A RECURRENT RBF NETWORK MODEL FOR NEAREST NEIGHBOR CLASSIFICATION

M. K. Muezzinoglu    J. M. Zurada

Electrical and Computer Engineering Department
University of Louisville

International Joint Conference on Neural Networks
Montreal, 2005
Outline

1 Motivation
   - Nearest Neighbor Classification
   - Previous Work

2 Design Procedure
   - Gradient RBF Network
   - Network Model

3 Experiments

4 Conclusions
Given a finite $M \subset \mathbb{R}^n$, the set of prototype patterns, and $x \in \mathbb{R}^n$, calculate

$$f(x) = \arg \min_{y \in M} d(x, y)$$

- Straightforward but intractable on digital computers.
- Requires $|M|$ comparisons
  - Sorting the distances
  - Revealing the nearest prototype
- Design of NN classifiers on recurrences is a key problem.
- There are many results and methods to make specific recurrent models approximate $f(\cdot)$. 
Brain-State-in-a-Box
Hopfield
M-Model
All maximize a quadratic energy landscape.
Independent of prototypes.
Proposal

- A gradient architecture adaptive to prototypes.
- A modular structure maximizing sum of RBFs.
- Grows (linearly) with the number of prototypes.
- Operates in continuous state-space.
- Converges to \( f(x) \) from the initial condition \( x \) along autonomous dynamics without comparisons.
- Allows real-valued prototypes.
Outline

1. Motivation
   - Nearest Neighbor Classification
   - Previous Work

2. Design Procedure
   - Gradient RBF Network
   - Network Model

3. Experiments

4. Conclusions
Energy Landscape
Sum of RBFs

- Assign a center to each element of $M$.
- Assume Gaussian RBFs.
- Construct the functional

$$E(x) = \exp\left(-\gamma_1 \|x - c_1\|^2\right) + \ldots + \exp\left(-\gamma_1 \|x - c_m\|^2\right).$$
What is the gradient system that maximizes $E(\cdot)$?
To preserve local maxima: $\gamma_1, \gamma_2 < \frac{2}{\|c^i - c^j\|^2}$
Network Dynamics

A continuous dynamics to maximize $E(\cdot)$

$$
\dot{x} = \nabla_x E(x)
$$

$$
= -2 \sum_{i=1}^{m} \gamma_i \left( x - c^i \right) \exp \left( -\gamma^i \| x - c^i \|^2 \right)
$$

$$
= -x \cdot 2 \sum_{i=1}^{m} \gamma^i \exp \left( -\gamma^i \| x - c^i \|^2 \right)
$$

$$
+ 2 \sum_{i=1}^{m} \gamma^i c^i \exp \left( -\gamma^i \| x - c^i \|^2 \right).
$$
Multiplication Sub-Blocks

For $a \in [0, 1]$ and $b \in [0, m]$, this sub-block gives $a \cdot b$ with error in the order of $10^{-7} m$. 

\[
\begin{array}{c}
\sum \tanh(\cdot) \\
\sum \tanh(\cdot) \\
\sum \tanh(\cdot) \\
\sum \tanh(\cdot)
\end{array}
\]

\[
\begin{array}{c|c|c}
 v_{11} & 0.02418 & \theta_2 \\
v_{21} & 16.34820 & \theta_3 \\
v_{31} & 0.12108 & \theta_4 \\
v_{41} & 0.09868 & w_1 \\
v_{12} & -0.10386/m & w_2 \\
v_{22} & -58.58570/m & w_3 \\
v_{32} & 0.12528/m & w_4 \\
v_{42} & -0.02288/m & \rho \\
\theta_1 & 0.68984 &
\end{array}
\]
quick brown fox jumps over the lazy dog
0123456789

quick brown fox jumps over the lazy dog
QUICK BROWN FOX JUMPS OVER THE LAZY DOG
0123456789
Grayscale Patterns
Summary

- Restriction on the parameters: $\gamma^i, \gamma^j < \frac{2}{\|c^i - c^j\|^2}$
- No restriction on $M$.
- Any initial condition $x$ is almost guaranteed to be mapped to the nearest prototype. $\Rightarrow$ Almost perfect NN classification
- Singular points at midpoints of centers are non-problematic in practice.
Thank you!

Questions/comments?