backgrounds. Exemplarily two resulting distributions of target and background areas are presented in Figure 9 left. For a nominal/actual comparison of these distributions, we distinguish four cases. The confusion matrix in Figure 9 shows the cases: TP as actual target for nominal target, FN as actual target for nominal background, FP as actual background for nominal target and TN as actual background for nominal background.

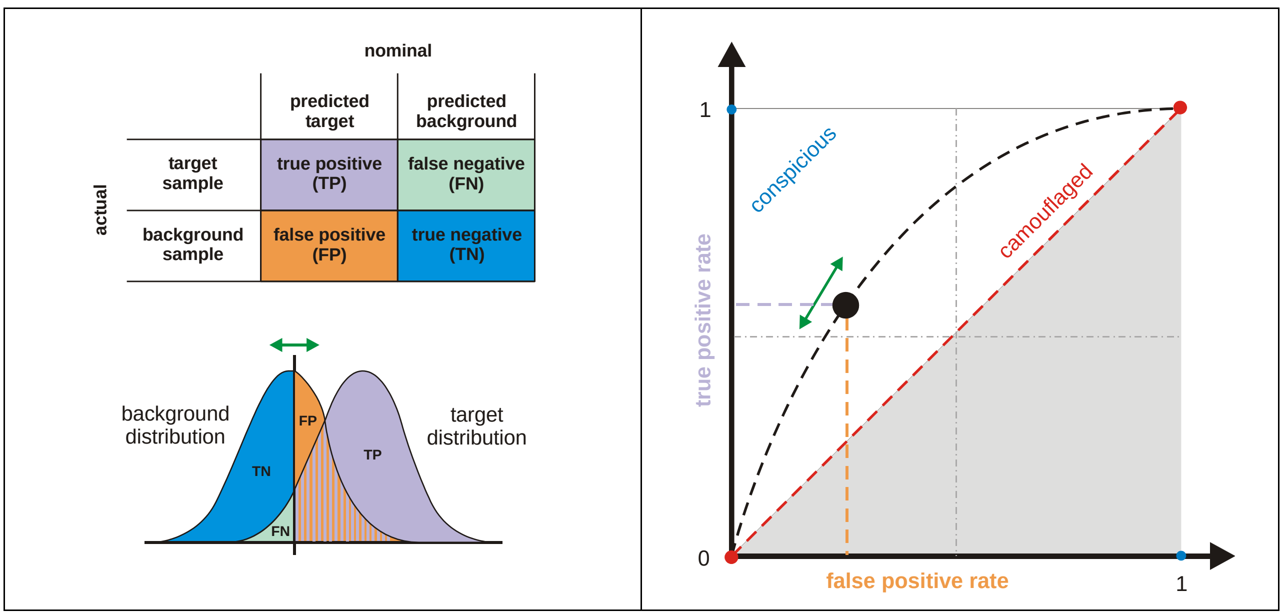

Figure 9: Confusion matrix with density functions (left) and ROC diagram (right).

Considering a given background and a target distribution w.r.t. the used features, as shown left in Figure 9, we can allocate a point (FP,TP) in the right diagram for a given threshold. For a moving threshold from the minimum to maximum of the domain, we get different rates of TP and FP, as outlined by green arrows in both figures. These are the coordinates of a curve on the so called receiver operating characteristic (ROC) curve [5] as shown on the right side of Figure 9.The ROC curve is a visualization of the discriminance between target and background for different parameter settings of a classification process. Figure 9 shows at the left side the basic features of a ROC diagram: The abscissa denotes the rate of false positives (FP), while the ordinate denotes the rate of true positives (TP). The upper left point (0,1) in the ROC diagram represents a perfect discrimination of target and background, the target appears to be *extremely conspicuous*. The diagonal line y=x coincides with a process of randomly determination. With respect to the considered characteristic, the target is *perfectly camouflaged*. The realistic cases lie in-between these outstanding cases and denote intermediate camouflage effectiveness.

A frequent interpretation for quality measure of feature evaluation is the area under the ROC curve [6]. The area under curve (AUC) is ratio of the normalized area, because its value is always between 0.0 and 1.0. If the area is close to 1 or close to 0, the target is extremely conspicuous. Conversely, in the case of an AUC close to ½ it is hint for a good camouflage. The AUC can be approximated by using of trapezoidal rule for integration and this is exactly the same quantity measured using the Mann Whitney U-test of ranks [7]. The Mann Whitney U-test is a non-parametric statistical test for assessing whether two independent samples of observations come from the same distribution [8]. Considering the samples B and T for a background and a target region, the two-sided U-test discriminates between

$$H_0 : P(T > B) = \frac{1}{2} \quad \text{versus} \quad H_A : P(T > B) \neq \frac{1}{2},$$

i.e. it is verified, if the probability of choosing T instead of B. So, the test is very sensitive for median difference between the distributions, but not in the case of variance differences. Also, the statistical test gives a significant decision, when the null hypothesis is rejected. This defines a point of rejection in the sequence to a given significance level α and so we get a further quality measure beside the effect size AUC, when the samples of targets and backgrounds are distinguished.

In the following AUC diagrams, we consider two general methods for the assessment of feature evaluation and target distinctiveness. The first method is called the absolute discrepancy and measures the discrimination between the targets and the background. These values are given as distances of the target AUC-values to ½. The other method is defined as relative discrepancy and determines the target distinctiveness, e.g. it measures the AUC-differences between the targets.

TESTS AND RESULTS

Some exemplary results of the automatic camouflage evaluation for two targets tested on an image sequence, as previously displayed in Figure 8, are discussed. We consider the bare (T1) and camouflaged object (T2) w.r.t. the background area (BG). In Figure 10 three ROC curves related to selected frames of the image sequence are displayed regarding to the hot spot operator. Approaching the ROC curves from left to the right, the discrimination of both targets is increasing due to the higher distance of the curves to the diagonal red line. In the right two images the ROC curve of the bare target T1 is running above the curve of T2. So, the target T1 is easier to distinguish from the background (BG) than the camouflaged object T2. Because the samples are including only a few hot spots in their sets, the majority of points of the ROC curves are close to the point (0,0).

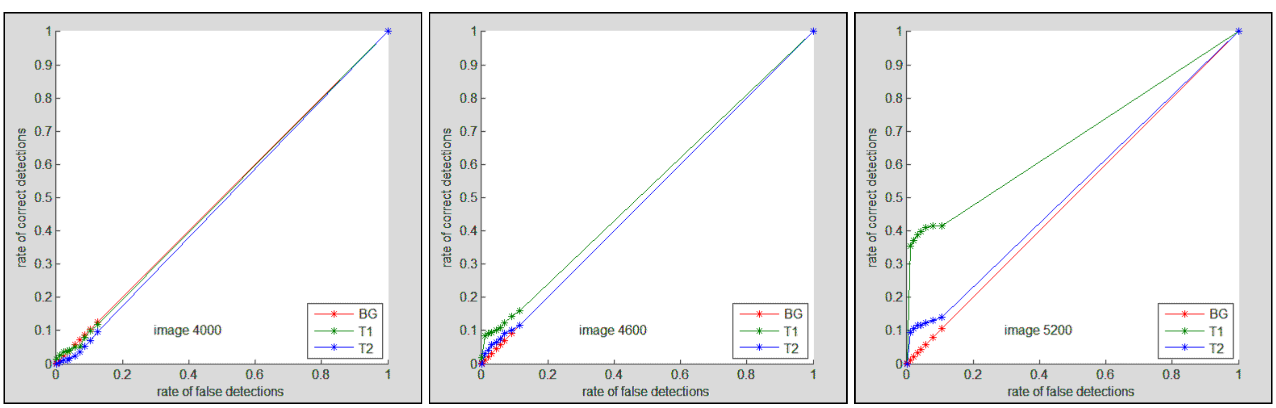

Figure 10: ROC curves to the bare and camouflaged targets T1 and T2 defined in Figure 8 for hot spot operator.

These three ROC diagrams are parts of a common AUC diagram over a sequence of images as displayed in Figure 11. The vertical lines show the positions of the selected ROC curves. Between the frames 3000 and 4500 the AUC values of both targets are close to ½, so the samples are undistinguishable. In the second part from [4500,5500] the runs of T1 and T2 are clearly different. Because the increasing distance of the AUC values to the constant curve in ½, both targets are easier to distinguish from the background. Furthermore, the AUC curve of target T1 is conspicuous above the run of T2. So, the bare target is more conspicuous than the camouflaged object.

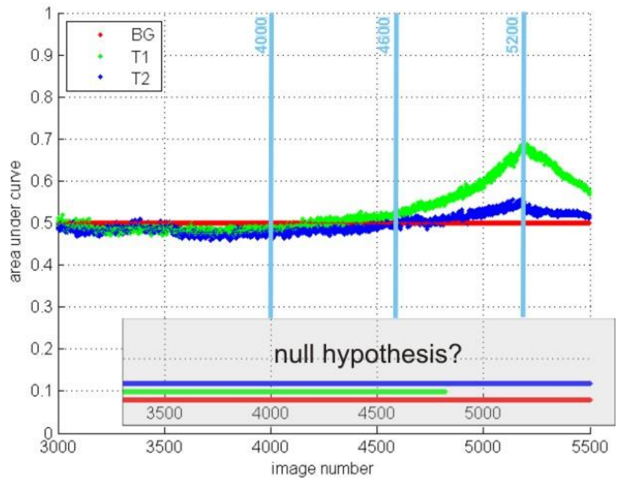

Figure 11: The AUC-values and the rejection of null hypothesis of Mann Whitney U-test for hot spot operator.

The overlaying rectangle in Figure 11 includes the significant rejections of the null hypothesis for α=0.05 in the Mann Whitney U-tests. The rejections are displayed with respect to the sample couples (BG, BG), (T1, BG) and (T2, BG). After the position 4800 it is visible, that the target T1 is rejected in contrast to T2. The rejection gives another qualitative discrepance information between the bare target and the background. Alltogether, for the hot spot operator the relative and the absolute discrepancy separate the targets T1 and T2. Here, the bare target is easier distinguished as the camouflaged object.

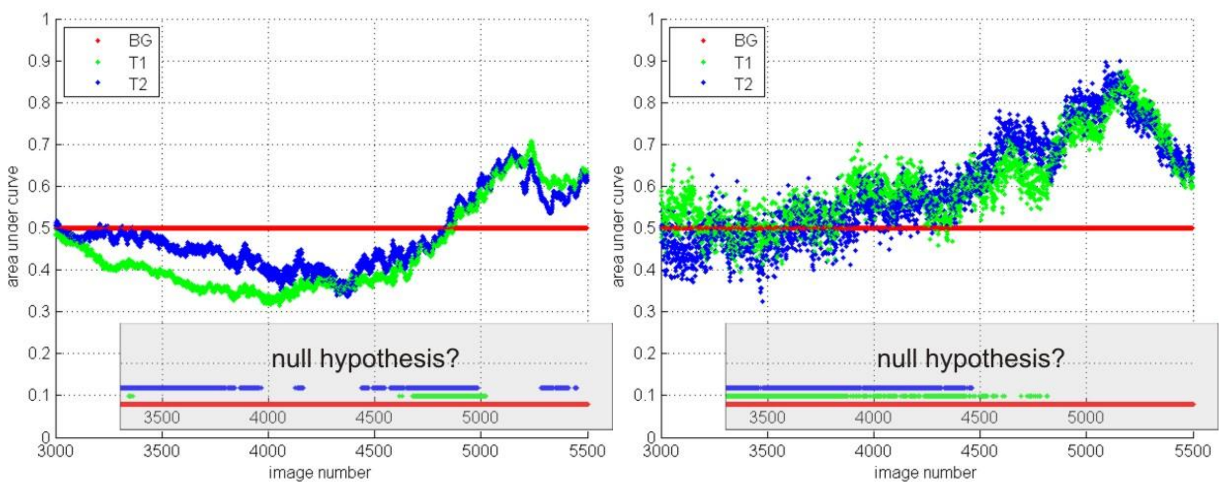

Figure 12: The AUCs and the rejection of null hypothesis of Mann Whitney U-test for intensity and Foerstner operator.

The AUC curves applying the intensity image and Foerstner operator over the same image sequence from 3000 to 5500 are presented in Figure 12. Two different behaviors of the curves in both diagrams are distinguishable. A frame around 4800 separates these two parts. In the first part of the intensity image the AUC values of both targets are obviously smaller than the red line, so the targets are distinguished to the background. Contrary, in the second phase a large part of the values are significantly greater than ½, that's why T1 and T2 are also discriminate to BG. The rejections of the null hypothesis for α=0.05 show the same behavior of absolute discrepancy. Because the curves of T1 and T2 are hardly distinguished in the complete sequence, the relative discrepancy between the two targets is very small. The intensity image shows also, that the AUC values of both curves are very strongly varying. Misplaced ROIs due to registration errors or statistical bias for small samples are possible reasons of these jumps.

For the diagram of the Foerstner operator in Figure 12 a non functional relation between the AUC values is assumable. The jumps of these values are extensively as in the case of the intensity image. The varied behavior is comparable to the above reasons. In the first interval [3000,4800] both targets are undistinguished to the background. Contrary in the second part the AUC values are significantly greater than ½, so T1 and T2 are easier distinguished to the BG. This observation is also visible by the rejections of the U-test. Because of the small differences between the AUC values of T1 and T2, the relative discrepancy of the targets is not discriminable. A further important message: the AUCs of Foerstner in both targets around frame 5200 are obviously greater than the values for the hot spot operator.

DISCUSSION AND OUTLOOK

The experimental results have shown that different features are produced by diversified evaluation influences. The three considered feature operators have demonstrated that an optimal feature extractor over the complete sequence does not exist. Further, we have considered the absolute and relative discrepancy for the assessment of feature evaluation. The decisions of both methods are not comparable. The tests demonstrate systematic errors in the automatic registration process, which must be analyzed in the future. Another exciting future task is the development of the three indicators - registration error, relative and absolute discrepancy - over the sequence.

A general method for feature evaluation for target/background discrimination based on a high accuracy automatic registration process has been demonstrated. Additional investigations on the determination and improvement of the derived reference values' accuracy are required. More complex feature extraction methods and their combination are topic of ongoing investigations. The goal is not to develop a new automatic target recognition (ATR) system but to be able to introduce the typical primitives of state of the art ATR's for best case evaluation. Additional efforts are undertaken to derive a summarized figure of merit of all considered features in the evaluation process. It is not expected that this figure of merit is directly related to human performance, but gives an objective statement of the background/target discrimination in different applications.

REFERENCES

[1] Hartley R., Zisserman A., “Multiple view geometry in computer vision”, Cambridge University Press (2000).

[2] Förstner W., Gülch E., “A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centers of Circular Features”, Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data, 281-305 (1987).

[3] Fischler A.M., Bolles R.C., “Random Sample Consensus: A Paradigm for Model fitting with application to image analysis and automatic cartography”, Journal of the ACM, Vol.24, No.6, 381-395 (1981).

[4] Kohnle, A., Neuwirth R., Schuberth, W., Stein, K., Höhn, D. H., Gabler, R., Hofmann, L., and Euing, W., “Evaluation of essential design criteria for IRST systems”, *Infrared Technology XIX*, SPIE Vol. 2020, 76-92 (1993).

[5] Hanley J., McNeil B. J. “The meaning and use of the area under a receiver operating characteristic (ROC) curve”, Radiology, vol. 143, no. 1, 29-36 (1982).

[6] Bradley A. P., “The use of the area under the ROC curve in the evaluation of machine learning algorithms”, Pattern Recognition vol. 30, 1145-1159 (1997).

[7] Bamber D., “The area above the ordinal dominance graph and the area below the receiver operating graph”, Journal of mathematical psychology 12, 387 - 415 (1975).

[8] Mann H., Whitney D., “On a test of whether one of two random variables is stochastically larger than the other”, Annals of mathematical Statistics 18, 50-60 (1947).