

# RECOMBINANT SPATIALIZATION FOR ECOACOUSTIC IMMERSIVE ENVIRONMENTS

*Matthew Burtner and David Topper,*

VCCM, McIntire Department of Music,  
University of Virginia  
Charlottesville, VA 22903 USA

mburtner@virginia.edu, topper@virginia.edu,

## ABSTRACT

An approach to digital audio synthesis is implemented using recombinant spatialization for signal processing. This technique, which we call Spatio-Operational Spectral Synthesis (SOS), relies on recent theories of auditory perception, especially research by Kubovy and Bregman. In SOS, the perceptual spatial phenomenon of objecthood is explored as an expressive musical tool. In musical applications of these theories, we observe the emergence of a "persistence of audition" exposing interesting opportunities for compositional development.

In essence, SOS, breaks an audio signal into salient components then recombines and spatializes them in a multichannel environment. Following an introduction to the technique and several examples demonstrating potential applications, this paper concentrates on some applications of the technique in ecoacoustic compositions by Matthew Burtner, *Anugi Unipkaa*, *Sikniq Unipkaa* and *Siku Unipkaa*. These works draw on environmental systems as models for multichannel processing.

## 1. INTRODUCTION

Spatial techniques in music composition can be traced at least to the 16th century. In the Venetian polychoral antiphonal tradition in the late 16th and early 17th centuries, composers composed for multiple choruses set around the space, creating a *cori spezzati* or *split chorus*. From the two choir works of Willaert, ca. 1580 the tradition of Cori Spezzati evolved into an elaborate practice in the music of Giovanni Gabrieli.

The electroacoustic multichannel tradition has roots back to Varese's *Poeme Electronique* (1958) in which over 400 loudspeakers routed multichannel sound throughout the Philips Pavilion in the Brussels World Fair. These techniques, including the more recent practices of electroacoustic music, have concentrated on the projection of coherent sound object or objects into a defined space.

Spatio-Operational Spectral Synthesis or SOS, is a

signal processing technique based on recent psychoacoustic research. The literature on auditory perception offers many clues to the psychoperceptual interpretation of audio objecthood as a result of streaming theory. Streaming describes audio objects as sequences displaying internal consistency or continuity (McAdams and Bregman 1979). Bregman has further defined a stream as, "a computational stage on the way to the full description of an auditory event. The stream serves the purpose of clustering related qualities (Bregman, 1999)." Thus it becomes the primary defining factor of an acoustic object.

SOS breaks apart an existing algorithm (ie, Additive Synthesis, Physical Modeling Synthesis, etc.) into salient spectral components, with different components being routed to individual or groups of channels in a multichannel environment. Due to the inherent limitations of audition, the listener cannot readily decode the location of specific spectra, and at the same time can perceive the assembled signal. In this sense, the nature of the auditory object is altered by situating it on the threshold of streaming, between unity and multiplicity.

The "Theory of Indispensable Attributes" (TIA) proposed by Michael Kubovy (Kubovy and Valkenburg, 2001) puts forth a framework for evaluating the most critical data the mind uses to process and identify objects. In the case of audio objects, TIA holds that pitch is an indispensable attribute of sound while location is not, simply put, because the perception of audio objects can not exist without pitch. His experiments have demonstrated that pitch is a discriminating factor the brain seems to use in distinguishing sonic objecthood, whereas space is not as critical.

Bregman notes that conditions can be altered to make localization easier or more difficult, so that, "conflicting cues can vote on the grouping of acoustic components and that the assessed spatial location gets a vote with the other cues. (Bregman p305)": "Curious about how Kubovy's and Bregman's theories could be utilized for signal processing, we began applying spatial processing algorithms to spectral objects.

When spectral parameters are spatialized in a certain manner the components fuse and it is impossible to localize the sound, yet when they are spatialized differently the localization or movement is predominant over any type of spectral fusion. Creatively modulating between fusion and separation is where SOS comes into being. One of our main questions is this: if the mind does not treat location as indispensable, can SOS force the signal into an oscillation between unity and multiplicity by exploiting spatialization of the frequency domain?

The technique exploits what might be called a "Persistence of Audition" insofar as the listener is aware that auditory objects are moving, but not always completely aware of where or how. This level of spatial perception on the part of the listener can also be controlled by the composer with specific parameters.

SOS is essentially a two-step operation. Step one consists of taking an existing synthesis algorithm and breaking it apart into logical components. Step two re-assembles the individual components generated in the previous step by applying various spatialization algorithms. Figure 1 illustrates the basic notion of SOS as demonstrated in the following example of a square wave.

## 2. SOS ADDITIVE SYNTHESIS

In initial experiments testing SOS we used simple mathematical audio objects such as a square wave generated by summing together sinusoids having odd harmonics and inversely proportional amplitudes. Formula (1) describes the basic formula used in this initial example:

$$x_s(t) = \sin(\omega_0 t) + 1/3 \sin(3\omega_0 t) + 1/5 \sin(5\omega_0 t) \dots \quad (1)$$

In this experiment the first eight sine components of the additive synthesis square wave model were separated out and assigned to a specific speaker in an eight-channel speaker array. Although the square wave is spatially separated, summation of the complex object is accomplished by the mind of the listener (Figure 1).

Separation need not be completely discrete however. Any number of sinusoids can be used and animated in the space, sharing speakers. In a simple extension of this example sinusoids were used to generate a sawtooth wave as shown in Formula (2).

$$x_s(t) = \sin(\omega_0 t) + 1/2 \sin(2\omega_0 t) + 1/3 \sin(3\omega_0 t) \dots \quad (2)$$

When the sinusoids were played statically, in separate speakers, the ear can identify the weighting of the frequency spectrum between different speakers. For

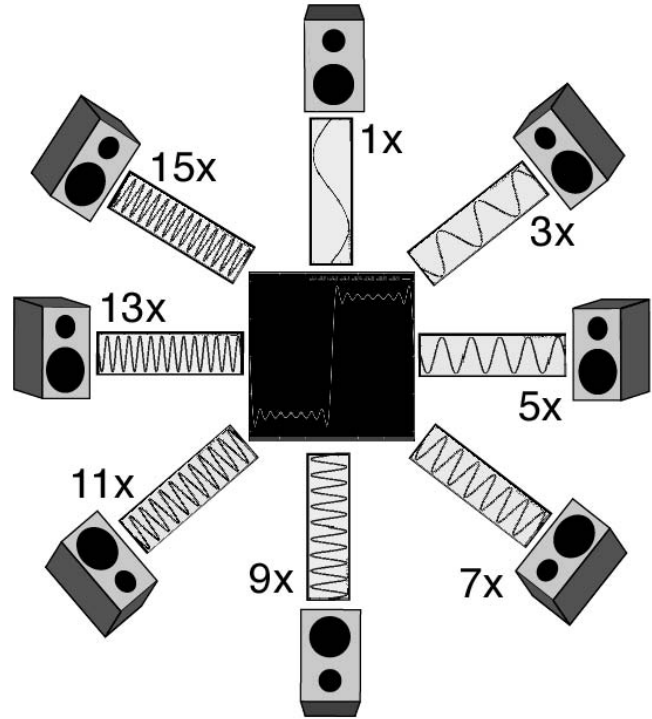


Figure 1. SOS Recombinant Principle.

example, if the fundamental is placed directly in front of the listener and each subsequent partial is placed in the next speaker clockwise around the array, a slight weighting occurs in the right front of the array. The First Wavefront law would of course suggest this, but in actuality the blending of the sinusoids into a square wave is more perceptible than the sense of separation into components. In fact, the effect is so subtle that a less well-trained ear still hears a completely synthesized square wave when listening from the center of the space.

Animating each of the sinusoids in a consistent manner exhibits a first example of the SOS effect. By assigning each harmonic a circular path, delayed by one speaker location in relation to each preceding harmonic, the unity of the square wave was maintained but each partial also began to exhibit a separate identity. This of course is the result, in part, of phase and shifting (eg., circularly moving) amplitude weights. The mind of the listener, tries to fuse the components while also attempting to follow individual movement.

This simple example illustrates how the Precedence Effect can be confused so that the mind simultaneously can cast conflicting cognitive votes for oneness and multiplicity in the frequency domain. This state of ambiguity, as a result of spatial modulation, is what we call the SOS effect.

We experimented with different rates of circular modulation of each sine component. Interestingly, each relationship was different but not necessarily more

pronounced than the similar, delayed motion. Using the same, non-time-varying signal, a time-varying frequency effect can be achieved due to spatial modulation using only circular paths in the same direction. Figure 2 illustrates this type of movement.

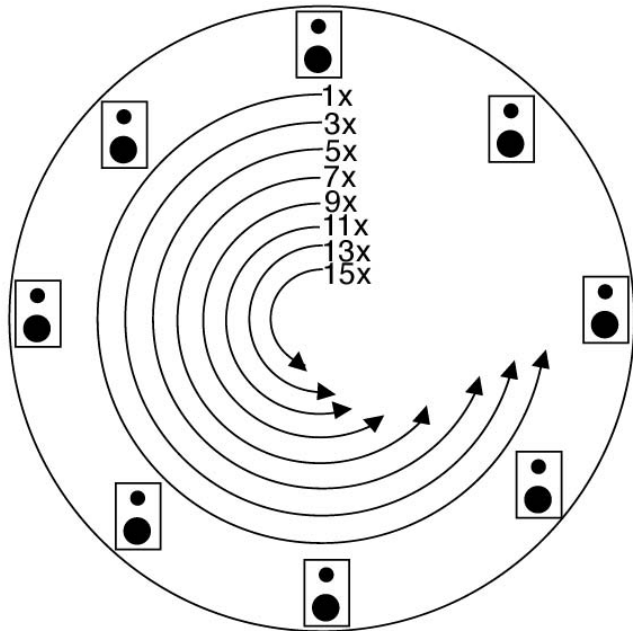


Figure 2. SOS with varying rate circular spatial path of the first eight partials of a square wave

An early example of spectral separation of this sort has been implemented in Roger Reynolds' composition, *Archeipelago* (1983) for orchestra and electronics (Bregman p296). In tests done at the IRCAM, Reynolds and Thierry Lancino divided the spectrum of an oboe between two speakers and added slight frequency modulation to each channel. If the FM were the same in both channels the sound synthesized, but if different FM were added to each channel, the sounds divided into two independent auditory objects.

In our later tests, we noticed similar results to Reynolds and Lancino, even within the context of animated partials. By exaggerating the movement of one partial, either by increasing its rate of revolution, or assigning it a different path, the partial in question stood out and the SOS effect was somewhat reduced. By varying the amount of oscillation and specific paths of different partials, the SOS effect can be changed subtly.

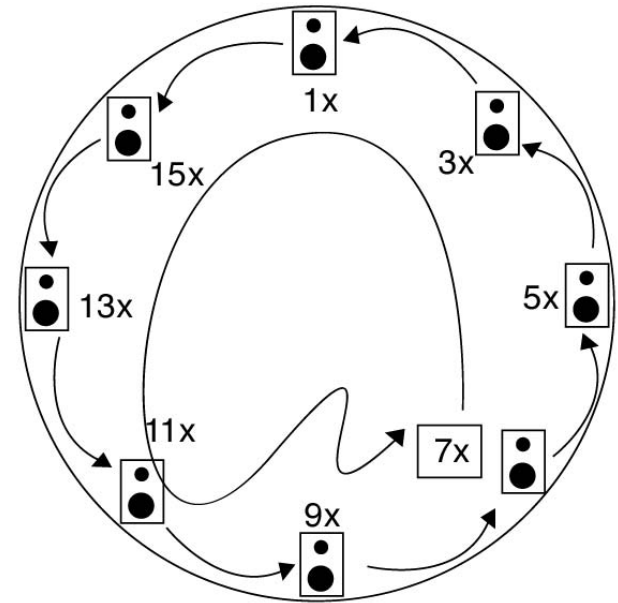


Figure 3. SOS with one partial moving against the others moving in a unified circular motion.

### 3. DEFINITIONS OF SOS SPATIAL ARCHETYPES

Any number of spatialization algorithms can be applied to the separated components' variables or audio stream. The types of spatialization employed by SOS can be thought of as having two attributes: motion and quality. A series of archetypal quality attributes were explored in a two dimensional environment.

Motion was divided into three categories:

- 1) static: no motion
- 2) smooth: a smooth transition between points
- 3) cut: a broken transition between points

Quality was divided into five archetypal forms:

- 1) circle: an object defines a circular pattern
- 2) jitter: an object wobbles around a point
- 3) across: an object moves between two speakers
- 4) spread: an object splits and spreads from one point to many points
- 5) random: an object jumps around the space between randomly varying points

These archetypes can be applied globally, to groups, or to individual channels. Each archetype has specific variables that can be used to emphasize or de-emphasize the SOS effect. Variables can also be mapped to trajectory or rate of change, defined by a time-varying function, or generated gesturally in real time.

#### 4. SOS FILTER SUBBAND DECOMPOSITION

The balance between frequency separation and sonic object animation became much more complicated when we attempted to apply our initial technique to an audio signal. Our initial tests assigned eight simple two pole IIR filter outputs to discrete speaker locations. Selection of the ration between the filters became a critical component in being able to achieve any effect at all. With filters set to frequencies that were not very strong in the underlying signal, the filters tended to blend together and sound as if some type of combined filtering were taking place rather than SOS. Similarly, when spatialization algorithms were applied with an improper filter weight, the underlying movement was more apparent than the separation.

We tested the filter technique with both white noise and live instrument (eg., Tenor Saxophone). The former of course offered much more flexibility with respect to frequency range and filter setup. The saxophone signal used, having the majority of its spectrum located between 150Hz and 1500Hz (with significant spectral energy up to approximately 8000Hz) suggested a filter/bandwidth weighting of: 32/5Hz, 65/15Hz 130/30Hz, 260/60Hz, 520/120Hz, 1000/240Hz, 2000/500Hz, 4000/1000Hz.

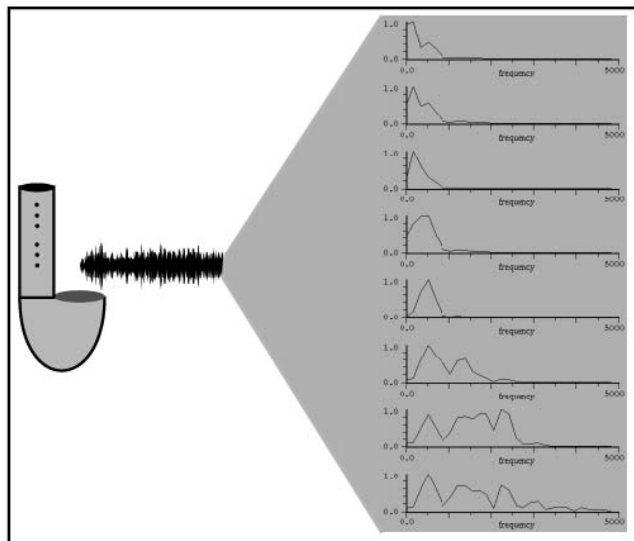


Figure 4: Saxophone signal subband filter decomposition for SOS.

#### 5. SOS ECOACOUSTIC EMMERSIVE ENVIRONMENTS=

Multichannel composition has a basis in acoustic ecology through Soundscape composition (Truax, 1978/99, 1994). Multichannel soundscape compositions reconstruct sonic environments through the sampling and redistribution of

distinct sounds to construct externally referential environments. A related area of research is ecoacoustics, an approach that derives musical procedures from abstract environmental systems, remapping data into structural musical material. it is a form of sonification for ecological models (Keller 1999, 2000).

In the most general sense, ecoacoustics is a type of environmentalism in sound, an attempt to develop a greater understanding of the natural world through close perception. In the field of composition, this takes the form of musical procedures and materials that either directly or indirectly draw on environmental systems to structure musical material.

In *Winter Raven* (Burtner 2001), a large scale work for instrumental ensemble, 8-channel computer-generated sound, three video projections, dance and theater, SOS techniques were implemented in a multimedia context. Each of the three acts of *Winter Raven* contains one *Unipkaaq* or “story” in Unupiaq Inuit language. Each of these pieces is scored for 8-channel computer-generated sound using SOS techniques, percussion, and a dancer wearing a specially constructed mask. The masked dancer represents a magical character playing a shamanic role in the evolution of the piece.

The Shaman character uses three different masks in *Ukiuq Tulugaq*, representing Sun, Ice and Wind. Each mask is distinguished by different choreography, music and video processing. An interface written with Isadora, processes the incoming live video and layers it with prerecorded video. The electronics from these three movements contain different SOS processing of the electronic sound. Each spatialization model corresponds to a dance mask with interactive video. The combination of video and multichannel audio evoke a personification of the environmental elements of sun, ice and wind. In Figure 8, the live video is shown above the corresponding staged scene.

In the first of these three pieces, *Siknik Unipkaaq* (*the story of sun*), a group of interlocking concentric planal paths were created (figure 5).

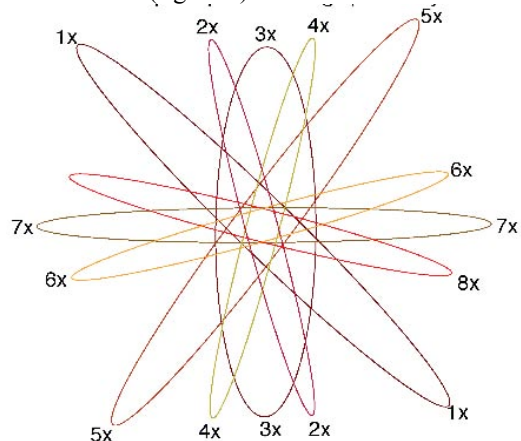


Figure 5: *Siknik Unipkaaq* SOS processing

Spatial modulation tempo ratios of 1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 were employed for the eight independent paths of audio. The base tempo of the structure was modulated globally, accelerating from a time base of 1 = 120" to a time base of 1 = 20". This yields a meta-tempo structure of 120" : 60" : 40" : 30" : 24" : 20" : 17" : 15" which is gradually collapsed into a mesa-tempo structure of 20" : 10" : 6.7" : 5" : 4" : 3.3" : 2.8" : 2.5".

In addition to the electronics, a battery of percussion helps articulate the perpetual motion of this composition. Two percussionists playing timpani and cymbals create slow crescendo/decrescendo pulses. Two other percussionists play congas, bass drum and floor toms, following a repetitive pattern derived from the spatial motion. Both the repeated dynamic changes of the timpani/cymbals and the repeated rhythmic patterns of the drums, help underscore the cyclical motion of the computer-generated sound.

In *Siku Unipkaa*q (the story of ice) a "shaking" algorithm was employed to model the freezing of motion in the spatial domain. Each component of the ice sound pans between two randomly selected points very rapidly and gradually reduces movement, increasing frequency. The panning occurs on the order of 600 to 20 milliseconds, varying for each particle of sound. The result is a feeling of gravity pulling the sound towards a single point between the two spatial anchors. Thus the sound is "frozen" into multifaceted crystals, continually spawning new paths that are again frozen. At any given time there are four simultaneous paths of shaking. In addition, the ice sound is played out of each speaker quietly to create a background into which the shaking algorithm can blend smoothly. Figure 6 depicts this motion type.

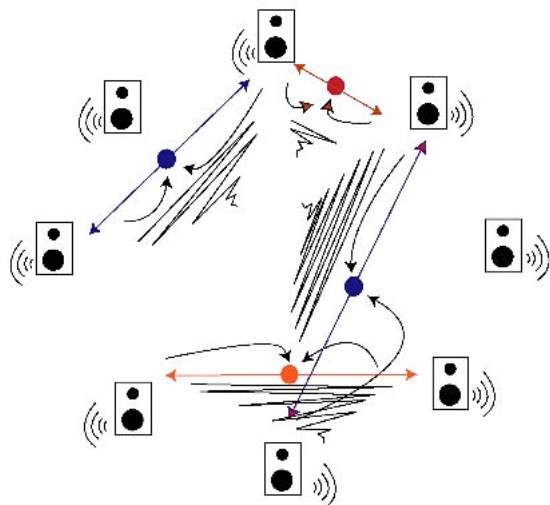


Figure 6: *Siku Unipkaa*q SOS "shaking" algorithm

A global freezing process is created by two glockenspiel played by four players. Over the course of the four minutes

of the piece, the density and variety of pitches are reduced, focussing the frequency energy into reduced bands of sound. Finally, the voices slow and freeze into individual points in the frequency spectrum.

*Anugi Unipkaa*q (the story of wind) most effectively captures the principle of SOS in this group. The source material of the work is the sound of wind recorded in Alaska. The wind is band pass filtered to isolate individual frequency regions of the sound. In this sense it is treated as the saxophone signal in the experiment discussed previously. Four such independent wind bands are created from the original source.

Each excerpted wind channel is panned rapidly between groups of randomly selected speakers. The path accelerates logarithmically, speeding up as it approaches its target point. In figure 7, each straight line represents this accelerating curve. Amplitude is tied to spatial change such that the wind sounds crescendo into each new location. The bands of wind rush simultaneously around the space, creating a kind of SOS blizzard of wind.

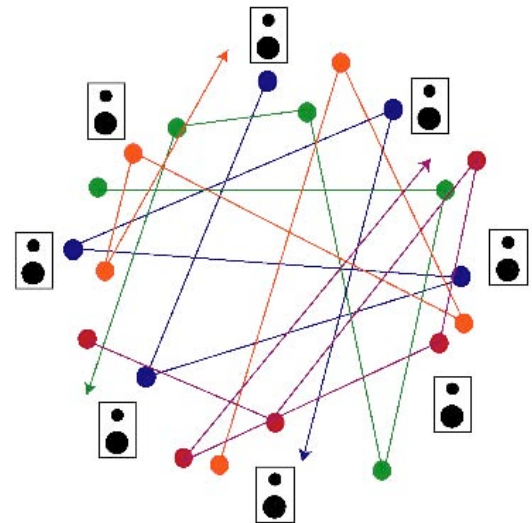


Figure 7: *Anugi Unipkaa*q SOS spatial motion "blizzard" algorithm

Accompanying the spatialized four winds are four percussionists. The piece is scored for a solo percussionist who plays a battery of toms and drums. The other three players are gathered around a single large bass drum, playing it simultaneously. At the end of the piece, as the rhythmic structure concentrates into a single common rhythm, the solo percussion joins the other players at the large bass drum and they end together. The four players focussed around a single point on the stage create a kind of focus for the four winds thrashing around the hall.

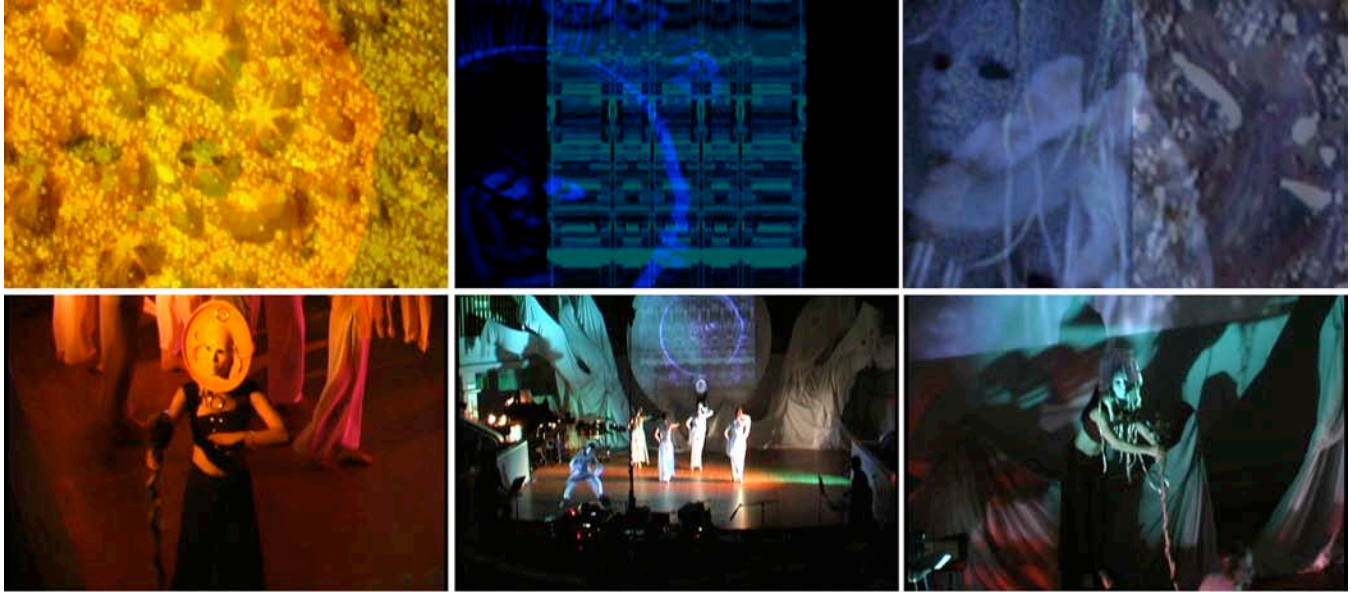


Figure 6. Each column above shows the processed video (above) and mask dancer (below). The rows from left to right show: *Siknik Unipkaa* (the story of sun) *Siku Unipkaa* (the story of ice) *Anugi Unipkaa* (the story of wind)

## 6. FUTURE DIRECTIONS

Current SOS research has been done primarily in a two dimensional environment. Exploring a three dimensional environment will increase the effect of spatialization algorithms and offer a greater means of separation for various models (ie, 3D waveguides).

So far, only the authors who agreed on the results have performed listening tests. Future work consists of testing more subjects, in order to see if the segregation of the synthesis algorithms is performed in the same way by human listeners.

Much of the psychoacoustic research that inspired SOS also looks at the related phenomenon of audio streaming, in sequential segregation. In addition to exploring SOS based on "spectral" separation, it would be interesting to explore sequential stream separation and granular synthesis.

With respect to the creative applications of SOS, the work described here has relied on macro-level procedures and more work on micro-level structures (eg particle-based synthesis) is anticipated. In addition, stronger and more concrete sonification algorithms will help articulate the ecoacoustic compositional strategies. Further integration of the video aspects of the works with SOS would also be advantageous.

## 7. REFERENCES

- [1] A. S. Bregman. *Auditory Scene Analysis: the perceptual organization of sound*. MIT Press, Cambridge, MA, 1999.
- [2] M. Burtner, *Ukiuq Tulugaq (Winter Raven)*. Doctoral of Musical Arts Thesis. Stanford University, Stanford, California. 2001.
- [3] M. Burtner, D. Topper, S. Serafin. *S.O.S. (Spatio-Operational Spectral) Synthesi*. Proceedings of the Digital Audio Effects (DAFX) Conference. Hamburg, Germany, 2002.
- [4] D. Keller. *Social and perceptual dynamics in ecologically-based composition*. Proceedings of the VII Brazilian Symposium of Computer Music, Curitiba, PN: SBC. 2000.
- [5] D. Keller, (1999). *touch'n'go: Ecological Models in Composition*. Master of Fine Arts Thesis. Burnaby, BC: Simon Fraser University. 1999.
- [6] M. Kubovy, D. V. Valkenburg. "Auditory and Visual Objects," *Cognition*. 80, p97-126. 2001.
- [7] S. McAdams, and A. Bregman. "Hearing Musical Streams." *Computer Music Journal*. vol. 3 num. 4. CA., 1979.
- [8] B. Garton, and D. Topper. "RTcmix -- Using CMIX in Real Time," Proc. of *International Computer Music Conference (ICMC)*, Thessalonika, Greece, 1997.
- [9] D. Topper. "PAWN and SPAWN (Portable and Semi Portable Audio Workstation)." Proc. of *International Computer Music Conference (ICMC)*, Berlin, Germany., 2001.

- [10] B. Truax. ed. "Handbook for Acoustic Ecology." Arc Publications, Cambridge Street Publishing, CD-ROM Edition, Version 1.1. 1978/1999.
- [11] B. Truax. "Discovering Inner Complexity: Time-Shifting and Transposition with a Real-time Granulation Technique," *Computer Music Journal*, 18(2), 1994, 38-48 (sound sheet examples in 18(1)).