Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 1112–1122
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Consistent parameter clustering: Definition and analysis

Ulrich Hillenbrand *

Institute of Robotics and Mechatronics, German Aerospace Center (DLR), 82234 Wessling, Germany

Received 9 June 2006; received in revised form 18 January 2007
Available online 30 January 2007

Communicated by R. Davies

Abstract

Parameter clustering is a popular robust estimation technique based on location statistics in a parameter space where parameter samples are obtained from data samples. A problem with clustering methods is that they produce estimates not invariant to transformations of the parameter space. This article presents three contributions to the theoretical study of parameter clustering. First, it introduces a probabilistic formalization of parameter clustering. Second, it uses the formalism to define consistency in terms of a symmetry requirement and to derive criteria for a consistent choice of parameterization. And third, it applies the criteria to the practically relevant cases of motion and pose estimation of three-dimensional shapes. Bias and error statistics on random data sets demonstrate a significant advantage of using a consistent parameterization for rotation clustering. Moreover, clustering parameters of analytic shapes is discussed and a real application example of circle estimation given.

Keywords: Robust estimation; Clustering; Hough transform; Statistical consistency

1. Introduction

Parameter estimation is generally based upon some parametric model whose best match to the data is sought. Parameter estimators may be categorized according to whether they optimize a robust or non-robust match criterion, whether they utilize statistics in the data space or in a parameter space, and whether or not they assume a parametric probability density for the data or parameters. Parameter clustering (PC) is a technique distinguished by computing robust location and perhaps dispersion statistics in a parameter space. Location and dispersion estimates do not require a density model and, hence, PC is usually realized without assuming specific densities as a so-called non-parametric technique.

The notion of clustering is used by some authors without relation to location statistics on parameters or features but with an emphasis on partitioning a data set into distinct groups. In the present context of parameter estimation, however, partitioning may be a side effect but not the goal.

The general strategy of PC has been exploited for a long time in numerous variations (Ballard, 1981; Stockmann et al., 1982, 1987; Illingworth and Kittler, 1988; Moss et al., 1999), perhaps its most popular incarnation being the many variants of the Hough transform. Common to all these approaches is that data samples are drawn from which parameter samples are computed, often called 'votes', that satisfy constraints posed by each data sample. The intuition is that significant data populations matching an instance of the model constraint will produce many parameter samples that coincide approximately, hence localize in a cluster.

For methods that are based directly on the data statistics (Stewart, 1999), such as least-squares or M-estimators, the choice of parameterization of the model constraint does not affect the parameter estimate, as long as the mapping between parameters and constraint is invertible and

* Tel.: +49 8153 283501; fax: +49 8153 281134.
E-mail address: Ulrich.Hillenbrand@dlr.de

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.01.006

sufficiently smooth within the relevant domain. The situation for parameter-space methods, however, is quite different in general. For example, the maximum-a-posterior estimate of a parameter does not transform as the parameters do when changing the parameterization of the model constraint. Hence, the estimates obtained in different parameterizations are not equivalent. Likewise, the result of PC is not invariant to parameter transformation.

This fact brings up the question of a proper choice of parameterization for clustering, which has not received adequate treatment in the literature. It is the purpose of this article to analyze the problem from the viewpoint of statistical consistency. The latter requires of an estimate that it matches some property of the underlying data population in the limit of having infinite data samples. If one assumes a specific parametric form of the data population, one requires that the estimated parameter should take the value underlying the given population. For PC as a nonparametric technique, on the other hand, consistency shall here be defined to require that symmetry in the data be reflected in the estimates. This will yield general criteria for choosing a consistent parameterization. The criteria will be applied to the practically relevant cases of motion and pose estimation of three-dimensional (3D) shapes, and of straight line and circle estimation.

In the next section, I point out the difference of the present analysis to some major previous studies on consistency or parameterization in PC. Section 3 introduces the general estimation problem considered and the mathematical framework for its analysis; the PC algorithm and the concept of consistency are defined; criteria for consistency are derived. In Section 4, the cases of motion and pose clustering are treated; bias and error statistics on random data sets demonstrate the advantage of using a consistent parameterization. Clustering parameters of analytic shapes is discussed in Section 5; a real application example of circle estimation is given. In Section 6, I note the main questions that are not addressed in the analysis. Section 7 summarizes and concludes this study.

2. Relation to previous work

By far most studies of systematic errors in PC relate to its most popular variant, the Hough transform. Among all these studies, however, only few address the issue of consistency or parameterization.

The work in (Stewart, 1997) discusses the inconsistency of a number of robust estimators for special mixture populations of data. Although the author includes the Hough transform in his analysis, he treats it as equivalent to random sample consensus (RANSAC) and, hence, does not capture effects of the parameter space. The inconsistency analyzed in the present work is related to the parameter space and not to the data population.

Assuming a specific distribution of background data additional to the structure of interest, one may find a parameterization or parameter-space quantization such that the background contribution is uniform in parameter space. This idea is pursued in (Cohen and Toussaint, 1977; Alagar and Thiel, 1981, 1986; Hu and Ma, 1996) for the detection of simple planar structures. In fact, the resulting algorithm can be viewed as a statistical significance test against the null hypothesis of a pure background-data distribution. In most applications, however, the actual distribution of background data in each individual data set will itself differ significantly from any a-priori assumption, and structure detection may thus not be enhanced; cf. Hu and Ma (1996). In the present work, the goal of parameterization is not optimal signal detection, but consistency of the parameter estimate. Adequate parameterizations do not depend on the background-data distribution, but on intrinsic properties of the parametric model constraint.

For the Hough transform, there is a relationship of the parameter-space quantization to an equivalent template in data space, which is analyzed in (Princen et al., 1992). The authors conclude that an adequate choice of parameterization and quantization should yield a template shape that matches the sought data populations. Their analysis does not extend to continuous parameter spaces. As above, the proposed criterion for a good parameterization is dependent upon the data population and not directly related to the consistency of a parameter estimate. In the present study, continuous parameter spaces are considered, while their uniform quantizations are trivially covered by the analysis as well. The derived criteria for a consistent parameterization are independent of the underlying data population.

The work most closely related to the present one is found in (Hu and Ma, 1995). The authors discuss the adequacy of various line parameterizations for yielding consistent estimates of line orientation in the plane by the Hough transform. However, they do not provide the mathematical framework required to generalize the analysis to other estimation problems. Here, we start out from the mathematical framework, provide general criteria for consistent parameterizations, and then derive some practically relevant cases.

3. Parameter clustering

In this section, the general estimation problem and the PC approach considered are formalized. Based on the formalization, a consistency requirement is defined and necessary and sufficient conditions for consistent parameterizations are derived.

3.1. The estimation problem

Suppose we want to estimate a transformation T from a model- or data-point set $X \subset \mathbb{R}^m$ to a data-point set $Y \subset \mathbb{R}^n$. The transformation of a point $x \in \mathbb{R}^m$ is assumed to have the general parametric form

$$T(x, \alpha) = F(G_\alpha(x)) = F \circ G_\alpha(x), \tag{1}$$