Pattern Recognition 85 (2019) 26–36

Contents lists available at ScienceDirect

Pattern Recognition

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

Hierarchical Bayesian image analysis: From low-level modeling to robust supervised learning☆

Adrien Lagrange[a,*], Mathieu Fauvel[b], Stéphane May[c], Nicolas Dobigeon[a,d]

[a] University of Toulouse, IRIT/INP-ENSEEIHT Toulouse, BP 7122, Toulouse Cedex 7 31071, France
[b] INRA, DYNAFOR, BP 32607, Auzeville-Tolosane, 31326 Castanet Tolosan, France
[c] CNES, DCT/SI/AP, 18 Avenue Edouard Belin, 31400 Toulouse, France
[d] Institut Universitaire de France, France

ARTICLE INFO

Article history:
Received 30 November 2017
Revised 11 May 2018
Accepted 22 July 2018
Available online 31 July 2018

Keywords:
Bayesian model
Supervised learning
Image interpretation
Markov random field

ABSTRACT

Within a supervised classification framework, labeled data are used to learn classifier parameters. Prior to that, it is generally required to perform dimensionality reduction via feature extraction. These preprocessing steps have motivated numerous research works aiming at recovering latent variables in an unsupervised context. This paper proposes a unified framework to perform classification and low-level modeling jointly. The main objective is to use the estimated latent variables as features for classification and to incorporate simultaneously supervised information to help latent variable extraction. The proposed hierarchical Bayesian model is divided into three stages: a first low-level modeling stage to estimate latent variables, a second stage clustering these features into statistically homogeneous groups and a last classification stage exploiting the (possibly badly) labeled data. Performance of the model is assessed in the specific context of hyperspectral image interpretation, unifying two standard analysis techniques, namely unmixing and classification.

1. Introduction

In the context of image interpretation, numerous methods have been developed to extract meaningful information. Among them, generative models have received a particular attention due to their strong theoretical background and the great convenience they offer in term of interpretation of the fitted models compared to some model-free methods such as deep neural networks. These methods are based on an explicit statistical modeling of the data which allows very task-specific model to be derived [1], or either more general models to be implemented to solve generic tasks, such as Gaussian mixture model for classification [2]. Task-specific and classification-like models are two different ways to reach an interpretable description of the data with respect to a particular applicative non-semantic issue. For instance, when analyzing images, task-specific models aim at recovering the latent (possibly physics-based) structures underlying each pixel-wise measurement [3] while classification provides a high-level information, reducing the pixel characterization to a unique label [4].

Classification is probably one of the most common way to interpret data, whatever the application field of interest [5]. This undeniable appeal has been motivated by the simplicity of the resulting output. This simplicity induces the appreciable possibility of benefiting from training data at a relatively low cost. Indeed, experts can generally produce a ground-truth equivalent to the expected results of the classification for some amount of the data. This supervised approach allows a priori knowledge to be easily incorporated to improve the quality of the inferred classification model. Nevertheless, supervised methods are significantly influenced by the size of the training set, its representativeness and reliability [6]. Moreover, in some extent, modeling the pixel-wise data by a single descriptor may appear as somehow limited. It is the reason why the user-defined classes often refer to some rather vague semantic meaning with a possible large intra-class variability. To overcome these issues, while simultaneously facing with theoretical limitations of the expected classifier ability of generalization [7], an approach consists in preceding the training stage with feature extraction [8]. These feature extraction techniques, whether parametric or nonparametric, have also the great advantage of simultaneously and significantly reducing the data volume to be handled as well as the dimension of the space in which the training should be sub-

☆ Part of this work has been supported Centre National d'Études Spatiales (CNES), Occitanie Region and EU FP7 through the ERANETMED JC-WATER program, MapInvPlnt Project ANR-15-NMED-0002-02.

* Corresponding author.
E-mail addresses: adrien.lagrange@enseeiht.fr (A. Lagrange), mathieu.fauvel@ensat.fr (M. Fauvel), stephane.may@cnes.fr (S. May), nicolas.dobigeon@enseeiht.fr (N. Dobigeon).

https://doi.org/10.1016/j.patcog.2018.07.026
0031-3203/

sequently conducted. Unfortunately, they are generally conducted in a separate manner before the classification task, i.e., without benefiting from any prior knowledge available as training data. Thus, a possible strategy is to consider a (possibly huge) set of features and selecting the relevant ones by appropriate optimization schemes [9].

This observation illustrates the difficulty of incorporating ground-truthed information into a feature extraction step or, more generally, into a latent (i.e., unobserved) structure analysis. Due to the versatility of the data description, producing expert ground-truth with such degrees of accuracy and flexibility would be time-consuming and thus prohibitive. For example, for a research problem as important and well-documented as that of source separation, only very few and recent attempts have been made to incorporate supervised knowledge provided by an end-user [10]. Nonetheless, latent structure analysis may offer a relevant and meaningful interpretation of the data, since various conceptual yet structured knowledge to be inferred can be incorporated into the modeling. In particular, when dealing with measurements provided by a sensor, task-related biophysical considerations may guide the model derivation [11]. This is typically the case when spectral mixture analysis is conducted to interpret hyperspectral images whose pixel measurements are modeled as combinations of elementary spectra corresponding to physical elementary components [12].

The contribution of this paper lies in the derivation of a unified framework able to perform classification and latent structure modeling jointly. First, this framework has the primary advantage of recovering consistent high and low level image descriptions, explicitly conducting hierarchical image analysis. Moreover, improvements in the results associated with both methods may be expected thanks to the complementarity of the two approaches. The use of ground-truthed training data is not limited to driving the high level analysis, i.e., the classification task. Indeed, it also makes it possible to inform the low level analysis, i.e., the latent structure modeling, which usually does not benefit well from such prior knowledge. On the other hand, the latent modeling inferred from each data as low level description can be used as features for classification. A direct and expected side effect is the explicit dimension reduction operated on the data before classification [7]. Finally, the proposed hierarchical framework allows the classification to be robust to corruption of the ground-truth. As mentioned previously, performance of supervised classification may be questioned by the reliability in the training dataset since it is generally built by human expert and thus probably corrupted by label errors resulting from ambiguity or human mistakes. For this reason, the problem of developing classification methods robust to label errors has been widely considered in the community [13,14]. Pursuing this objective, the proposed framework also allows training data to be corrected if necessary.

The interaction between the low and high level models is handled by the use of non-homogeneous Markov random fields (MRF) [15]. MRFs are probabilistic models widely-used to describe spatial interactions. Thus, when used to derive a prior model within a Bayesian approach, they are particularly well-adapted to capture spatial dependencies between the latent structures underlying images [16,17]. For example, Chen et al. [18] proposed to use MRFs to perform clustering. The proposed framework incorporates two instances of MRF, ensuring consistency between the low and high level modeling, consistency with external data available as prior knowledge and a more classical spatial regularization.

The remaining of the article is organized as follows. Section 2 presents the hierarchical Bayesian model proposed as a unifying framework to conduct low-level and high-level image interpretation. A Markov chain Monte Carlo (MCMC) method is derived in Section 3 to sample according to the joint posterior distribution of the resulting model parameters. Then, a particular and illustrative instance of the proposed framework is presented in Section 4 where hyperspectral images are analyzed under the dual scope of unmixing and classification. Finally, Section 5 concludes the paper and opens some research perspectives to this work.

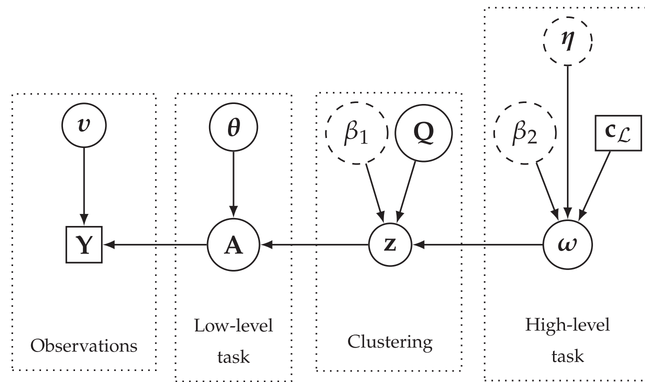

Fig. 1. Directed acyclic graph of the proposed hierarchical Bayesian model. (User-defined parameters appear in dotted circles and external data in squares).

2. Bayesian model

In order to propose a unifying framework offering multi-level image analysis, a hierarchical Bayesian model is derived to relate the observations and the task-related parameters of interest. This model is mainly composed of three main levels. The first level, presented in Section 2.1, takes care of a low-level modeling achieving latent structure analysis. The second stage then assumes that data samples (e.g., resulting from measurements) can be divided into several statistically homogeneous clusters through their respective latent structures. To identify the cluster memberships, these samples are assigned discrete labels which are a priori described by a non-homogeneous Markov random field (MRF). This MRF combines two terms: the first one is related to the potential of a Potts-MRF to promote spatial regularity between neighboring pixels; the second term exploits labels from the higher level to promote coherence between cluster and classification labels. This clustering process is detailed in Section 2.2. Finally, the last stage of the model, explained in Section 2.3, allows high-level labels to be estimated, taking advantage of the availability of external knowledge as ground-truthed or expert-driven data, akin to a conventional supervised classification task. The whole model and its dependences are summarized by the directed acyclic graph in Fig. 1.

2.1. Low-level interpretation

The low-level task aims at inferring P R-dimensional latent variable vectors $\mathbf{a}_p$ ($\forall p \in \mathcal{P} \triangleq \{1, \ldots, P\}$) appropriate for representing P respective d-dimensional observation vectors $\mathbf{y}_p$ in a subspace of lower dimension than the original observation space, i.e., $R \leq d$. The task may also include the estimation of the function or additional parameters of the function relating the unobserved and observed variables. By denoting $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_P]$ and $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_P]$ the $d \times P$- and $R \times P$- matrices gathering respectively the observation and latent variable vectors, this relation can be expressed through the general statistical formulation

$$\mathbf{Y}|\mathbf{A}, \boldsymbol{\upsilon} \sim \Psi(\mathbf{Y}; f_{\text{lat}}(\mathbf{A}), \boldsymbol{\upsilon}), \tag{1}$$

where $\Psi(\cdot, \boldsymbol{\upsilon})$ stands for a statistical model, e.g., resulting from physical or approximation considerations, $f_{\text{lat}}(\cdot)$ is a deterministic function used to define the latent structure and $\boldsymbol{\upsilon}$ are possible additional nuisance parameters. In most applicative contexts aimed by this work, the model $\Psi(\cdot)$ and function $f_{\text{lat}}(\cdot)$ are separable