Efficient 3D Audio Processing on the GPU

Emmanuel GALLO[1,2] and Nicolas TSINGOS[1]

[1]REVES / INRIA Sophia-Antipolis
http://www-sop.inria.fr/reves

[2]CSTB
http://www.cstb.fr

Audio processing applications are among the most compute-intensive and often rely on additional DSP resources for real-time performance. The widespread availability and increasing processing power of GPUs could offer an alternative solution.

GPU features, like multiply-accumulate instructions or multiple execution units, are similar to those of most DSPs used for audio processing. Besides, 3D audio rendering applications require a significant number of geometric calculations, which are a perfect fit for the GPU.

Storing audio in texture memory

To store PCM audio data in texture memory, we subdivide it into frames :
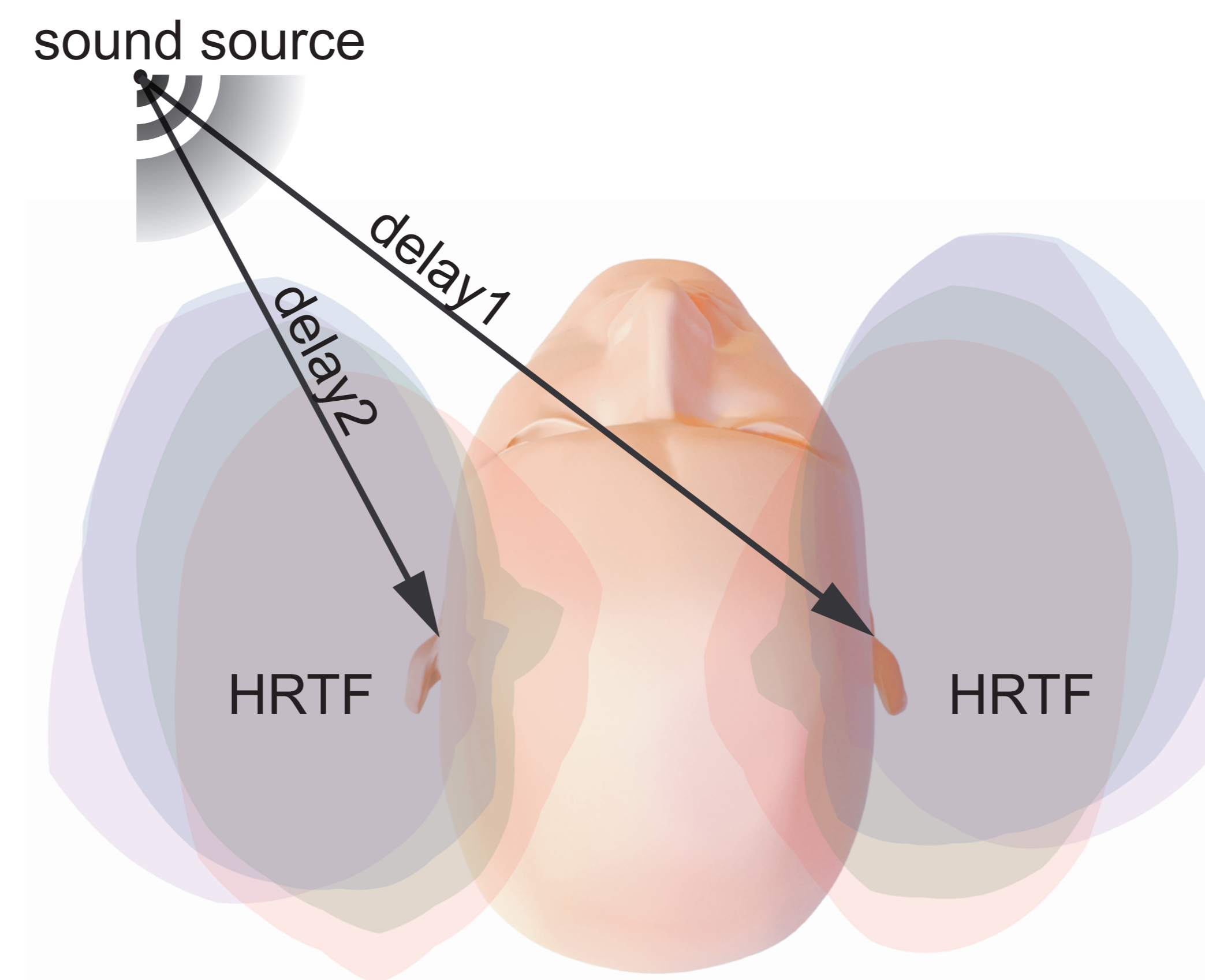
input signal   time →
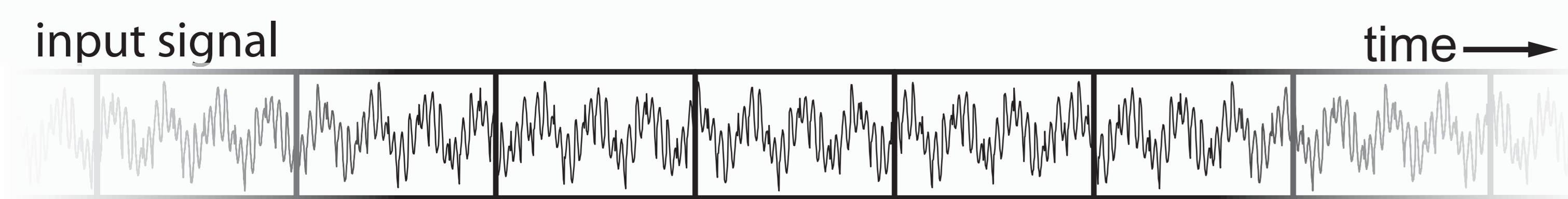
Then, we split the signal into 4 frequency sub-bands (using a band-pass filter bank) :

- 0-1000Hz
- 1000-2000Hz
- 2000-6000Hz
- 6000-22050Hz
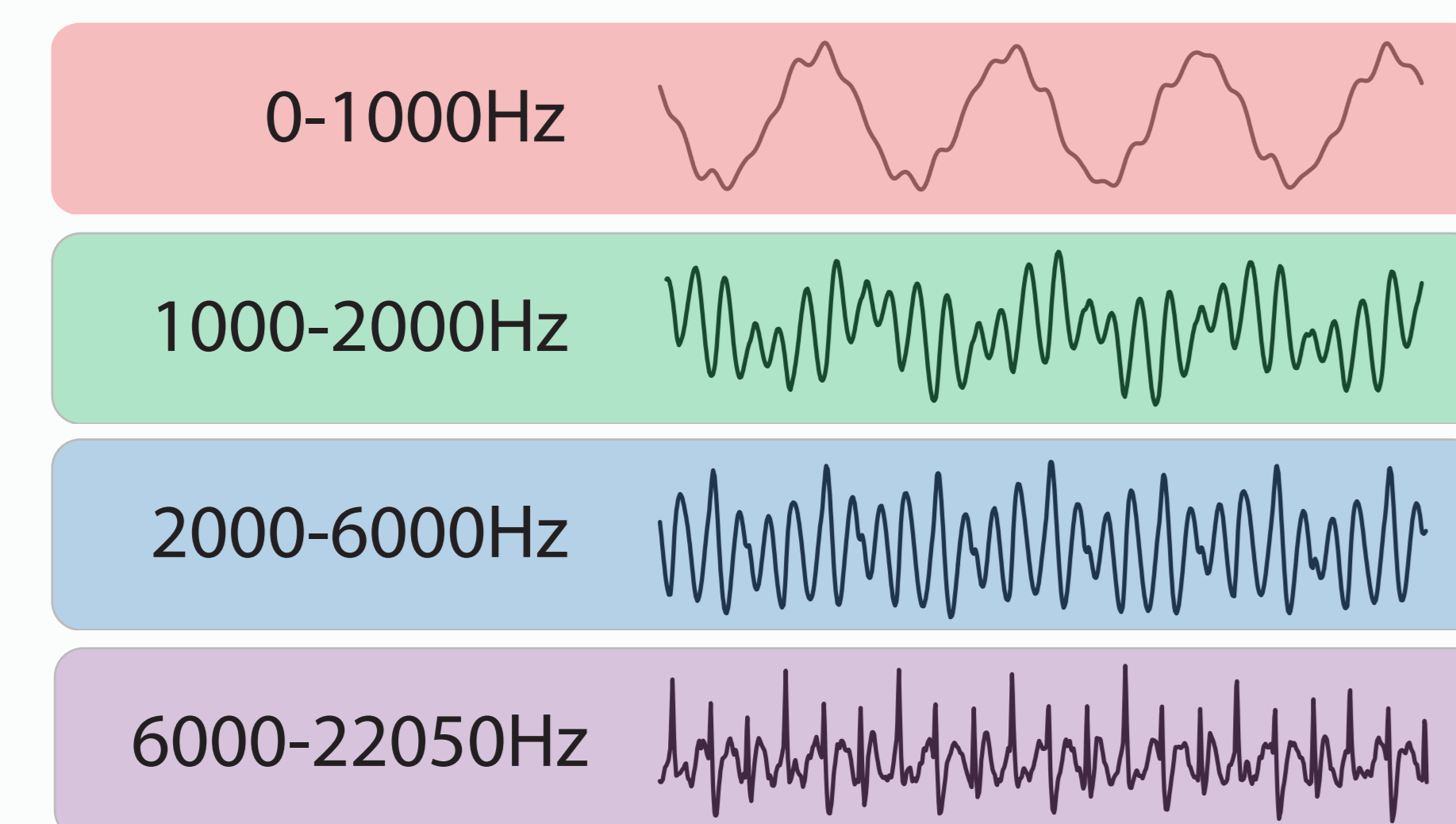
Finally, we store the result as a 1D RGBA textures :

Audio Processing

sound source

delay1

delay2

HRTF   HRTF
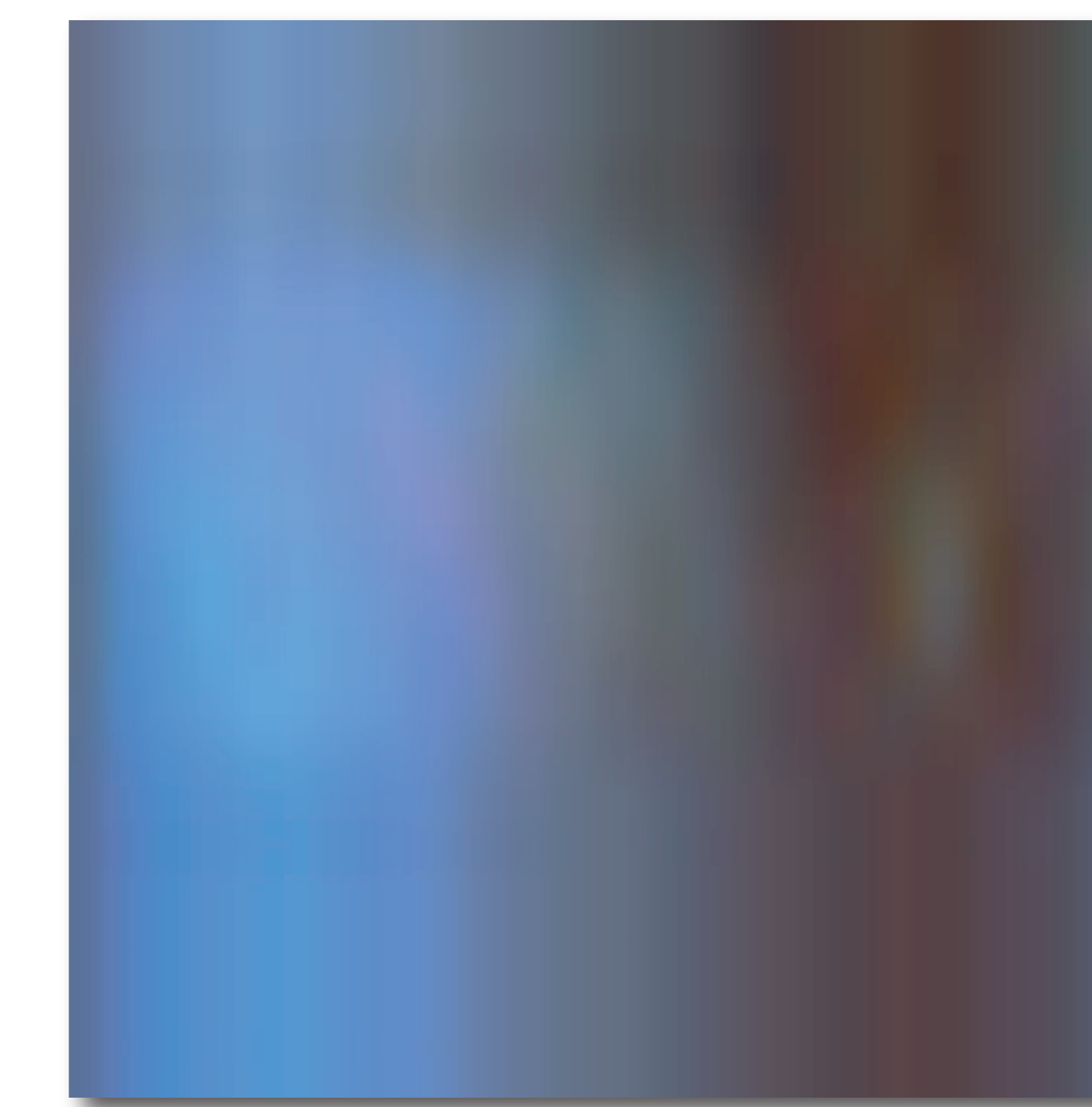
Spatialisation is achieved by a direction dependent delay and a binaural equalization derived from Head Related Transfer Functions (HRTFs).
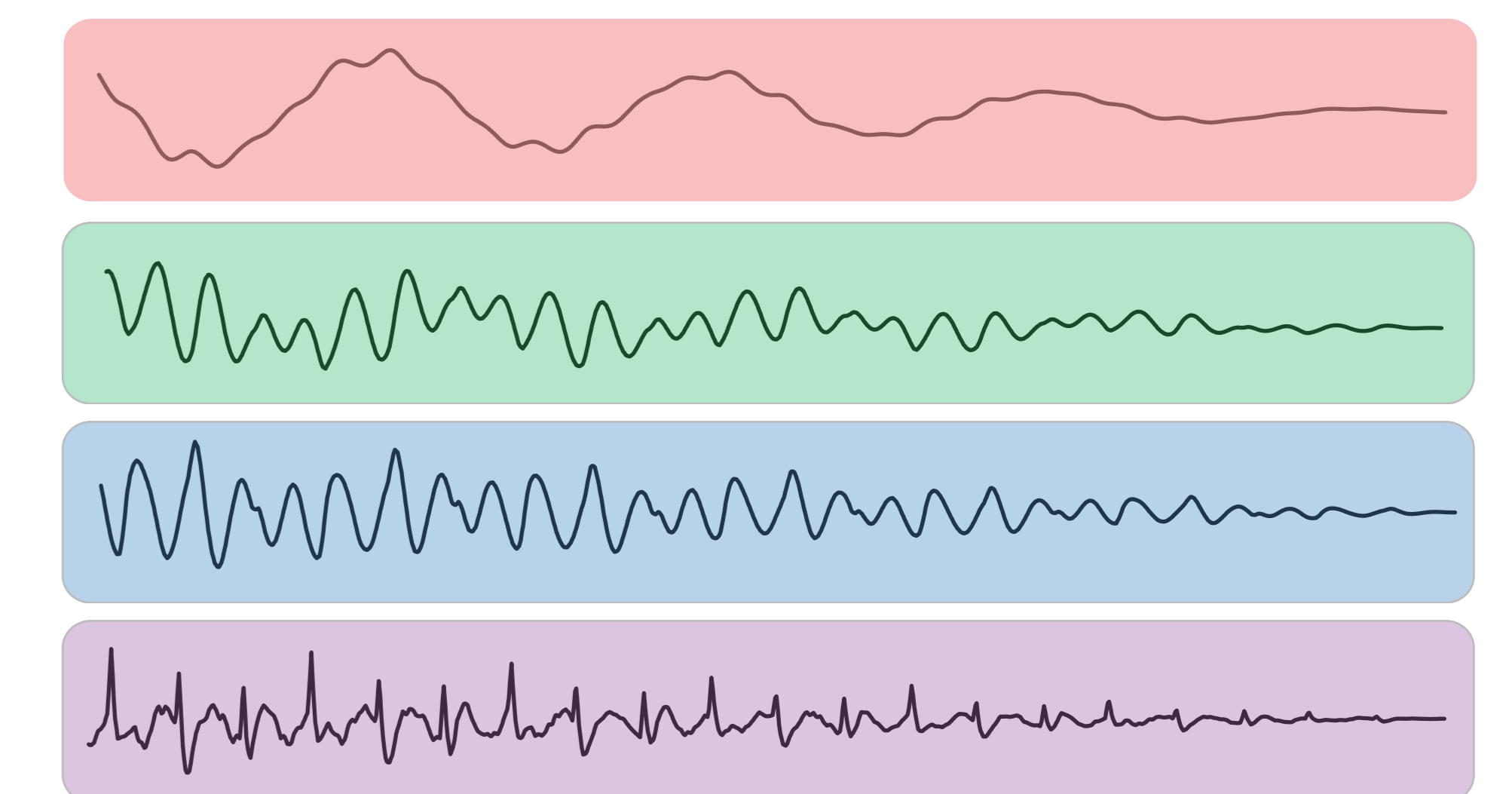
HRTFs account for head, torso and pinnae filtering effects on sound propagation. They can reproduce our 3D perception of sound.
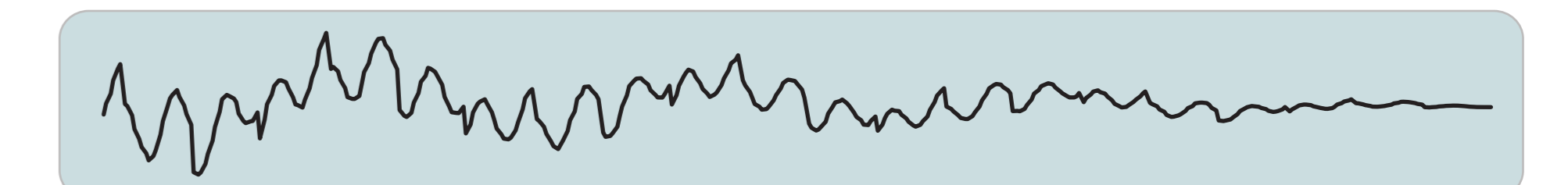
HRTF function for the right ear encoded as a longitude / latitude map.

Texture resampling is used to simulate doppler shift effects due to propagation delays

resulting in the following time-stretched signals :

Color modulation is also used to alter the frequency components of the signal, yielding the final reconstructed signal :

Can the GPU be a good audio DSP ?

Our first experiments suggest that GPUs can be used for 3D audio processing with similar or increased performance compared to optimized software implementations running on the latest CPUs.

However, several shortcomings still prevent the use of GPUs for mainstream audio processing (texture limitations, no accumulation across fragments).

Performance

sound sources

Legend:
- SSE on P4 3GHz
- Nvidia FX5950 Ultra
- Nvidia FX5950 Ultra no texture resampling

X-axis: binaural processing, mono processing
Y-axis: 0, 500, 1000, 1500, 2000, 2500