

Research Report

IMAGE, TEXT, AND AUDIO INVERSIONS, CONVERSIONS, AND CONTORTIONS IN *EPIGLOTTIS*

Alexander SIGMAN
iCLA, Yamanashi Gakuin University

ABSTRACT

Based on Croatian artist Damir Ocko's video work *Spring* (for which the author provided musical material), which was included on a solo exhibition at the Palais de Tokyo in Paris called *The Kingdom of Glottis, epiglottis* (2013) is a piece in three parts for two voices, ensemble, live electronics, and video. The title *Spring* refers not to the season, but rather to a spring coil, to contents or components under pressure or other physical constraints. In response to this theme, Ocko's video consists of three primary elements: the recitation of poems written by the artist, performances by a contortionist, tight-rope walker, and fire-eater, and documentation of the volcanic Mount Stromboli in Sicily at various stages of eruption.

The primary intention behind *epiglottis* was to employ the text, image, and physical motion from *Spring* as raw source-materials for the instrument/voice parts, electronics, and visuals. Rather than simply reacting to the visible surface of the images and imagery in the texts utilized in *Spring*, the author sought to construct deeper, and perhaps more contorted relationships between the original video and its "setting" in *epiglottis* through analyzing and cross-modally re-synthesizing the physical properties of the still and moving images in the video. In so doing, the hierarchy between primary (image and text) and secondary (music/audio) elements in *Spring*, and by extension, the collaborative process between Ocko and the author, were inverted. In this paper, the image-to-sound and sound-to-image processing and conversion techniques, the translation of physical movements of the performers in the video onto physical parameters of musical performance, and approaches to text in *epiglottis* are discussed and illustrated.

1. THE KINGDOM OF GLOTTIS

In 2010-2012, the author collaborated with the Croatian visual artist Damir Ocko [1] on two video works entitled *The Moon shall never take my Voice*[2] and *Spring*[3]. In both cases, layers of instrumental and electroacoustic music samples were combined with image and narration. The pieces were exhibited at the Palais de Tokyo in Paris as part of a solo exhibition of Damir's called *The Kingdom of Glottis* in 2013.

The title *Spring* does not (necessarily) refer to the season... but rather to a spring-coil mechanism, springing into action, to matter and mechanisms under pressure and subject to various physical constraints. It is to these conno-

tations of the word "spring" that Ocko integrates images of the volcanic Mount Stromboli [4] at various stages of eruption, performances on a black-box set by a contortionist, tight-rope walker, and fire-eater—representing the human body operating under precarious conditions—and the recitation of poems written by the artist.

2. EPIGLOTTIS

In 2013, the Chicago-based Fonema Consort[5] commissioned the author for a work for two high voices and ensemble, with the possibility of including staged elements. In response, the author composed *epiglottis*—the title referring both to cartilage/valve at the entrance of the larynx and to the literal meaning of "over the glottis," as in "over/on (*The Kingdom of*) *Glottis*." Scored for two sopranos, flute(s), cello, contrabass, live electronics, and video, *epiglottis* was intended as a sort of convoluted commentary on *Spring*. Rather than simply reacting to the visible surface of the images and imagery in the texts utilized in the video piece, the author sought to construct deeper, and perhaps more contorted relationships between the original video and its "setting" in *epiglottis* through analyzing and cross-modally re-synthesizing the physical properties of the still and moving images in the video. In so doing, the hierarchy between primary (image and text) and secondary (music/audio) elements in *Spring*, and by extension, the collaborative process between Ocko and the author, were inverted. The techniques by which this cross-modal re-synthesis was achieved are described below.

epiglottis consists of three songs—each corresponding to an Ocko text—separated by brief audiovisual interludes, and concluding with a brief audiovisual postlude, for a total duration of ca. 13-14 minutes. The second song, entitled *Meat Puppet*, incorporates staging. Thus far, the piece has received two performances: at Constellation in Chicago by the Fonema Consort in February 2014[6], and by the Paxos Ensemble at the International Computer Music Conference (ICMC) in Athens in September 2014. Due to time and resource constraints, only the first and third songs—entitled (!) and *Nickering*, respectively—were presented on these occasions, interspersed with audiovisual interludes, for a duration of ca. nine minutes.

3. IMAGE-SOUND/SOUND-IMAGE

Using PhotoSounder[7], an image processing and analysis program, the author created several representative still im-

ages from the video of the contortionist, tight-rope walker, fire-eater, and volcano, which were analyzed for color density levels. This data was then converted into (audio) spectral information and used to synthesize audio samples. In turn, the newly generated audio files were analyzed and re-synthesized into images. While the produced sounds themselves—filtered noise of varying duration, sound levels, central frequencies, bandwidth, resonance, and contour—became the basic ingredients of the electronics component, the spectral data associated with these sounds determined the pitch content of the vocal and instrumental parts.

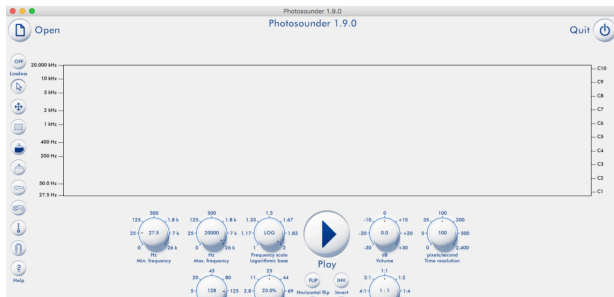


Figure 1. PhotoSunder interface.

As is made evident in Figure 1, the PhotoSunder interface enables the user to control several audio parameters such as frequency range and scale, audio file duration (in seconds), sound level, semitone quantization, and EQ. Malleable visual dimensions include pixel spray width, intensity, and gamma (envelope expansion/compression). Color inversion and image rotation are also possible. Spray and drawing tools allow for manual alteration of image color values. Furthermore, through layer duplication (or working with multiple distinct layers), one may construct time-varying filters, or perform cross-convolution. PhotoSunder spectrograms may be exported as soundfiles (AIFF, WAV, FLAC, or OGG), or as BMP image files. Audio and image files in most formats in most compressed and uncompressed formats may be imported. As such, this software may be employed purely for the editing/transformation of audio files, for image processing and/or conversion, or for the exporting of (audio) spectrograms as images. The author has explored a range of image-sound relationships both in *epiglottis* and in his *scremap* series, in which visual and audio data derived from an animation work become the raw data for electronic and instrumental material.

For *epiglottis*, the author experimented with manipulations of the aforementioned parameters, in the interest of producing contrasting audio output. At the same time, audio (and spectral images) related by inversion—of image orientation or color—were generated, such that there would be intrinsic structural connections between audio (and image) source materials.

In Figures 2-7, examples of PhotoSunder representations of images extracted from *Spring* are provided:

The spectral rendering of an audio sample used in the *Spring* soundtrack is illustrated in Figure 8:

For the electronics, audio derived from analyses of both Stromboli and performer images underwent further processing in Max 6[8]. As was mentioned previously, a subset of this audio was also employed to generate harmonic



Figure 2. Original image of Mount Stromboli.

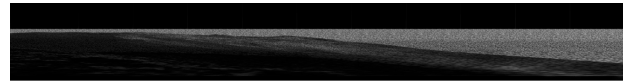


Figure 3. Mount Stromboli image, rendered and transformed in PhotoSunder.

structures applied to the vocal and instrumental parts. In these cases, frequency bandwidths spanning the ranges of the voices and instruments (respectively) were assigned to the image-derived source-sounds in question. These source-sounds were then imported into AudioSculpt[9], and subjected to partial tracking and chord sequence analyses.[10] The analyses were exported in SDIF (Sound Description Interchange Format), and rendered in notation via an OpenMusic[11] patch, and quantized accordingly (to the nearest semi-tone, in the case of the voices, and quarter-tone, in the case of the instruments).

With respect to image-to-sound conversion, PhotoSunder certainly presents limitations. First of all, it is only sensitive to color values, which can only be interpreted as spectral data, and lacks the “intelligence” to identify patterns, objects, faces, or other features. Secondly, images can only be scanned left-to-right. Given, however, that images utilized in *epiglottis* are in fact stills derived from a video, one could contend that the software functions as a kind of camera, panning across a fixed surface. By the same token, the author considered it necessary to somehow address and render not only still image attributes, but also the movements captured in *Spring*. These movements could then, in turn, be applied to the mobility of instrument and voice physical parameters, as well as the kinesis of the performers in general.

4. MOVEMENT GESTALTEN

In the first two principal sections of *Spring*, a contortionist performs a series of complex, repeated motions. E.g.:
<https://vimeo.com/84058909> and
<https://vimeo.com/84058910>.[12]

These motions were transcribed (by hand), resulting in a collection of four contours, of abstractions from the physical actions, referred to henceforth as *Gestalten*. Pairs of physical parameters were mapped to these contours, which were assigned to both voices and instruments. In effect, the performers reconstitute the anatomy, the moving parts of the human contortionist. Due to the changes



Figure 4. Volcanic clouds.

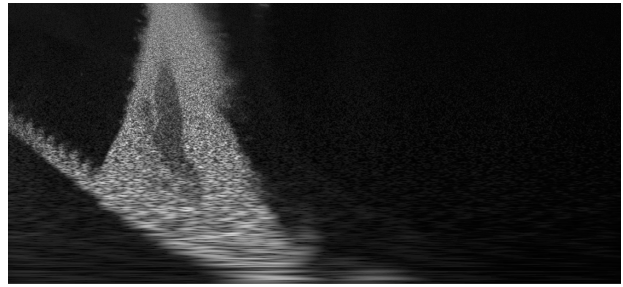


Figure 6. Tight-rope walker image.

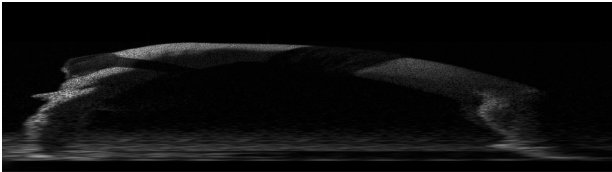


Figure 5. Contortionist image.

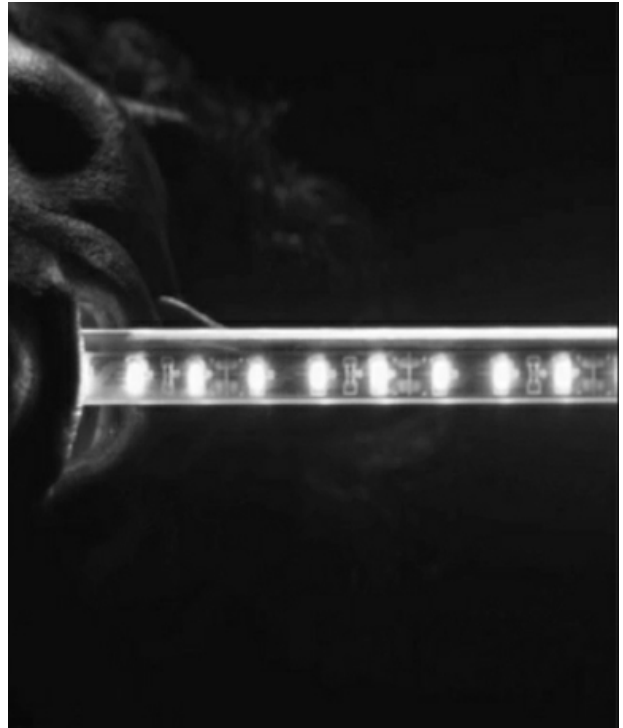


Figure 7. Fire-eater image.

in and rate of change of the parameter mapping and shifts in layer density, this “anatomy” takes on a volatile, fragile, and unpredictable character.

Figure 9 indicates the instrumental physical parameters to which these Gestalten have been mapped, and the attack types by which these physical parameter pairings are articulated: as isolated attacks, as dense collections of attacks, steady-state, sustained actions/resonances (with little or no detectable attack), or pulsating sustained actions/resonances:

In Figure 10, a realization of dynamic changes across pairs of physical parameters in the instruments from the middle of the first song is depicted. All attack-types are represented, and extend to the voices. Each instrument cycles through the four it Gestalten in a different order, at a different rate.

These Gestalten are applied to the staging and vocalizations of the singers in *Meat Puppet*, and the physical parameters of the instruments in *Nickering*. In the third song, pitch and dynamic contours are controlled via “unfolded,” linear versions of the original four Gestalten, in service of reflecting the relative linearity of the actions of the fire-eater in *Spring*, as compared to those of the contortionist.

5. TEXT AND VOICE

The *Spring* narrator recites four poems that Damir Ocko himself wrote for the project. In *epiglottis*, the author set three of these texts, as well as one that was originally intended for the video, but was ultimately discarded.

The first poem makes several allusions to resonance, ringing, melting, trembling, and cracking. The subsequent text consists of an “instruction manual” for constructing a vocalizing meat puppet using a pig’s lungs and larynx [13], and is set in the eponymous second song. Poems 3 and 4 appear in *Nickering*, and depict a schizophrenic state, while making reference to the glottis and to positions within the vocal apparatus. Through the gradual increases in tempo, rate of change, intensity, and exchange-

ing of texts between the singers in *Nickering*, comprehensibility progressively diminishes, enabling the physicality, the sonic properties of the poetry, to percolate to the foreground.

Recordings of the poem corresponding to *Meat Puppet*, recited by the *Spring* narrator and Ocko, are blended with the electronics, which include distortions of Ocko’s narration, in the interest of introducing a disfigured echo of the source-text narrative, in a manner akin to generating and editing spectrogram shadows of *Spring* images in PhotoSounder and incorporating them into the video (see below). However, the two layers are subjected to distinct live transformation. While the narration is ring-modulated by the live flute input, the non-narration electronics are cross-convolved with the cello and contrabass material. The live electronics components of the work are further addressed below.

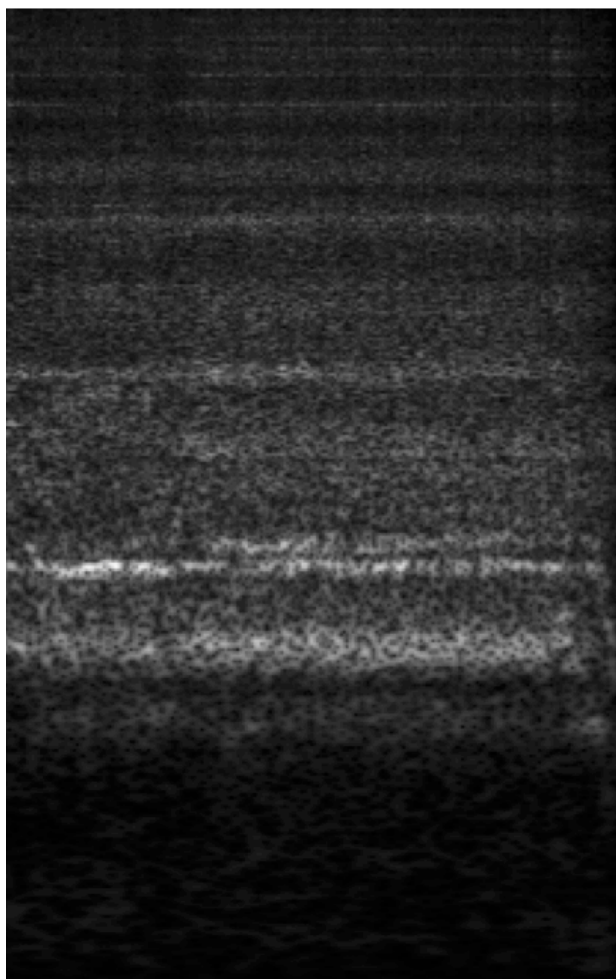


Figure 8. PhotoSounder spectral representation of edited audio sample derived from *Spring* soundtrack.

6. LIVE ELECTRONICS

In (!), the voices and instruments undergo live processing. Reflecting both the recurring imagery in the first poem and the Stromboli volcano footage, a resonant penumbra of varying harmonic content, density, and intensity surrounds the source-sounds. Pre-recorded video and electronics run concurrently.

The distinctions between real-time and records sources collapse in *Meat Puppet*. Besides the voice/instrument and audio interactions mentioned above, live footage of the staged actions of the singers is superimposed upon a series of pre-assembled images.

All live elements, as well as the triggering of audiovisual interludes, are controlled within a Max patch, presented in Figure 11.

7. VIDEO

The video component consists of four of types of material: 1) moving images of the contortionist; 2) stills derived from the tight-rope walker and fire-eater segments of *Spring*; 3) stills and video of Mount Stromboli at various levels of lava activity; and 4) black-and-white stills

A. Flute:

- 1) Vibrato speed (n.v./p.v./m.v.)
- 2) Pressure (minimal/normal/overblowing)
- 3) Pitch Content (breath/aeolian/normal)

C. VC/CB

- 1) Vibrato speed (n.v./p.v./m.v.)
- 2) Bow pressure
- 3) Finger pressure
- 4) Bow position (AST-> DSP)
- 5) Bow angle (crini -> c.l.t.)

II. Attack-Types:

- 1) Point
- 2) Granular
- 3) Steady-state
- 4) Pulsation

Figure 9. Instrument physical parameters and attack types.

“resynthesized” from the visual-audio-visual conversion process described and illustrated above. All video is manipulated and triggered using Jitter objects within the aforementioned Max patch.

Images conveying stages of volcanic eruption are distributed across the audiovisual segments of the piece (introduction, two audiovisual interludes, and audiovisual postlude). These eruption stages are depicted in Figures 12-15:

8. STAGING

In keeping with the content of the poem associated with *Meat Puppet*, one of the sopranos is requested to operate a crack-operated mechanical larynx—designed for laryngectomy patients [14]—and a plastic larynx model, which is typically employed in biomedical laboratories for educational/demonstration purposes. While holding up the mechanical larynx to the throat, the vocalist forms a series of vowels in the mouth. The plastic larynx model is activated at various points along its surface with a contact microphone in the scraping direction indicated. Meanwhile, the second soprano inflates and deflates a pair of either swine lungs or a balloon-lung model with a bellows. The audio output of these actions is amplified and balanced with the electronics.

9. FUTURE WORK

Besides a presentation of *epiglottis* in its entirety in the near future, it is intended in the coming years to complete a ca. 45-minute chamber opera encompassing the trilogy of video pieces represented on *The Kingdom of Glot-*



Figure 13. Audiovisual Interlude 1.



Figure 15. Audiovisual Postlude.



Figure 14. Audiovisual Interlude 2.

12. AUTHOR'S PROFILE

Alexander SIGMAN

Alexander Sigman's award-winning instrumental, electroacoustic, multimedia, and installation works have been featured on major international festivals, exhibitions, institutions, and venues across Europe, Asia, Australia, and the US. In June 2007, Sigman was Composer-in-Residence at the Musiques Démesurées festival in Clermont-Ferrand, France. Subsequently, he was awarded residency fellowships by the Akademie Schloss Solitude (Stuttgart, Germany), the Djerassi Foundation, the Paul Drescher Ensemble Artists Residency Center, and Gullkistan (Laugarvatn, Iceland). In 2013-2014, he undertook a musical research residency at IRCAM. He is currently Associate Professor and Chair of the Music program at the International College of Liberal Arts (iCLA) of Yamanashi Gakuin University in Kofu, Japan.

More information may be found here:
www.lxsigman.com

- [6] Here is a video of the premiere in Chicago: <https://vimeo.com/88258047> (accessed 10 September 2016).
- [7] <http://photosounder.com/> (accessed 10 September 2016).
- [8] <https://cyclimg74.com/products/max/#.V9UhOpN95PM> (accessed 10 September 2016).
- [9] <http://anasynt.ircam.fr/home/english/software/audiosculpt> (accessed 10 September 2016).
- [10] Although it may seem redundant to export from one spectral analysis program and import into another, PhotoSounder does not export SDIF files.
- [11] <http://repmus.ircam.fr/openmusic/home> (accessed 10 September 2016).
- [12] Accessed 10 September 2016.
- [13] Inspired by this video: <https://www.youtube.com/watch?v=2WNKSJQdbac> (accessed 13 September 2016).
- [14] Or a ratchet, if the crank-operated larynx is unavailable.