

Acoustics researchers decompose sound accurately into its three basic components

July 13 2023



Leonardo Fierro and Vesa Välimäki of the Aalto Acoustics Lab demo the SiTraNo app for sound decomposition. Credit: Anna Berg/Aalto University

Researchers have been looking for ways to decompose sound into its basic ingredients for more than 200 years. In the 1820s, French scientist Joseph Fourier proposed that any signal, including sounds, can be built using a sufficient number of sine waves. These waves sound like whistles, each have their own frequency, level and start time, and are the

basic building blocks of sound.

However, some sounds, such as the flute and a breathy human voice, may require hundreds or even thousands of sines to exactly imitate the original waveform. This comes from the fact that such sounds contain a less harmonical, more noisy structure, where all frequencies occur at once. One solution is to divide [sound](#) into two types of components, sines and noise, with a smaller number of whistling sine waves and combined with variable noises, or hisses, to complete the imitation.

Even this "complete" two-component sound model has issues with the smoothing of the beginnings of sound events, such as consonants in voice or drum sounds in music. A third component, named transient, was introduced around the year 2000 to help model the sharpness of such sounds. Transients alone sound like clicks. From then on, sound has been often divided into three components: sines, noise, and transients.

The three-component model of sines, noise and transients has now been refined by researchers at Aalto University Acoustics Lab, using ideas from auditory perception, fuzzy logic, and perfect reconstruction. Their research is published in the *Journal of the Audio Engineering Society*.

Decomposition mirrors the way we hear sounds

Doctoral researcher Leonardo Fierro and professor Vesa Välimäki realized the way that people hear the different components and separate whistles, clicks, and hisses is important. If a click gets spread in time, it starts to ring and sound noisier; by contrast, focusing on very brief sounds might cause some loss of tonality.

This insight from auditory perception was coupled with fuzzy logic: at any moment, part of the sound can belong to each of the three classes of sines, transients or noise, not just one of them. With the goal of perfect

reconstruction, Fierro optimized the way sound is decomposed.

In the enhanced method, sines and transients are two opposite characteristics of sound, and the sound is not allowed to belong to both classes at the same time. However, any of two opposite component types can still occur simultaneously with noise. Thus, the idea of [fuzzy logic](#) is present in a restricted way. The noise works as a fuzzy link between the sines and transients, describing all the nuances of the sound that are not captured by simple clicks and whistles. "It's like finding the missing piece of a puzzle to connect those two parts that did not fit together before," says Fierro.

This enhanced decomposition method was compared with previous methods in a listening test. Eleven experienced listeners were individually asked to audit several short music excerpts and the components extracted from them using different methods.

The new method emerged as the winning way to decompose most sounds, based on the listeners' ratings. Only when there is a strong vibrato in a musical sound, such as in a [singing voice](#) or the violin, all decomposition methods struggle, and in these cases some previous methods are superior.

A test use case for the new decomposition method is the time-scale modification of sound, especially slowing down of music. This was tested in a preference listening test against the lab's own previous method, which was selected as the best academic technique in a comparative study a few years ago. Again, Fierro's new method was a clear winner.

"The new sound decomposition method opens many exciting possibilities in sound processing," says professor Välimäki. "The slowing down of sound is currently our main interest. It is striking that for

example in sports news, the slow-motion videos are always silent. The reason is probably that the [sound quality](#) in current slow-down audio tools is not good enough. We have already started developing better time-scale modification methods, which use a deep neural network to help stretch some components."

The high-quality sound decomposition also enables novel types of music remixing techniques. One of them leads to distortion-free dynamic range compression. Namely, the transient component often contains the loudest peaks in the sound waveform, so simply reducing the level of the transient component and mixing it back with the others can limit the peak-to-peak value of audio.

More information: Leonardo Fierro et al, Enhanced Fuzzy Decomposition of Sound Into Sines, Transients, and Noise, *Journal of the Audio Engineering Society* (2023). [DOI: 10.17743/jaes.2022.0077](https://doi.org/10.17743/jaes.2022.0077)

Provided by Aalto University

Citation: Acoustics researchers decompose sound accurately into its three basic components (2023, July 13) retrieved 28 April 2024 from <https://phys.org/news/2023-07-acoustics-decompose-accurately-basic-components.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.