Int J Comput Vis (2013) 103:100–126
DOI 10.1007/s11263-012-0596-6

A Sequential Topic Model for Mining Recurrent Activities from Long Term Video Logs

Jagannadan Varadarajan · Rémi Emonet · Jean-Marc Odobez

Received: 6 October 2011 / Accepted: 2 November 2012 / Published online: 19 December 2012

Abstract This paper introduces a novel probabilistic activity modeling approach that mines recurrent sequential patterns called *motifs* from documents given as word × time count matrices (e.g., videos). In this model, documents are represented as a mixture of sequential activity patterns (our motifs) where the mixing weights are defined by the motif starting time occurrences. The novelties are multi fold. First, unlike previous approaches where topics modeled only the co-occurrence of words at a given time instant, our motifs model the co-occurrence and temporal order in which the words occur within a temporal window. Second, unlike traditional Dynamic Bayesian networks (DBN), our model accounts for the important case where activities occur concurrently in the video (but not necessarily in synchrony), i.e., the advent of activity motifs can overlap. The learning of the motifs in these difficult situations is made possible thanks to the introduction of latent variables representing the activity starting times, enabling us to implicitly align the occurrences of the same pattern during the joint inference of the motifs and their starting times. As a third novelty, we propose a general method that favors the recovery of sparse distributions, a highly desirable property in many topic model applications, by adding simple regularization constraints on the searched distributions to the data likelihood optimization criteria. We substantiate our claims with experiments on synthetic data to demonstrate the algorithm behavior, and on four video datasets with significant variations in their activity content obtained from static cameras. We observe that using low-level motion features from videos, our algorithm is able to capture sequential patterns that implicitly represent typical trajectories of scene objects.

Keywords Unsupervised · Latent sequential patterns · Topic models · PLSA · Video surveillance · Activity analysis

List of Symbols

$\mathcal{D}$	Dataset, count matrices of the form $n(w, t_a, d)$
z	Motif index
w	Word index (SLA patterns for real data. See Sect. 5.)
d	Temporal document index
t_a	Absolute time index in temporal documents
t_s	Start time of a motif
t_r	Relative time from the start of the motif
Θ	Model parameters $\{P(z\|d), P(t_s\|z, d), P(w, t_r\|z)\}$
N_z	Number of motifs
N_w	Vocabulary size (number of different words)
D	Number of temporal documents
T_z	Maximum duration of a motif
T_d	Duration of the temporal document
T_{ds}	Number of motif start time indices in a temporal document
$\lambda_{z,d}$	Sparsity constraint weight
λ_{bic}	Penalty term weight in BIC equation

1 Introduction

Immense progress in sensor and communication technologies has led to the development of devices and systems

J. Varadarajan (✉)· R. Emonet · J.-M. Odobez
Idiap Research Institute, Martigny, Switzerland
e-mail: vjagann@idiap.ch

J. Varadarajan · R. Emonet · J.-M. Odobez
École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
e-mail: remonet@idiap.ch

J.-M. Odobez
e-mail: odobez@idiap.ch