FUZZY INHIBITED NETWORK

*M. João Rendas**

I3S, CNRS URA 1376, Université de Nice Sophia Antipolis
250 Rue Albert Einstein, Les Lucioles 1, 06560 Valbonne, France

rendas@mimosa.unice.fr

Résumé

Dans cet article, on présente FIN (Fuzzy Inhibited Network), une structure connexioniste nouvelle specialement conçue pour réaliser un partitionnement dynamique non-supervisé. Le réseau résoud le problème de partitionnement en employant des mécanismes d'inhibition (compétition) pour séparer les différentes classes, et un comportement collectif (coopératif) pour les représenter. On applique FIN au problème de poursuite multi-cible, et on présente des simulations qui permettent d'évaluer sa performance.

Abstract

In this paper we present FIN (Fuzzy Inhibited Network), a novel connectionist structure specially designed to perform adaptive unsupervised clustering. The network solves the clustering problem using fuzzy competitive (inhibitive) mechanisms to separate distinct classes, and cooperative (collective) behavior to represent each one. We apply FIN to the multiple target tracking problem, presenting simulations that demonstrate its performance.

1. INTRODUCTION

Since the definition of fuzzy sets in the 60's, several fuzzy clustering techniques have been proposed in the literature [6, 3, 4]. More recently, some authors have combined this approach to clustering with the adaptive and distributed characteristics of neural networks defining fuzzy/neural clustering architectures [7, 5, 2]. Compared to existing fuzzy/neural clustering architectures, the structure proposed in this paper, FIN, presents a set of features that collectively make it more convenient in situations where classes present gradually changing membership functions of arbitrary shapes with considerable degree of overlapping.

To represent each class, FIN does not rely in the definition of a single prototype, as the Fuzzy ART [2] or Adaptive Fuzzy Leader Clustering [5] networks; instead, each class is represented by a set of points that *collectively* describe its variability, allowing in this way membership functions of arbitrary shapes, and increasing robustness against outliers.

The Fuzzy Min-Max network [7] also develops a distributed fuzzy representation of the classes membership function; however, it is based on the definition of hard boundaries for each class, and thus is not adequate for situations where classes overlap.

FIN is based on an inhibition mechanism, such as [2] and the Fuzzy Kohonen network of Bezdeck *et al*, but contrary to these architectures, it's inhibition structure has itself a fuzzy interpretation, and is determined by the geometry of the membership functions, which allows a more appropriate behavior under class overlapping. Its name reflects this distinctive feature.

In this paper we apply FIN to the problem of multiple target tracking (RADAR). In this case, the limited resolution of the observer may not be able to separate close objects (targets) producing a single observation that is related to both targets. We show that FIN allows the update of elements assigned to more than one class (target) by the same piece of data, maintaining in this way track continuity during target crossings. Since hard decisions are never made inside FIN, outliers only slightly and temporaryly deform the membership function of some classes by increasing its value in some neighborhood. The fuzzy distributed representation that is maintained for each class absorbs in this way sporadic erroneous information (outliers).

---

* On leave from ISR, Instituto Superior Técnico, Lisboa, Portugal.

The paper is organized as follows. In Section 2 we present our fuzzy approach to unsupervised clustering. In Section 3 we describe the network structure and updating procedure. Finally, this structure is applied, in Section 4, to the multiple target tracking problem. Results showing the behavior of the network for realistic radar scenarios are shown, under situations of strong noise, high probability of false alarm, and low probability of detection, that demonstrate the capability of FIN to maintain track continuity under target crossings in adverse conditions.

2. FUZZY UNSUPERVISED CLUSTERING

The clustering problem is to define, from a set of data points, a partition $\mathcal{P}$ that agregates subsets of these points as having a set of common characteristics, i.e. as belonging to the same class. In problems of dynamical clustering like the one we address here, the temporal evolution of the features of each class may itself be an important factor in partitioning the data. We assume, consequently, that the input space $\mathcal{Y}$ (where the input data y takes values) and the clustering space $\mathcal{W}$ (where the membership functions are defined) are not necessaryly the same. This means that we will be searching not only for an *explanation* of the isolated data points themselves, but rather for dynamical models for the observed (or infered) set of trajectories.

We assume (which is equivalent to the definition of a dynamical representation) that there is a known one-to-one map $\Phi(\cdot)$ from $\mathcal{W}$ to $\mathcal{Y}$:

$$\begin{array}{rccc} \Phi: & \mathcal{W} & \longrightarrow & \mathcal{Y} \\ & w & \longrightarrow & y = \Phi(w) \end{array}$$

We also admit that a known injective restriction of Φ is known, and denote it by Φ_o^{-1}:

$$\begin{array}{rccc} \Phi_o^{-1}: & \mathcal{Y} & \longrightarrow & \mathcal{W}_o \\ & y & \longrightarrow & w = \Phi_o^{-1}(y) \end{array}$$

In fuzzy approaches to clustering problems, the solution to the clustering problem, i.e., the partition mentioned above, is equivalent to the definition of a set of *fuzzy* sets,

$$\mathcal{P} = \left\{ \underset{\sim}{A}_1, \underset{\sim}{A}_2, \ldots, \underset{\sim}{A}_K \right\}$$

each one being defined by a membership function

$$\begin{array}{rccc} \mu_{\underset{\sim}{A}_i}: & \mathcal{W} & \longrightarrow & [0,1] \\ & w & \longrightarrow & \mu_{\underset{\sim}{A}_i}(w) \end{array}$$

The set of membership functions $\{\mu_i\}_{i=1}^{K}$ [1] defines a vector field μ over $\mathcal{W}$ that is the fuzzy solution of the clustering problem.

Based on this vector field, we define two other quantities relevant for our solution of the adaptive clustering problem: the *network evidence*, which is a vector field defined over the input space $\mathcal{Y}$, and a *compatibility relation* between the elements of the clustering space $\mathcal{W}$.

The *network evidence* is an extrapolation of the membership vector μ into the input space. For each set $\underset{\sim}{A}$ (each component of the membership field), we define the network evidence for association of point y in $\mathcal{W}$ to that set as the logical value of the proposition "y is close to a good prototype of the set A":

$$Evid_A(y) = \max_{w \in \mathcal{W}} \min\left[d(y, \Phi(w)), \mu_A(w)\right]$$

where $0 \leq d(\cdot,\cdot) \leq 1$ is a fuzzy normalized distance in $\mathcal{Y}$ (linguistic variable "close to"), which precise definition depends on the application.

Besides the network evidence, we define a *compatibility relation* between the elements of the clustering space, $R(w_1, w_2)$ that measures to which degree points w_1 and w_2 are members of the same fuzzy set(s). We define this relation as the truth value of the proposition "for all sets, if w_1 is an element of A then w_2 is also a member of A," which results in the following mathematical expression in terms of the membership field μ:

$$\mu_R(w_1, w_2) = 1 - \max_A \max(0, \mu_A(w_1) - \mu_A(w_2))$$

From this definition, we see that $\mu_R(w_1, w_2) \simeq 1$ if and only if w_2 is a good representation of all the sets for which w_1 is a good prototype. We note that this compatibility relation is not symmetric, meaning that a point w_2 may be compatible with another point w_1 (if it is has high membership for all sets for which w_1 has) and at the same time w_1 not be compatible with w_2 (meaning that for at least one set A^* of which w_2 is a member, $\mu_{A^*}(w_1) \simeq 0$).

These two quantities, both dependent on the vector field μ, form the basis for our clustering procedure as will be explained next.

The goal of the dynamical clustering procedure is to gradually increase the membership field in the regions of the clustering space that are systematically observed. Assume that such a membership field, of dimension K, has been defined. For each data point y presented at time t, the network evidence expresses the possibility of assigning y to each fuzzy partition $\underset{\sim}{A}_i$. Consider now each data point y. If $Evid_{\underset{\sim}{A}}(y)$ is very small for all sets, then the dimension of the membership field is increased, by adding a new element $\underset{\sim}{A}_{K+1}$,

[1] In the following, we will use the short representation μ_i for the membership function defining the fuzzy set $\underset{\sim}{A}_i$

corresponding to the creation of a new fuzzy set. On the contrary, if for at least one set $\underset{\sim}{A}_*$ this evidence is high, then there will be neighborhoods $W_i(y) \subset \mathcal{W}$ such that $\forall w \in W_i(y), \mu(w)$ is high. In this case, y is associated to $\underset{\sim}{A}_*$ by letting $\mu_*(w)$ increase in the neighborhoods $\tilde{W_i}$.

3. FIN ARCHITECTURE

FIN consists of a single layer of identical inter-connected (by an inhibition matrix C) elements, with the input data being distributed to all nodes.

Each node n of the network is associated to a point in clustering space, $w_n \in \mathcal{W}$. In this way, the network defines a dynamic grid over the clustering space $\mathcal{W}$.

The strength of the inhibitive connection between two nodes depends on their compatibility, as measured by the relation R defined in the previous section. Let n and m be two generic nodes. Then the value of the inhibitive connection from node m to node n is given by

$$\begin{aligned} C_{n,m} &= -\mu_{\overline{R}}(w_n, w_m) \\ &= -\max_A \max(0, \mu_A(w_1) - \mu_A(w_2)) \end{aligned}$$

We can thus see that $C_{n,m}$ is always negative, with small absolute values if n and m are members of the same classes – meaning that they can become active for the same input– while nodes that belong to distinct classes strongly inhibit each other.

For each new data vector, y, the network does one of two distinct functions: *resizing of the network*: corresponding to the *detection* of a new class, not compatible with existing ones; *updating of existing nodes*, where a subset of network nodes is updated in response to the observed data.

Resizing of the network is done whenever the current input vector y cannot be assigned to one of the existing classes with sufficient confidence, i.e., whenever $Evid_A(y)$ is small for all A. In this case a new class is created, by adding a cluster of nodes to the network in a neighborhood of $\Phi_0^{-1}(y)$, new elements being taken as representatives of the new class.

Update of existing nodes corresponds to normal operation of the network, and is described by a set of differential equations, that implement the cooperative/-inhibitive mechanisms previously mentioned, meaning that only a subset of the network node will actually respond to the input data.

At each node, an auxiliar short term variable is defined, a_n the activation level of the node. At each presentation, this variables are set to a small negative value (constant accross the network). From this initial assignement, its evolution is described by the following equation:

$$\begin{aligned} \frac{\partial a_n}{\partial t} &= -a_n + g(a_n) + E_n + O_n \\ E_n &= min\,(d(y, \Phi(\mathbf{w}_n)), max_i \mu_o(w_n)) \geq 0 \\ O_n &= min_m\,(C_{n,m} a_m) \leq 0 \end{aligned}$$

where the term E_n measures the driving force of the input plot for node n, and O_n measures the inhibition of other nodes.

When a node becomes active ($a_n \geq 0$), its state is allowed to change, updating its position in $\mathcal{W}$, as well as its membership values.

Membership updating is a balance between the general decaying of membership values whenever nodes do not become active (meaning that the classes they represented are not present in the input data) and reinforcement of membership for the classes of largest evidence when a node becomes active.

All network variables are updated during this phase, including the values of the connection matrix, as a consequence of the variation of the membership field. We are able to maintain inhibitive weights that vary gradually in the interval $[0, 1]$ and are consistent with the instantaneous value of the membership field μ.

FIN can be considered as a generalization of other inhibitive networks, for which the connection matrix C is fixed. In FIN, the inhibition weights are a function of the classes to which the nodes belong, and are determined by the past trajectories of the corresponding nodes, i.e., the connection matrix is also part of the memory of the network.

As a result of network evolution for each input, two *outputs* are produced: the *identity* (class) of the data, obtained by finding the class that shows a larger network evidence ofr the current input, and an estimate of the *best prototype* of that class, corresponding to the position of the best representative of that class.

4. MULTIPLE TARGET TRACKING

For the tracking problem, $\mathcal{Y}$ is usually 2 or 3 dimensional, representing the position of the target, while $\mathcal{W}$ has increased dimension, including components for velocity and, possibly, acceleration. In the simulation presented below, we considered a simple linear uniform model for target motion, and observations in the horizontal plane, meaning that the clustering space is four dimensional, with two components (w_x, w_y) for position, and two components $(\dot{w}_x, \dot{w}_y)$ for velocity.

The fuzzy measure of compatibility between the observed plots and the network nodes is based on the Euclidean distance between the plot y and the predicted position at the observation time

$$\Phi(\mathbf{w}) = \begin{bmatrix} w_x + \Delta_t \dot{w}_x \\ w_y + \Delta_t \dot{w}_y \end{bmatrix}.$$

where Δ_t is the interval of time since last update of node n. For node n, $d(y, \Phi(\mathbf{w}_n))$ is, according to the assumed motion model, given by:

$$d(y, \mathbf{w}_n) = \begin{cases} 1, & ||y - \Phi(\mathbf{w})|| \leq \gamma \\ \left(\frac{\gamma}{||y-\Phi(\mathbf{w})||}\right)^{\theta}, & ||y - \Phi(\mathbf{w})|| > \gamma \end{cases}$$

where γ and θ are network parameters. State updating is done using the estimate of simple $\alpha - \beta$ filters [1].

To illustrate the performance of FIN, we applied it to simulated data. The case of two closed targets performing two crossings and with trajectories that are close to each other during a considerable period has been considered. The observations were simulated directly in range and bearing, and conversion to rectangular coordinates was performed prior to the presentation of the plots to the network.

Figure 1 illustrates FIN's ability to track the two closed targets, maintaining a good representation under crossings.

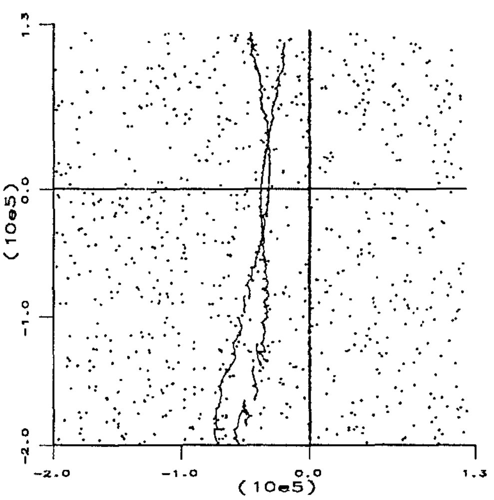

Figure 1: Detected targets.

We see that no false tracks have been formed, the network having successfully absorbed all false alarms. While the track initially (at the bottom of the figure) at the left is maintained during most of the trajectory), a number of distinct hypothesis for the other less well defined track have been propagated by the network. Most often, these tracks have been recognised as alternative representations of the same target, which is evidenced by track merging. Notice that track individuality is maintained even during the long period where the two trajectories are close to each other, and in particular during the two crossings.

Acknowledgements

We acknowledge the collaboration of Dr. V. Schmidlin, which provided the data used for the simulation presented.

5. REFERENCES

[1] Yaakov Bar-Shalom and Thomas E. Fortmann. *Tracking and Data Association.* Academic Press, 1988.

[2] G. Carpenter, S. Grossberg, and D. Rosen. Fuzzy art: An adaptive resonance algorithm for rapid, stable classification of analog patterns. *Proc. 1991 Int. Joint Conf. Neural Networks*, Vol II:pp: 441–416, (Seattle) 1991.

[3] R. N. Dave. Fuzzy shell clustering and application to circle detection in digital images. *Int. J. General Systems*, Vol. 16, pp 343–355 1990.

[4] R. Krishnapuram and J. Keller. A possibilistic approach to clustering. *IEEE Trans. on Fuzzy Systems*, Vol. 1, No. 2, May 1993.

[5] S. Newton, S. Pemmaraju, and S. Mitra. Adaptive fuzyy leaer clustering of complex data sets in pattern recognition. *IEEE Trans. on Neural Networks*, Vol. 3, September 1992.

[6] E. Ruspini. A new approach to clustering. *Information and Control*, Vol. 15, 1969.

[7] Patrick Simpson. Fuzzy Min-Max neural networks – part 1: Classification. *IEEE Trans. on Neural Networks*, Vol. 3, September 1992.