Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 70 (2007) 1198–1206
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Adaptive sensor modelling and classification using a continuous restricted Boltzmann machine (CRBM)

Tong Boon Tang*, Alan F. Murray

School of Engineering and Electronics, The University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JL, UK

Available online 13 December 2006

Abstract

A probabilistic, "neural" approach to sensor modelling and classification is described, performing local data fusion in a wireless system for embedded sensors using a continuous restricted Boltzmann machine (CRBM). The sensor data clusters are non-Gaussian and their classification is non-linear. A CRBM is shown to be able to model complex data distributions and to adjust autonomously to measured sensor drift. Performance is compared with that of single layer and multilayer neural classifiers. It is shown that a CRBM can resolve the problem of catastrophic interference that is typical of associative memory based models.

Keywords: Embedded sensor; Sensor drift; Probabilistic computation

1. Introduction

Driven by the recent advances in microsensor technology, multisensor microsystems have been developed for diverse applications. For example, in the wireless integrated network sensors (WINS) [15], the *SMARTDUST* [21] and the *picoNode* [16] arrays, low cost sensing devices are deployed for environmental and machinery monitoring. To overcome the restrictions of communication network bandwidth, sensor signals are often partly processed locally. Events of interest trigger an alarm to a basestation for further decision making. Local data fusion (typically classification) must allow data fusion at all levels (signal, pixel, feature and symbol) and neural solutions are often chosen [11].

We focus on local data fusion using a non-linear, probabilistic generative model, the "continuous restricted Boltzmann machine" (CRBM) [2] and a linear classifier in the form of a single layer perceptron (SLP). The CRBM extracts salient features by modelling the data distribution of the integrated sensors, while the SLP performs binary classification on the extracted features. The CRBM was developed to fuse continuous-valued (analogue) data locally and to have a hardware-amenable architecture [2].

The CRBM has one visible and one hidden layer with a symmetric inter-layer weight matrix $\{\mathbf{W}\}$. Each stochastic neuron j takes the following form:

$$s_j = \tanh\left[a_j\left(\sum_i w_{ij}s_i + \sigma \cdot N_j(0,1)\right)\right], \tag{1}$$

where

- s_i is the input from neuron i,
- $N_j(0,1)$ is a zero-mean Gaussian noise source with amplitude of 1,
- $\sigma \cdot N_j(0,1)$ allows the CRBM to perform probabilistic analogue computation via Gibbs sampling [4],
- a_j is a parameter which controls the slope of the sigmoid function, such that the behaviour of a neuron j can vary from deterministic (small a_j) through continuous-stochastic (moderate a_j) to binary-stochastic (large a_j).

Both $\{a_j\}$ and $\{w_{ij}\}$ are trained by a "minimizing contrastive divergence" (MCD) learning rule [6]. The simplified MCD learning rule requires only addition and multiplication, and is therefore more hardware-amenable [2].

*Corresponding author. Tel.: +44 131 6505664; fax: +44 131 6506554.
E-mail address: tbt@ee.ed.ac.uk (T.B. Tang).

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2006.11.014

The objectives of this paper are

(1) to demonstrate that the CRBM can model complex non-Gaussian distributions,
(2) to illustrate the advantage of adaptive sensor fusion in a dynamic environment.

There are two major challenges in our chosen sensing application (digestive system monitoring with an array of ion sensitive field effect transistor (ISFET) pH sensors [19]). Sensor drift, the first challenge, is caused by unknown dynamic processes such as poisoning or ageing of the sensor. The CRBM is able to adapt ''on-line'' to changing data distribution and can thus compensate for sensor drift. The second challenge is that, at any time during normal usage, the sensor array will produce data from only one area of data space. The data available for on-line training are not, therefore, representative of the full distribution from which training data are drawn (i.e. only data drawn from a sub-class of the full data space are likely to be available at one time, under normal operating conditions). Under these conditions, most associative memory based systems simply learn to model the available sub-class and effectively ''forget'' the existence of all other sub-classes. Such a system loses the ability to classify data. This problem is commonly known as 'catastrophic interference' (CI) [12].

This paper is organized as follows. Section 2 shows that a CRBM can both model non-Gaussian distributions and avoid CI. In this case, the learning is evaluated based on the reconstruction model and a binary classification with a SLP connected to the hidden units of the CRBM (Section 3). Section 4 examines the CRBM's modelling capability in both static and dynamic environments. Finally, we summarize the meaning of these results and draw conclusions.

2. Sensor modelling

In an embedded sensor fusion/classification system, an initial classification of the data from multiple sensors is useful, in order to decide whether onward transmission of the sensors' outputs is worthwhile. For example, it may be desirable that the wireless communication system be switched off until a particular set of sensor outputs is detected, in order to minimize power consumption. The noise that is present in most real data and their measurement systems is often non-Gaussian. For example, strong non-Gaussian radio frequency interference is unavoidable in landmine detection [18], and the emitted signals by sources are often non-Gaussian in localization of multiple sources [8]. It is therefore necessary to extend our previous work on Gaussian data distributions [20] to non-Gaussian modelling problems.

Modelling non-Gaussian distributions with the Gaussian ''experts'' that underlie a CRBM is non-trivial. To illustrate this, we modelled a 2D non-Gaussian data distribution (Fig. 1(a)) using a CRBM with five (Gaussian)

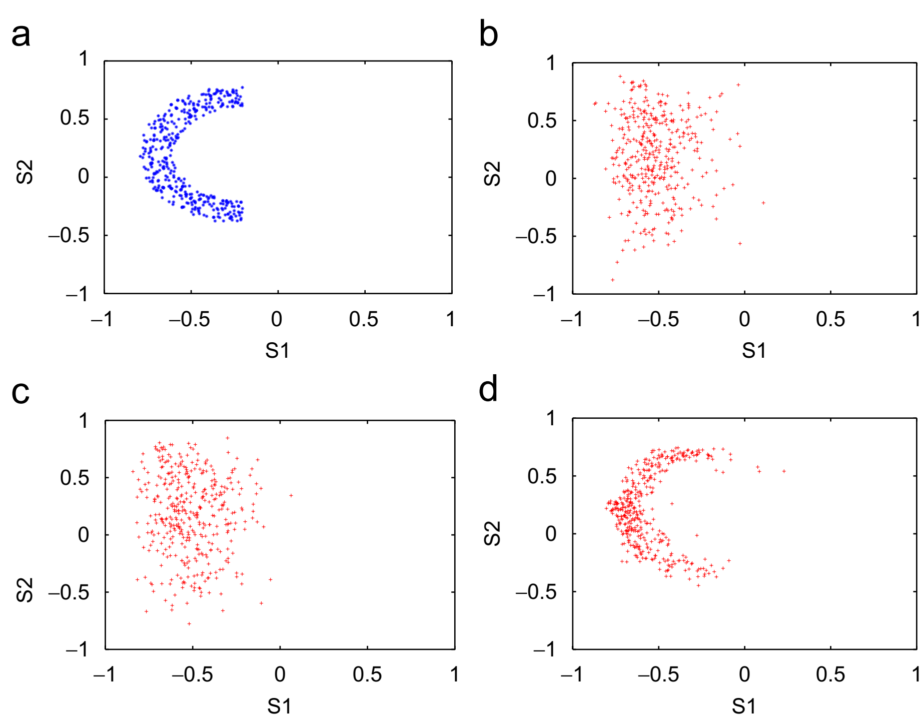

Fig. 1. (a) Non-Gaussian training data; (b) reconstruction data after training with a_j unconstrained and 1-step Gibbs sampling; (c) reconstruction data after training with a_j unconstrained and 20-step Gibbs sampling; and (d) reconstruction data after training with a_j fixed at 0.1. The training duration for (b)–(d) is 5000 epochs.