Information Fusion 14 (2013) 311–326

Contents lists available at SciVerse ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

Information geometry of target tracking sensor networks

Yongqiang Cheng [a], Xuezhi Wang [b,*], Mark Morelande [b], Bill Moran [b]

[a] School of Electronic Science and Engineering, National University of Defence Technology, Changsha, Hunan, China
[b] Department of Electrical and Electronic Engineering, University of Melbourne, Australia

ARTICLE INFO

Article history:
Received 23 June 2011
Received in revised form 19 February 2012
Accepted 24 February 2012
Available online 6 March 2012

Keywords:
Information geometry
Target tracking
Sensor networks
Integrated Fisher information distance
Kullback–Leibler divergence
Energy functional
Ricci curvature

ABSTRACT

In this paper, the connections between information geometry and performance of sensor networks for target tracking are explored to pursue a better understanding of placement, planning and scheduling issues. Firstly, the integrated Fisher information distance (IFID) between the states of two targets is analyzed by solving the geodesic equations and is adopted as a measure of target resolvability by the sensor. The differences between the IFID and the well known Kullback–Leibler divergence (KLD) are highlighted. We also explain how the energy functional, which is the "integrated, differential" KLD, relates to the other distance measures. Secondly, the structures of statistical manifolds are elucidated by computing the canonical Levi–Civita affine connection as well as Riemannian and scalar curvatures. We show the relationship between the Ricci curvature tensor field and the amount of information that can be obtained by the network sensors. Finally, an analytical presentation of statistical manifolds as an immersion in the Euclidean space for distributions of exponential type is given. The significance and potential to address system definition and planning issues using information geometry, such as the sensing capability to distinguish closely spaced targets, calculation of the amount of information collected by sensors and the problem of optimal scheduling of network sensor and resources, etc., are demonstrated. The proposed analysis techniques are presented via three basic sensor network scenarios: a simple range-bearing radar, two bearings-only passive sonars, and three ranges-only detectors, respectively.

1. Introduction

Advanced technologies for sensing, computing and networking create enormous opportunities for handling, gathering and processing measurement information via various sensor networks. It is desirable to assess the performance of a sensor network effectively in many application fields, where the statistical properties of sensor networks are crucial. Information geometry, which is gradually gaining significance as it allows the analysis of statistical properties of sensor networks from a unified perspective, has been identified as a sophisticated and powerful tool for this purpose [1,2].

Information geometry is the study of intrinsic properties of manifolds of probability distributions [2], where the ability of data to discriminate those distributions is translated into a Riemannian metric.[1] Specifically, the Fisher information provides a local measure of discrimination of the distributions that translates immediately into a Riemannian metric on the parameter manifold of the distributions. The main tenet of information geometry is that many important notions (e.g. Fisher information, testing, estimation, and estimation accuracy) in probability theory, information theory and statistics can be treated as structures (e.g. metric, divergence, projection, and embedded curvatures) in differential geometry by regarding the space of probabilities as a differentiable manifold endowed with a Riemannian metric and a family of affine connections, including, but not exclusively, the canonical Levi–Civita affine connection [3]. By providing the means to analyse the Riemannian geometric properties of various families of probability density functions, information geometry offers comprehensive results about statistical models simply by considering them as geometrical objects.

This geometric theory of statistics was pioneered in the 1940s by Rao [4], who first interpreted the Fisher information matrix as a Riemannian metric on the space of probability distributions. Since then many scholars have contributed to the development of this theory for statistical models. In 1972, Chentsov in [5] introduced a family of affine connections and proved the uniqueness of the intrinsic metric and the one-parameter family of affine connections. Meanwhile, Efron [6] undertook pioneering work in a slightly different direction. He defined a concept of curvature called

* Corresponding author.
E-mail addresses: nudtyqcheng@gmail.com (Y. Cheng), xwang@unimelb.edu.au (X. Wang), mmorelande@unimelb.edu.au (M. Morelande), wmoran@unimelb.edu.au (B. Moran).

[1] A Riemannian metric is an inner product defined on the tangent space of a manifold. It encodes how to measure distances, angles and area at a particular point on a manifold by specifying a scalar product between tangent vectors at that point.

1566-2535/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.inffus.2012.02.005

statistical curvature and described the basic role of curvature in the high-order asymptotic theory of statistical inference. Since then, several different groups have brought to maturity the theoretical framework of statistical geometry. Of particular note is the work of Amari and his collaborators [3,7,8] who have developed a duality structure theory and have unified all of these theories in a differential–geometrical framework which not only enriches the theory of information geometry but also provides opportunities for a wide range of applications. Amari's major motivation is in learning of neural networks. Here, we study the theory from a statistical signal processing perspective.

Information geometry has found many applications in the asymptotic theory of statistical inference [9], semiparametric statistical inference [10], the study of Boltzmann machine [11], the Expectation–Maximization (EM) algorithm [12], and learning of neural networks [13], all with certain degree of success. In the last two decades, its application has spanned several discipline areas such as information theory [14,15], systems theory [16,17], mathematical programming [18], and statistical physics [19,20]. It also played a central role in the multi-terminal estimation theory [21]. In neuroscience it has been used to extract higher-order interactions among neurons [22]. Many researchers around the world are applying information geometry to new applications and formulating new interpretations. An example of the former is the derivation of the intrinsic Cramér–Rao bound for the subspace tracking problem on manifolds given in [23].

Information geometry can also provide new viewpoints in the analysis of sensing systems. While important, understanding information geometry theory is nontrivial. Sensor networks for target tracking form an important class of information networks. It is well understood that the performance of target detection and tracking depends heavily on the sensing ability of the underlying sensor network, which may consist of sensors ranging from large like radars to small like motes. The advances in engineering and sensing technologies enable more complex sensor networks to be built for target detection and tracking. The evaluation of sensor network performance becomes increasingly important, in particular, for sensor network design, configuration and optimization. We believe that information geometry is able to offer advanced tools to allow us to explore and therefore understand the structures of sensor systems. This work is motivated to explore such potential in a simple and sensible way, using basic sensor problems as exemplars.

In our recent work in [24,25], the Integrated Fisher information distance (IFID) between two targets was approximately calculated and used to measure target resolvability in the region of interest covered by a sensor network. Nevertheless, the proposed approximation for calculating IFID is only valid for closely spaced targets and the exact IFID must be evaluated by computing the integral along the geodesic connecting the two target states, which is generally nontrivial.

In this paper, the connections between information geometry and the performance of sensor networks for target tracking are explored in an attempt to gain a better understanding of sensor network measurement issues. The exact calculation of IFID and Ricci curvatures for the sensor networks with a joint likelihood are presented and analyzed. The interpretation of the geometry of statistical manifolds for sensor networks is illustrated via the affine immersion. The analysis is presented via three typical sensor network scenarios: (1) a simple range-bearing radar, (2) two bearings-only passive sonars, and (3) three ranges-only detectors, respectively. In these scenarios it is shown how information geometry can be used to address system measurement issues such as evaluating sensor capability to distinguish closely spaced targets, measuring the amount of information collected by sensors and solving the problem of optimal scheduling of network sensor and resources. Although simple synchronized sensor networks with sensors of the same type are considered in the demonstrative examples, the analysis method can be applied to a more general case where dissimilar sensors are involved as long as the likelihood and Fisher information matrix of the measurement system are available.

The major contributions of this paper are summarized as below.

1. The IFID between the states of two targets is computed by solving the geodesic equations and is used to measure the ability of a sensor network to resolve targets. The differences between IFID and the well known Kullback–Leibler divergence are described.The relationship with the energy functional, which is the integrated differential Kullback–Leibler divergence, and the differences between it and the other two measures of divergence are described.
2. The structures of statistical manifolds are elucidated by computing the canonical Levi–Civita affine connection as well as Riemannian and scalar curvatures. The relationship between the Ricci curvature tensor field and the amount of information achievable by the network sensors is highlighted.
3. An analytical presentation of statistical manifolds as immersions in Euclidean space for the distributions of the exponential family is given.

The rest of the paper is organized as follows. In the next section, the problem of interest and the motivations of this work are described. The principles of information geometry are then introduced in Section 3. In Section 4, sensor network information, as measured by the IFID, is analyzed for three basic types of sensor network problems; the canonical Levi–Civita affine connection as well as Riemannian and scalar curvatures are calculated to elucidate the structure of the statistical manifold; an interpretation of Ricci curvature tensor field related to information issues is discussed at the end of this section. The affine immersions of manifolds corresponding to sensor networks are presented in Section 5, which is followed by the conclusions in Section 6.

2. Target tracking in sensor networks

Let target state at time k be denoted as an n dimensional vector,[2] i.e., $\boldsymbol{\theta}_k = [\theta_{1,k}, \cdots, \theta_{n,k}]^T \in \mathbb{R}^n$, where the superscript T is the matrix transpose. Target dynamics are assumed to follow a Markov process with additive Gaussian noise.

$$\boldsymbol{\theta}_{k+1} = f(\boldsymbol{\theta}_k) + \boldsymbol{v}_k, \quad \boldsymbol{v}_k \sim \mathcal{N}(0, \mathbf{Q}_k) \tag{1}$$

where f is the system transition (dynamical) model and $\boldsymbol{v}_k$ represents process noise, which is assumed to be a zero-mean Gaussian distribution with covariance matrix $\boldsymbol{Q}_k$. The measurement of the system at time k is modelled as

$$\boldsymbol{x}_k = \boldsymbol{\mu}(\boldsymbol{\theta}_k) + \boldsymbol{w}_k, \quad \boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{C}_k) \tag{2}$$

where $\boldsymbol{\mu}$ is the measurement-to-target state space transition function and $\boldsymbol{w}_k$ is the measurement noise approximated by a zero mean Gaussian distribution with covariance matrix $\boldsymbol{C}_k$. The problem of target tracking is to find the posterior probability density of target state based on a sequence of measurements, i.e., $p(\boldsymbol{\theta}_k|\boldsymbol{x}_{1:k})$, where $\boldsymbol{x}_{1:k}$ stands for a sequence of measurements up to time k.

[2] In this paper, a symbol in bold face is used to denote a vector and the subscript k refers to time index. Sometimes, the time index is dropped and the subscript is subsequently used to index the location of an vector without causing confusion.