Empirically Generated Metric Spaces for ATR in Clutter

D. Bitouk
Center for Imaging Science
Johns Hopkins University
Baltimore, MD 21218

M. I. Miller
Center for Imaging Science
Johns Hopkins University
Baltimore, MD 21218

L. Younes
CMLA
Ecole Normale Superieure de Cachan
Cachan CEDEX, France F-94 235

Abstract

One of the central problems in Automated Target Recognition is to accommodate the infinite variety of clutter in real military environments. The principle focus of our paper is to construct metric space where the metric measures the distance between objects of interest invariant to to the infinite variety of clutter. Such metrics are formulated using second-order random field models.

1 Introduction

One of the central problems in Automated Target Recognition is to accommodate the infinite variety of clutter in real military environments. In model-based approaches, identification/classification rates are largely determined by the accuracy of models used to represent real-world scenes. In heavy cluttered environments, it is impractical to explicitly represent each of the objects in the scene by a deterministic 3D model whereas applying low-dimensional statistical description of clutter may improve ATR performance in clutter. The modeling of clutter can approached from the point of view of empirical statistics. Similar approaches have been successfully carried out under the context of modeling object signatures using principal component analysis [1], [2], [3], [4].

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 rigid motions, representing geometric variations.

In a continuous setting, the set of images $\mathcal{I}$ is defined as a space of functions on the background space X, $I : X \to V$, with V the value space. In this paper, we consider $X \subset \mathbb{R}^k, k = 2,3$ and $V = \mathbb{R}$. The group of geometric transformations G acts on the background space X according to

$$g \in G : x \in X \to g \circ x \in X. \quad (1)$$

The group action of G on the image space $\mathcal{I}$ is defined by its action of the background space

$$g \in G : I \in \mathcal{I} \to I \circ g \in \mathcal{I}. \quad (2)$$

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

$$G.I_{temp} = \{I \in \mathcal{I} : I = I_{temp} \circ g,\ g \in G\}. \quad (3)$$

For the purposes of automated target recognition, images I_{temp} correspond to the targets.

Many automated target recognition problems involve comparison of the observed image I_D to a template I_{temp} in the presence of a group action. For instance, the group G may correspond to rigid motions of the template I_{temp}. Considering images as elements of some functional space, one can readily equip this space with a functional distance $d_0\left(I^{(0)}, I^{(1)}\right) = N_{\mathcal{I}}\left(I^{(1)} - I^{(0)}\right)^2$, where $N_{\mathcal{I}}$ is a functional norm (L^p norm, Sobolev norm, etc). In the situation, in which the group action does not alter the essence of objects, those functional distances are poor candidates for object comparison, since one can find a rigid motion g, for which $d_0(I, I \circ g)$ will be large. In this paper, we are interested in designing metric distances d, which are robust with respect to the action of the group G. More precisely, if the group G is the group of rigid motions, one of the desired properties of the distance would be $d(I, I \circ g) = 0$ for all $g \in G$ and all $I \in \mathcal{I}$. The next subsection follows the development by Miller and Younes [5], presenting the basic setup for constructing metrics on the space of images $\mathcal{I}$.

Construction of a metric distance on $\mathcal{I}$, which takes into account the action of G, usually requires using a variational approach on the product space $G \times \mathcal{I}$, in which the metric would be defined through a geodesic distance. The principle of this construction relies on concepts of differential geometry [6] and consists of defining the energy of paths on $G \times \mathcal{I}$ with subsequent computation of the distance by minimizing these energies.

A differentiable path on $G \times \mathcal{I}$ is a continuous function $(g_t, I_t) : [0;1] \to G \times \mathcal{I}$, for which the time derivative $\left(\frac{dg_t}{dt}, \frac{dI_t}{dt}\right)$ is defined for all $t \in [0;1]$. Given a collection of functional norms $\mathcal{N}_G$ and $\mathcal{N}_\mathcal{I}$, the associated energy of the path (g_t, I_t) is

$$E(g_t, I_t) = \int_0^1 \mathcal{N}_G \left(\frac{\partial g_t}{\partial t} \circ g_t^{-1}\right)^2 dt + \tag{4}$$

$$+ \int_0^1 \mathcal{N}_\mathcal{I} \left(\frac{\partial I_t}{\partial t} \circ g_t^{-1}\right)^2 dt \tag{5}$$

Then, the distance $d(I^{(0)}, I^{(1)})$ on $\mathcal{I}$ is defined through the infimum energy over all the paths g_t starting at the identity, $g_0 = id$, and all the paths I_t, which take $I^{(0)}$ to $I^{(1)} \circ g_1$

$$d(I^{(0)}, I^{(1)}) = \inf_{\substack{g_t : g_0 = id, \\ I_t : I_0 = I^{(0)}, I_1 = I^{(1)} \circ g_1}} \sqrt{E(g_t, I_t)}. \tag{6}$$

In equation 5, there are two penalty terms for the deformation, which takes $I^{(0)}$ to $I^{(1)}$. The first integral penalizes transformations in the background space associated with the group G, and the second one penalized photometric deformations.

In general, the choice of the norms $\mathcal{N}_G$ and $\mathcal{N}_\mathcal{I}$ depends on the application. In this work, the group G, associated with the pose of the objects of interest, is considered to be the group of rigid motions, $G = SO(k) \otimes \mathbb{R}^k$. The natural requirement would be to set the penalty, which is due to the rigid motions of the objects, to zero, $\mathcal{N}_G(\cdot) \equiv 0$. On the other hand, another desired property of the metric distance, in this case, is robustness with respect to photometric variations in the observed images associated with natural clutter. This property can be achieved by designing norms $\mathcal{N}_\mathcal{I}$ using statistical models of natural clutter. In section 2, such norm are formulated as Sobolev norms associated with the inverse spatial covariance of clutter.

2 Covariance Norms

In this section we build a Hilbert space of photometric variability with the norm, induced by the empirical covariance kernel.

Assuming the covariance kernel $K(x, y)$ to be Hilbert-Schmidt

$$\int\int K^2(x, y) dx dy < \infty, \tag{7}$$

the Hilbert-Schmidt theorem implies that there exists a complete orthonormal sequence of eigenfunctions $\{\psi_n(x)\}$ with corresponding eigenvalues $\{\lambda_n\}$ such that

$$\int K(x, y) \psi_n(y) dy = \lambda_n \psi_n(x). \tag{8}$$

Definition 2.1 *Given a positive-definite covariance kernel $K(x, y)$, define a Hilbert space $H \subseteq L_2$ with the norm $\|\cdot\|_H$, induced by the kernel $K(x, y)$, according to*

$$H = \{f \in L_2 : \|f\|_H^2 < \infty\}, \tag{9}$$

where the norm on H is defined by

$$\|f\|_H^2 = \sum_n \frac{|\langle f, \psi_n \rangle_{L_2}|^2}{\lambda_n} \tag{10}$$

Consider an element g in the group of rigid motions $G = SO(k) \otimes \mathbb{R}^k$. The kernel $K(g \circ x, g \circ y)$ is also Hilbert-Schmidt and, according to definition 2.1, induces a corresponding Hilbert space $g.H$. The following lemma establishes the connection between the norms $\|\cdot\|_H$ and $\|\cdot\|_{g.H}$.

Lemma 2.1 *For any $f \in g.H$,*

$$\|f\|_{g.H} = \left\|f \circ g^{-1}\right\|_H. \tag{11}$$

3 Metrics induced by empirical covariance

In this section, we design a metric distance on the set of images $\mathcal{I}$ under the action of the group of rigid motions defined as

$$g \circ x = Rx + b \tag{12}$$

where $R \in SO(k), b \in \mathbb{R}^k$.

Following the basic framework of section 1, we set the penalty term for geometric transformations G to zero, $\mathcal{N}_G \equiv 0$ and associate the covariance norms $\|\cdot\|_{g.H}$, defined in the previous section, with the functional norm $\mathcal{N}_\mathcal{I}$. Then, the energy of the path (g_t, I_t) is defined by

$$E(g_t, I_t) = \int_0^1 \int_0^1 \left\|\frac{\partial I_t}{\partial t} \circ g_t^{-1}\right\|^2_{g_t^{-1}.H} dt. \tag{13}$$

First, we define a measure of similarity between two images $d(I^{(0)}, I^{(0)})$ and, then, we provide the conditions, under which $d(I_0, I_1)$ is a metric distance.

Definition 3.1 *Given two images $I^{(0)}, I^{(0)} \in \mathcal{I}$, the function $d : \mathcal{I} \times \mathcal{I} \to [0; +\infty)$ given by*

$$d^2(I^{(0)}, I^{(1)}) = \inf_{\substack{g_t : g_0 = id, \\ I_t : I_0 = I^{(0)}, I_1 = I^{(1)} \circ g_1}} \tag{14}$$

$$\int_0^1 \left\|\frac{dI_t}{dt} \circ g_t^{-1}\right\|^2_{g_t^{-1}.H} dt \tag{15}$$

The function $d(I^{(0)}, I^{(1)})$ is defined as a variational problem. The following lemma relates the definition 3.1 to the variational problem over all the differentiable paths g_t and all the paths J_t connecting the image $I^{(0)}$ to $I^{(1)}$.

Lemma 3.1 *Defining $J_t = I_t \circ g_t^{-1}$,*

$$d(I^{(0)}, I^{(1)}) = \inf_{\substack{\Omega(t),\beta(t) : \Omega_0 = 0, \beta_0 = 0 \\ J_t : J_0 = I^{(0)}, J_1 = I^{(1)}}} \tag{16}$$

$$\sqrt{\int_0^1 \left|\left| \frac{dJ_t}{dt} + \left\langle \frac{\partial J_t}{\partial x}, \Omega(t)x + \beta(t) \right\rangle_{\mathbb{R}^k} \right|\right|^2_{g_t^{-1}.H} dt}, \tag{17}$$

where $\Omega(t)$ is a skew-symmetric matrix defined by

$$\frac{dR(t)}{dt} = \Omega(t)R(t), \tag{18}$$

and

$$\beta(t) = \frac{db(t)}{dt} - \Omega(t)b(t). \tag{19}$$

3.1 Statistical interpretation

The following theorem establishes the connection of the variational problem for $d\left(I^{(0)}, I^{(1)}\right)$ to the maximum likelihood estimation.

Theorem 3.1

$$d^2(I^{(0)}, I^{(1)}) = \inf_{g_1} \left|\left| I^{(0)} - I^{(1)} \circ g_1 \right|\right|^2_H \tag{20}$$

Theorem 3.1 implies that, informally, in the Bayesian statistical framework, the variational problem for $d(I^{(0)}, I^{(1)})$ corresponds to the imaging model

$$I^{(0)} = I^{(1)} \circ g_1 + W, \tag{21}$$

where $I^{(0)}$ is an observation of the template $I^{(1)}$ transformed by the rigid motion g_1 and corrupted by an additive stationary second-order random field W, which covariance is specified by the kernel $K(x, 0)$.

3.2 Invariance properties

For an arbitrary covariance $K(x, 0)$, a measure of similarity $d(I^{(0)}, I^{(1)})$, defined by equation 3.1, is not a metric distance on $\mathcal{I}$, since it is not symmetric and does not satisfy the triangular inequality. In this subsection the condition are provided, under which $d(I^{(0)}, I^{(1)})$ is a metric distance.

The following lemma establishes the connection between invariance properties of $d(I^{(0)}, I^{(1)})$ and $K(x, y)$.

Lemma 3.2 *Suppose for all $h \in G = SO(k) \bigotimes \mathbb{R}^k$*

$$K(h \circ x, h \circ y) = K(x, y). \tag{22}$$

Then, for any $h \in G = SO(k) \bigotimes \mathbb{R}^k$ the function $d(I^{(0)}, I^{(1)})$ is invariant under the action of G, i.e

$$d^2(I^{(0)} \circ h, I^{(1)} \circ h) = d^2(I^{(0)}, I^{(1)}) \tag{23}$$

Theorem 3.2 *Given a Toeplitz covariance $K(x, 0)$ which is invariant under the action of $SO(k)$,*

$$K(r \circ x, 0) = K(x, 0), \quad \forall r \in SO(k), \tag{24}$$

the function $d : \mathcal{I} \times \mathcal{I} \to \mathbb{R}^+$, defined by Eq. 3.1, is a metric distance on $\mathcal{I}$.

Properties of the metric distance $d(I^{(0)}, I^{(1)})$ allow an efficient alternating minimization algorithm for computing this metric distance (see [7] for details). Figure 1 displays the minimum energy path J_t connecting an image $I^{(0)}$ of a tank in trees to its template $I^{(1)}$.

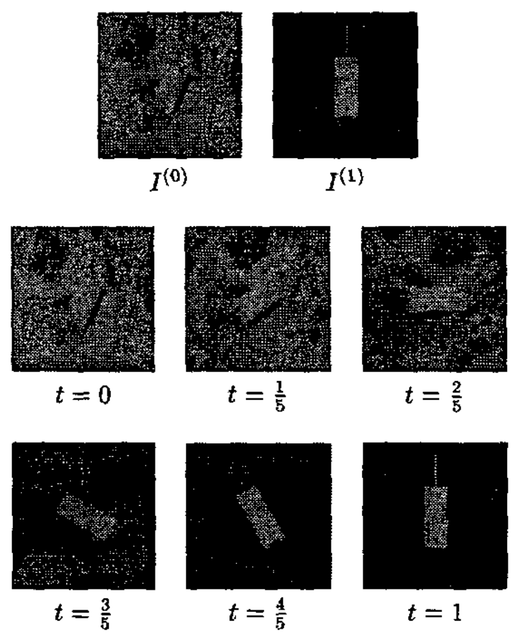

Figure 1: Path J_t connecting $I^{(0)}$ to $I^{(0)}$ at $t = 0, \frac{1}{5}...1$.

4 Target detection/identification in EO imagery

This section presents numerical results on target detection/identification in electro-optical images of targets in clutter. For the numerical experiments presented below, we synthesized a dataset of ray-traced target chips in natural clutter. Ray-tracing [8] 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 oclusion, shadows and lighting variation as well as other effects usually encountered in natural scenes.

In a statistical framework, target detection and identification are formulated as a hypothesis testing problem and tackled using classic detection theory [9]. The generalized likelihood ratio can be rewritten in terms of metric distances between the observed image I_D and the template

$$d^2(I_D, I_{temp}^{(0)}) - d^2(I_D, I_{temp}^{(1)}) \underset{H_0}{\overset{H_1}{\gtrless}} \nu \qquad (25)$$

Figure 2 displays ROC curves for target detection and identification in clutter.

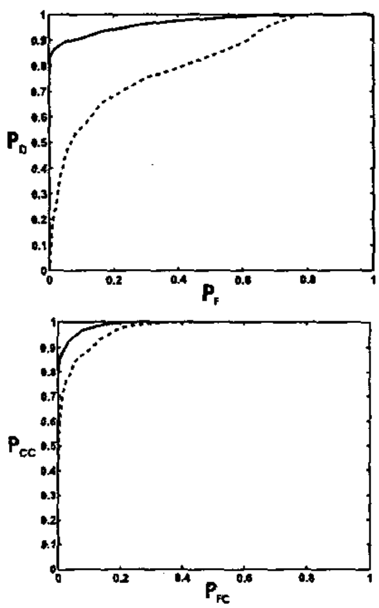

Figure 2: Top: ROC curve for detection of T72 in clutter; bottom : ROC curve for classification of T72 tank versus Jeep. Solid lines correspond to covariance norm; dashed lines - Euclidean norm.

5 Conclusions

We have introduced a metric distance for automated target recognition, which is robust with respect to geometric variability, associated with the pose of targets, as well as to the infinite variety of natural clutter in the observed images. While this model is based on the second-order statistics of natural clutter, it significantly improve detection/identification rates.

Acknowledgments

This work was supported by ONR N00014-00-1-0327, ARO DAAD19-99-1-0012, ARO DAAD19-01-10644

References

[1] M.I. Miller M.L. Cooper, U. Grenander and A. Srivastava. Accomodationg geometric and thermodynamical variability for forward-looking infrared sensors. *In Proceedings of the SPIE, Algorithms for Synthetic Aperture Radar IV*, 3070:162–172, 1997.

[2] A. Yuille. Deformable templates for face recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991.

[3] M.L. Cooper and M.I. Miller. Information measures for object recognition accommodating signature variability. *IEEE Transactions on Information Theory*, 46(5):1896–1906, 2000.

[4] Peter N. Belhumeur, Joao Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In *ECCV (1)*, pages 45–58, 1996.

[5] M.I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1/2):61–84, 2001.

[6] Manfredo P. Do Carmo. *Differential Geometry of Curves and Surfaces.* Prentice Hall, Upper Sadle River, New Jersey, 1976.

[7] M.I. Miller L. Younes D, Bitou. Clutter metrics for atr. *IEEE PAMI (submitted)*.

[8] POVRay Team. Persistency of vision ray tracer, 1993.

[9] Harry L. van Trees. *Detection, Estimation, and Modulation Theory : Part I.* John Wiley and Sons, New York, 1968.