

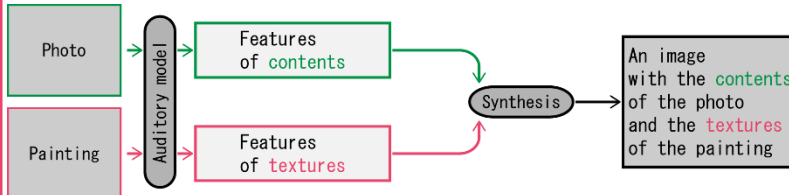
- Sound texture conversion with an auditory model -

Abstract

Natural scenes such as bubbling water and rustling trees give us specific perception of **sound textures**. We developed a method to **artificially give such textures** to speech and music. Inspired by research on manipulating image textures, we improved the method so that we can apply it to sound. A computational model that takes into account our hearing mechanism enabled **effective control of sound textures** in terms of hearing sensation. The method is realized in the same framework as the image texture manipulation. This indicates that, in the brain, seen and heard textures are processed by the similar mechanisms. From a scientific viewpoint, this study leads to **understanding of the mechanisms of sound texture perception** by comparing the model's internal states with the brain activities induced by hearing sounds. From an application perspective, the proposed method enables us to speak in a voice that does not actually exist or to play music with an instrument that does not exist.

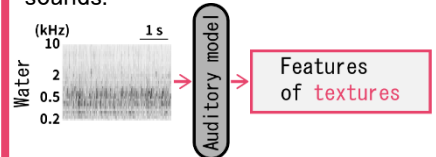
Prior study: image texture conversion

By combining **contents** in a photo and **textures** in a painting, an image that has both of them is synthesized. (Converting the **textures** of the photo to that of the painting while preserving its **contents**.)



Prior study: representation of sound textures

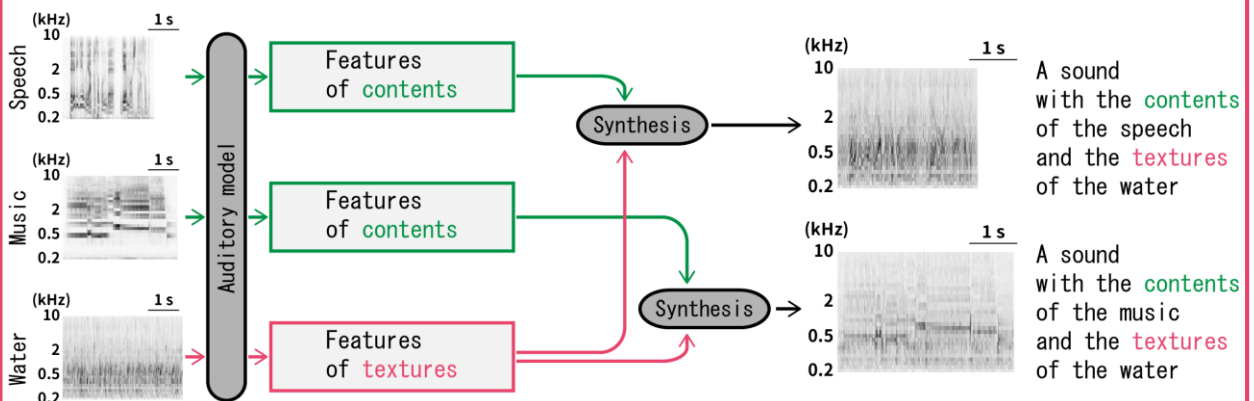
A model of the auditory system calculates the features that represent **textures** of environmental sounds.



This study: sound texture conversion

A model of the auditory system calculates the features that represent **textures** of environmental sounds and the features that represent **contents** of speech and music.

By combining **contents** in a speech or music and **textures** in an environmental sound, a sound that has both of them is synthesized. (Converting the **textures** of the speech or music to that of the environmental sound while preserving its **contents**.)



References

- [1] T. Koumura, H. Terashima, S. Furukawa, "Chimeric sounds with shuffled "texture" and "content" synthesized by a model of the auditory system," in Proc. International Symposium on Universal Acoustical Communication, 2018.
- [2] T. Koumura, H. Terashima, S. Furukawa, "Sound texture transfer using a model of the auditory system," in Proc. Annual Conference of the Japanese Society for Artificial Intelligence, 2017.

Contact

Takuya Koumura Email: cs-liaison-ml at hco.ntt.co.jp
Sensory Resonance Research Group, Human Information Science Laboratory



Innovative R&D by NTT
Open House 2019