Spectral Size Does Matter

Lars Engeln
Rainer Groh
firstname.lastname@tu-dresden.de
Chair of MediaDesign, Technische Universität Dresden
Dresden, Germany

ABSTRACT

Nowadays spectrograms are often used for the analysis of audio signals. Usually, process and analysis tools are separated. Although manipulations within the visualization of the analysis results are more understandable. Spectral editing does directly manipulate the frequency-domain within a spectrogram. Here, affine transformations especially scaling is highlighted for the creative use in spectral editing.

CCS CONCEPTS

• **Human-centered computing → User interface design**; *Sound-based input / output*.

KEYWORDS

spectral editing, visual audio

1 INTRODUCTION

Transposition, restoration and compression are the popular spectrogram manipulations. In addition, visual manipulation can also be used for more advanced spectral processing [4]. Klingbeil [4] emphasized the *control problem* next to other challenges occuring in spectral editing, pointing out that for music composition a huge amount of frequencies should be finely controlable.

Spectral editing is classically implemented as a kind of painting program with tools for selecting and moving [1, 2]. Often these applications use a single audio source file on a sonagram. Therefore, TAPESTREA [6] showed for the first time the usage of different audio sources, which are blended in the frequency-domain to create sound collages. Besides to that, SPEAR [5] simplifies the frequency-domain to a line visualization of partials.

The VisualAudio-Design (VAD) [3] also uses a Sonagram with freely arrangeable sounds and images, that are used
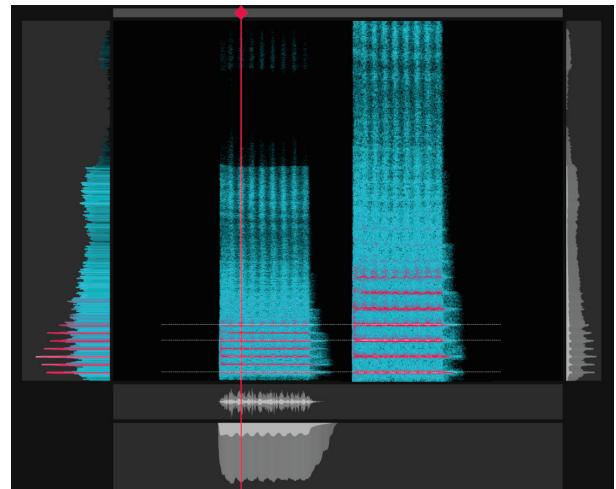
Figure 1: Scaling by 2 (scaling by one octave) according to the y-axis results in matching overtones again.

for resynthetisation. In VAD the rotation as an affine transformation is briefly introduced. Here, VAD is extended and scaling for node based spectral editing is shown.

2 SCALING SPECTRALS

Scaling in the frequency-domain results in either time-stretching for scalings in direction of time, or in overtone expansion or chrunching for scalings in direction of the frequencies. A uniform-scale does both. Time-stretching in frequency-domain comes on the fly by just stretching the magnitudes. The phases are reassigned via a phase-vocodor or reestimated with Griffen&Lim's algorithm [7]. Scaled sound-nodes can be precisely arranged to other sounds. Scaling the frequencies leads to beats or skewed tones if scaled freely with decimal numbers. But the scaling will be definitely in tune with itself, if quantized integers as scaling factor are used (see Figure 1).

Scaling Tool in VAD

The tool for scaling in the VisualAudio-Design (VAD) [3] lets the user the oppertunity to scale the time or the frequencies individually, or the whole sound-node uniformly. Therefore, an overlay with distinct interaction areas are displayed (see Figure 3). By simply click & drag or touch & move on an area the scaling is performed.
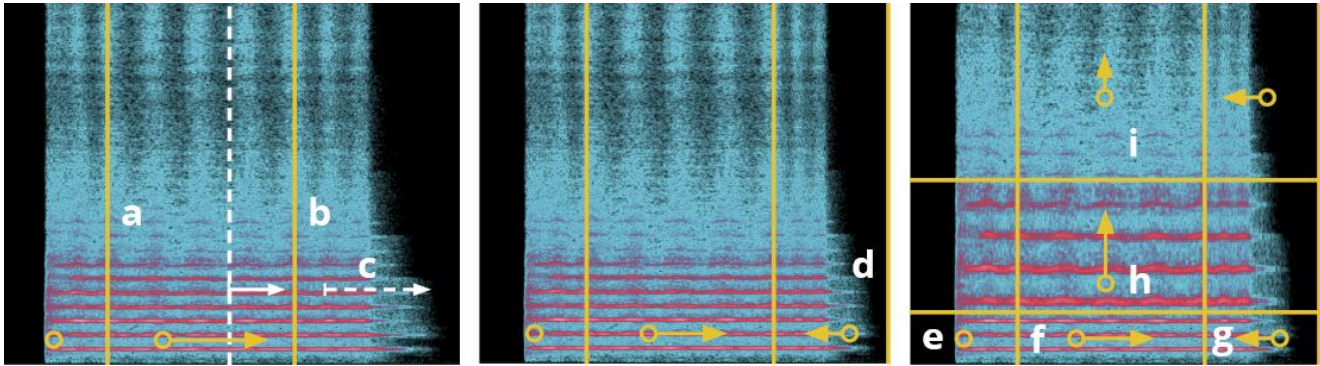
Figure 2: Partly scaling of sound-nodes: fixated time (a), fixated and dragged time (b) (timestretch from (a) to (b)), remaining is translated (c), fixated ending to prevent translation (d) (*counterstretched* to (b)), no scaling (e), fixated time (f), fixated and dragged time with fixated ending (g), fixated frequency (h), fixated and dragged frequency (i) (expansion from (h) to (i)).
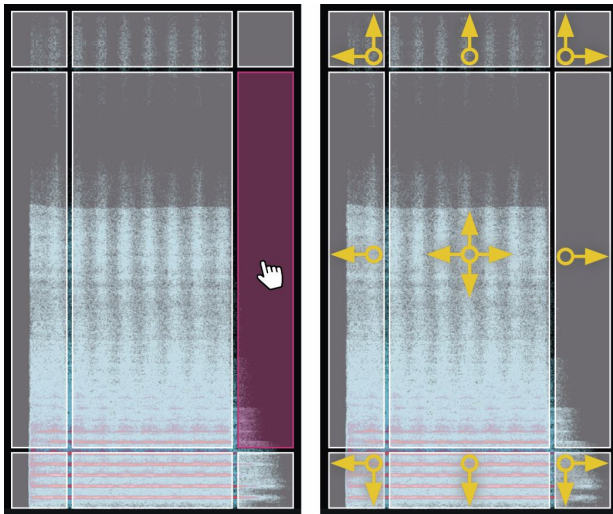

Figure 3: Overlay with areas for scaling directions: uniform scaling in the center, directional on the sides, ratio preserving directional in the corners.

In addition, in future work more complex scalings will be offered. Hereby, the user can select and fixate two timepoints (see *a & b* in Figure 2). By moving a fixated timepoint the spectrals between these fixated timepoints are scaled. Remaining spectrals are translated in the direction of scaling. To prevent the translation the sounds ending (d) can be fixated. Thereby, the time is *counterstretched* according to (b). Not only timepoints can be fixated, but also frequencies. To expand the frequencies in between of two fixated frequencies, the procedure is the same as for the time.

A grid system made of transients and overtones of the sound-nodes can assist the user while arranging and transforming. Moreover, fixating and dragging a overtone should lead into a snap to the overtone grid (see Figure 4).
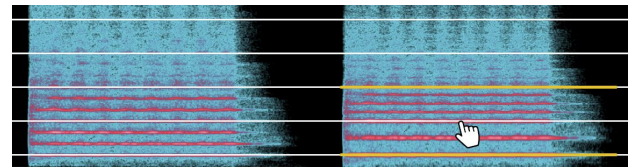

Figure 4: Scaling with snap to overtone: fixated overtones (yellow), dragging a overtone which snaps to an other overtone in the grid.

3 CONCLUSION

We briefly showed of scaling as an affine transformation to enrich the manipulation capabilities of spectral editing. Therefore, the way the user is performing a scale was shown. Additionally, more complex scalings and assistance with grids while manipulating were pointed out.

REFERENCES

[1] Niels Bogaards, Axel Roebel, and Xavier Rodet. 2004. Sound Analysis and Processing with AudioSculpt 2. *International Computer Music Conference (ICMC)* (Nov 2004), 1.

[2] G Eckel. 1992. Manipulation of Sound Signals Based on Graphical Representation-A Musical Point of View. In *Proceedings of the International Workshop on Models and Representations of Musical Signals, Capri, Italia*.

[3] Lars Engeln and Rainer Groh. 2019. VisualAudio-Design – Towards a Graphical Sounddesign. In *Proceecings of the 22th International Conference on Digital Audio Effects (DAFx-19) (in press)*.

[4] Michael Klingbeil. 2005. Software for spectral Analysis, Editing, and synthesis.. In *ICMC*.

[5] Michael Klingbeil. 2012. SPEAR: Sinusoidal Partial Editing Analysis and Resynthesis. *retrieved from on Nov* 12, 3.

[6] Ananya Misra, Perry R. Cook, and Ge Wang. 2006. TAPESTREA: Sound Scene Modeling by Example. In *ACM SIGGRAPH 2006 Sketches (SIGGRAPH '06)*. ACM, New York, NY, USA, Article 177.

[7] Nathanaël Perraudin, Peter Balazs, and Peter L Søndergaard. 2013. A fast Griffin-Lim algorithm. In *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 1–4.