City Research Online

City, University of London Institutional Repository

Citation: Einbond, A. (2017). Mapping the Klangdom Live: Cartographies for piano with two performers and electronics. Computer Music Journal, 41(1), pp. 61-75. doi: 10.1162/COMJ_a_00397

Aaron Einbond
Department of Music
City, University of London
Northampton Square,
London EC1V 0HB, UK
Aaron.Einbond@city.ac.uk

Mapping the Klangdom Live: *Cartographies* for Piano with Two Performers and Electronics

Abstract: The use of high-density loudspeaker arrays (HDLAs) has recently experienced rapid growth in a wide variety of technical and aesthetic approaches. Still less explored, however, are applications to interactive music with live acoustic instruments. How can immersive spatialization accompany an instrument already with its own rich spatial diffusion pattern, like the grand piano, in the context of a score-based concert work? Potential models include treating the spatialized electronic sound in analogy to the diffusion pattern of the instrument, with spatial dimensions parametrized as functions of timbral features. Another approach is to map the concert hall as a three-dimensional projection of the instrument's internal physical layout, a kind of virtual sonic microscope. Or, the diffusion of electronic spatial sound can be treated as an independent polyphonic element, complementary to but not dependent upon the instrument's own spatial characteristics. *Cartographies* (2014), for piano with two performers and electronics, explores each of these models individually and in combination, as well as their technical implementation with the Meyer Sound Matrix3 system of the Südwestrundfunk Experimentalstudio in Freiburg, Germany, and the 43.4-channel Klangdom of the Institut für Musik und Akustik at the Zentrum für Kunst und Media in Karlsruhe, Germany. The process of composing, producing, and performing the work raises intriguing questions, and invaluable hints, for the composition and performance of live interactive works with HDLAs in the future.

Background

The richness and irreproducibility of acoustic instrumental sound comes from complex interdependencies between timbre and space (Otondo et al. 2002). Only recently, however, do computer tools make it possible for composers to strategize a high-level control of these dimensions for spatialized sound synthesis, an *écriture* of space and timbre (Warusfel and Misdariis 2001). The recent confluence of tools for analysis and synthesis based on audio descriptors, compositional control of spatialization, and high-density loudspeaker arrays (HDLAs) for rendering spatialized sound in three-dimensions encourages synergies of timbre and spatial synthesis that can be exploited for musical expression.

Instrument Directivity

Research into the spatial directivity patterns of acoustic instruments goes back a half century; for a review see Frank Zotter's PhD dissertation (2009, p. 89). Research in reproducing instrumental diffusion patterns is more recent, however. In the ORA project (d'Alessandro et al. 2009), researchers at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) explored the projection of the interior space of the pipe organ into the space surrounding the audience by applying techniques from higher-order Ambisonics (HOA). In further case studies, instrumental directivity patterns were measured with a spherical microphone array and recreated using Wave-Field Synthesis (WFS) and spherical loudspeaker arrays (Noisternig, Zotter, and Katz 2011). Another precedent is the use of a controlled directivity source to reproduce instrumental sound, such as the spherical 120-loudspeaker array developed at the Center for New Music and Audio Technologies (CNMAT), Univeristy of California, Berkeley, in collaboration with Meyer Sound (cf. Avizienis et al. 2006). This allows for the detailed mapping of sound to a three-dimensional focused directivity pattern, in emulation of an acoustic instrument, implemented using WFS techniques (Schmeder and Noisternig 2010). A compositional application of WFS and HOA combined interactively with a live instrument is Rama Gottfried's *Flouresce* for violoncello and electronics (Gottfried 2012). Although this project considered compositional analogies between instrumental timbre and space, it did not use spatial microphone

Computer Music Journal, 41:1, pp. 61–75, Spring 2017
doi:10.1162/COMJ_a_00397

placement to reproduce the instrument's geometry electroacoustically.

Timbre Space

In contrast to instrumental directivity patterns, one can give a different account of spatial and timbral perception as a cognitive construct. Studies by Grey (1977) and Wessel (1979) have pointed to the multidimensional nature of timbral hearing, and to spatial models as possible representations of this aural reasoning. Although questions remain open about the details of timbral perception and its variation over musical and cultural contexts, timbre space can nevertheless be taken as a fruitful model to guide specific creative situations. *Cartographies* embraces these dual relationships of timbre and space, both as objective physical interaction and subjective mental map.

Concatenative Synthesis

To retrieve fine timbral details of its sonic materials, *Cartographies* takes advantage of the technique of corpus-based concatenative synthesis, permitting high-level control of sound synthesis based on audio features. Corpus-based concatenative synthesis has been implement in the software CataRT, with its associated signal processing package FTM & Co., by Diemo Schwarz and the Sound Music Movement Interaction team at IRCAM (Schwarz 2007). CataRT allows for the analysis and segmentation of a database of samples recorded live or in deferred time, the *corpus*, and resynthesis through to a variety of control paradigms. It gives access to the full richness of time-domain audio, combined with the fine control of sonic *descriptors*, defined as any characteristic extracted from the source sound or higher-level features attributed to it.

Sounds are initially segmented algorithmically or manually into short *units*, or grains. During synthesis, units are selected from the corpus based on their descriptors, usually according to Euclidean distance from desired target values. Units may be further manipulated using granular synthesis parameters before being overlapped or *concatenated*, and sent to output. When the target descriptor values are derived from the analysis of another, longer target sound, then the resulting synthesis can be described as an *audio mosaic*. This mosaic may resemble the target to varying degrees, as a visual mosaic may resemble its subject. The user interface of CataRT includes a multidimensional plot of units organized by selected descriptors (see Figure 1). An analogy can be drawn between this representation of timbral features and the notion of timbre space.

Previous work extends corpus-based concatenative synthesis as a tool for computer-assisted composition and real-time treatment. In *What the Blind See* (2009), for five performers and electronics, each instrument is amplified with contact microphones (harp, piano, bass drum, and tam tam) or proximate microphones (viola and bass clarinet) to isolate its delicate inner sound world. These live instrumental signals are compared by a variety of descriptors to a prerecorded corpus of instrumental samples, also recorded by closely placed microphones. The samples are then concatenated during performance to produce a shadow of live instrumental timbre (Einbond, Schwarz, and Bresson 2009). The same process can also be transcribed into instrumental notation to be reinterpreted acoustically, a technique termed "corpus-based transcription." In this way the timbre of a recorded sound—for example a field recording—is mapped to a score for live performance. The resulting "instrumental audio mosaic" can fuse seamlessly with the source recording; but the mosaic also can be used to mask or oppose its source, navigating different degrees of relation and reference.

What the Blind See also introduces the process of "corpus-based spatialization" to map timbral descriptors to spatial trajectories (Einbond and Schwarz 2010). The goal is to guide real-time spatialization for recorded sounds, or even live sounds whose descriptor values are not known in advance, along preplanned descriptor templates. Samples are compared to existing spatialization schemata according to their descriptor values and placed in space at the appropriate location. As the instrumentalists perform, the microscopic timbral details of their actions are analyzed and mapped

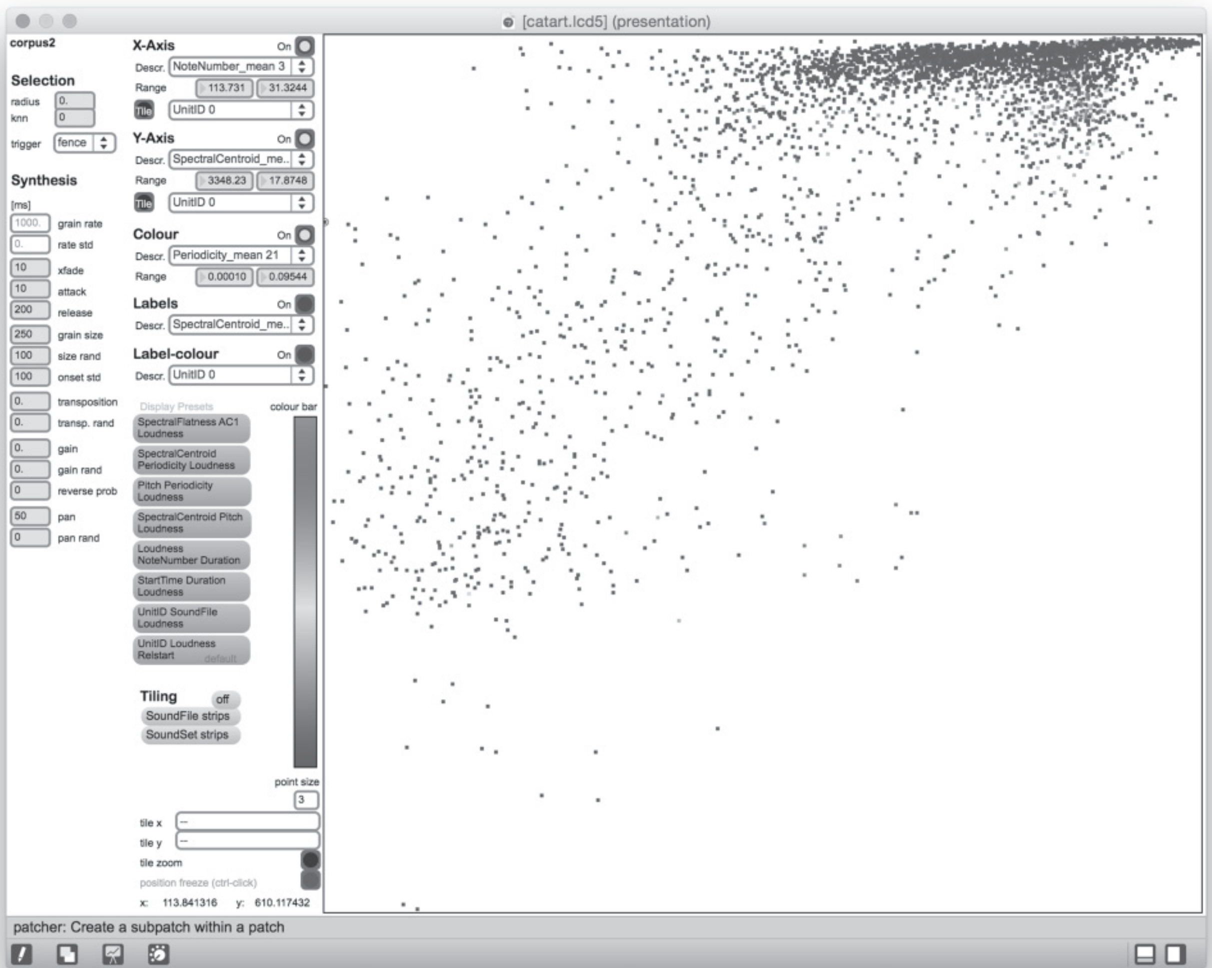


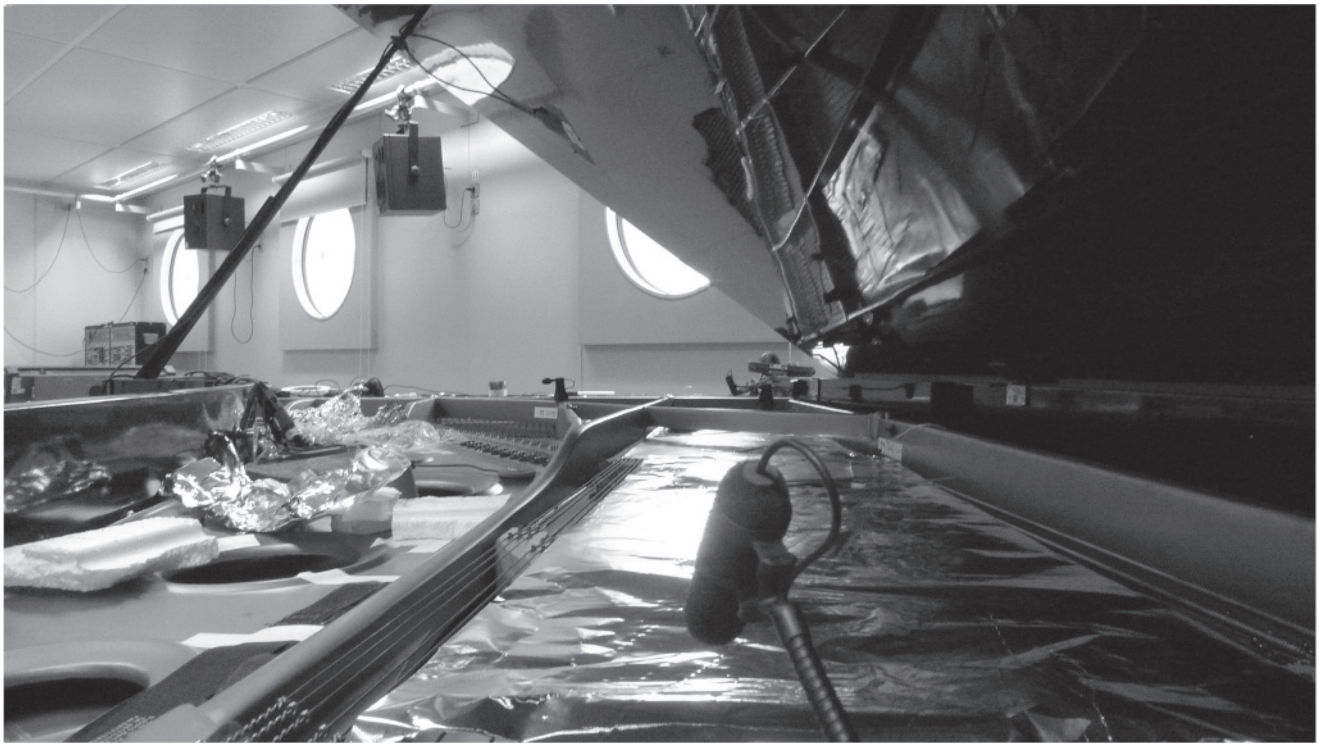
Figure 1. Screenshot of CataRT showing a sample corpus (a database of samples) used in Cartographies. *Points represent grains of piano samples that are plotted along axes representing audio descriptors, in this case MIDI note number (*x*-axis) and spectral centroid (*y*-axis). The plot is used to organize both timbral structure and spatialization strategies in the composition.*

to a spatial trajectory in the concert hall, which may differ from one performance to another. Like acoustic instrumental directivity patterns, this presents a correlation between timbral features and spatial distribution, albeit in a more abstract way.

Motivation for *Cartographies*

Cartographies extends corpus-based spatialization by projecting the inside of the piano out to immerse the public in a magnified view of the instrument. A key dimension is the isolation of sonic details from their physical sources through recording and amplification. This could be compared to Pierre Schaeffer's "reduced listening," or Helmut Lachenmann's *"musique concrète instrumentale,"* in which playing techniques approach independent sound objects with their proper forms and morphologies. The effect is enhanced by contact or proximity microphones placed in the piano interior, where the performers act with their hands,

Figure 2. Microphones and materials in the piano interior at the Experimentalstudio. Materials used include aluminum foil, knitting needles, Velcro, styrofoam, felt, and plastic wrap.



percussion mallets, and objects to elicit sounds barely audible without amplification. The minute scale serves both to reduce the effect of masking of amplified and spatialized sound by the live source, and to defamiliarize the sonic identity of this most "conventional" of instruments.

Unlike Lachenmann, for whom the physicality of sonic gestures relates back to their "mechanical origins" (Ryan 1999), the sound world of *Cartographies* focuses on nearly invisible actions that belie their sources. Bent over the instrument, the performers produce microscopic gestures that are obscured from the public's view; the results are transmitted through microphones and projected into the concert space to reproduce a larger-than-life map of the piano interior. See Figure 2 for a view of the installation of the microphones and some of the materials sampled and played by the performers inside the piano, including aluminum foil, knitting needles, Velcro, styrofoam, felt, and plastic wrap. Although some objects resemble John Cage's piano preparations in the *Sonatas and Interludes* (1948), they differ in acting directly on the piano strings and frame, not mediated by the keyboard and hammers. They are more indebted to the inside-piano improvisation practice of Andrea Neumann (Haenisch 2013).

Cartographies makes reference to an extramusical source as well: the cosmic microwave background (CMB), the faint electromagnetic radiation that reaches the earth from distant space, possible evidence of the first moments of the inflation of the universe after the Big Bang. Although this radiation surrounds us in all directions, it only allows us dimly to apprehend its source: like the sounds of the piano that are only partially revealed through their capture and projection by a network of microphones and dome of loudspeakers. To underline this metaphor, a sonification of CMB data is incorporated as a "found sound object" at several points in the work.

Spatialization Models

During the composition and production of *Cartographies* three spatialization models were explored and each incorporated at different points in the work: projection of the piano interior through analog amplification, corpus-based spatialization of piano samples, and independent spatial trajectories convolved with piano resonances.

Projection

The first approach to spatialization is accomplished in the simplest possible way: Twelve microphones are installed in the grand piano at the Zentrum für Kunst und Media (ZKM), and each is mapped to a virtual source position in the Klangdom. (The piano is a Yamaha seven-foot model S6, similar in dimensions to a Steinway B.) There are several differences from other, earlier examples of three-dimensional instrumental recording and diffusion. First, such recording-diffusion scenarios are often applied in deferred time, whereas this was a real-time application. Second, although Ambisonic recording involves regularly spaced microphone placement, here the microphones were positioned subjectively, based on the instrument's geometry, tone color, and accessibility for the performers (without obstructing their hand movements or sight lines). For this reason compact lavalier and transducer microphones were used, as listed in Table 1. Last, although multichannel microphone placement usually attempts to capture the instrument's diffusion pattern through air, for this project microphones were positioned either in extreme proximity to or in contact with their sound sources. Rather than capturing a "realistic" sonic image of the piano to be reproduced virtually, this approach favored a more abstract mapping that could be manipulated artificially.

The chosen microphone positions and virtual source positions are shown in Figures 3 and 4. Initial experimentation revealed that, when the internal positions of the microphones are mapped consistently to diffusion positions, recognizable spatial effects can be reproduced. For example, when "circular bowing" with a contrabass bow *col legno*—using the wood of the bow to design large circular arcs left and right across the metal stress bar of the instrument between transducers T2, T3, and T4—a corresponding left–right motion is heard

Table 1. Microphones and Output Routing

Microphone	*Position*	*Output*
Schaller	T1 (mobile)	L1-6
DYN-P	T2	L6
DYN-H	T3	L7
DYN-P	T4	L2
DYN-P	T5	L5
DYN-P	T6	L4
DPA4099	M7	L1
DPA4060	M8	L2
DPA4099	M9	L3
ME104	M10	L6
ME104	M11	L5
BLM03C	M12 (floor or lid)	L8

The microphones used were: Schaller Oyster piezo; Schertler DYN-P and DYN-H transducers; DPA 4060, 4099, and Sennheiser ME104 lavalier microphones; and Schoeps BLM03C boundary microphone.
Refer to Figures 3 and 4 for microphone layout and output positions.

from the corresponding source positions, L6, L7, and L2.

This approach to sound projection takes advantages of the unique collaboration between ZKM and the Südwestrundfunk Experimentalstudio (EXP), and the particular technological and human capabilities available through this collaboration. The Experimentalstudio has developed a unique approach to sound diffusion, eschewing regularly spaced circular loudspeaker arrays in favor of tailor-made loudspeaker arrangements for each project. An emblematic example is Luigi Nono's . . . *sofferte onde serene* . . . where stereo loudspeakers are complemented by a third speaker under the instrument for maximal fusion with piano samples in the tape part. In *Cartographies*, even though the final diffusion system was the symmetrical HDLA of the Klangdom, the virtual source positions were treated more subjectively and could be adjusted during rehearsals based on listening criteria.

The approach is enhanced by the electronic performance practice of EXP, with three computer music engineers each following a score of the work. One engineer is responsible for advancing the electronic

Figure 3. Microphone positions inside the Yamaha S6 piano used in Cartographies.

Figure 4. Positions of virtual loudspeakers, piano, and mixing console. The virtual loudspeakers correspond to adjustable positions in the Klangdom, interpolated using Zirkonium software. See Table 1 for the mapping of the piano microphone positions to the positions of the virtual loudspeakers.

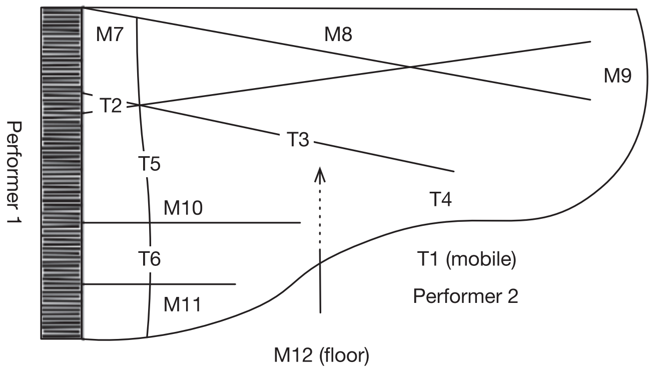


Figure 3

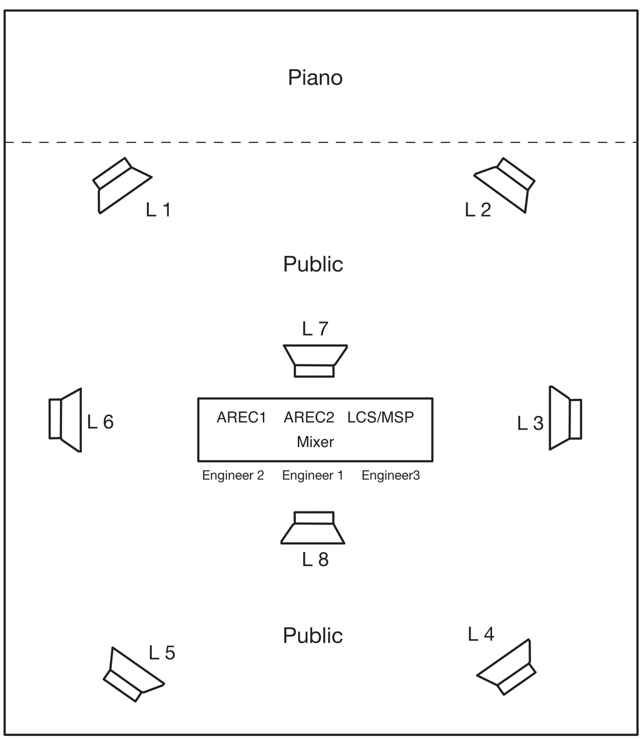


Figure 4

cues, one for the input and output levels of the live processing, and one for the input and output levels of the instrumental amplification. This permits a level of electroacoustic craft not easily achieved otherwise. In the case of *Cartographies*, the score indicates changes in the levels of different microphones at specific moments based on their locations and timbres. For example, near the end of the work, the levels of the five Schertler transducers affixed to the piano frame are raised while other microphones are silenced, producing a specific color filtered by the metal frame and limited frequency response of the transducers. As the microphones are mapped to fixed virtual locations in the Klangdom, these performance instructions also affect the spatialization pattern, creating a dynamically shifting map throughout the performance.

Corpus-Based Spatialization

A contrasting approach is corpus-based spatialization, where descriptor values analyzed from the live performers are used to pilot spatial trajectories of prerecorded samples. This is promising for its potential to create spatial motions responding to timbral descriptors analyzed dynamically in performance without relying on preprogramed trajectories. Different models for the synthesis of spatial motion were explored: In the simplest case, grains are concatenated in a monophonic audio channel that is then moved to a new virtual source location depending on the most recent grain synthesized. Although this leads to a clear perception of spatial motion, it produces a jerky effect by suddenly displacing grains as they are sounding. Alternately, each grain is treated as a separate polyphonic voice, remaining at a constant spatial location for its duration, regardless of the location of subsequent grains. This fills the space more vividly, yet still allows for the perception of virtual motion between grains. An acoustic analogy could be to a percussionist surrounded by a large collection of small instruments: Although each instrument remains stationary, we experience the performer's spatial gesture through the sequence of interactions with the instruments.

In *Cartographies*, the performers trigger corpus-based spatialization through a process of live audio "mosaicking." As the performers interact with the piano, their sound is captured through different combinations of microphones whose input levels to the software are regulated by the audio engineers, once again permitting fine control over microphone position and color. This input signal is used to trigger audio mosaicking by CataRT based on the prerecorded database of piano samples. The corpus samples were recorded using the same microphone setup as in the performance to ensure that they are as similar as possible. The same sample database was also used to compose the instrumental score of the work through the process of corpus-based transcription: Target sounds, drawn from longer piano samples as well as the found sound of the CMB, were transcribed into music notation readable by the performers, producing an audio mosaic to be reinterpreted instrumentally. When the source sound, for example Performer 1 rubbing a piezo microphone against a piece of aluminum foil, is reinterpreted in performance along with its mosaicked transcription by Performer 2, the result is a close timbral fusion between the performers. A detailed account of the transcription process using CataRT and the Bach package for Max is given in an earlier paper (Einbond et al. 2014).

Independent Trajectories

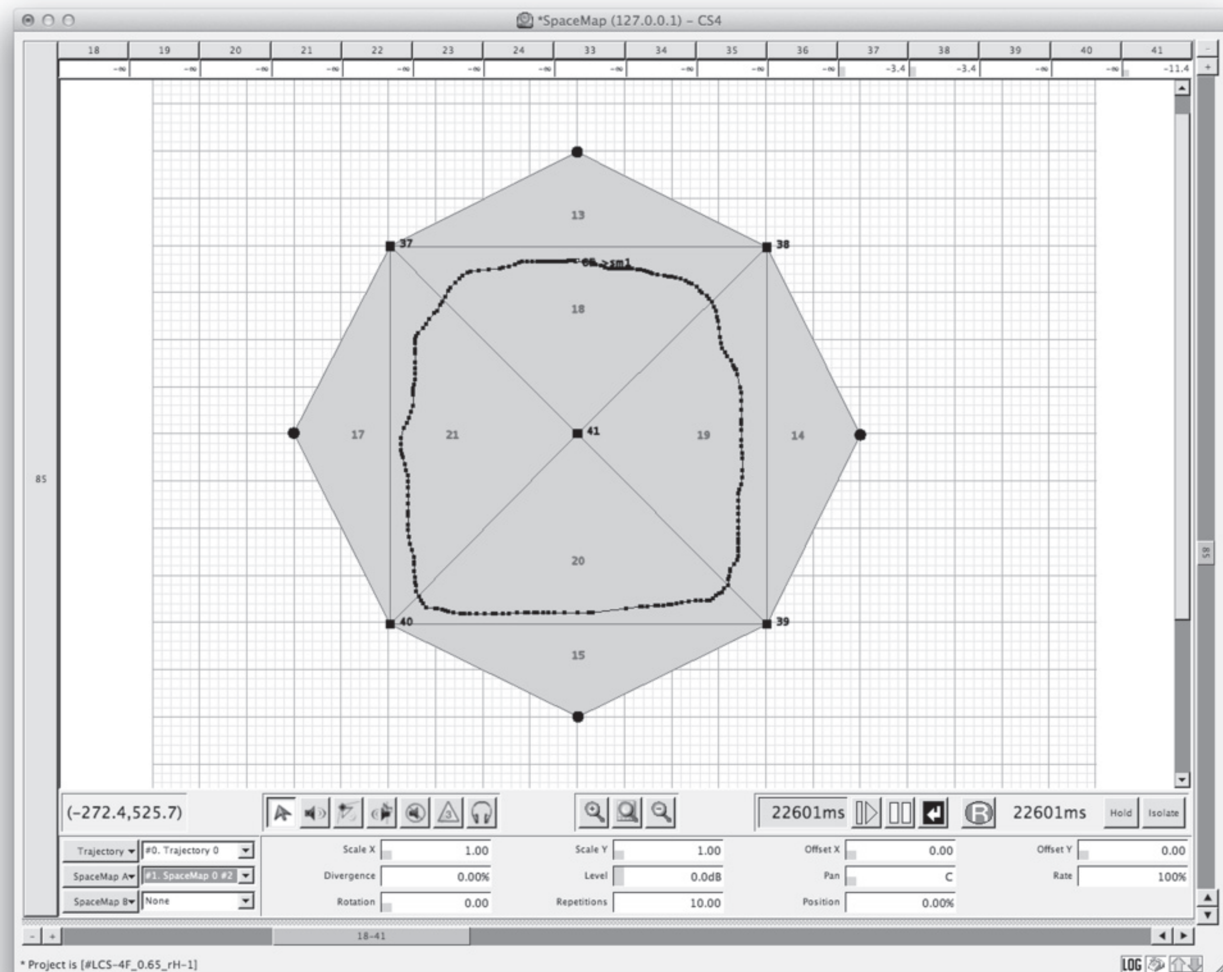
Finally, a third spatialization model was explored in which the spatial trajectory of the electronic sound was composed in counterpoint to the acoustic piano. This could be thought of as the default configuration for spatialization in music for instruments and electronics. As an example, *Cartographies* begins with an analog input signal panned in a circular motion around the outer ring of virtual output positions (L1–6), realized with two Halophones rotating at different speeds. (This technique, named after hardware built by longtime director of EXP Hans Peter Haller, is now implemented with a Max patch.) *Cartographies* takes this idea further, however. The Meyer Sound Matrix3 system incorporates the software tool SpaceMap (Ellison 2013), permitting a high level of customization and control of spatial trajectories. A spatial path can be recorded in real time using the mouse, then further edited in deferred time, and played back on cue in performance (see Figure 5). At the suggestion of Reinhold Braig (sound director and music computer engineer at EXP responsible for the production of *Cartographies*), this tool was used not only to create virtual spatial trajectories, but also to pilot signal processing by routing the outputs of SpaceMap to other real-time treatments. As an input channel is moved along a trajectory, its relative output levels to different processes are gradually cross-faded, and their outputs may, in turn, be positioned spatially.

This technique was used to apply different shadings of an impulse-response (IR) convolution reverberator to live input. The idea was to fuse the input sound with the resonance of the piano by convolving it with IR models sampled from impulses on the metal frame. Impacts of different percussion mallets (snare drum stick, yarn mallet, and bass drum beater) were recorded using four of the Schertler transducers mounted on the frame (T2, T3, T4, and T6 in Figure 3). Sounds convolved with these IR samples took on some of the coloring of the metal frame and the Schertler microphones themselves, similar to the sounds amplified using the same microphones. The `convolvezerolatency.maxpat` abstraction, from the AHarker library, was used for real-time processing, as it offers an ideal balance of sound quality, low latency, and manageable CPU cost. Please see the HISSTools Impulse Response Toolbox (cf. Harker and Tremblay 2012) for the updated external `multiconvolve~`.

The outputs of the four IR treatments were then routed to virtual source positions L6, L3, L4, and L5. In addition to the four convoluted signals, virtual source positions L1 and L2 were reserved for the untreated input signal, with levels carefully balanced by ear with the quieter, treated outputs. Trajectories were recorded and edited in SpaceMap to route the input signal in a circular motion to these six output channels, four treated and two untreated, producing a gradual cross-fade of untreated and convoluted signal accompanying the spatial motion.

For the live input here, as well as at the work's opening, the Schaller Oyster piezo microphone (T1) was selected as Performer 2 used it to interact with various materials (aluminum foil, plastic wrap, felt, scrub brush, and sponge) as well as the piano interior. Because of the coloring of the piezo and these "foreign" objects, this was one of the least "pianistic" sounds in the work. By convolving it with the IR samples, however, it was brought into

Figure 5. Screenshot of Meyer Sound's SpaceMap. Squares at the corners represent virtual loudspeaker positions, and the curve represents a hand-drawn spatial trajectory that can be played back to simulate the motion of virtual sources.



a closer fusion with the sound world of amplified piano.

Implementation

The concert patch for *Cartographies* builds on previous approaches to live treatment and spatialization (Einbond, Schwarz, and Bresson 2009; Einbond and Schwarz 2010), with novel features such as the use of a custom, expanded version of CataRT, communication between Max and Zirkonium software, and communication between the Matrix3 system of EXP and the ZKM Klangdom.

Matrix

The Experimentalstudio makes use of Meyer Sound's Matrix3 Audio Show Control System (www.meyersound.com/products/matrix3), combining low-latency hardware inputs and outputs with Open Sound Control (OSC) communication. It is controlled by custom-built fader surfaces that allow access to faders for physical input and output channels, as well as programmable "virtual faders" to send OSC messages to Max. Owing to the practical limitations of the collaboration between EXP and ZKM, a single piece of Meyer LX-300 hardware was used, limiting the input to the Klangdom to 16 channels. To get the most out of the 43.4 channels of the Klangdom (the configuration in 2014), these 16 channels were mapped to 16 virtual sources that could be positioned dynamically using ZKM's Zirkonium software (Brümmer et al. 2014). The 16 channels of audio were sent over a Dante digital audio network, and spatialization information was sent as OSC messages from Max to Zirkonium over a local network using the Max objects `udpsend` and `udpreceive`. Eight channels were mapped to fixed virtual source locations within the Klangdom, to be used for live amplification (see Figures 3 and 4) and some live processing, and the other eight channels were mapped to dynamically moving virtual sources, to be used for corpus-based spatialization with CataRT. Rather than controlling the levels of the eight shifting sources individually, their master levels were instead controlled by two "virtual faders" sent from the control surface back to the concert patch using OSC messages. Other treatments, such as IR convolution, were mapped to the fixed eight output positions.

CataRT

Cartographies adds to previous musical work with CataRT by using a custom version with a modular descriptor analysis framework as described by Schwarz and Schnell (2010). It includes an extended descriptor list based on the `ircamdescriptors~` analysis module, allowing for more fine control of a large selection of timbral and other features, as defined by Peeters (2004).

Choice of Corpora and Descriptors

Two corpora of piano samples were chosen for use in the concert patch: Corpus 1, recorded with the Schaller Oyster piezo microphone, and Corpus 2, recorded with the Schoeps BLM03C boundary microphone. The former provided a filtered, compressed, "electronic-sounding" version of the piano, and the latter gave a broadband, "naturalistic" sound. The input from any combination of microphones could be routed to either or both corpora to trigger live audio mosaicking. This provides a wide range of coloristic choices, as well as the possibility to assign treatments to only one of the two performers by selecting the microphones closest to the performer's range of action inside the piano. For example, summing the signals from M7, M10, and M11 triggers synthesis primarily in response to Performer 1, whereas M8 and M9 respond primarily to Performer 2 (for reference see Figure 3).

For corpus-based spatialization, two descriptor axes were chosen, taking into account the distribution of the grains in the corpus, the perceptual salience of the timbral axes, and the fusion of this timbral distribution with the amplified piano sound. After testing several possibilities, the same pair of axes was used for both sample corpora: pitch along the *x*-axis (specifically, the MIDI note number analyzed by the Yin algorithm; cf. de Cheveigné and Kawahara 2002), and brightness (spectral centroid) along the *y*-axis. These axes correspond well to descriptor variance within the sample corpora, which contains both pitched sounds (such as high piano strings strummed with wire brushes) and band-limited noise (such as low strings, prepared with aluminum foil and knitting needles, and hit with the pianist's palm). These axes are not necessarily the same as the axes chosen for live-audio mosaicking, for which more than two descriptors could be used to select timbres with greater precision, and could be changed to suit the audio content at different points in the work. In addition to pitch and spectral centroid, descriptors include spectral spread (a measure of bandwidth), level (in dB), and periodicity (a measure of noise content output by the Yin algorithm).

Spatial Mapping

The spatialization axes were oriented with low values toward the front right of the room and high values toward the rear left. This choice, along with the approximately triangular distribution of the grains (see Figure 1), suggests a loose analogy to the shape of the spatialized piano itself: its low strings projected across the front of the room and its high strings projected to the rear left (compare to Figures 3 and 4). The decision to use a single mapping of descriptor space for the entire work echoes the fixed projection of the amplified piano. In both cases the listener is surrounded by a stable, coherent virtual timbre space that is slowly revealed over the course of listening.

The concert patch allows for the range of spatial x- and y-positions to be scaled to minimum and maximum descriptor values, to the mean $\pm$ standard deviation, or to an arbitrary interval. Initial listening trials revealed that mean $\pm$ standard deviation gave the best results, mapping the most dense part of the corpus to a central location. Source distance is not taken into account by the spatialization algorithm used in the Zirkonium software, so only two spatial degrees of freedom are left: azimuth and elevation. Therefore, to map the two-dimensional descriptor space from CataRT to the three-dimensional Klangdom, the z-position of each grain was projected to the surface of the unit sphere $x^2 + y^2 + z^2 = 1$ for $x^2 + y^2 < 1$, and 0 elsewhere, as given by $z = \sqrt{\max(1 - x^2 - y^2, 0)}$.

Virtual Polyphony

Under other circumstances, CataRT could be used for spatial positioning in an arbitrary number of channels: The module `catart.synthesis.multi` communicates with the external object `gbr.ola~` from library FTM & Co. to control an efficient overlap-add algorithm. For this production, however, because only eight mobile output channels were available to send to Zirkonium, a different approach was used to take advantage of the full 43.4 channels of the Klangdom. The objects `catart.synthesis.multi` and `gbr.ola~` are still used, but each grain is mapped to a separate channel with an amplitude of one, while spatial information for that grain is simultaneously sent to Zirkonium. A "busy map" is generated at the output of the concert patch, returning a list of available channels to `catart.synthesis.multi` before it synthesizes the next grain. Conceptually similar to a `poly~` object in "voice-stealing" mode, it limits synthesis to only eight simultaneous virtual spatial sources. Because the grains used in the piece are typically short (250–1000 msec), however, this limitation was not perceived to be significant and was outweighed by the benefit of using the entire Klangdom. According to the voice-stealing algorithm, if all eight channels are already occupied, the next grain is overlapped with the grain that has been playing the longest—the entry effectively masking the motion of the previously sounding grain.

Following EXP practice, the live electronic treatments other than CataRT are sent as separate stems to the Matrix3 system, so that each level can be controlled manually before being routed to the eight fixed output channels in Zirkonium. Eight premix outputs from the concert patch are needed: four IR convolutions, a bandpass filter using `biquad~`, distortion using `overdrive~`, and two channels of sound file playback (serving for both monophonic and stereo files). See Figure 6 for a screenshot of the concert patch.

Zirkonium

The Zirkonium software was developed at ZKM as a free OSX-based spatialization interface (Brümmer et al. 2014). It uses the vector base amplitude panning (VBAP) algorithm (Pulkki 1997) to map virtual source positions to triples of loudspeakers. Zirkonium responds to OSC messages specifying positions, allowing for flexible and legible control. In *Cartographies*, the 16 channels received from the Matrix3 system are mapped to 16 input channels in Zirkonium. To limit possible rapid position changes, all OSC messages to Zirkonium are filtered with a `speedlim` object set to 20 milliseconds, balancing smooth spatial motion and CPU cost.

The eight fixed virtual source positions (channels 1–8) are stored in a text file loaded into a `coll` object for easy editing. This allows them to be positioned by ear during rehearsal in the concert space. Rather than relying on theoretical positions, the particular spatial and acoustic characteristics of the room are taken into account in deciding how to

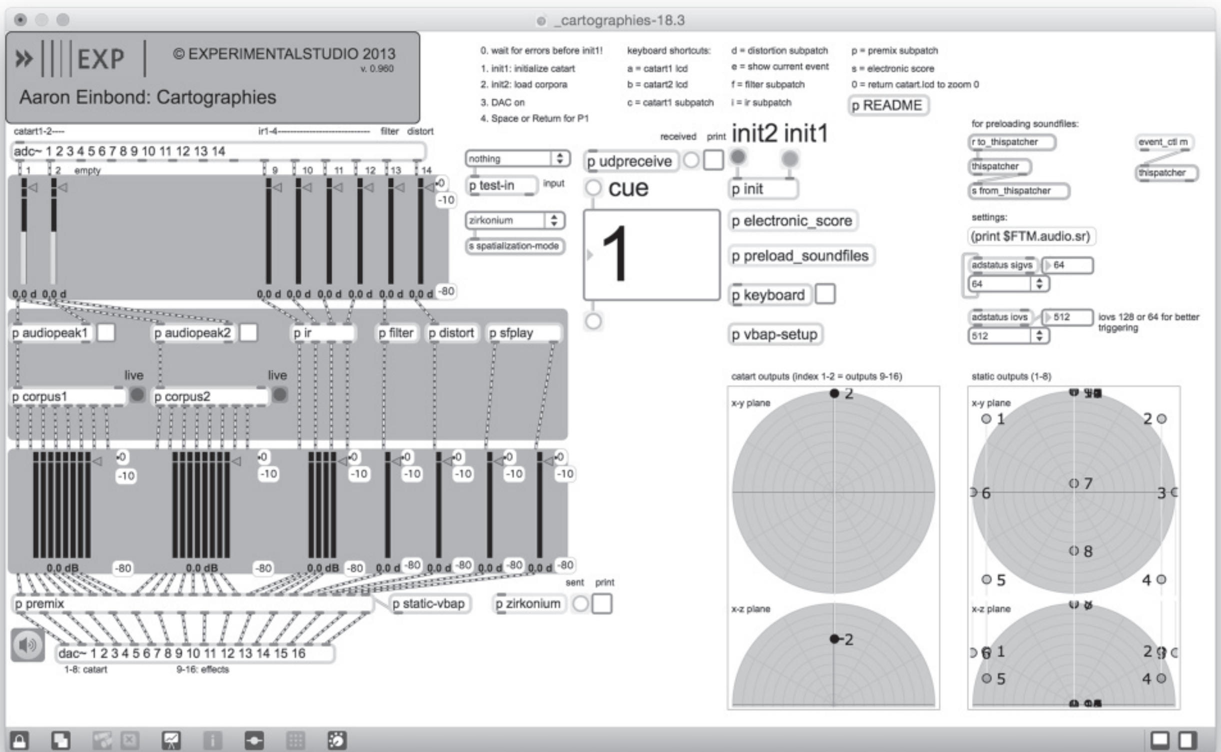


Figure 6. Screenshot of Cartographies *concert patch showing audio inputs (upper left) and outputs (lower left) for real-time treatments. Spatialization controls sent via* `udpsend` *to the Zirkonium software are visualized using the* `ambimonitor` *objects on the lower right.*

project the amplified piano in an immersive way. On initialization of the concert patch, these eight positions are recalled and sent to Zirkonium and channels 9–16 are positioned at the origin, serving as visual confirmation that OSC communication between Max and Zirkonium is functioning properly. Figure 7 shows the final positions chosen for channels 1–8 as well as the default positions for channels 9–16.

Discussion

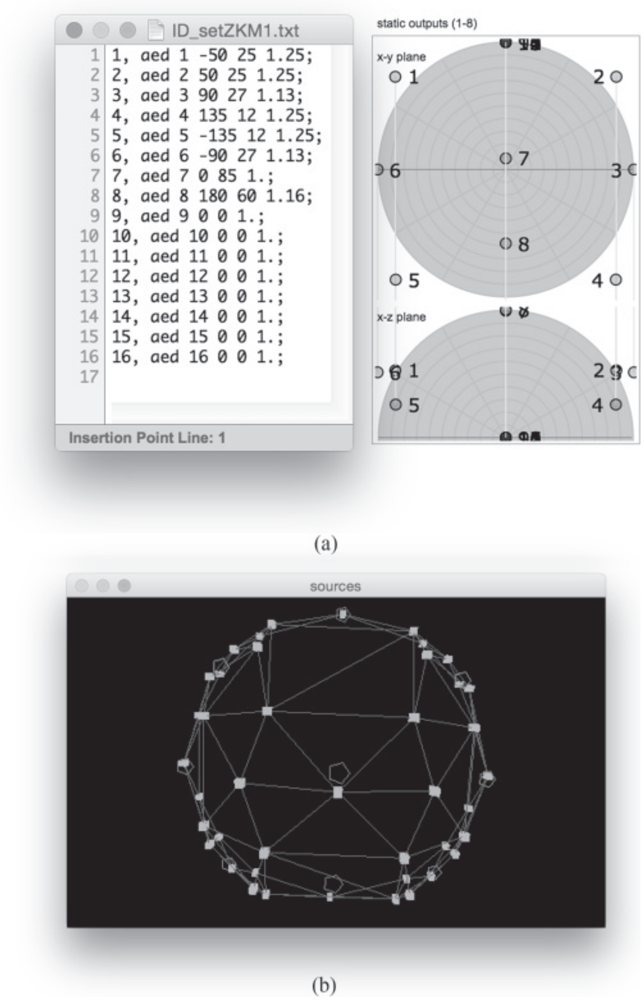
As a work for live instruments and interactive electronics projected over an HDLA, *Cartographies* presents an uncommon scenario, and the particularities of the collaboration between EXP and ZKM played a large part in the work's successful realization. Yet given these special conditions, it is worth considering how this or a similar work could be carried forward to other performance situations or realized in a scaled-down version.

[Editor's note: a binaural mix with video of *Cartographies* is available at https://youtu.be/XbmGgYoEezU]

Listening Observations

Although certainly subjective, and localized to a single performance realization, the author's listening observations at the premiere may nonetheless be instructive for future implementations and extensions of corpus-based spatialization. The

Figure 7. Concert patch (details) showing initial positions of outputs 1–16 listed as rows of a `coll` *and displayed with the* `ambimonitor` *object (a). (The* `ambimonitor` *detail as seen on the lower right of Figure 6.) Screenshot of Zirkonium showing source positions (pentagons) superposed over the 43 speakers (squares) of the ZKM Klangdom (b).*



choice of descriptor axes of pitch and brightness for spatialization worked well to spread the corpus sounds across the Klangdom, as did descriptor scaling to the range defined by mean ± standard deviation, so that the center of weight of the distribution was shifted to the center of the room. Still, as visible in Figure 1, more samples were gathered toward the front of the room, coinciding with the location of the stage. Yet this density had the positive effect of concentrating additional listening focus in the direction of the piano and performers. The importance of the subjective choice of this mapping should, therefore, not be underestimated. A possible future direction would be to use principal component analysis to identify the descriptors, or linear combinations of descriptors, that produce the most variance across the corpus, and therefore best distribute units across the space. Such a calculation must still be evaluated subjectively through listening, however.

A related observation is that corpus-based spatialization produces a productive reinforcement between audio parameters. When, for example, a granular gesture from the right to the left of the room is accompanied by a rise in brightness, the trajectories reinforce each other, making the effect more robust perceptually. This redundancy helps to ensure that spatial and timbral effects are salient for listeners positioned at different positions in the room, not only those at the "sweet spot." This mitigates one possible drawback of HDLAs: During the rehearsals and performance, listening positions at many different locations in the Klangdom were found to yield satisfactory timbral and spatial aural experiences.

The "voice-stealing" algorithm, limiting synthesis to eight channels of virtual polyphony, was relatively successful in communicating an immersive granular texture without audible artifacts. Monitoring the patch during performance revealed that some voice-stealing was taking place, in the sense of grains being moved to new locations while still sounding (not in the sense of grains being interrupted). These displacements were not readily audible, however, likely due both to the fast decay of the relatively short grains and to the psychoacoustic precedence effect. Owing to technical limitations, more than eight channels of virtual polyphony were not tested in the Klangdom. This restriction, necessitated by communication between the EXP and ZKM systems, would be lifted, however, if the work were realized using a single system (discussed subsequently in the section on "Portability Beyond the Klangdom").

Live Interpretation and Spatialization

The unique affordances of EXP performance practice, with its attention to live diffusion of both amplified and synthesized sound, became an important part of the work's identity. The two instrumental performers and three electronic performers must adjust their interpretation in response to each other and the acoustic conditions of the space. The details of the score and concert patch, although leaving enough latitude for inter-performance variation, also ensures enough predictability for effective rehearsal. The goal is a "chamber music–like" interaction between the instrumentalists and engineers. If the work were realized by fewer engineers, especially if they were brought in at a late stage of the rehearsal process, the same level of control would be difficult to master. Also, programming more level changes into the patch would have the disadvantage of adding latency, as well as reducing sensitivity to timing and level variations by the live instrumentalists. Nevertheless, given the possible practical constraints of future performances, alternative solutions must be considered.

Portability Beyond the Klangdom

By realizing *Cartographies* with Zirkonium software, a degree of portability is already assured: Zirkonium's VBAP algorithm is independent of specific loudspeaker setup. The first step after opening the Zirkonium Spatialization Server is to load a predefined speaker configuration as an XML file. Separate files describe the speaker positions of ZKM's Klangdom, the 24-channel "minidome" (where some of *Cartographies* was prototyped), or other systems. In theory, the production could be taken to any HDLA, the appropriate speaker configuration could be loaded, and the spatial setup would be transparently adapted to the new system.

A greater degree of portability is needed to realize rehearsal and performance versions of the work with a reduced loudspeaker setup, however. The distribution of eight virtual sources (shown in Figure 4) was conceived with this possibility in mind, to be replaceable by eight physical loudspeakers in the same positions. This was the setup used for rehearsals at EXP before traveling to ZKM. The amplification and live treatments (IR, filtration, distortion, sound-file playback) are easily adapted: Instead of sending eight channels from the Matrix3 to Zirkonium, they are sent directly to the loudspeakers. Channels 1–6 are positioned in a ring around the space, and channels 7 and 8 are positioned near the center, raised, and pointed upward to simulate the upper speakers of the Klangdom.

The only part of the concert patch that requires adjustment is concatenative synthesis with CataRT, which is now used for spatialization over the same eight physical outputs as for the other treatments and amplification. This is easily achieved with the built-in `catart.synthesis.multi` and `gbr.ola~` modules, returned to their usual function to perform spatialization with an overlap-add algorithm. An onboard VBAP object is used to calculate spatial coefficients based on the eight loudspeaker positions, sent to `catart.synthesis.multi` and `gbr.ola~` before each grain is synthesized throught the eight outlets. The eight amplitude coefficients from VBAP are applied independently to each grain, so an arbitrary number of virtual spatial positions can be superposed. Unlike the Zirkonium version of the patch, now corpus-based spatialization is realized with unlimited virtual spatial polyphony, an advantage despite the smaller loudspeaker setup.

Finally, a still "lighter" version of the work is planned for performances at which EXP is not present, dispensing with the Matrix3 system in favor of a single Max concert patch. In this case, the built-in CataRT and VBAP modules will be used for spatialization as described earlier. The remaining treatments, rather than being sent to Matrix3, will be routed directly in the patch using a second `VBAP` module to replace SpaceMap. This will have the disadvantage that live electronic treatments and amplification can no longer be controlled independently from the mixer. But, as a compromise solution, a MIDI controller such as the Behringer BCF2000 can be added to adjust premix levels in the patch during performance. Although this light version is imagined for the eight-channel setup described herein, thanks to the built-in `VBAP` modules, it could be transparently scaled to larger or smaller loudspeaker setups.

Future Directions

CataRT-style concatenative synthesis has now been implemented in the MuBu package for Max (Schnell et al. 2009), including many of the capabilities of the custom version used in *Cartographies*. Rather than `catart.synthesis.multi`, spatialization is performed by `mubu.concat~`, which can be set to generate an arbitrary number of channels, to which amplitude coefficients derived from `VBAP` can be applied for corpus-based spatialization. A modular analysis framework permits the use of the full range of `ircamdescriptors~` for fine timbral control. Based on Max externals rather than FTM, MuBu benefits from greater stability and portability as well as access to 64-bit computing to permit larger database sizes.

Recent research combines MuBu with a computer improvisation algorithm based on the PyOracle library for Python (Surges and Dubnov 2013). The resulting tool, CatOracle (part of Forum IRCAM's MuBu package: http://forumnet.ircam.fr/product/mubu-en/), can concatenate novel sequences of grains based on shared context in descriptor space (Einbond et al. 2016). Combining it with corpus-based spatialization would have powerful consequences as a tool for spatialized improvisation and composition, complementing spatial logic with learning and generation of musical structure. A first application, *Xylography* for cello and electronics (2015), experiments with projecting the signal of the amplified cello, captured by four contrasting microphones (DPA 4060 miniature omni microphone, AKG C411 piezo microphone, and two Schertler DYN-P transducers), across a four-channel loudspeaker system.

A longer-term goal is to incorporate research in instrumental directivity patterns more directly into corpus-based spatialization. Rather than focusing on the physical geometry of the instrument or a simple timbral-spatial mapping, as in *Cartographies*, this would imply directly modeling the directivity pattern projected from the instrument into the concert space. By measuring, in a controlled environment, different timbral descriptors as functions of space, these distributions could be used as templates for mapping concatenated grains to spatial locations according to their descriptor values. In contrast to the linear mapping used in *Cartographies*, this would permit a more fine-grained distribution of timbre in space, and promise an even closer fusion of acoustic instrumental directivity with live corpus-based spatialization.

Acknowledgments

I thank Reinhold Braig and Gary Berger, sound directors and music computer engineers responsible for the production of *Cartographies*; Dominik Kleinknecht and Simon Spillner, music computer engineers for the premiere of *Cartographies*; Detlef Heusinger and the team of Südwestrundfunk Experimentalstudio; Götz Dipper, David Wagner, Ludger Brümmer, and the team of the Institut für Musik und Akustik at ZKM, Karlsruhe; Diemo Schwarz and the Sound Music Movement Interaction team at IRCAM; and Rei Nakamura and Olaf Tzschoppe, collaborators and performers of the premiere of *Cartographies* with the ZKM Klangdom on 25 November 2014.

References

Avizienis, R., et al. 2006. "A Compact 120 Independent Element Spherical Loudspeaker Array with Programmable Radiation Patterns." In *Proceedings of the 120th Convention of the Audio Engineering Society*. Available online at www.aes.org/e-lib/browse.cfm?elib=13587 (subscription required). Accessed November 2016.

Brümmer, L., et al. 2014. "New Developments for Spatial Music in the Context of the ZKM Klangdom: A Review of Technologies and Recent Productions." *Divergence Press* 3. Available online at dx.doi.org/10.5920/divp.2015.36. Accessed November 2016.

d'Alessandro, C., et al. 2009. "The ORA Project: Audio-Visual Live Electronics and the Pipe Organ." In *Proceedings of the International Computer Music Conference*, pp. 477–480.

de Cheveigné, A., and H. Kawahara. 2002. "YIN, a Fundamental Frequency Estimator for Speech and Music." *Journal of the Acoustical Society of America* 111(4):1917–1930.

Einbond, A., and D. Schwarz. 2010. "Spatializing Timbre with Corpus-Based Concatenative Synthesis." In *Proceedings of the International Computer Music Conference*, pp. 72–75.

Einbond, A., D. Schwarz, and J. Bresson. 2009. "Corpus-Based Transcription as an Approach to the Compositional Control of Timbre." In *Proceedings of the International Computer Music Conference*, pp. 223–226.

Einbond, A., et al. 2014. "Fine-Tuned Control of Concatenative Synthesis with CataRT Using the Bach Library for Max." In *Proceedings of the International Computer Music Conference*, pp. 1037–1042.

Einbond, A., et al. 2016. "Introducing CatOracle: Corpus-Based Concatenative Improvisation with the Audio Oracle Algorithm." In *Proceedings of the International Computer Music Conference*, pp. 140–146.

Ellison, S. 2013. "SpaceMap: 20 Years of Audio Origami." *Lighting and Sound America* (April):80–88.

Gottfried, R. 2012. "Studies on the Compositional Use of Space." Technical report. Paris: Institut de Recherche et Coordination Acoustique/Musique. Available online at articles.ircam.fr/textes/Gottfried13a/index.pdf. Accessed November 2016.

Grey, J. M. 1977. "Multidimensional Perceptual Scaling of Musical Timbres." *Journal of the Acoustical Society of America* 61(5):1270–1277.

Haenisch, M. 2013. "Materiality and Agency in Improvisation: Andrea Neumann's 'Inside Piano'." In A. Cassidy and A. Einbond, eds. *Noise in and as Music*. Huddersfield, UK: University of Huddersfield Press, pp. 147–170.

Harker, A., and P. A. Tremblay. 2012. "The HISSTools Impulse Response Toolbox: Convolution for the Masses." In *Proceedings of the International Computer Music Conference*, pp. 148–155.

Noisternig, M., F. Zotter, and B. F. G. Katz. 2011. "Reconstructing Sound Source Directivity in Virtual Acoustic Environments." In Y. Suzuki, D. Brungart, and H. Kato, eds. *Principles and Applications of Spatial Hearing*. Singapore: World Scientific, pp. 357–373.

Otondo, F., et al. 2002. "Directivity of Musical Instruments in a Real Performance Situation." In *Proceedings of the International Symposium on Musical Acoustics*, pp. 312–318.

Peeters, G. 2004. "A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project." Technical report. Paris: Institut de Recherche et Coordination Acoustique/Musique. Available online at recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf. Accessed November 2016.

Pulkki, V. 1997. "Virtual Sound Source Positioning Using Vector Base Amplitude Panning." *Journal of the Audio Engineering Society* 46(6):456–466.

Ryan, D. 1999. "Composer in Interview: Helmut Lachenmann." *Tempo* (210):20–24.

Schmeder, A., and M. Noisternig. 2010. "Spatialization Schemata for High-Order Source Directivity." Paper presented at the GDIF/SpatDIF Meeting, Institut de Recherche et Coordination Acoustique/Musique, Paris.

Schnell, N., et al. 2009. "MuBu and Friends: Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP." In *Proceedings of the International Computer Music Conference*, pp. 423–426.

Schwarz, D. 2007. "Corpus-Based Concatenative Synthesis." *IEEE Signal Processing Magazine* 24(2):92–104.

Schwarz, D., and N. Schnell. 2010. "A Modular Sound Descriptor Analysis Framework for Relaxed-Real-Time Applications." In *Proceedings of the International Computer Music Conference*, pp. 76–79.

Surges, G., and S. Dubnov. 2013. "Feature Selection and Composition Using PyOracle." Paper presented at the AIIDE Workshop on Musical Metacreation, October 14–15, Northeastern University, Boston, Massachusetts. Available online at www.aaai.org/ocs/index.php/AIIDE/AIIDE13/paper/view/7452/7685. Accessed November 2016.

Warusfel, O., and N. Misdariis. 2001. "Directivity Synthesis with a 3D Array of Loudspeakers, Application for Stage Performance." In *Proceedings of the International Conference on Digital Audio Effects*. Available online at www.csis.ul.ie/dafx01/proceedings/papers/warusfel.pdf. Accessed November 2016.

Wessel, D. 1979. "Timbre Space as a Musical Control Structure." *Computer Music Journal* 3(2):45–52.

Zotter, F. 2009. "Analysis and Synthesis of Sound-Radiation with Spherical Arrays." PhD dissertation, Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz.