SOMs AS A REPRESENTATION INFRASTRUCTURE FOR SIGNAL PROCESSING AND CONTROLS

Jose C. Principe, Ph.D.
Distinguished Professor of Electrical and Computer Engineering
University of Florida, USA
principe@cnel.ufl.edu

The self-organizing map (SOM) is well known in many statistical applications ranging from clustering to data mining. However, the SOM is also an extremely useful methodology in time series modeling and system's control. Because these application domains are perhaps less well known, this talk will summarize SOM based designs being pursued in the Computational NeuroEngineering Laboratory for both signal processing applications and controls.

The first question to be addressed is how to cope with time, since the original SOM is static, i.e. it adapts the weights with the present data sample. The problem of time varying inputs can be addressed using the delay embedding theorem [1], which states that it is always possible to create a one-to-one mapping with a unique inverse between the original time series and a multidimensional space of sufficiently large dimension ($d > 2L+1$, where L is the dimension of the original deterministic system that produced the time series). Even in the case of noise contamination, this is approximately correct if filtered embeddings are utilized. With this result, it is possible to self-organize SOMs in signal (or state) space. Once this is done from a training set, time series modeling and system controls can be tremendously simplified from the trained SOM, due to its clustering properties. For this reason, we should think of the SOM as a representation infrastructure for signal processing and control.

One of the fundamental properties of the SOM that has been exploited in our work is the neighborhood preservation property. Indeed neighborhood preservation allows local linear modeling, which is the key modeling infrastructure that we have exploited [2]. The idea is very simple. Complex nonlinear mappings, if smooth, can always be locally approximated by linear models in a sufficiently small region around the present operating point. The practical difficulties that the designer must solve are (1) how to select the neighborhood that is relevant, (2) how to derive the parameters of the linear models, (3) how to avoid discontinuities between the models. The SOM enables very elegant answers to all these 3 questions, as we will present during the talk. Indeed, the winner take all operation of the SOM provides immediately the location in state space that is relevant. If one associates a linear model to each of the SOM processing elements (PEs), the selection of the linear model is also done easily. As for the training, the *weights* of the SOM PE and its neighbors can be thought as an approximation to the data in the neighborhood, and by simple weighted least squares with the desired response we can train the optimal weights. Finally, since the SOM preserves neighborhoods, the local linear model coefficients vary smoothly, and so switching transients are avoided.

Deriving controllers from local linear models is also rather simple, in particular with winner take all PEs [3]. Indeed, under the inverse model control framework, from each local linear model, a corresponding local linear controller is easily derived by inversion, for a high performance, robust, and easy to implement model based control methodology. We will be presenting results from the control of a NASA UAV (unmanned aerial vehicle) we are working on, comparing the SOM results with alternative methodologies. One of the difficulties of the SOM based approach is the poor scaling up of the technique to large input spaces.

Another application of the SOM that will be discussed is an infrastructure for the estimation of distances in probability spaces. Indeed, clustering is a form of density estimation, and as such it may contain the building blocks to estimate distances between probability density functions. Once again, the idea is to train the SOM for the application and work in the SOM space as a low dimensional projection of the input data space, which preserves neighborhoods. The Kulback-Leibler divergence is perhaps the best well known “distance” measure in probability spaces, but it is very hard to estimate nonparametrically (i.e. directly from data). An alternative distance proposed by Diks [4] and independently also by us [5] is able to approximate the KL divergence. Recently, we have utilized a SOM to perform this estimation with very interesting results. Basically, the method trains a SOM and builds an histogram of the fired PEs. The comparison of the SOM’s histogram during training with the histogram of the same SOM applied to unknown data is able to estimate very accurately the differences between the data distributions. In a sense this method extends the concept of the matched filter to much larger signal segments, and can be considered a nonparametric similarity index in probability spaces.

References

[1] Takens, F., “Detecting strange attractors in turbulence,” in Dynamical Systems and turbulence, ed. by D.A. Rand and L.-S. Young, Springer Lecture Notes in Mathematics, 898, pp. 365-381, Springer-Verlag, New York, 1980.

[2] J.C. Principe, L. Wang and M.A. Motter, “Local Dynamic Modeling with Self-Organizing Maps and Applications to Nonlinear System Identification and Control,” *Proc. of IEEE*, 86(11):pp.2240-2258, 1998.

[3] Cho J., Principe P., Erdogmus D. and Motter M., "Modeling and Inverse Controller Design for an Unmanned Aerial Vehicle Based on the Self-Organizing Map," accepted in IEEE Transactions on Neural Networks.

[4] C. Diks, W.R. van Zwet, F. Takens and J. DeGoede, Phys. Rev. E 53,2169, 1996.

[5] Principe, J., Xu D., Fisher J., Information Theoretic Learning, in Unsupervised
Adaptive Filtering, Simon Haykin Editor, 265-319, Wiley, 2000.