Translation as Technique: Collaboratively Creating an Electro-Acoustic Composition for Saxophone and Live Video Projection

Christopher Jette, Kelland Thomas, Javier Villegas, Angus Graeme Forbes
Creative Coding Lab, School of Information, Science, Technology & Arts, University of Arizona
chrisotpherjette@gmail.com, {kelland, jvillegas, forbes}@email.arizona.edu

ABSTRACT

This paper describes the development of an electro-acoustic composition, titled *v→t→d*, which positions the act of translation as the focal point of the creative process. The work is a collaboration between a saxophonist, a composer, and two multimedia artists. We begin with a saxophone improvisation which is translated into a composition featuring solo saxophone and fixed electronics. The performance of this sound composition is translated again with live image processing, where the audio and visual content serve as control signals. We locate translation as an effective tool for composition and describe work that has inspired this approach to our creation processes.

1. INTRODUCTION

v→t→d is a work that assumes that the bias of a transcriptionist will impart a distinctive perspective on the material translated from one medium to another. In this work, a saxophone improvisation is the seed for translations common to spectralism. The translations serve as materials used to compose a new work. This new work for saxophone and fixed electronics is the source for live image processing. In the live performance the saxophonist and the processed images of the performer are juxtaposed, presenting the audience with a transcription of the visual experience. Each collaborator's transcription across mediums is an opportunity to impart an interpretation. In this paradigm, each phase of the process leaves a mark upon the final work. The spectral translation necessarily navigates from the acoustic to digital to domain; and the image translation maps both the live sound data and video stream as control signals for the image processing. This paper will present an aesthetic context in which the work is located, followed by works that have inspired our position. The next component of this paper will be a description of the processes involved in creating the sound composition, including the compositional strategies and a discussion of improvisation in this context. The last section will provide a description of the live image processing.

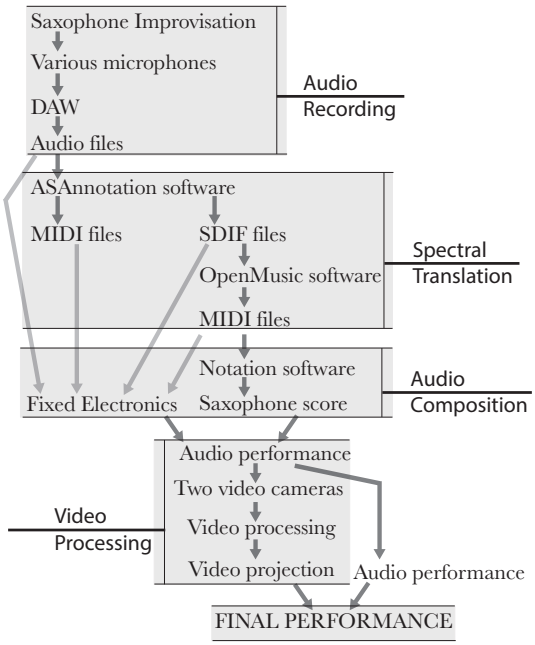


Figure 1. The stages of translation. This figures shows the four stages that lead to the final performance.

2. MOTIVATION

The motivation for this work is the realization of a collaborative process where various artists contribute to a final work. The function of *translation* in this context varies with each artist. This section will first describe the collaborative process and the artists involved, followed by an indication of philosophical influences.

2.1 Collaboration

The team includes: Kelland Thomas, Christopher Jette, Angus Forbes and Javier Villegas. Kelland Thomas created the seed improvisation material for the work and performs the notated score. Christopher Jette translated the recording, creating the saxophone score and the fixed electronic sounds. Angus Forbes and Javier Villegas developed a video processing technique [1] that used the audio as input for creating a visual interpretation of the piece. Javier Villegas also performed the video processing live, adjusting control parameters of the visuals during live performances of *v→t→d*.

2.2 Translation

In his essay *The Task of the Translator*, Walter Benjamin notes that translation fails due to "the inaccurate transmission of inessential content" [2]. The inessential content that he references is the material beyond the statement, the artistic method, not the message. This inessential content constitutes the voice of an artist, be it the unique sonic language of an improvising pianist or the cinematographic choices of a director. Every stage in the creation of $v{\rightarrow}t{\rightarrow}d$ is a translation of earlier material. We leverage the idiosyncratic perspective of each artist as a compositional tool. We cast translation as a means of codifying the aesthetic perspective of the translator. Each collaborator contributes something to the final work. $v{\rightarrow}t{\rightarrow}d$ is an assemblage of perspectives collapsed into a single work.

3. RELATED WORK

$v{\rightarrow}t{\rightarrow}d$ begins with a supposition that spectral translation is a process whereby the the act of transforming the material reshapes the output. The basic techniques of spectralism are well documented [3] and in our context serves as one of the compositional tools. The process of translation refines spectral translation techniques previously reported by the first author [4]. The composed saxophone work is sculpted from multiple translations of the seed material. Our approach is informed by the perspective of Gerard Grisey, who suggested, "What, for me, is very important is to have a sort of ecological attitude toward different sounds, to just accept them as they are and try to find the right place or right function for them in the context of the piece." [5] This ecological approach to the function of sounds informs our compositional process. The primacy of the sonic character of the source material surpasses other compositional concerns during the act of composing.

The formal design of $v{\rightarrow}t{\rightarrow}d$ is a series of translations. The first translation is that of the improviser's ideas into sound, and this acoustic sound, which is translated to a digital recording (discussed further in sections 3.1 and 4.2.2). The audio is edited by the composer to highlight principal sonic components. This approach of reductive editing is inspired by techniques employed by the acoustic ecology movement. Barry Truax reports on recordings where short segments from each hourly recording over a 24-hour period were transparently edited together to create a one-hour experience of that specific soundscape [6]. In $v{\rightarrow}t{\rightarrow}d$ our goal in editing is to distill the improvisation to a few soundbites that encapsulate the different stages of the improvisation.

The emphasis on the effects of transcription in each stage elevates the process to an aspect of the creative process. The role of transcription moves away from aspiring to realize a faithful recreation and approaches a location within the artists technique. Berio suggests this notion when he discusses the way in which "forms of transcription" can become assimilated into the process of creation. He writes:

> Here we are no longer dealing with transcription as a genre, but as part of the ups and downs of creativity: when, that is, you have a single musical vision going through different and self-sufficient formulations before arriving at the definitive realization, decanted from (or destroying) all the others [7].

Our work embodies Berios proposition; each collaborator in $v{\rightarrow}t{\rightarrow}d$ translates the previous work with a different aesthetic intention. Each of the artists concentrates on translation as a creative tool when creating their portion of the work. This emphasis on translation underlies the work and through discourse among the involved artists, translation became the focal point of artistic exploration.

3.1 Improvistational Context

Improvisation is inherently a live act, composed of plans and born of the immediacy of the moment. An improvising performer must bring to bear neuromuscular conditioning, control, and flexibility on a given instrument, coupled with a corpus of musical ideas, passages, and structures culled from a history of practicing and listening. Whether alone or in a group, musicians generally structure the act of improvisation around constraints of some kind. Saxophonists are often embedded in jazz ensembles, where the constraints are the result of the chord changes and melodic structure that define the work. Chord changes are often a key constraint (though not necessarily the only one) for structuring ones improvisation in a jazz context. In the absence of predefined constraints, whether by convention or dialog among participants, the performer constructs constraints as part of the process of improvisation itself in order to create a coherent musical outcome. These constraints can be cast as feedback loops. A paper by Forbes and Odai discusses improvisation as a means to create "a network of nested feedback loops." In order to encourage the emergence of new concepts during improvisation they invite musicians to think of themselves as "guiders" of a fluid performance that has its own agency rather than solely as the creators of it [8].

In the best case, improvisation results in a performed utterance that, even though based on material that has been previously learned and practiced, is sufficiently recombined, reworked and is relevant to the moment of creation so as to seem genuinely novel. The formal architecture of jazz and the curatorial guidance approach informed the creation of the seed material. In the creation of the seed material, the saxophonist strove to create an improvisation within this paradigm.

4. THE SOUND OF $V{\rightarrow}T{\rightarrow}D$

4.1 Compositional Strategy

With the improvisation created, compositional work proceeds in two directions: the creation of a new saxophone melody via translation of the improvisation; and the creation of an accompanying fixed electronics part. The fixed electronics are designed to contextualize and expand the sonic characters of the saxophone part.

The original compositional plan was the creation of a through-composed work that united the components by virtue of reassembly. While composing there was a shift toward the creation of distinct small works. This happened for two reasons. First, the myopic focus on the character of each of the twelve segments created great independence of material. Second, we were inspired to utilize sectional design patterns from a radio interview with John Zorn where he referenced his work *Speedfreaks* [9] in order to highlight the challenge of creating short and independent compositions. The unique states of the live image processing further reinforces this sectional autonomy. The confluence of the radio interview, the autonomous states of the live image processing and the hyper-focus on sections spawned a formal structure where the resulting work is a series of twelve short movements (the average length is 50 seconds).

The role of the fixed electronics is to provide an accompaniment to the solo saxophone line. A compositional constraint is the creation of fixed electronics, not live audio processing. Fixed electronics are employed to simplify the performance set-up and ensure that the final translation would be exclusively in the visual realm.

Relating the electronics to the saxophone provided a sonic framework that amplifies the sonic character defining each section. With the saxophone line preliminarily in place, each movement's electronic component was created. The sonic character of the original material, the new saxophone line, and the various materials from the translation provided a starting point. A conscious effort was undertaken to create a kind of chamber music, where the saxophone and electronics have equal footing. The fixed nature of the electronics enables embedded sonic cues for the player as well as sections where there is metric freedom for the saxophonist.

4.1.1 Spectral Translation Discussion

The intention of the spectral translation process in $v{\rightarrow}t{\rightarrow}d$ is the amplification of salient traits in the source signal. This process begins with listening to the source and identifying important features. These sonic characters serve as a guide in the compositional phase. The spectral translation process produces a realization that is mathematically precise yet different than the human impressions. The compositional process blends these precise (yet curated) mathematical descriptions of sound with intentions to amplify particular aspects of the sound, yielding a new saxophone line.

In $v{\rightarrow}t{\rightarrow}d$, spectral translation is uniquely utilized to create a sonic snapshot of the source, in this case the particular sound of a performer and the saxophone. Just as a photograph captures not only the qualities of light but also the framing and perspective of the photographer, here spectral translation reflects the input signal, the curatorial hand of the translator, and the tendencies of the software. Additionally, from the standpoint of an artist, the software output can be said to "fail in interesting ways." These "failure figures" populate the final composition when they are relevant to the character being portrayed.

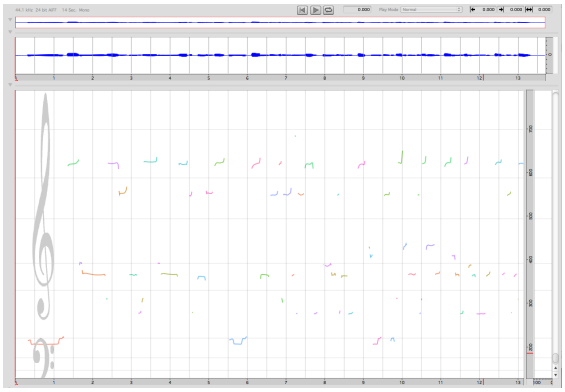


Figure 2. A screen shot of ASAnnotation showing a single partial trace (the multicolored lines). The frequency (vertical axis) is constrained from 150 to 800 Hz.

4.1.2 Spectral Translation Process

The original 8:48 solo saxophone improvisation is punctuated with breaths, these provide a convenient point for segmentation. The resulting twenty-nine components contain twelve unique sonic concepts that serve as the sections of the final piece. The workflow for the spectral translation process utilized features of *ASAnnotation* and *OpenMusic* (see the Spectral Translation section of Figure 1). The *ASAnnotation* program provides partial tracking with the ability to limit the amount of partials to a specified quantity. Partial tracking leverages analysis using the SuperVP kernel for phase vocoder based processing and the Pm2 kernel for additive modeling of partials [10].

This workspace enabled several approaches to extracting a single line. To highlight different sonic features, various analyses were produced by manipulating the FFT Settings and the Partial Connection parameters. Figure 2 shows a single partial tracking analysis which is output as an *sdif* file (compare with the sonogram analysis in Figure 3). Using the MIDI Note Annotation feature (see Figure 4), a rough sketch could be developed in *ASAnnotation* and rendered as a midi file.

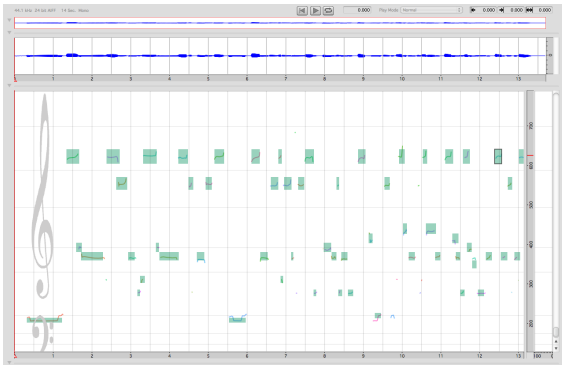


Figure 4. A screen shot of ASAnnotation where green midi notes are hand drawn over the existing analysis. This is rendered as a separate MIDI file The frequency (vertical axis) is constrained from 150 to 800 Hz.

The results of the Partial Tracking analysis were exported

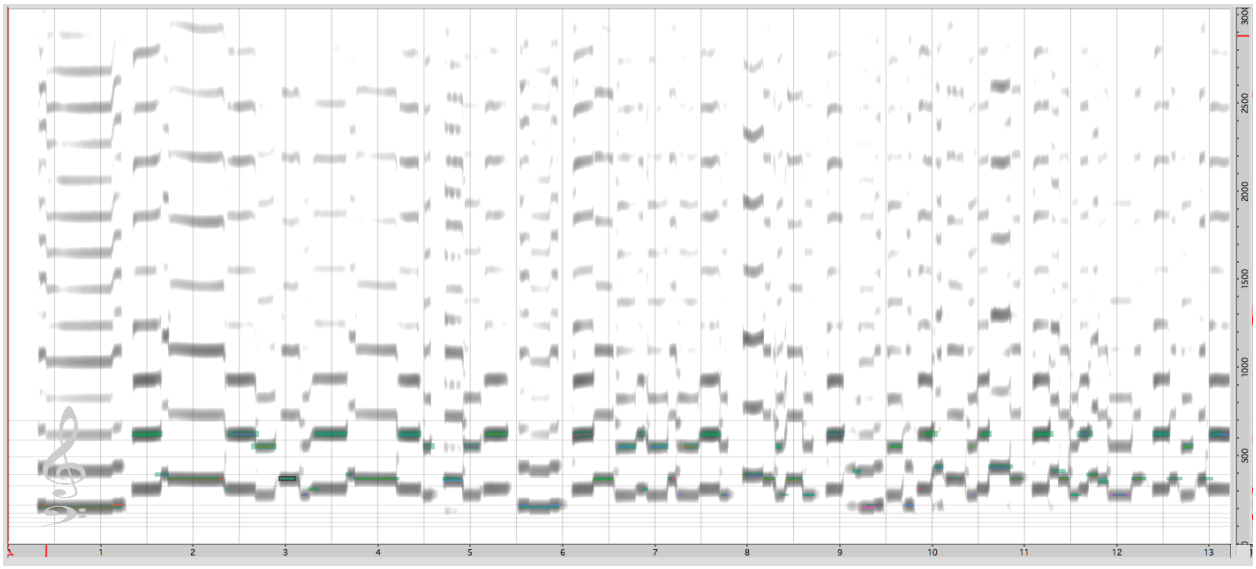


Figure 3. Compare this screen shot of the more familiar sonogram analysis in ASAnnotation with partial tracking of Figure 2. The frequency (vertical axis) is constrained from 0 to 3000 Hz.

as *sdif*s and brought into *OpenMusic*, where further filtering and shaping of the data could occur before rendering MIDI files. The various files generated from a single input collectively served as the the source material for the final solo saxophone line. With the collected variation of melodic lines, the composers recombined these multiple perspectives into a single melody that amplified the sonic features of the original improvisation and the relevant anomalies of the translation process.

Concurrent with the composition of the solo saxophone line was the creation of electronic accompaniment material. The fixed electronics were assembled in a DAW environment utilizing the large array of synthesis and processing tools available to the acousmatic composer. SuperCollider and ACToolBox figured prominently as production and composition tools. Each section demanded a unique approach in the creation of the fixed electronics. Often, the original audio file provided a structural template, where the amplitude, spectral or pitch envelopes suggest a formal design shape. In some cases audio from the original recording are embedded of the fixed electronics. In other instances, the new saxophone material offered a point of departure, the form of the section being an extension of the solo saxophone.

In creating the larger work, the autonomy of the individual section within the larger context emerged as a prominent compositional goal. To increase the autonomy of each section, the electronic material often amplifies spectral or rhythmic components of the saxophone line. For instance, the highly rhythmic nature of the saxophone line in the sixth section is intensified with percussive electronics. In contrast, the tenth section is comprised of a melodic line of longer tones and the wash of drones in the electronics extends this aesthetic. Each section demands a unique sound and variations in timbre, density, spectral identity and rhythmic figuration are parameters where contrast among sections is achieved.

4.2 Improvisation as Seed

4.2.1 Why Translate Improvisation?

Kelland Thomas, the saxophonist who initially commissioned *v→t→d*, performs both improvised and notated music. This provided a unique opportunity to utilize the performers improvisation as source material for a composed work. The saxophonist created several improvisations serving as a musical introduction to his improvisational sound world. By way of reply, the composer selected an improvisation that best encapsulated the range of material presented and/or reflected the particular sonic interests of the composer.

4.2.2 Recording Chain

The act of recording solidifies the sonic component of an acoustic event. Each recording encodes many things, including; the particular sonic imprint of an instrument, the distinct approach and improvisational choices of an individual performer, the convolution of the instrument and the space and the unique signature of the microphone and electronics used in the recording chain. Each aspect of this chain shapes the sound that is captured on the recording. To limit the influence of the physical space and recording chain, the saxophone was recorded in a dry recording studio with several different microphones. While this set-up influences the sound of the recording, the locus of these recordings are the sound of the horn and the improvisational choices of the performer, not the recording technique.

4.2.3 Performer on Improvisation

The specific improvisations recorded for *v→t→d* were free improvisations. Although the performer did not decide on any specific constraints in advance of recording them, discussions between the composer and performer influenced

choices made in the studio. In particular, timbral exploration and variation as an improvisational approach was a consequence of the participants' shared interest in the inner workings of sound. Several structural elements emerged during the improvisations. These constraints (which served as improvisational strategies) are described below.

Small rhythmic motives consisting of several notes were subsequently repeated, varied, and extended, lending local coherence to sections and more extended forms of coherence when these motivic ideas were recalled from earlier in the performance. Several long sections were developed from a single central pitch or pedal tone, along with the exploration of sonic variation on that tone. Variations included such timbral effects as key clicks, timbre trills, vibrato modulation, microtones, diffusion of tone with air or fuzz, and overtone trills.

Choosing a musical parameter and highlighting contrasts in its use is often a very effective way to structure free improvisation. In this case, contrast was created between definitive melodic statements (phrases having a sequence of pitches with clear rhythmic profiles) and non-melodic sonic textures or sound effects. Additionally, sections with clearly defined rhythmic character were contrasted with sections with little or no discernible rhythm. Other explored parameters include contrasts between having sounds or silences, using pentatonic or chromatic scales, and using discrete or continuous elements.

5. LIVE IMAGE

The live performance of $v \rightarrow t \rightarrow d$ presents the audience with a saxophone and fixed electronics work, as well as a image projection of the saxophonist. There are two video cameras that capture the saxophonist and act as input to video processing software. One camera is mounted on the bell of the instrument and the other captures the entirety of the performer. Figures 5(b) and 5(d) show the bell mounted camera, while Figure 5(c) illustrates the remote camera, and Figure 5(a) shows a mixing of the two cameras. The video processing techniques are extended from previous works by the authors on creative visualization, including the use of motion textures [11] and the use of musical parameters to manipulate visual patterns [12].

5.1 About Image Processing

The manipulation of the videos is inspired in the audio processing technique of granulation. Small fragments of the input (grains) are re-arranged in space and time to create the video effects. The audio and two video signals of the performance are used to control how the output video is generated. Akin to the repurposing of the audio material from the saxophone for compositional purposes, the video repurposes the live performance as a control signal.

5.2 Three decisions in Image Processing

5.2.1 Manipulate Which Grains?

The full video sequence can be interpreted as a three dimensional array of small overlapping videos, where time provides the z dimension. All the effects can be interpreted as different ways of mapping the two-input arrays of grains to create an output array that will be the output sequence (see Figure 5(a)). Individual grain modifications at the pixel level and changes in size and orientation are also allowed in this system. The automatic selection of what grains will be mapped to a different place in the output can be done in different ways. In some instances of the processing all the grains are manipulated, in others, the grains to be modified are selected at random. Additionally, there are some cases where only the properties of grains in areas of motion activity are changed. There are also cases where the energy of the audio input determines how far from the center a grain must be in order to be considered.

5.2.2 Manipulation Type

The most common manipulation we used was changing the spatiotemporal position of the grains (see Figure 5(c)). The spatial position of the grains can be used as a narrative element moving from a random distribution to perfect organization. Non-linear deformations of the input image can be generated by programmatically by changing the spatial placement of the grains in each frame. In addition to these more visually oriented techniques, visual analogs to well known audio granular techniques are possible (see Roads [13]). An example is *cloning* which can be created by repeating and skipping grains in the output array (see Figure 5(b)). Changes in position can also be used to generate transitions between the two input/video streams (see Figure 5(a)). Other manipulation that were used include grain rotation around the temporal axis (as in Figure 5(d)) and changing the grain size.

5.2.3 Controlling Manipulations

The final stage in each video manipulation is deciding what control parameters to map to the processing algorithm. In most of the cases, properties of the grain : the amount of geometric distortion, grain rotation, repeating factor (for the cloning effect) were linked to the amplitude energy of the audio signal. In other cases the amount of motion in the input sequence was used as controller. The decisions regarding control parameter mapping were taken on a section by section basis, employing intuition and attempting to achieve a visual translation of the auditory impression.

6. CONCLUSION

$v \rightarrow t \rightarrow d$ is an attempt to locate translation within the creative process through the creation of a working strategy. To achieve this end, the collaborators each perform the creation of the next step in the final work by translating a previous step. Each artist utilizes a known technique and locates their creative work in the act of translation. The work grew out of the conversation around this topic and the saxophonist was tasked with bringing something of this conversation into the improvisations. The composer uses the recorded audio as a source for the creation of a new saxophone work, endeavoring to characterize components of the original material using techniques of reductive editing, spectral translation and synthesis. The performance of the

saxophone and electronics work is the stage for the translation of the visual setting. The sonic and visual presence of the performer become control parameters in the translation of the visual scene. Each collaborator brought a creative skill to bear and endeavored to focus on the task of translation as creative act. This act brought creative preferences and assumptions to the attention of each artist. Future work will continue to explore different types of translations as a potentially rich area for productive investigation.

7. REFERENCES

[1] A. G. Forbes and J. Villegas, "Creative applications of microvideos," in *MMEDIA 2014, The Sixth International Conferences on Advances in Multimedia*, 2014, pp. 108–111.

[2] W. Benjamin, *Illuminations: Essays and Reflections*, H. Arendt, Ed. Harcourt Brace Jovanovich, 1968.

[3] J. Fineberg, "Spectral music," *Contemporary Music Review*, vol. 19, no. 2, pp. 1–5, 2000.

[4] C. Jette and K. Kirchoff, "Wii play piano: Composing for piano and wii remote," in *Proceedings of the International Computer Music Conference (ICMC)*, Huddersfield, England, 2011, pp. 354–359.

[5] D. Bundler. (1996) Interview of Gerard Grisey. [Online]. Available: http://www.angelfire.com/music2/davidbundler/grisey.html

[6] B. Truax, "Genres and techniques of soundscape composition as developed at Simon Fraser University," *Organised Sound*, vol. 7, no. 1, pp. 5–14, 2002.

[7] R. Dalmonte, *Two Interviews With Berio*. Marion Boyars Publishers Ltd, 1985.

[8] A. G. Forbes and K. Odai, "Iterative synaesthetic composing with multimedia signals," in *Proceedings of the International Computer Music Conference (ICMC)*, Ljubljana, Slovenia, 2012, pp. 573–578.

[9] T. Gross, "At 60, 'challenges are opportunities' for John Zorn," Audio Interview, National Public Radio, September 3rd 2013. [Online]. Available: http://www.npr.org/2013/09/03/217195249

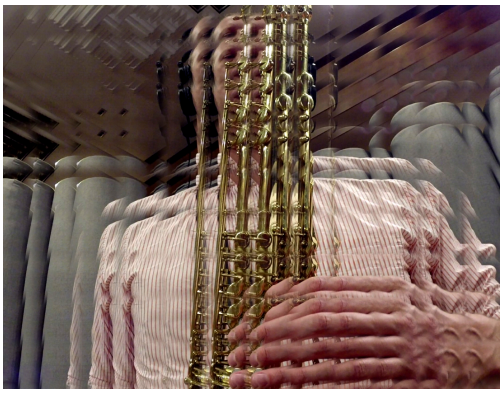
[10] N. Bogaards, C. Yeh, and J. Burred, "Introducing ASAnnotation: A tool for sound analysis and annotation," in *Proceedings of the International Computer Music Conference (ICMC)*, Belfast, Northern Ireland, 2008.

[11] A. G. Forbes, C. Jette, and A. Predoehl, "Analyzing intrinsic motion textures created from naturalistic video captures," in *Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP)*, 2014, pp. 107–113.

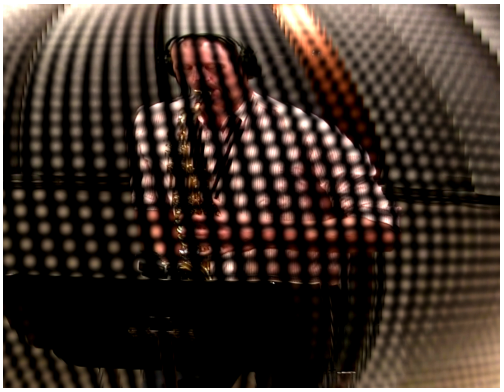
[12] A. G. Forbes, T. Höllerer, and G. Legrady, "Generative fluid profiles for interactive media arts projects," in *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, 2013, pp. 37–43.

[13] C. Roads, *Microsound*. MIT Press, 2004.

(a) Mixing two signals together



(b) Cloning fragments of the image



(c) Changing spatiotemporal positions of image fragments

(d) Rotating fragments of the images

Figure 5. Examples of various video processing effects.