An Unbiased Clustering Algorithm based on Self-organisation Processes in Spiking Networks

Thomas Ott, Markus Christen and Ruedi Stoop
Institute of Neuroinformatics
ETHZ/UNIZH Zurich, Switzerland
e-mail: `tott@ini.phys.ethz.ch`
WWW: `http://www.stoop.net/group`

Abstract—*We show that the combination of short-range excitatory interaction and Hebbian-like learning in integrate-and-fire networks constitutes a robust self-organising clustering process. This process is utilised for designing an unbiased data clustering algorithm for which neither the number of clusters nor their shapes need to be known. The usefulness and the success of this novel algorithm is discussed on the basis of toy-systems and real-world applications.*

I. Introduction

In the last decades, an enormity of clustering algorithms has been developed (e.g, [1]). On the one hand, this reflects the need for analysis and structuring techniques to process the ever growing amount of (mainly digital) data. On the other hand, the interest in clustering acknowledges the important role that the identification of classes plays for perception and cognition. Clustering is an unsupervised classification approach. Whereas supervised methods are trained on the basis of a training set with known classification, clustering methods have to find a reasonable classification on the basis of a given data distribution only. Completely unbiased methods do not even require prior knowledge about the number of clusters, their shape or their structure [2]. Due to the absence of a supervisor, clustering tasks are often associated with a self-organisation process (e.g., self-organised maps SOM [3]). Typically, clusters result from such a process by means of some low-level or local mechanism. This mechanism can be based on an unsupervised learning rule, as in the case of SOM, or on local interaction rules that may lead to clusters of "synchronised" elements [2], [4], [5].

Networks of simple integrate-and-fire neurons (IF, see next section) with short-range excitation, and possibly long-range inhibition, are an example of the latter. Fig. 1b demonstrates how groups of neurons in an IF-network can immediately engage in synchronisation. The synchronised groups correspond to the spatial neuron clusters in the toy system of Fig. 1a. This observation inspires the design of a clustering algorithm. However, the results shown in Fig. 1b are in some sense trivial as only the underlying network connectivity is reflected. Fig. 1c shows that the graph of short-range excitatory connections consists of three large components. The excitation-dominated dynamics quickly leads to synchronised firing within each component, resulting in Fig. 1b. Therefore, in terms of a clustering algorithm, besides reflecting the large connectivity components, the neuronal dynamics does not reveal any more insight into possible structures of the data set (Fig. 1a). The components, however, are the consequence of the special choice of the connectivity parameters; in the given example, an interneuron distance smaller than $r = 0.5$ results in an excitatory connection (strength $J = 1$). For different, arbitrary, data sets, this parameter choice will not necessarily yield any useful results. Hence the basic question is: How to make the best choice of connectivity parameters for a given data set?

In this contribution, we show that by combining the dynamics with a local Hebbian-like learning rule, IF networks are able to evolve their structure in a self-organised way reflecting a useful choice of connectivity parameters. Hence the combination of local excitatory interaction with unsupervised learning constitutes a robust network self-organisation process that can be used as a basis for unbiased clustering of arbitrary data sets.

II. Clustering with IF neurons

A. IF equations

The neuron elements we consider are simple leaky integrate-and-fire neurons whose membrane potentials u_i are described by

$$C\frac{du_i}{dt} = I_i(t) - \frac{u_i}{R}. \qquad (1)$$

If $u_i(t)$ reaches the threshold θ, a spike is emitted and u_i is reset to 0. The input current $I_i(t) = I_{ext} + I_{int}$ consists of a (time-independent) external current I_{ext}

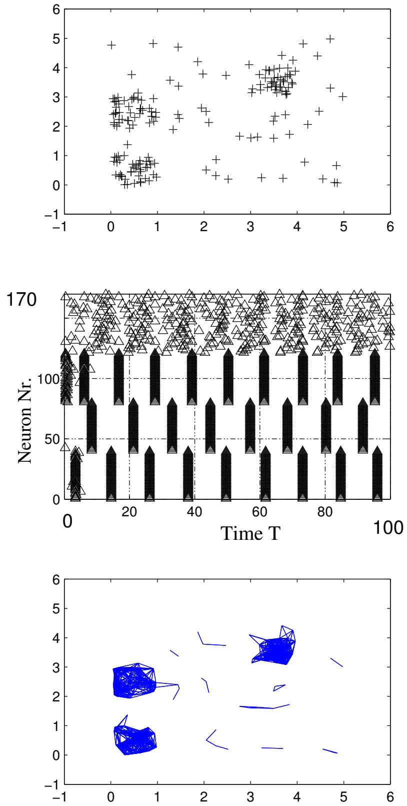

Fig. 1. a) A 2-dimensional distribution of $N = 170$ points representing IF neurons. b) The spike activity diagram reveals three clusters of synchronised neurons and a group of non-synchronised neurons (background points). c) The connectivity graph of excitatory connections shows three large connectivity components.

and the contributions of the presynaptic neurons

$$I_{int} = \sum_j J_{ji} \sum_k \delta(t - t_j^k), \quad (2)$$

where t_j^k is the emission time of the kth spike at neuron j and J_{ji} is the synaptic efficacy.

The simplicity of the used IF model allows for an exact integration.If the potentials at time t are given as $(u_1, u_2, ..., u_N)$, the time interval to the next spike is given by $T_k = \min_j\{T_j\}$, where

$$T_j = RC \, \ln\left(\frac{u_j(t) - I_{ext}R}{\theta - I_{ext}R}\right). \quad (3)$$

The resulting potentials for $j \neq k$ at time $t + T_k$ are $u_j = I_{ext}R(1 - \exp(\frac{-T_k}{RC})) + u_j(t)\exp(\frac{-T_k}{RC}) + J_{ij}R$. As there is no transmission delay in this model, an outgoing spike can provoke a whole bunch of simultaneous spikes at further sites. This gives the possibility of quick synchronisation among tightly coupled neurons, as observed in Fig 1b. One neuron, however, can only fire once at a given time.

The internal parameters were chosen according to $RC = 8ms$ and $\theta = 16mV$. The external current was chosen as $I_{ext}R = 25mV$.

B. Clustering

For the purpose of clustering, we translate a given data set into an IF network, i.e., each data point is represented by an IF neuron. We define data clusters by groups whose neurons are strongly coupled (mutually maximal connection strength) and show synchronised firing. The network's task is to develop its structure, starting with relatively weak connections and ending with strong connectivity components that reflect the natural data clusters. The initial network connectivity is given by weak (symmetric) excitatory connections between each neuron and its k nearest neighbours. The initial connection strengths $J_{ij} = J_{ji}$ are given by a decreasing function of the data point distance d_{ij}, reflecting the data structure, i.e.,

$$J_{ij} = J_{ji} = \exp\left(-\frac{d_{ij}^2}{a^2}\right). \quad (4)$$

The parameter a shall be related to the average distance between connected sites $\bar{d}$. E.g., for the choice $k = 10$, we take $\bar{d}/4$. The exact choice of a, in fact, is of considerable importance as it strongly influences the global dynamics. For very large a, all synaptic strengths are close to one and global synchronisation occurs quickly (one connectivity component). For smaller a, such as $\bar{d}/4$, the synaptic connections are weak and synchronisation can only be established very slowly (Fig. 2a).

However, even for small a the potentials u_i and u_j of two connected neurons tend to be more correlated if the connection J_{ij} is stronger. This can be used to "breed" clusters by means of Hebbian-like learning. In this paradigm, a synapse is strengthened if the firing at the presynaptic and the postsynaptic neuron is more or less coincident. As the synapses are symmetric, we define the learning rule in the following way

$$J_{ij} = J_{ji} \to \min\{2J_{ij}, 1\}, \quad \text{if } G_{ij}^{\tau}(t) = 1, \quad (5)$$

where $G_{ij}^{\tau}(t) = 1$ only if both of the two neurons i and j have emitted a spike within $[t - \tau, t]$. τ defines the width of the learning window and is chosen in relation to the interspike interval of independently spiking neurons $T = RC\ln(I_{ext}R/(I_{ext}R - \theta)) = 8.17\ ms$.

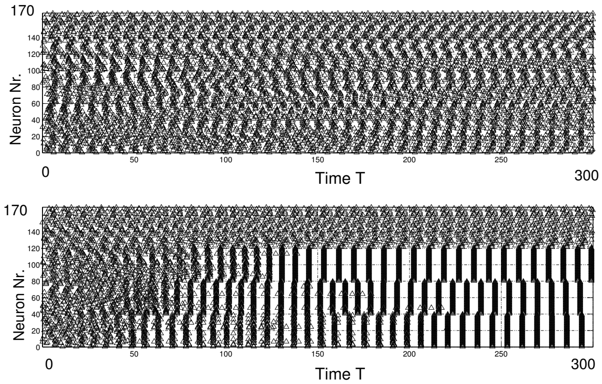

Fig. 2. Spike diagram a) without learning and b) with learning.

If τ is too small, learning is slow. For computational reasons, τ should be chosen as large as possible. However, if τ is too large, learning is too fast quickly leading to global synchronisation. For the given parameters, we chose $\tau = T/4 = 2\,ms$.

The effect of Hebbian learning for the toy system can be observed in Fig. 2b. The neurons of a cluster are completely synchronised after about $100T \approx 200\,ms$. This is in stark contrast to the situation without learning (Fig. 2a). In Fig. 3, the evolution of the corresponding grid structure is illustrated. While at the very beginning only a few strong connections are present (the seed of the clusters), after 200 ms the network has developed three large connectivity components, reflecting the three data clusters (Fig 1a).

C. Hierarchical clustering and inhibition

In a purely excitatory network (as used so far), the learning rule (5) can always lead to global synchronisation. This is because all synaptic connections will potentially be strengthened after some time, albeit on very different time scales, so that finally all weights become strong enough to support global synchronisation. E.g. in our toy system, for $\tau = 5\,ms$ global synchronisation is practically reached after about 900 ms.

The different learning speed results in a merging process - successively smaller clusters merge to larger units. This resembles a hierarchical clustering mechanism where time plays the role of the resolution parameter. In comparison to simpler hierarchical methods, such as single linkage clustering (SLC) [1], the obtained clustering solutions are more robust. Being the consequence of local neural cooperation, clustering is not only based on information provided by the single weights (as in the case of SLC). Particularly clear cluster structures, such as the three clusters in Fig. 1a, are easy to detect as they persist unmodified over a long time period. If desired, they can be stabilised by establishing inhibitory connections between them (the establishing of the inhibitory connections can only be described by a semi-local learning rule.). Such long-range inhibitory connections not only make global synchronisation impossible and preserve the stable clusters, they also lead to a phase shift between the different synchronised clusters. Depending on the data analysis problem, hierarchical clustering or clustering with inhibition can be preferred.

D. Applications

(I) In the visual scene analysis toy example given by Fig. 4, the task is to segregate the “natural “ objects (relatively homogeneous regions) from the inhomogeneous background. The given grey-scaled image is described by an intensity-encoding matrix $f(x, y)$, where x and y are the pixel coordinates and $f(x, y) \in [0, 1]$ is the pixel intensity. Due to the background inhomogeneity, the problem cannot necessarily be solved by simple thresholding techniques as often used for segregation problems [6]. For a clustering solution, the pixel neighbourhood grid is translated into the grid structure of an IF network. I.e., each neuron is connected to its direct neighbours (thus $k = 4$).

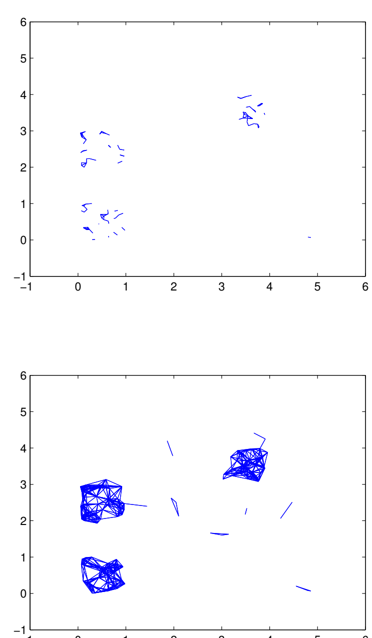

Fig. 3. Structure of strong connections with $J_{ij} > 0.2$ at time a) $t = 0\,ms$ and b) $t = 300\,ms$.

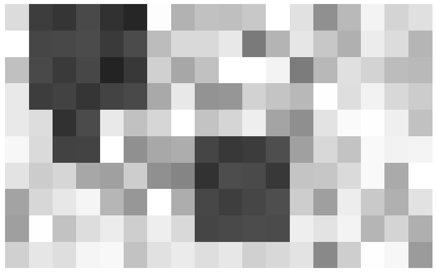

Fig. 4. Visual scene toy example with $N = 180$ pixels. "Natural" objects are identified with relatively homogeneous regions.

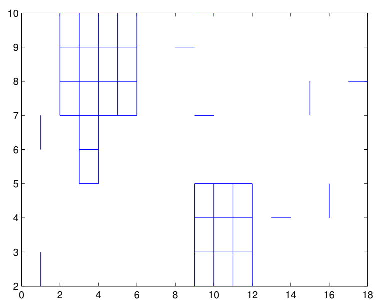

Fig. 5. Result for Fig. 4: The structure of strong connections ($J_{ij} = 1$) after learning ($t = 300\,ms$) reflects the "natural" objects. The connection ends correspond to the pixel centers.

The initial connection strengths J_{ij} are given by (5), where $d_{ij} = |f(i) - f(j)|$ for neighbouring pixels i and j. In this case, a is $\bar{d}/20$.

In principle, the problem could be tackled by means of graph theory-based methods. In this approach, the image is described as a weighted graph (e.g., as given by the connection grid with J_{ij}) and all connections below a certain threshold strength are cut. The main difficulty of this approach, however, is to find a reasonable criterion for the threshold. In contrast, our IF network approach does not require to fine-tune any threshold parameter. The network shows a very robust evolution towards a solution that reflects the most natural objects (Fig. 5). The obtained graph of strong connections consists of two large components corresponding to the two dark rectangular structures that are easily recognised in Fig. 4.

(II) In a further test, we applied our algorithm (with $k = 10$ and $a = \bar{d}/4$, no inhibition) to the Iris data set which is a famous benchmark for clustering algorithms [1]. This data set consists of three classes with 50 elements each. We found that our algorithm can perfectly identify one class (Iris-setosa). The remaining two classes are correctly recognised up to about 10-15 misclassifications (early in the time hierarchy, the number of detected clusters is larger than three, but the clusters are pure, i.e, they contain only elements of the same class. Later, the number of clusters is three -as expected- but one cluster is not pure any more). The classification thus is not perfect, nevertheless, the performance is comparable to that of standard algorithms [1] that, however, typically rely on the specification of prior parameters.

III. Summary and Conclusion

We presented a novel clustering method based on self-organisation processes of integrate-and-fire networks. By combining the neural firing dynamics with a Hebbian-like learning rule, the network is able to evolve its structure in such a way that it represents the "natural" clusters within a given data set. The important advantage of our method is that it is completely unbiased, i.e., it does not require the specification of any parameters (number of clusters, cluster shape, cluster size etc.). Rather, the self-organisation process automatically leads to a suitable parameter choice. In this respect, the method is superior to most standard algorithms as it also shows a similar performance for typical benchmark problems such as the Iris data set. However, there is still some room for further improvements by optimising the initial parameters or by modifying and optimising the learning rules for both excitatory and inhibitory connections. For very large data sets, the algorithm's applicability might be limited due to its limited speed performance. Yet from a conceptual standpoint, the algorithm seems an interesting example of how structural self-organisation could function in the brain - although our method is not primarily based on biological plausibility.

Acknowledgements: *This research has been supported by a ZNZ (neuroscience center Zurich) grant.*

References

[1] R. Xu, D. Wunsch. Survey of clustering algorithms. In *IEEE Trans. Neural Netw.* 16, pp. 645–678, 2005.

[2] T. Ott, A. Kern, W.-H. Steeb, R. Stoop. Sequential clustering: tracking down the most natural clusters. In *J. Stat. Mech.* P11014, doi:10.1088/1742-5468/2005/11/P11014, 2005.

[3] T. Kohonen. Self-organizing maps. Springer-Verlag, New York, 1997.

[4] M. Blatt, S. Wiseman, E. Domany. Superparamagnetic clustering of data. In *Phys. Rev. Lett.* 76, pp. 3251–3254, 1996.

[5] L. Angelini, F. De Carlo, C. Marangi, M. Pellicoro, S. Stramaglia. Clustering data by inhomogeneous chaotic map lattices. In *Phys. Rev. Lett.* 85, pp. 554–557, 2000.

[6] N. R. Pal, S. K. Pal. A review on image segmentation techniques. In *Pattern Recognition* 26, pp. 1277–1294, 1993.