Category and Session number: NL3

Preferred Mode of Presentation: Oral

Probabilistic Markovian Models for Fractal Data Sequences

RITA SINGH

School of Computer Science, Carnegie Mellon University, USA

Geophysical phenomena are usually manifestations of myriad processes that interact in complex ways over time. It is only very rarely, if at all, that any of these interacting processes are linear in nature. The interaction of several non-linear processes generally leads to chaotic behavior and it is for this reason that most geophysical measurements display fractal characteristics. Although these are not the perfect and elegant structures seen in computer simulated images generated from fractal geometry or images of strange attractors, they are nevertheless complex enough that they can be closely characterized by fractal geometric measurements. These measurements may then be used to discover patterns, identify component processes, classify, compress, denoise or analyze in other ways. These approaches ubiquitously assume that the data are fractal (or approximately so) and attempt to measure fractal-geometric parameters or, attempt to infer the dynamics of the underlying chaotic process.

This paper offers a different approach. Rather than derive fractal geometric measurements from the data directly, it proposes a probabilistic statistical model whose structure is fractal, can be pre-specified, and whose parameters can be automatically learned from the observed data. Inferences about the data can be made using this model. The model is a hidden Markov model (HMM), where each state in the model is recursively a transform of the model, or the model itself. The fractal structure imposes constraints on the state output distributions of the HMM. The structure of the HMM, and the transforms in the recursion of its distributions are characteristic of the kinds of fractals it can model. The paper outlines the constraints imposed on learning algorithms that can be used to learn the parameters of such an HMM from real data. Once the HMM parameters are learned, the model can be used to generate synthetic data with the appropriate statistical characteristics. Although this is a statistical fractal characterization of the data, rather than a deterministic one, it does not contradict the nature of the underlying process, which is governed by deterministic chaos. This is because the approach avoids making the assumption that the data themselves are fractal, and merely attempts to best describe the data with a fractal model. The paper examines the potential constraints on, and uses of, such synthetic data in analyzing geophysical signals.