An Efficient Clustering Algorithm Using Stochastic Association Model and Its Implementation Using Nanostructures

Takashi Morie, Tomohiro Matsuura, Makoto Nagata, and Atsushi Iwata
Graduate School of Advanced Sciences of Matter, Hiroshima University
Higashi-hiroshima, 739-8526 Japan.
http://www.dsl.hiroshima-u.ac.jp
morie@dsl.hiroshima-u.ac.jp

Abstract

This paper describes a clustering algorithm for vector quantizers using a "stochastic association model". It offers a new simple and powerful softmax adaptation rule. The adaptation process is the same as the on-line K-means clustering method except for adding random fluctuation in the distortion error evaluation process. Simulation results demonstrate that the new algorithm can achieve efficient adaptation as high as the "neural gas" algorithm, which is reported as one of the most efficient clustering methods. It is a key to add uncorrelated random fluctuation in the similarity evaluation process for each reference vector. For hardware implementation of this process, we propose a nanostructure, whose operation is described by a single-electron circuit. It positively uses fluctuation in quantum mechanical tunneling processes.

1 Introduction

Vector quantization (VQ) techniques are used in a wide range of applications, including speech and image processing, data compression. VQ techniques encode a data manifold \( V \subseteq \mathbb{R}^D \) using only a finite set of reference vectors \( \mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_N) \). A data vector \( \mathbf{v} \in V \) is represented by the best-matching or "winning" reference vector \( \mathbf{w}_c \), which minimizes the average distortion error:

\[
E = \int \| \mathbf{v} - \mathbf{w}_c \|^2 p(\mathbf{v}) d^D \mathbf{v},
\]

where \( p(\mathbf{v}) \) is the probability distribution of data vectors over manifold \( V \).

Various clustering algorithms to obtain the best reference vectors have been reported. Here, we treat on-line training, in which the data point distribution is not given a priori, but instead a stochastic sequence of incoming sample data points drives the adaptation procedure.

The straightforward approach is the well-known on-line K-means clustering algorithm, in which only the nearest reference vector to the sample vector is adjusted:

\[
\Delta \mathbf{w}_i = \varepsilon \cdot \delta_{ic} \cdot (\mathbf{v}(t) - \mathbf{w}_i),
\]

where
where, $\varepsilon$ is the step size and $\delta_{ij}$ is the Kronecker delta. However, this simple clustering algorithm is often stuck in a local minimum. To avoid this difficulty, a common approach is to introduce a “soft-max” adaptation rule that not only adjusts the “winning” reference vector but affects other reference vectors depending on their proximity to $\nu$.

The maximum-entropy (ME) algorithm [1] adjusts all reference vectors $w_i$ depending on the Euclidean distance to $\nu$:

$$\Delta w_i = \varepsilon \cdot \frac{e^{-\beta(\nu - w_i)^2}}{\sum_{j=1}^{N} e^{-\beta(\nu - w_j)^2}} \cdot (\nu(t) - w_i),$$

where parameter $\beta$ defines the proximity.

The Kohonen’s self-organization map (SOM) algorithm [2] is another well-known model:

$$\Delta w_i = \varepsilon \cdot h_\sigma(i, c) \cdot (\nu(t) - w_i).$$

In this model, every reference vector is assigned to a site of a lattice. Each time a sample vector is presented, not only the “winning” reference vector is adjusted but also the reference vectors assigned to the lattice sites adjacent to the winner are updated according to function $h_\sigma(i, c)$, which is typically chosen to be a Gaussian:

$$h_\sigma(i, c) = e^{-\|\nu - w_i\|^2/(2\sigma^2)},$$

where $\sigma$ is a parameter that defines the proximity.

The neural-gas (NG) clustering algorithm [3] is a powerful soft-max adaptation rule, in which all reference vectors are adjusted depending on the “neighborhood ranking”:

$$\Delta w_i = \varepsilon \cdot h_\lambda(k_i(\nu, w)) \cdot (\nu(t) - w_i),$$

where $k_i(\nu, w)$ is the ranking, which depends on $\nu$ and the whole set $w$. The function $h_\lambda(k_i)$ is typically as follows:

$$h_\lambda(k) = e^{-k/\lambda},$$

where parameter $\lambda$ defines the proximity. This algorithm exhibits faster convergence to smaller distortion errors, however consumes higher computational power especially for sorting. An efficient version of the NG clustering that adjusts only several reference vectors having upper ranking was also proposed [4].

In the next section, we propose a new efficient soft-max adaptation algorithm. It employs the stochastic association model that we have proposed related to single-electron circuits [5], [6]. In Sec. 3, it is demonstrated from simulation results that this new clustering algorithm is as powerful as the other algorithms. In Sec. 4, we propose a nanostructure based on a single-electron circuit for implementing the stochastic association model.

## 2 Stochastic association algorithm

A usual associative memory is defined as a system that deterministically extracts the vector most similar to the input vector from the stored reference vectors. This just corresponds to the process choosing the winning reference vector for a certain data vector in all conventional clustering algorithms.

In our stochastic association (SA) model, the association probability depends on the similarity between the input and the reference vectors. The SA algorithm extracts not only the reference vector most similar to the input but also other similar reference vectors with the probability depending on the similarity.

In the SA algorithm, stochastic fluctuation is added in the evaluation process of distortion error $D_t$ between data vector $\nu$ and reference vector $w_i$. We propose this algorithm inspired