ACOUSTICS2008/246
A computational model for auditory scene analysis

Maria Niessen, Ronald Van Elburg, Dirkjan Krijnders and Tjeerd Andringa
University of Groningen, Dept. Artificial Intelligence, P.O. Box 407, 9700 AK Groningen, Netherlands

Primitive auditory scene analysis (ASA) is based on intrinsic properties of the auditory environment. Acoustic features such as continuity and proximity in time or frequency cause perceptual grouping of acoustic elements. Various grouping attributes have been translated into successful signal processing techniques that may be used in source separation. A next step beyond primitive ASA is source identification through schema-based ASA. We present a computational model for ASA that is inspired by models from cognitive research. It dynamically builds a hierarchical network of hypotheses, which is based on (learned) knowledge of the sources. Each hypothesis in the network, initiated by bottom-up evidence, represents a possible sound event. The network is updated for each new input event, which may be any sound in an unconstrained environment. The analysis of new input events is guided by knowledge of the environment and previous events. As a result of this adaptive behavior, information about the environment increases and the set of possible hypotheses decreases. With this method of continuously improving sound event identification we make a promising advance in computational ASA of complex real-world environments.