Contents lists available at ScienceDirect

Pattern Recognition

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

Modeling and recognizing human trajectories with beta process hidden Markov models

Shiliang Sun*, Jing Zhao, Qingbin Gao

Shanghai Key Laboratory of Multidimensional Information Processing, Department of Computer Science and Technology, East China Normal University, 500 Dongchuan Road, Shanghai 200241, PR China

ARTICLE INFO

Article history:
Received 12 July 2013
Received in revised form
8 February 2015
Accepted 27 February 2015

Keywords:
Human activity recognition
Trajectory classification
Beta process
Markov chain Monte Carlo
Hidden Markov model

ABSTRACT

Trajectory-based human activity recognition aims at understanding human behaviors in video sequences, which is important for intelligent surveillance. Some existing approaches to this problem, e.g., the hierarchical Dirichlet process hidden Markov models (HDP-HMM), have a severe limitation, namely the motions are shared among trajectories from the same activity and not shared among activities (classes). To overcome this shortcoming, we propose a new method for modeling human trajectories based on the beta process hidden Markov models (BP-HMM) where the motions are selectively shared among trajectories. All the trajectories from different activities can be jointly modeled with a BP-HMM, which allows motions being shared among activities. Using our technique, the number of available motions and the sharing patterns can be inferred automatically from training data. We develop an efficient Markov chain Monte Carlo algorithm for model training. Experiments on both synthetic and real data sets demonstrate the effectiveness of our approach.

1. Introduction

Effective human activity recognition (HAR) is crucial for the successful application of intelligent surveillance systems, which has drawn growing interest during recent years. The purpose of HAR is to understand what people are doing from their position [1], figure [2], motion [3], or other spatio-temporal information derived from video sequences. Research in this field mainly includes three stages: (a) object segmentation and tracking, (b) feature extraction, and (c) activity classification, which are also typical steps of designing a surveillance system. In this paper, we focus on the activity classification tasks, or more specifically, learning to recognize human behaviors from trajectory data [4–6].

An example of this kind of problem can be found in a shopping mall. In this scene, customers take different classes of normal activities such as “entering the shop”, “leaving the shop”, “passing the shop”, etc. Meanwhile, some abnormal activities may also occur, such as “wandering” and “fighting”. In order to detect the abnormal behaviors, a surveillance system should have the ability to characterize and classify the activities. Due to a growing need from the public safety and security, this requirement is also prevalent in many other scenes, such as airports, railway stations, and banks.

From daily experience we know that a human activity can be modeled by transitions among simple motions, and moreover, multiple trajectories may share the same motions. For example, in a certain shopping mall, the activity of a customer “entering the shop” may be decomposed into “moving east” first and then “moving north”, and the activity of a customer “leaving the shop” may be decomposed into “moving south” first and then “moving east”. In this simple case, the motion “moving east” is shared while the motions “moving north” and “moving south” are class-specific.

Hidden Markov models (HMMs) are a popular class of models in this field, by which the invisible motions can be modeled as hidden states and the hidden state sequence corresponding to a human trajectory forms a Markov chain. Moreover, the transitions between the hidden states in an HMM represent the transitions between the motions from a human activity. Therefore, HMMs are capable to capture intrinsic structures of activities. However, a serious limitation of the standard HMMs is that one needs to define an appropriate number of hidden states. Models with too few states lack flexibility while too many states lead to the overfitting problem. Another challenge is that the standard HMMs do not have the capability to jointly model multiple activities with shared motions. Nevertheless, in the real HAR practice, motions are often shared among different activities. Eliminating these limitations can result in more flexible models, which is also our desideratum.

In this paper, we propose a method for trajectory-based HAR tasks. We first try to discover a set of latent motions, including the shared motions and the unique motions, in specific trajectories. Then we model the trajectories by transitions of different motions. A Markov chain Monte Carlo (MCMC) algorithm is developed for efficient model training. The final classifier for HAR tasks is given by maximizing the

---

* Corresponding author. Tel.: +86 21 54345183; fax: +86 21 54345119.
E-mail address: slsun@cs.ecnu.edu.cn (S. Sun).

http://dx.doi.org/10.1016/j.patcog.2015.02.028
0031-3203/© 2015 Elsevier Ltd. All rights reserved.

Please cite this article as: S. Sun, et al., Modeling and recognizing human trajectories with beta process hidden Markov models, Pattern Recognition (2015), http://dx.doi.org/10.1016/j.patcog.2015.02.028

log-likelihood of a test trajectory. Our work is inspired by the beta process hidden Markov models (BP-HMM) [7], which is concerned with the patterns of sharing motions. In particular, motions shared among the multiple time series or unique can be inferred from data. This is quite different from other HMM-based approaches [1,12–15] where the state spaces, the transitions among these states and the emission parameters are exactly shared in multiple time series from one kind of activity. The characteristics of soft-sharing motions in BP-HMM help to jointly model all the trajectories from different activities with one model, which allows motions being shared among activities. In contrast, in other methods [1,12–15] trajectories from different activities are separately modeled with different parameters for classification, which causes that motions cannot be shared among activities. Important distinctions between (BP-HMM) [7] and our method are as follows. First, the main dynamical system in the BP-HMM is a vector autoregressive process while we use an HMM with Gaussian emissions which is more appropriate for trajectory-based HAR tasks. Second, our sampling method is partially different from the BP-HMM, which is more suitable for our model. Third, we employ BP-HMM on human trajectory recognition through jointly modeling all the trajectories for training and performing an additional procedure on trajectory-specific transitions before classification. A preliminary result was reported previously [6]. This paper substantially extends that work in terms of additional knowledge presentations including Dirichlet process, hierarchical Dirichlet process and beta process, more understandable model descriptions, additional detailed sampling algorithm, more adequate experiments and additional performance and complexity analysis.

The remainder of this paper is organized as follows. Section 2 gives the overview of related HAR work and statistical models. Section 3 presents our method for modeling human trajectories and Section 4 introduces the adopted parameter inference techniques. Section 5 gives a detailed algorithm with complexity analysis for sampling. After providing the classification rule in Section 6, Section 7 reports experimental results on both synthetic and real data sets including comparisons with other methods. Finally, Section 8 gives conclusions and future research directions.

2. Related HAR work and statistical models

HMMs are widely adopted methods for trajectory recognition, which model the temporal evolution of motions in a human trajectory as a Markov chain. The real challenge of these methods is to model the hidden states. Bashir et al. [8] considered the hidden states as subtrajectories. They segmented a trajectory at points of change in curvature and represented the subtrajectories by their principal component analysis coefficients. Nascimento et al. [1] considered the hidden states as motions to be estimated from a model selection method. In addition, conditional random fields (CRFs) have also been applied to trajectory recognition [9,10,4]. For example, Liao et al. [9] showed how to extract a person's activities and significant places from trajectories of GPS data using a hierarchical CRF, and Gao and Sun [4] proposed a method for trajectory-based human activity recognition using a hidden CRF (HCRF) [19]. In this paper, we build our model based on HMMs due to the potential Markov property of a human trajectory. It is a reasonable assumption since in human trajectories the next motion is weakly correlated with (even independent of) the previous motions, given the current motion.

In order to form the foundation of the proposed human activity model, we now briefly review several related statistical models.

2.1. Dirichlet process

The Dirichlet process (DP) is the infinite-dimensional generalization of the Dirichlet distribution and takes two inputs: a positive real scalar α_0 and a base distribution G_0. A draw, G, from a DP can be denoted by $G \sim \mathrm{DP}(\alpha_0, G_0)$. Sethuraman [11] provided a stick-breaking construction which facilitates the interpretation of the DP. Consider two infinite collections of random variables $\theta_k \sim G_0$ and $v_k \sim \mathrm{Beta}(1, \alpha_0)$ for $k = \{1, \ldots, \infty\}$. By introducing a proportion variable $\pi_k = v_k \prod_{j=1}^{k-1} (1 - v_j)$, the stick-breaking representation of G has the form $G = \sum_{i=1}^{\infty} \pi_i \delta_{\theta_i}$. We simplify the notation of $(\pi_k)_{k=1}^{\infty}$ by $\boldsymbol{\pi} = (\pi_k)_{k=1}^{\infty} \sim \mathrm{GEM}(\alpha_0)$.

The DP is commonly used as a prior on the parameters of a mixture model with an unbounded number of mixture components. If we model a set of observations $\{x_1, \ldots, x_n\}$ using a set of latent parameters $\{\theta_1, \ldots, \theta_n\}$, the Dirichlet process mixture model can be expressed as the following generative process:

$$G \sim \mathrm{DP}(\alpha_0, G_0), \quad \theta_i \sim G, \quad x_i \sim F(\theta_i). \tag{1}$$

2.2. Hierarchical Dirichlet process

Consider modeling groups of data, where each group is modeled via a mixture model and clusters are shared across groups. A set of original DPs with the same base distribution, namely $\{G_j\} \sim \mathrm{DP}(\alpha_0, G_0)$, is impractical for this problem. Clusters may not be shared between groups in this simple model, since for continuous G_0, different G_j have no atoms in common. To resolve this issue, Teh et al. [12] proposed the HDP, in which the base distribution is itself drawn from a DP:

$$G_0 \sim \mathrm{DP}(\gamma, H), \quad G_j \sim \mathrm{DP}(\alpha_0, G_0). \tag{2}$$

Here, the base distribution H is continuous, while G_0 is discrete with probability one. γ governs the variability of G_0 around H and α_0 governs the variability of G_j around G_0. In this hierarchical model, G_0 is atomic and is a base measure for the draw $G_j \sim \mathrm{DP}(\alpha_0, G_0)$, which implies that only the atoms of G_0 can appear in G_j. Thus, these atoms can be shared among the random measures $\{G_j\}$.

The constructive definition of (2) is as follows. Since G_0 is distributed as a DP, it can be alternatively described via the stick-breaking representation $G_0 = \sum_{k=1}^{\infty} \beta_k \delta_{\phi_k}$, where $\phi_k \sim H$ and $\boldsymbol{\beta} = ((\beta_k)_{k=1}^{\infty} \sim \mathrm{GEM}(\gamma)$. Similarly, G_j can be written as $G_j = \sum_{k=1}^{\infty} \pi_{jk} \delta_{\phi_k}$ where $\pi_{jk} = v_{jk} \prod_{l=1}^{k-1} (1 - v_{jl})$ and $v_{jk} \sim \mathrm{Beta}(\alpha_0 \beta_k, \alpha_0 (1 - \sum_{l=1}^{k} \beta_l))$.

The HDP can be used to develop an HMM with an infinite state space, namely HDP-HMM or iHMM [12,13]. The model employs an HDP prior over the transitions:

$$\begin{aligned} &\boldsymbol{\beta} = (\beta_k)_{k=1}^{\infty} \sim \mathrm{GEM}(\gamma), \\ &\boldsymbol{\pi}_k \sim \mathrm{DP}(\alpha_0, \boldsymbol{\beta}). \end{aligned} \tag{3}$$

Using this technique, the number of hidden states can be inferred from data rather than fixed in advance.

2.3. Beta process

The beta process (BP) is a distribution on positive random measures, which is defined on the product space $\Theta \otimes [0, 1]$, with Θ a measure space. A draw $B \sim \mathrm{BP}(c, B_0)$ can be described by

$$B = \sum_{k=1}^{\infty} \omega_k \delta_{\theta_k}. \tag{4}$$

Here, c is the concentration parameter and B_0 is the base measure. θ_k are atoms in B, ω_k are the corresponding weights of θ_k. If B_0 is a discrete distribution in which the mass of the kth atom θ_k is $q_k \in (0, 1)$, ω_k is generated by a beta distribution $\mathrm{Beta}(c q_k, c(1 - q_k))$. If B_0 is a continuous distribution with mass parameter $\gamma = B_0(\Theta)$, (θ_k, ω_k) are draws from a Poisson process with the Lévy measure $\nu(d\theta, d\omega)$ [15]. The beta process is conjugate to the Bernoulli process (BeP) [15]:

$$\boldsymbol{f} \mid B \sim \mathrm{BeP}(B), \tag{5}$$

where $\boldsymbol{f}$ is a finite sequence of binary variables. Given the BP prior for B, and the BeP likelihood for $\boldsymbol{f}$ conditional on B, the posterior

Please cite this article as: S. Sun, et al., Modeling and recognizing human trajectories with beta process hidden Markov models, Pattern Recognition (2015), http://dx.doi.org/10.1016/j.patcog.2015.02.028