

## Multimodal Signal and Image Processing

**J.-P. Thiran**

*School of Engineering*

*Swiss Federal Institute of Technology (EPFL)*

*Lausanne, Switzerland*

multimodal signal is by definition a set of signals coming from different modalities but emanating from the same physical scene, i.e. from the same reality. Multimodal signals or images can be found in many areas such as medical imaging, where a Computed Tomography (CT) and a Magnetic Resonance Imaging (MRI) of the same patient can be acquired, that show complementary information about that patient. Similarly, in multimedia applications, an audio-video sequence is made of sound and images of the same scene. Intuitively multimodal signals share some common information, as they represent the same reality, but have also their marginal information, specific to each modalities. How to extract, describe and use the common and marginal information contained in a multimodal signal is the purpose of this talk. Therefore we will introduce an information theoretic framework, based on multimodal feature extraction, from which we can derive similarity measures between the different modalities of a multimodal signal. Applications will be shown in multimodal medical image analysis and in speaker tracking in an audio-video sequence.