CRASH COURSE · 5 LECTURES · SCIENCE & ENGINEERS

How Much Does a Measurement Know?

Crash course in Information Theory for science and engineering students

Every point in a measurement is a unit of information about the sample, every stream of data is a message. Information theory tells us how much each one carries, how much an instrument can deliver at best, and how close our reconstructions come to that limit. This crash course gives a working command of the theory and algorithms with emphasis in applications. As much as possible preference will be given to applications in optics: information theory methods of analysis and data reduction. Hands on sesion will focus in the use of a practical software suite *lzcomplexity* for analysing real data.

1 Surprise, Uncertainty and Entropy

Shannon entropy from its axioms; what it measures and what it does not; entropy vs. variance; entropic uncertainty relations in physics; estimating entropy from finite data and its biases.

In the lab: entropy of intensity histograms; measuring spatial homogeneity and anisotropy in images; entropy-based autofocus.

2 Relative Entropy, Mutual Information and the Logic of Measurement

Joint and conditional entropy, Kullback–Leibler divergence, mutual information; the data-processing inequality (no algorithm creates information); Fisher information and the Cramér–Rao bound; maximum-entropy inference.

In the lab: mutual-information registration of multimodal sources; maximum-entropy reconstruction

3 The Instrument as a Channel: Noise, Dose and Capacity

Channels and capacity; the Gaussian and Poisson (photon-counting) channels; shot noise and the Rose criterion; sampling and the degrees of freedom of an system (space–bandwidth product).

In the lab: information per count, resolution as an information limit.

4 Compression, Sparsity and Reconstruction from Less

Source coding and the typical set; rate–distortion theory; transform coding and sparsity; compressed sensing and why undersampling can work; redundancy in phase retrieval.

In the lab: Image sampling under a time and space budgeta: undersampled recovery, inpainting.

5 Structure, Disorder and Complexity in Data

Sources and stationarity; block entropies and the entropy rate; how fast estimates converge; excess entropy as a measure of structure; Lempel–Ziv complexity and information distance; a glimpse of Kolmogorov complexity.

In the lab: quantifying disorder and structure; comparing spectra and time series without a model.

Format

- 5 sessions of 90 min: theory first, then a worked application.
- *lzcomplexity* software suite, for participants to keep.
- Participants are invited to bring their own data; the last session leaves room to discuss them.
- Prerequisites: undergraduate probability and Fourier analysis. No prior information theory.

What participants take away

- A clear sense of what "information" means in a measurement.
- Tools to compute and estimate entropies and mutual information on real data, and to know when not to trust them.
- A way to reason sampling and resolution as an information budget.

Lecturer. Prof. Ernesto Estévez Rams, Faculty of Physics, University of Havana; Humboldt Fellow; collaborator of the Max Planck Institute for the Physics of Complex Systems, Dresden. He trained in electron microscopy (PhD, TU Wien) and worked for two decades on quantitative diffraction and TEM analysis of disordered crystals before turning to information theory and complexity. He has taught information theory for several years and develops entropy-based methods for data from materials, the heart, the brain and language.