Cascaded Amplitude Modulations in Sound Texture Perception

Sound textures, such as crackling fire or chirping crickets, represent a broad class of sounds defined by their homogeneous temporal structure. It has been suggested that the perception of texture is mediated by time-averaged summary statistics measured from early auditory representations. In this study, we investigated the perception of sound textures that contain rhythmic structure, specifically second-order amplitude modulations that arise from the interaction of different modulation rates, previously described as “beating” in the envelope-frequency domain. We developed an auditory texture model that utilizes a cascade of modulation filterbanks that capture the structure of simple rhythmic patterns. The model was examined in a series of psychophysical listening experiments using synthetic sound textures-stimuli generated using time-averaged statistics measured from real-world textures. In a texture identification task, our results indicated that second-order amplitude modulation sensitivity enhanced recognition. Next, we examined the contribution of the second-order modulation analysis in a preference task, where the proposed auditory texture model was preferred over a range of model deviants that lacked second-order modulation rate sensitivity. Lastly, the discriminability of textures that included second-order amplitude modulations appeared to be perceived using a time-averaging process. Overall, our results demonstrate that the inclusion of second-order modulation analysis generates improvements in the perceived quality of synthetic textures compared to the first-order modulation analysis considered in previous approaches.