# A8: Sound Transformations

## Audio Signal Processing for Music Applications

## Introduction

This is a peer assessed assignment where you will use the HPS model to creatively transform sounds. There are two questions in this assignment. In the first one you should perform a natural sounding transformation on the speech sound that you used in the previous assignment (A7). In the second question you should select a sound of your choice and do a "creative" transformation. You will have to write a short description of the sound and of the transformation you did, giving the link to the original sound and uploading several transformed sounds.

## Guidelines

For this assignment, you can use the sms-tools transformations GUI, `transformations_GUI.py` (in `software/transformations_interface/`) and using the HPS window, or you can modify the HPS transformation code to try other types of transformations (it is quite easy to add new transformations to the existing code). You should save the synthesized output sounds to upload them here after the explanation text. If you use the GUI, the output sounds get stored in the folder `sms-tools/software/transformations_interface/output_sounds`.

After the evaluation phase is over, we encourage you to upload the transformed sounds to Freesound.org and put a link in the discussion forum thread of the assignment for everyone to listen to your sounds and give you feedback. We will all learn from that. To comply with the guidelines of Freesound that the sounds should be useful to others, describe the sound and tag them accurately. If you used the sounds in sms-tools repository, use the Creative Commons Attribution (CC BY 4.0) license while uploading sounds. If you used a different sound, please check the license of the original sound and make sure you use a compatible Creative Commons license (CC BY 4.0 license should work for most cases, though there might be exceptions).

In order to perform a good/interesting transformation you should make sure that you have performed an analysis that is adequate for the type of transformation you want to do. Not every HPS analysis representation will work for every type of sound transformation. There will be things in the analysis that when modified will result in undesired artifacts. In general, for any transformation, it is best to have the harmonic values as smooth and continuous as possible and an stochastic representation as smooth and with as few values as possible. It might be much better to start with an analysis representation that does not result in the best reconstruction in exchange of having smoother and more compact data.

To help you with the assignment, we give a brief description of the transformation parameters used by the HPS transformation function:

- **freqScaling**: frequency scaling factors to be applied to the harmonics of the sound, in time-value pairs (where value of 1 is no scaling). The time values can be normalized, from 0 to 1, or can correspond to the times in seconds of the input sound. The scaling factor is a multiplicative factor, thus a value of 1 is no change. Example: to transpose an octave the sound you can specify [0, 2, 1, 2].

- **freqStretching**: frequency stretching factors to be applied to the harmonics of the sound, in time-value pairs (value of 1 is no stretching). The time values can be normalized, from 0 to 1, or can correspond to the times in seconds of the input sound. The stretching factor is a multiplicative factor whose effect depend on the harmonic number, higher harmonics being

more affected that lower ones, thus resulting in an inharmonic effect. A value of 1 results in no transformation. Example: an array like [0, 1.2, 1, 1.2] will result in a perceptually large inharmonic effect.

- **timbrePreservation**: 1 preserves the original timbre, 0 does not. It can only have a value of 0 or of 1. By setting the value to 1 the spectral shape of the original sound is preserved even when the frequencies of the sound are modified. In the case of speech it would correspond to the idea of preserving the identity of the speaker after the transformation.

- **timeScaling**: time scaling factors to be applied to the whole sound, in time-value pairs (value of 1 is no scaling). The time values can be normalized, from 0 to 1, or can correspond to the times in seconds of the input sound. The time scaling factor is a multiplicative factor, thus 1 is no change. Example: to stretch the original sound to twice the original duration, we can specify [0, 0, 1, 2].

All the transformation values can have as many points as desired, but they have to be in the form of an array with time-value pairs, so of even size. For example a good array for a frequency stretching of a sound that has a duration of 3.146 seconds could be: [0, 1.2, 2.01, 1.2, 2.679, 0.7, 3.146, 0.7].

## Question 1. Perform natural sounding transformations of a speech sound

Use the HPS model with the sound speech-female.wav, available in the sounds directory of sms-tools, to first analyze and then obtain a natural sounding transformation of the sound. The synthesized sound should sound as different as possible to the original sound while sounding natural. By natural we mean that it should sound like speech, that it could have been possible to be produced by a human, and by listening we should consider it as a speech sound, even though we might not be able to understand it. You should first make sure that you start from a good analysis, then you can do time and/or frequency scaling transformations. The transformation should be done with a single pass, no mixing of sounds coming from different transformations. Since you used the same sound in A7, use that experience to get a good analysis, but consider that the analysis, given that we now want to use it for applying a very strong transformation, might be done differently than what you did in A7.

Write a short paragraph for every transformation, explaining what you wanted to obtain and explaining the transformations you did, giving both the analysis and transformation parameter values (sufficiently detailed for the evaluator to be able to reproduce the analysis and transformation).

After the parameter explanation, upload a maximum of three transformed sounds (e.g., `speech-transformed-1.wav`, `speech-transformed-2.wav`, `speech-transformed-3.wav`). You should compress them into one zip file which you upload.

## Question 2. Perform creative transformations with a sound of your choice

Pick any natural and harmonic sound from freesound.org and use the HPS model to do the most creative and interesting transformation you can come up with. Sounding as different as possible from the original sound.

It is essential that you start with a natural harmonic sound. Examples include (but not limited to) any acoustic harmonic instrument, speech, harmonic sound from nature, etc. As long as they have a harmonic structure, you can use it. You can even reuse the sound you used in A7-Part2 or upload your own sound to freesound and then use it.

The sound from freesound to use could be in any format, but to use the sms-tools software you will have to first convert it to be a monophonic file (one channel), sampling rate of 44100, and 16bits samples.

You can do any interesting transformation with a single pass. It is not allowed to mix sounds obtained from different transformations. The transformed sound need not sound natural. So, time to show some creativity!

Write a short paragraph for every transformation, explaining what you wanted to obtain and explaining the transformations you did, giving both the analysis and transformation parameter values (sufficiently detailed for the evaluator to be able to reproduce the analysis and transformation).

Upload a maximum of three transformed sounds (name them `a8p2-transformed-1.wav`, `a8p2-transformed-2.wav`, ...). You should compress them into one zip file which you upload. Your peers will choose the "most interesting" among those submitted for evaluation.