AudioFlux: A Proposal for interactive Visual Audio Manipulation

Lars Engeln, Rainer Groh

TU Dresden, Faculty of Computer Science, Chair of Mediadesign

Abstract

AudioFlux is a proposal for a creative manipulation interface respectively sound design. The spectral data within the frequency domain is construed as liquid material. Due to its flow across a surface it can be directly disturbed via multitouch or tangible interaction. With the aid of a fluid simulation and a resynthesis afterwards, an audio signal is shaped. This work shows off a wide variety of creative manipulation techniques for the frequency domain.

1   Introduction

While working with audio digitally, many tools still imitate the look and handling of audio processing hardware. Working with virtual knobs for instance are not that intuitiv as it could be. Especially missing visual feedback may lead to a vagueness how strong a parameter is affecting the audio signal. Manipulating a certain frequency band is not achieved in a simple way. For this a combination auf band-rejection and band-pass filters are needed. But this action decreases the flexibility of the design process, because even more parameters have to be automated for a complex project.

TouchNoise (Berndt et al., 2015) demonstrates a creative workflow for noise generation and manipulation. Little points moving above the screen. Each point has it frequency defined by its y-position. All frequencies sumed together results in the generated noise. The points are manipulated with different tools via multitouch. For example they are attracted to or dispersed from the interaction location to shape the noise. In this way noise with differently weighted frequency bands are generated.

Adding the time at the x-axis and the energy at the z-axis to the frequency at the y-axis results into a spectrogram as a representation of the frequency domain. Now this spectrogram can be modified with image processing. All changes to it are understandable visually and the modification are freely applied to different bands at different time. One method is to treat the spectrogram as a fluid like proposed here. Afterwards it is resynthesized. Fluid simulation was already used for an audio-visual dance performance (AJ Johnston, 2013). The dancer influences the fluid with their movement and the fluid itself has an impact on the audio synthesis.
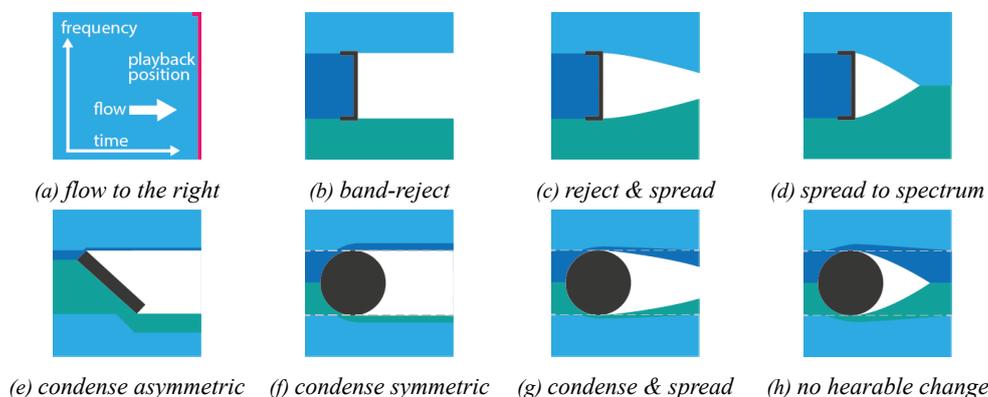
(a) flow to the right   *(b) band-reject*   *(c) reject & spread*   *(d) spread to spectrum*

(e) condense asymmetric   *(f) condense symmetric*   *(g) condense & spread*   *(h) no hearable change*

Figure 1: Basic band-reject and condense techniques for AudioFlux

In a further work (Bluff and Andrew Johnston, 2014) a instrumentalist is adding grains to a granular synthesis while playing. The grains are moving via a fluid simulation, which is influenced again due to the movement of the performer. Only grains in certain areas for different speakers are used for the synthesis.

Additionally there is a vast area of so called *sound textures* (Strobl et al., 2006). The theory of audio synthesis via texture generation is not discussed here. In this work we concentrate on manipulating the frequency domain.

2   Audio as Fluid

First of all the audio data is transformed into the frequency domain. Therefore the *short time fourier transformation (STFT)* (Allen, 1977) is used. The spectrogram is handled as a image with the time and frequency at x- and y-axis and the energy expressed as intensity of the pixel. The pixels are now interpreted as particles to achieve a *particle-based fluid simulation* (Premžoe et al., 2003). They are flowing on the screen form the left to the right (see fig.1a). At the most right the leaving pixels are resynthesized with the inversed STFT. To optimize the performance of the fluid simulation several adjacent pixels can pooled together to one particle. Doing that it is similar to *granular synthesis* in a visual way.

Every fluid has a certain liquidity (compare *Reynolds number*). A low liquidity means, that the fluid is sluggish and needs time to spread out. Whereas one with high liquidity fills up empty space fast. In this case fluids with high and low liquidities have the same speed, because audio is time related. Only the propagation within the frequencies at the y-axis deqends on the liquidity. So is the liquidity near to zero, audio particles does not move vertically, but they are moving horizontally just in time. Is the liquidity very high, the particles are spreading empty frequency bands and could fill up the whole spectrum.

The user disturbes the audio particles via multitouch tools. The same tools can be built as tangibles or expressed with different shapes of a hand. Altough pure virtuall tools are more flexible in size and quantitiy. Looking at the propagation of fluids, following tools are conceivable.
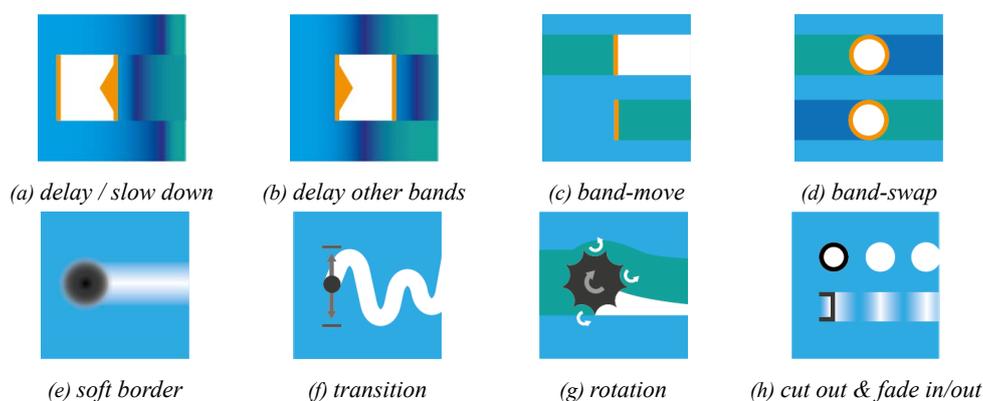
(a) delay / slow down	*(b) delay other bands*	*(c) band-move*	*(d) band-swap*
(e) soft border	*(f) transition*	*(g) rotation*	*(h) cut out & fade in/out*

Figure 2: Advanced interaction techniques for AudioFlux

2.1 Interaction

A bowl-tool, which holds back the particles, acts like a band-reject filter, if the liquidity L is very low (see fig.1b). By increasing L the frequencies spread apart (fig.1c) and fills up to the whole spectrum (fig.1d). If the tool is flat or convex, it will deflect the particles. Using a flat stroke the particles are deflected asymmtrically (see fig.1e). Conversely with a circle they are deflected symmetrically (fig.1f). With a low L the frequency band *behind* the tool is empty, but those frequencies are not lost. They are condensed to adjacent frequency bands. A higher L results in a coalescence to the empty bands (compare fig.1g). Finally with a very high L the frequencies are growing together and no hearable change is affecting the signal (see fig.1h).

2.2 Advanced Interaction

Next to solid shapes other tools with advanced functionalities are proposed. Certain freqeuncy bands can be delayed (see fig.2a) or copied to a time in the future (fig.2b). In this way a resyn-cronisation within the frequencies can be achieved. A band-copy tool moves a whole band to an other band like a pitch shift (fig.2c). The moved band can be added or multiplied to the existing band or it replaces the band completely. In addition to this a similar tool can entirely swap the frequency bands (fig.2d).

Furthermore several modifications can apply to all imaginable tools. Every tool can have a soft border (fig.2e), applying their effect only in a certain strength. The tools can move freely or in a certain path over time (see fig.2f), as well as scaling themself. They are pushing and pulling the particles across frequency and time, which affects in turbulences and blending. Even more turbulences appear while a tool is rotating (fig.2g), especially the tools shape is a gear for example. Whirls will mix up the particles and will create vast effects within the signal. At last not only the tools border can be soft, but also the whole tools strength can change over time. Due to this a cut-out stamp, aswell as a fade in and out effect is applied to a tool (see fig.2h). Applying multiple tools builds up a cascade of manipulations.

3   Conclusions

In this work we showed off a creative method to manipulate audio construed as fluid. Processing the frequency domain of audio visually like an image, sound design is more understandable and easier to achieve. For the work with AudioFlux a set of basic interactions, aswell as advanced modulation techniques are proposed. The exact meaning of the techniques has to be evaluated with a prototyp and a user study accordingly. It should be examined, whether the visual manipulations are compliant to the auditory perception. So if the interactions are understandable natively, or if the resulting auditory effects of the visual interactions has to be learned.

AudioFlux-like interaction techniques can be applied within synthesiser, for generating complex sound structures. Therefore a simple input signal like noise or a certain overtone spectrum is needed. In addition the techniques can be utilized in new interfaces regarding sound design, especially for creative estranging, deforming and sound creation. In this case the input can reach from single audio sources to whole arrangements.

References

Allen, J. (1977). Short term spectral analysis, synthesis, and modification by discrete fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *25*(3), 235–238. doi:10.1109/TASSP.1977.1162950

Berndt, A., Al-Kassab, N., & Dachselt, R. (2015). TouchNoise: A New Multitouch Interface for Creative Work with Noise. In *Proceedings of the Audio Mostly 2015 on Interaction With Sound* (p. 4). ACM.

Bluff, A. & Johnston, A. [Andrew]. (2014). Creative control of granular synthesis using fluid simulation & motion tracking. In *Proceedings of the 2014 international workshop on movement and computing* (150:150–150:153). MOCO '14. Paris, France: ACM. doi:10.1145/2617995.2618022

Johnston, A. [AJ]. (2013). Fluid simulation as full body audio-visual instrument. In *International conference on new interfaces for musical expression*. Korea Advance Institute of Science and Technology.

Premžoe, S., Tasdizen, T., Bigler, J., Lefohn, A., & Whitaker, R. T. (2003). Particle-based simulation of fluids. *Computer Graphics Forum*, *22*(3), 401–410. doi:10.1111/1467-8659.00687

Strobl, G., Eckel, G., Rocchesso, D., & Le Grazie, S. (2006). Sound texture modeling: A survey. In *Proceedings of the 2006 Sound and Music Computing (SMC) International Conference* (Vol. 61, pp. 1–7).