Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 123 (2018) 314–320

Procedia Computer Science

www.elsevier.com/locate/procedia

8th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2017

Context-Dependent Robust Text Recognition using Large-scale Restricted Bayesian Network

Hidemoto Nakada and Yuuji Ichisugi

Abstract We have been proposing a computational model of the cerebral cortex called BESOM, which models the cerebral cortex as restricted Bayesian networks based on recent findings in the neuroscience area. Since BESOM is based on Bayesian network, it inherently allows bi-directional information flow, meaning that it can naturally merge information extracted from concrete data with highly-abstract prior knowledge. As an example of such kind of tasks, we report robust text recognition task with context information. We show that word spelling knowledge and word n-gram could be represented as a part of the network and they contribute the text recognition accuracy with noisy text images. We also show that the computational cost is approximately linear with the number of characters and words.

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures

1 Introduction

Robust text recognition under highly noisy environment requires utilizing prior knowledge such as word spelling or word n-gram knowledge. One of the well know method is a technique that use weighted acyclic automata, also known as lattice, that represents various hypotheses form multiple stages [4]. The best path from top to bottom of the lattice represents the most probable candidate. To adopt this method for text recognition, first, we get candidate character list for each character image using some character recognition mechanism. And then we can construct a lattice using the candidate character list and prior knowledge. This is not ideal, because character recognition stage and text recognition stage are completely separated. This means that characters once omitted from the candidate character list will never considered under the prior knowledge.

Hidemoto Nakada, Yuuji Ichisugi
National Institute of Advanced Science and Technology, 2-3-26 Aomi, Koto-ku, Tokyo, 135–0064 Japan, e-mail: {hide-nakada, y-ichisugi}@aist.go.jp

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures
10.1016/j.procs.2018.01.049

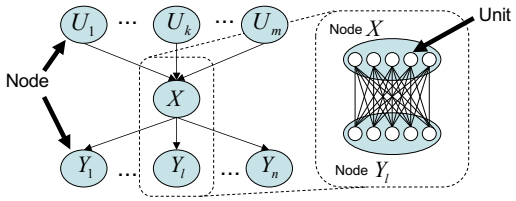

Fig. 1 BESOMNetwork. The ovals indicate the node and the small circles in oval stand for units. Typical BESOM network forms multi-layered structure as shown in this diagram.

We believe that Bayesian network based model is suitable for this kind of tasks, and have been developing BESOM model (BidirEctional Self Organizing Maps) [7][6]. BESOM is a machine learning model that models cerebral cortex as Bayesian network. One of the characteristics of BESOM is the bi-directional information flow in the model, that enables hybrid-style inference using both of the concrete data and highly-abstracted prior knowledge [5]. This means that BESOM can interpret low-level sensor information using background knowledge which is extracted with other means.

In this paper, we discuss robust text recognition task using prior knowledge. We demonstrate that, 1) BESOM can easily represent prior knowledge such as word spell and word 2-gram as network structure, 2) BESOM can achieve better text-recognition accuracy, leveraging the prior knowledge, than state-of-the-art CNN, and 3) The computation cost is linear to the number of words and characters in the text.

2 Text recognition using context information with BESOM

2.1 BESOM

According to the recent studies in computational neuroscience, Bayesian network [8] can be the underlying mechanism of the cerebral cortex. Various neuroscientific phenomena are successfully reproduced with models based on Bayesian networks. The cerebral cortex shares many aspects, not only in functions but also in structures, with Bayesian networks. Based on this assumption, we have been developing BESOM, a Bayesian network based cerebral cortex model. BESOM represents macro columns in cerebral cortex as *nodes*, and mini columns as *units* in nodes. In Bayesian network terminology, nodes stand for probabilistic variables and units stand for specific values that the variable can take.

Naive Bayesian network is notorious for huge memory footprint and computation hungry nature. We have carefully restricted the conditional probability model to reduce the memory footprint and computational burden.