



Contrastive Ecoacoustic Indices (CEI) : large-scale soundscape analysis via contrastive inference on natural audio recordings

Dr. Yann Teytaut, University of Lille (UL)

Facultad de Informática

Sala de Reuniones 1 - 28 de abril – 13:00

Entrada libre hasta completar el aforo

Abstract:

Among the various tasks in automatic audio signal analysis, there has been a recent interest in the study of environmental soundscapes, focusing on the sounds produced by living organisms (biophony), human activities (anthrophony and technophony), and natural physical processes (geophony) in order to monitor ecosystems, understand biodiversity, and assess environmental changes. Ecoacoustics is the scientific domain that mainly aims at analyzing and monitoring such soundscapes by means of non-invasive protocols.

Despite the widespread adoption of machine and deep learning techniques, existing ecoacoustic models predominantly rely on supervised learning and, consequently, face two primary limitations: (1) the necessity of annotated data; and (2) the restriction to fixed predefined classes. In this work, we leverage recent advances in Contrastive Language-Audio Pre-training (CLAP) and envision its first application to large-scale soundscape analysis. As trained on extensive datasets of paired audio and global text descriptions using contrastive learning, CLAP allows computing similarity scores between audio and text prompts without the constraints of predefined categorical lists. This flexibility enables a comprehensive investigation of various elements within the recordings from coarse-grained (e.g., mammals, weather, humans, vehicles) to fine-grained (e.g., dog, rain, speech, airplane) descriptions.

Here, we first conducted a preliminary experiment on a calibration dataset, featuring audio events likely to occur in soundscapes, which is shared with the community and constituted from an online, free sound library. Then, we developed a methodology to define reproducible, bounded, independent, and interpretable Contrastive Ecoacoustic Indices (CEI), which can characterize the prevalence of the four primary sound categories (biophony, geophony, anthropophony, and technophony) in soundscapes. We finally computed these new CEI on 9-month field recordings (189,137 1-min excerpts) monitoring both tropical (Ecuador) and temperate (France) soundscapes, portraying an anthropic gradient from protected forests to urban city centers. This experiment reveals clear

soundscape patterns associated with human population density suggestion that the CEI could be used in other ecological contexts. Finally, we briefly mention how such CEI could be used in a creative context, through the theoretical lense of a sound design interface allowing dynamic transformation of « virtual soundscapes » in video games.

About Dr. Yann Teytaut:

Yann Teytaut is an Associate Professor at the University of Lille, teaching programming and data science in the Computer Science Department and carrying out audio & AI research in the Algomus ("Algorithmic Musicology") team of the CRISAL laboratory. His recent research project is focused on AI-based sound analysis and transformation, notably for Video Game Music. His research interests also encompass ecoacoustics, sound ecology, voice analysis, computational musicology and, more broadly, most Music Information Retrieval tasks.