

# Asymptotic Performance Analysis for Object Recognition in Clutter

Dmitri Bitouk <sup>a</sup>, Michael I. Miller <sup>a</sup> and Laurent Younes <sup>b</sup>

<sup>a</sup>Center for Imaging Science, Whiting School of Engineering, The Johns Hopkins University,  
3400 North Charles Street, Baltimore, MD 21218 USA

<sup>b</sup>Ecole Normale Supérieure de Cachan, 61, avenue du Président Wilson, F-94 235 Cachan  
CEDEX, France

## ABSTRACT

This paper analyzes the performance of ATR algorithms in clutter. The variability of target type and pose is accommodated by introducing a deformable template for every target type, with low-dimensional groups of geometric transformations representing position and pose. Signature variation of targets is taken into account by expanding deformable templates into robust deformable templates generated from the template and a linear combination of PCA elements, spanning signature intensities. Detection and classification performance is characterized using ROC analysis. Asymptotic expressions for probabilities of recognition errors are derived, yielding asymptotic error rates. The results indicate that the asymptotic error probabilities depend upon a parameter which characterizes the separation between the true target and the most similar but incorrect one. It is shown that the asymptotic expressions derived almost accurately predict performance of detection and identification of targets occluded by natural clutter.

**Keywords:** Pattern Theory, Nuisance Integration, Automatic Target Recognition

## 1. INTRODUCTION

In automated target recognition (ATR) systems, objects of interest are observed at arbitrary locations and orientations via remote sensors. Depending on the relative distance, orientation and position, henceforth called pose, between the object and sensor, the observed imagery may vary dramatically. Over the last decade, a number of model-based template matching approaches, which are invariant to pose, has emerged. The central challenge to such approaches is move from isolated targets to the full complexity of scenes in real military environments. In real scenes, ATR systems have to accommodate the infinity of representational shapes that the cluttered world provides. The difficulty is that the geometries of natural clutterers are highly complex. Variation is the rule, not the exception. Yet, in model-based approaches, target detection and identification rates are largely determined by the accuracy of the underlying models. One of the possible approaches to address this issue is to construct metric distances between observed images and robust deformable templates, which accommodate variation of target pose as well as signature variation, associated with natural clutter.<sup>1</sup>

The focus of this paper is on asymptotic analysis of ATR performance in clutter. Many researchers have pointed out that any measure of performance in clutter must be related to the targets of interest and has to quantify discriminability of objects detected and identified.<sup>2,3</sup> Our paper formalizes this intuitive statement. We show that detection, identification and false alarm rates are determined by the distance between the target models detected and recognized and all of the alternative hypothesis.

The paper is organized as follows. In section 2 we describe the basic setup for object detection and identification via discreet hypothesis testing in presence of nuisance parameters. Section 3 presents asymptotic ROC analysis of target recognition in confounding background clutter. In section 4, we introduce robust deformable templates to model target signature variation due to occlusion and generalized the results of section 3.

---

Further author information: send correspondence to D. Bitouk, e-mail: dima@cis.jhu.edu.

## 2. HYPOTHESIS TESTING

In a statistical framework, target detection and identification are performed via hypothesis testing,<sup>4</sup> while the unknowns, such as target pose and signature, are treated as nuisance parameters. Given an observed image  $I_D$ , the likelihood ratio is compared to the value of threshold

$$\frac{p(I_D|H_1)}{p(I_D|H_0)} \underset{H_0}{\overset{H_1}{>}} \nu. \quad (1)$$

In the presence of nuisance parameters  $\theta$ , such as pose and target signature, the likelihoods  $p(I_D|H_i)$ ,  $i = 0, 1$  are defined as the Bayes integrals

$$p(I_D|H_i) = \int p(I_D|H_i, \theta)p(\theta|H_i)\gamma(d\theta). \quad (2)$$

In most practical cases, the complicated integrands make the analytical analysis difficult. One possibility is to evaluate maximum-likelihood estimators of the nuisance parameters and perform generalized likelihood ratio test

$$\frac{p(I_D|H_1, \theta_1^*)}{p(I_D|H_0, \theta_0^*)} \underset{H_0}{\overset{H_1}{>}} \nu, \quad (3)$$

where

$$\theta_i^* = \arg \max_{\theta \in \Theta} p(I_D|H_i, \theta) \quad (4)$$

are the maximum-likelihood estimates of the nuisance parameter  $\theta$  under corresponding hypothesis.

Generalized likelihood ratio test (3) is equivalent to the following log-likelihood ratio test

$$\Lambda(I_D) = \inf_{\theta \in \Theta} \log p(I_D|H_1, \theta) - \inf_{\theta \in \Theta} \log p(I_D|H_0, \theta) \underset{H_0}{\overset{H_1}{>}} \log \nu. \quad (5)$$

## 3. RIGID TEMPLATE MODEL

In automated target recognition, nuisance parameters are usually employed to model variability in the observed images. The variability of target type and pose can be accommodated using rigid template approach, defining for every target type a template with a group of geometric transformations  $G$ , specifying pose and location of the target. The group of geometric transformations  $G$  includes both rotations and translation, and can be represented by the special Euclidean group  $G = SE(n) = SO(n) \otimes \mathbb{R}(n)$ .

The rigid template corresponds to the orbit  $G.I_{temp}$  under the group action  $G$  of one selected and fixed instance  $I_{temp}$

$$G.I_{temp} = \{I : I = I_{temp} \circ g, g \in G\}. \quad (6)$$

The framework of deformable templates captures both pro-typical features of the object of interest as well as typical variations of the prototype in the observed images. Rigid templates incorporate geometric transformations, representing the pose of objects. Variability in the observed images due to natural background clutter can be accommodated by introducing an additive random field  $W$ . Then, the imaging model becomes

$$I_D = \hat{T}(I_{temp} \circ g) + W. \quad (7)$$

In this paper, we will assume that the background clutter is a stationary Gaussian random field, specified by the covariance  $K(x, y)$ . Let  $H_K$  be the reproducing kernel space of  $K(x, y)$  with the inner product  $\langle f, h \rangle_{H_K}$  and norm  $\|f\|_{H_K}$  defined as

$$\langle f, h \rangle_{H_K} = \sum_n \frac{\langle f, \psi_n \rangle_{L^2} \langle \psi_n, h \rangle_{L^2}}{\lambda_n}, \quad (8)$$

$$\|f\|_{H_K}^2 = \langle f, f \rangle_{H_K}, \forall f, g \in H_K, \quad (9)$$

where  $\{\lambda_n\}, \lambda_1 \geq \lambda_2 \geq \dots > 0$  and  $\{\psi_n\}$  are eigenvalues and eigenfunctions of  $K(x, y)$ , and  $\langle f, h \rangle_{L^2} = \int_{\mathbb{R}^2} f(x) h(x) dx$ .

Assuming that targets are seen by the remote sensor which performs the transformation  $\hat{T}$  from 3D to the two-dimensional measurement space  $\mathcal{I}_D$

$$\hat{T} : I \rightarrow \hat{T}I, \hat{T}I \in \mathcal{I}_D \quad (10)$$

the likelihood  $p(I_D | H_i, g)$ ,  $i = 0, 1$  becomes

$$p(I_D | H_i, g) \propto \exp \left[ - \frac{\|I_D - \hat{T}(I_{temp}^{(i)} \circ g)\|_{H_K}^2}{2} \right] \quad (11)$$

### 3.1. Asymptotic ROC analysis

Our goal is to obtain asymptotic approximations for the probability of false classification  $P_{FC} = \mathcal{P}[H_1 | H_0]$  and the probability of correct classification  $P_{CC} = \mathcal{P}[H_1 | H_1]$  as the largest eigenvalue of  $K(x, y)$ ,  $\lambda_1 \rightarrow 0$ . Asymptotic analysis in this section follows the approach of Grenander et al.<sup>5</sup> In sections 3.2 and 4 we generalize this analysis to model mismatch and target signature variations.

In the case of the rigid template model, the generalized likelihood ratio test takes the form

$$\inf_g \left\| I_D - \hat{T}(I_0 \circ g) \right\|_{H_K}^2 - \inf_g \left\| I_D - \hat{T}(I_1 \circ g) \right\|_{H_K}^2 \underset{H_0}{\overset{H_1}{>}} \mu = 2 \log \nu \quad (12)$$

Under the hypothesis  $H_0$ , the observed image  $I_D = \hat{T}(I_0 \circ g_0^t) + W$ , where  $g_0^t$  is the true pose of the target I. Then, as  $\lambda_1 \rightarrow 0$ ,

$$g_0^* = \arg \min_g \left\| \hat{T}(I_0 \circ g_0^t) + W - \hat{T}(I_0 \circ g) \right\|_{H_K}^2 \rightarrow g_0^t \quad (13)$$

and

$$g_1^* = \arg \min_g \left\| \hat{T}(I_0 \circ g_0^t) + W - \hat{T}(I_1 \circ g) \right\|_{H_K}^2 \rightarrow g_{01} = \arg \min_g \left\| \hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g) \right\|_{H_K}^2. \quad (14)$$

The log-likelihood ratio takes the form

$$\Lambda(I_D) \sim - \inf_g \left\| \hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g) \right\|_{H_K}^2 + 2 \langle W, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^t) \rangle_{H_K} \quad (15)$$

Let us consider the probability of false classification  $P_{FC} = \mathcal{P}[H_1 | H_0] = \mathcal{P}[\Lambda(I_D) > \mu | H_0]$  The generalized likelihood ratio test yields

$$P_{FC}(\mu) = \mathcal{P} \left[ 2 \langle W, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^t) \rangle_{H_K} > \mu + \|I_0 \circ g_0^t - I_1 \circ g_{01}\|_{H_K}^2 \right] = \quad (16)$$

$$= \mathcal{P} \left[ t(I_D) > \frac{\mu}{2 \|\hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g_{01})\|_{H_K}} + \frac{\|\hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g_{01})\|_{H_K}}{2} \right], \quad (17)$$

where

$$t(I_D) = \frac{\langle W, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^t) \rangle_{H_K}}{\|\hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g_{01})\|_{H_K}} \quad (18)$$

is the test statistic.

**Proposition 1.** The test statistic  $t(I_D) \sim \mathcal{N}(0, 1)$ .

Let us define a parameter  $\beta_0$  to be the distance from  $I_0 \circ g_0^t$  to the orbit  $G.I_1$

$$\beta_0^2 = \inf_g \|\hat{T}(I_0 \circ g_0^t) - \hat{T}(I_1 \circ g)\|_{H_K}^2. \quad (19)$$

Then, the probability of false classification is given by

$$P_{FC} = \mathcal{P} \left[ t(I_D) > \frac{\mu}{2\beta_0} + \frac{\beta_0}{2} \right] = \quad (20)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{\mu}{2\beta_0} + \frac{\beta_0}{2}}^{\infty} \exp[-x^2/2] = \text{erfc} \left[ \frac{\mu}{2\beta_0} + \frac{\beta_0}{2} \right] \quad (21)$$

Similarly, the probability of correct classification is given by

$$P_{CC} = \text{erfc} \left[ \frac{\mu}{2\beta_1} - \frac{\beta_1}{2} \right], \quad (22)$$

where

$$\beta_1^2 = \inf_g \|\hat{T}(I_1 \circ g_1^t) - \hat{T}(I_0 \circ s)\|_{H_K}. \quad (23)$$

Thus, we arrive to the following result.

**Theorem 1.** As  $\lambda_1 \rightarrow 0$ , the probability of false classification  $P_{FC} = \mathcal{P}[H_1|H_0]$  and the probability of correct classification  $P_{CC} = \mathcal{P}[H_1|H_1]$  are asymptotically given by

$$P_{FC} = \text{erfc} \left[ \frac{\mu}{2\beta_0} + \frac{\beta_0}{2} \right], \quad P_{CC} = \text{erfc} \left[ \frac{\mu}{2\beta_1} - \frac{\beta_1}{2} \right]. \quad (24)$$

Note that the quantities  $\beta_0$  and  $\beta_1$  are quantifying the fundamental separation between the true object and most similar alternative.

### 3.2. Clutter model mismatch

Let us consider an interesting special case when the clutter process  $W$  is Gaussian with the Toeplitz covariance  $K^{(W)}(x, 0) \neq K(x, 0)$ .

Similarly to the previous section,

$$P_{FC}(\mu) = \mathcal{P} \left[ 2 \langle W, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^{(t)}) \rangle_{H_K} > \mu + \|\hat{T}(I_0 \circ g_0^{(t)}) - \hat{T}(I_1 \circ g_{01})\|_{H_K}^2 \right] \quad (25)$$

Then, Karhunen-Loeve expansion yields (in q.m. sense)

$$W = \sum_n X_n \sqrt{\lambda_n^{(W)}} \psi_n, \quad (26)$$

where  $\{\lambda_n^{(W)}\}$  and  $\{\psi_n^{(W)}\}$  are the eigenvalues and the eigenfunctions of  $K^{(W)}(x, 0)$ ;  $X_n \sim \mathcal{N}(0, 1)$  are i.i.d.

Let us define the test statistic  $t(I_D)$  to be

$$t(I_D) = \frac{\langle W, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^{(t)}) \rangle_{H_K}}{\sqrt{\sum_n \frac{\lambda_n^{(W)}}{\lambda_n^2} \left| \langle \psi_n, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^{(t)}) \rangle \right|_{L_2}^2}}. \quad (27)$$

It is easy to show that  $t(I_D) \sim \mathcal{N}(0, 1)$ . Repeating the arguments from the previous section gives,

$$P_{FC} = \text{erfc} \left[ \frac{\mu}{2\gamma_0} + \frac{\beta_0^2}{2\gamma_0} \right], \quad P_{CC} = \text{erfc} \left[ \frac{\mu}{2\gamma_1} - \frac{\beta_1^2}{2\gamma_1} \right], \quad (28)$$

where

$$\gamma_0 = \sqrt{\sum_n \frac{\lambda_n^{(W)}}{\lambda_n^2} \left| \langle \psi_n, \hat{T}(I_1 \circ g_{01}) - \hat{T}(I_0 \circ g_0^{(t)}) \rangle \right|_{L_2}^2}, \quad (29)$$

$$\gamma_1 = \sqrt{\sum_n \frac{\lambda_n^{(W)}}{\lambda_n^2} \left| \langle \psi_n, \hat{T}(I_1 \circ g_1^{(t)}) - \hat{T}(I_0 \circ g_{10}) \rangle \right|_{L_2}^2}, \quad (30)$$

and

$$g_{01} = \arg \min_g \left\| \hat{T}(I_0 \circ g_0^{(t)}) - \hat{T}(I_1 \circ g) \right\|_{H_K}^2, \quad (31)$$

$$g_{10} = \arg \min_g \left\| \hat{T}(I_1 \circ g_1^{(t)}) - \hat{T}(I_0 \circ g) \right\|_{H_K}^2. \quad (32)$$

To establish the correspondence between asymptotic expressions and empirically estimated performance of target identification, we computed ROC curves. The numerical probabilities of false and correct classification were computed from multiple realizations of the additive clutter process  $W$ , specified by the known Toeplitz covariance  $K^{(W)}(x, 0)$  and random target pose. The solid line on Fig. 1 corresponds to the analytical asymptotic probabilities of false and correct classification computed according to Theorem 1. Circles present numerically evaluated performance. The dashed line show analytical asymptotic performance with  $\|\cdot\|_{H_K} = \|\cdot\|_{L_2}$  corresponding to the clutter model mismatch. The squares show the empirical performance.

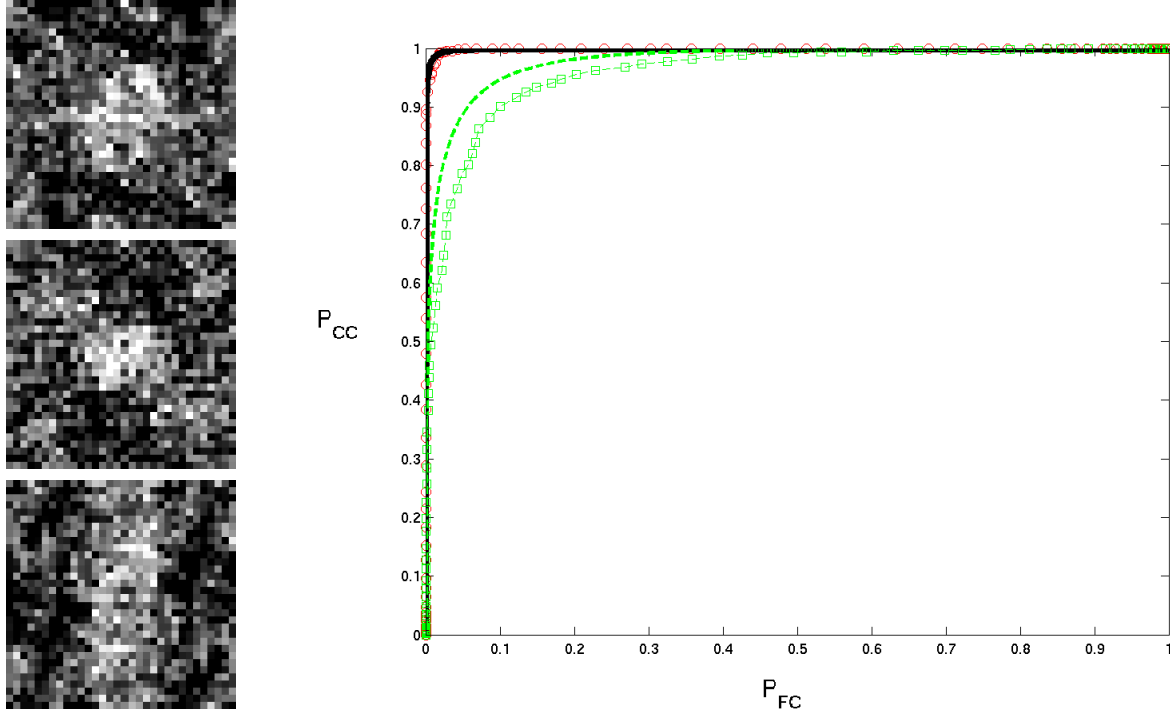
#### 4. ROBUST DEFORMABLE TEMPLATE MODEL

Variability in the observed imagery of target in natural is not solely associated with variation in target pose and background clutter, but as well as with occlusion and obscuration, which represent so-called photometric variations. Without any doubts, the asymptotic results obtained in the previous section using rigid templates will not accurately predict the performance in obscuring clutter. To address this issue, we extend the rigid templates to the robust deformable template, which explicitly include photometric variations.

The robust deformable template  $A.I_{temp}$  is the orbit under the action of the group of geometric transformations  $G$  and an additive action of the signature basis functions  $\{\phi_n\}$

$$A.I_{temp} = \{I \in \mathcal{I} : I = \left( I_{temp} + \sum_n I_n \phi_n \right) \circ g, g \in G\}. \quad (33)$$

The generation the signature basis  $\{\phi_n\}$  can be approached using principal component analysis (PCA).<sup>6</sup> Figure 2 shows a tank template and a few eigenfunction generated via PCA from a database of simulated target images in clutter.



**Figure 1.** Left panel show example images. Right panel displays ROC curves : solid line - analytical, circles - empirical, squares - empirical covariance norm, dashed line - analytical Euclidean norm, squares - empirical Euclidean norm.

For robust deformable templates, the likelihood  $p(I_D|H_i, g, \{I_n\})$  becomes

$$p(I_D|H_i, g, \{I_n\}) \propto \exp \left[ -\frac{\left\| I_D - \hat{T} \left( I_{temp}^{(i)} + \sum_n I_n \phi_n^{(i)} \right) \circ g \right\|_{H_K}^2}{2} \right] \quad (34)$$

The log-likelihood takes the form

$$\Lambda(I_D) = \inf_{g, \{I_k\}} \left\| I_D - \hat{T} \left( I_{temp}^{(0)} + \sum_k I_k \phi_k^{(0)} \right) \circ g \right\|_{H_K}^2 - \inf_{g, \{I_k\}} \left\| I_D - \hat{T} \left( I_{temp}^{(1)} + \sum_k I_k \phi_k^{(1)} \right) \circ g \right\|_{H_K}^2 \quad (35)$$

Repeating the analysis of the previous section and omitting cumbersome details, we arrive to the following result.

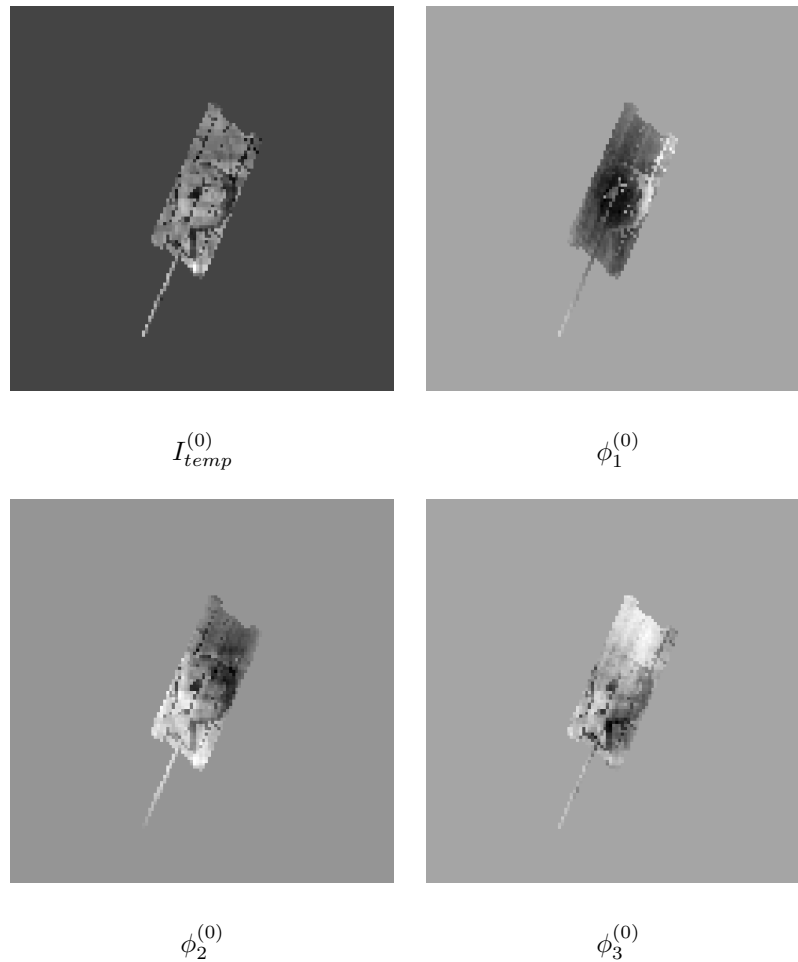
**Theorem 2.** As  $\lambda_1 \rightarrow 0$  and  $\lambda_1^{PCA} \rightarrow 0$ , the probability of false classification and the probability of correct classification are asymptotically given by

$$P_{FC}(\mu) = \text{erfc} \left( \frac{\mu}{2\beta_0} + \frac{\beta_0}{2} \right), \quad P_{CC}(\mu) = \text{erfc} \left( \frac{\mu}{2\beta_1} - \frac{\beta_1}{2} \right) \quad (36)$$

where

$$\beta_0^2 = \inf_{g, \{I_k\}} \left\| I_{temp}^{(0)} \circ g_t^{(0)} - \left( I_{temp}^{(1)} + \sum_k I_k \phi_k^{(1)} \right) \circ g \right\|_{H_K}^2, \quad (37)$$

$$\beta_1^2 = \inf_{g, \{I_k\}} \left\| I_{temp}^{(1)} \circ g_t^{(1)} - \left( I_{temp}^{(0)} + \sum_k I_k \phi_k^{(0)} \right) \circ g \right\|_{H_K}^2. \quad (38)$$



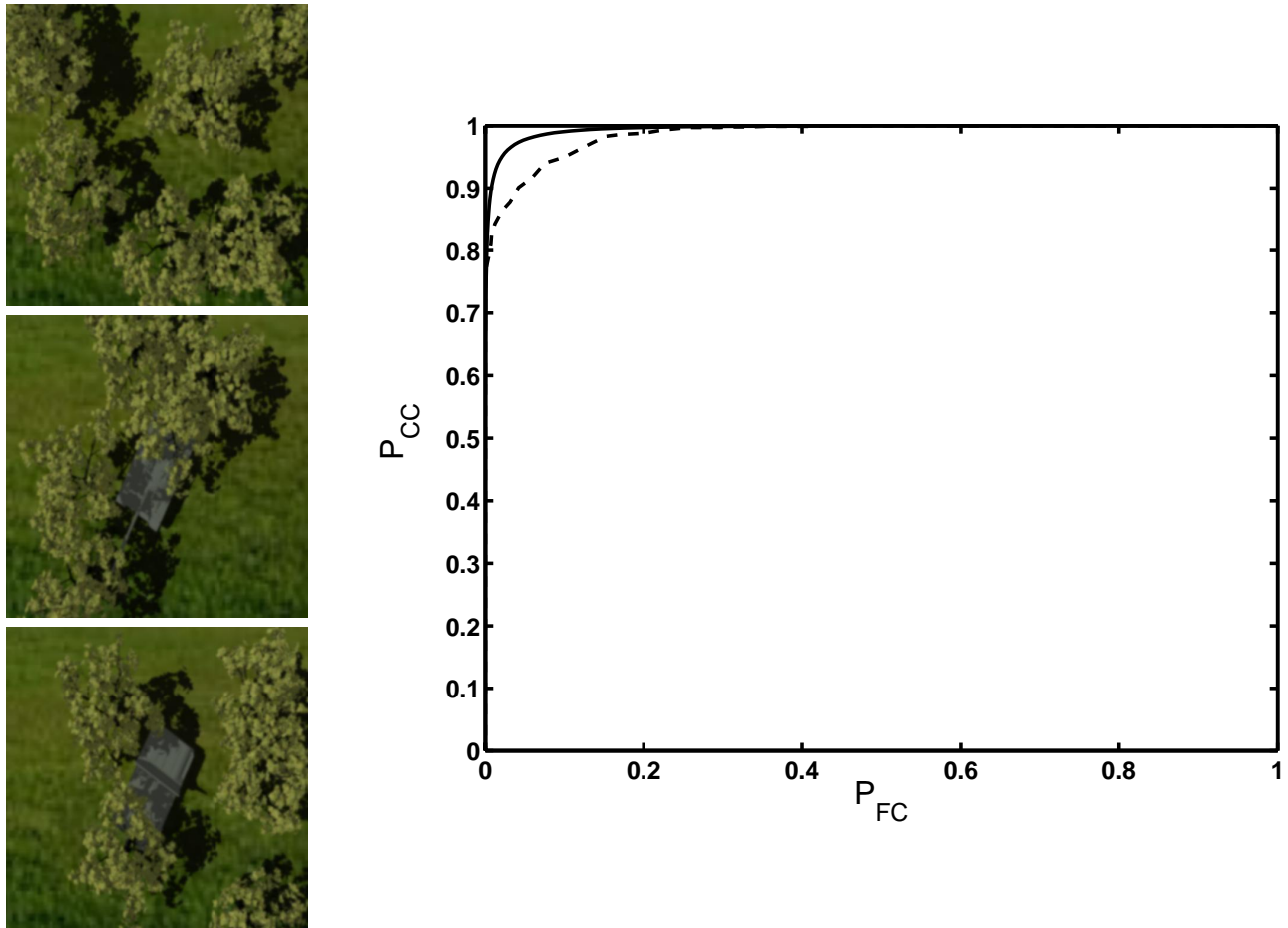
**Figure 2.** Template and principle components of a tank

Notice that similarly to the result of theorem, the probabilities of false and correct classification are determined by the distance between the true target and the most similar alternative. Clearly, including photometric variation explicitly into the model increases the invariance and decreases the discriminability. In the same time, performance is fully characterized by the fundamental separation between targets  $I_0$  and  $I_1$ .

For the numerical experiments presented on figure 3, we synthesized a dataset of ray-traced target chips in natural clutter. Ray-tracing appears to be especially attractive rendering technique, since it closely resembles the formation process of natural images. The dataset consists of more than 10,000 images. The ray-traced images in the dataset contain targets in a randomly synthesized terrain taking into account occlusion, shadows and lighting variation as well as other effects usually encountered in natural scenes. Terrain was generated by random placement of trees on grass backgrounds according to the specified distribution densities. The trees were synthesized using a random growing process controlled by physical parameters. The bottom panel of figure 3 displays example images from the synthetic dataset. The solid line present analytical ROC curve computed according to the results of theorem 2. The dashed line shows empirical performance of target identification via robust deformable template model.

## ACKNOWLEDGMENTS

This work was supported by ONR N00014-00-1-0327.



**Figure 3.** Left panel shows synthesized target chips. On the right panel solid line shows analytically predicted asymptotic performance of target identification; dashed line displays empirical results.

## REFERENCES

1. D. Bitouk, M. Miller, and L. Younes, "Clutter metrics for ATR." *IEEE Transaction on Pattern Analysis and Machine Intelligence*. submitted.
2. J. Ratches, C. Walters, R. Buser, and B. Guenther, "Aided and automateic target recognition based upon sensory inputs from image forming systems," *IEEE Transaction on Pattern Analysis and Machine Intelligence* **19(9)**, 1997.
3. S. R. F. Sims, "Signal to clutter measurement and ATR performance," *The SPIE Conference on Automated Target Recognition VII*, pp. 398–403, 1998.
4. H. L. van Trees, *Detection, Estimation, and Modulation Theory : Part I*, John Wiley and Sons, New York, 1968.
5. U. Grenander, A. Srivastava, and M. Miller, "Asymptotic performance analysis of baesyian target recognition," *IEEE Transaction on Information Theory* **46(9)**, 2000.
6. M. Cooper and M. Miller, "Information measures for object recognition accommodating signature variability," *IEEE Transactions on Information Theory* **46(5)**, pp. 1896–1906, 2000.