

Mapping and dimensionality of a cloth-based sound instrument

David Birnbaum*, Freida Abtan†, Sha Xin Wei†, Marcelo M. Wanderley*

*Input Devices and Music Interaction Lab/CIRMMT, McGill University, Montreal, Canada

†Topological Media Lab/Hexagram, Concordia University, Montreal, Canada

Abstract—This paper presents a methodology for data extraction and sound production derived from cloth that prompts “improvised play” rather than rigid interaction metaphors based on preexisting cognitive models. The research described in this paper is a part of a larger effort to uncover the possibilities of using sound to prompt emergent play behaviour with pliable materials. This particular account documents the interactivation of a stretched elastic cloth with an integrated sensor array called the “Blanket”.

Keywords— pliable interface, gesture tracking, mapping, emergent play.

I. INTRODUCTION

Interactivity can be characterized in many different ways. Structural models of gesture have informed efforts to design interactive sound systems, particularly for music performance [2]. Principles derived from human-computer interaction and music cognition studies considering task complexity [17], cognitive load [7], and engineering considerations [10] are common bases for design. At the level of design theory, a conceptual framework for digital musical instrument design proposed by two of the authors applies human-machine interaction and semiotics to musical performance [8].

At the same time, theories of embodiment suggest modes of interaction behaviour may emerge outside the cognitive or linguistic realms. The design of a gestural interface with sound feedback need not rely on semiotics or an assumption of an agent actor. Avoiding such assumptions eliminates a reliance on intentionality for control, and promotes emergent play behaviours [18].

Incorporating aspects of both of these paradigms, this project is part of *Wearable Sounds Gestural Instruments (WYSIWYG)*, a research effort aiming to create a suite of soft, cloth-based controllers that transform freehand gestures into sounds. These “sound instruments” [19][5] can be embedded into furnishings or rooms, or used as props in improvised play. Sound responds to diverse input variables such as proximity, movement, and history of activity. The interactions are designed in the spirit of games such as hide-and-seek, blind-man’s buff, and simon-says, working well with a variable number of players in live, ad hoc, co-present events. The design goal for the “wysiwyg” described in this paper was to use a simple fabric-based interface to explore the way synthesis methods may be used to represent interactions with cloth.

Fusing fabric art with digital feedback systems holds many possibilities because the basic interaction characteristics of fabric are so commonly experienced. Fabric

is malleable, tangible, textural, and material. It promotes multisensory, haptic modes of exploration and manipulation. It carries a pre-existing context of gesture and expression that need not rely on linguistic tokens to represent interaction modes; instead, the surface dynamics of cloth generate recognizable states based on structural similarities. For example, a “fold” is recognizable even though the set of all possible transformations that could be called a fold is infinite. Characteristics such as these are independent of their specific physical manifestation in cloth.

Textile environments have been widely used in art and sound installations. Electronics have been embedded into articles of clothing as a platform for interactive performance [9][12], textiles themselves have been used as a malleable physical interface [19][11][13], and fabric-based installations have been created on an architectural scale [16][14]. A detailed physical model of textile motion has been created for musical control [3], however the user interface is graphical rather than physical.

II. THE SYSTEM

To explore these concepts, a physical interface was constructed. The **Blanket** is a 3x3-meter square piece of elastic fabric. Sewn into its top surface is a sensor system made up of 25 light-dependent resistors (*LDRs*) arranged in a 5x5 array. The control surface of the interface consists of the entirety of the cloth. In exhibition, the fabric is stretched by its corners and elevated about 1.5 meters off the ground. It may be touched, stretched, pushed up, pulled down, shaken, scrunched, or interacted with in any way the human body might be applied. Most notably, it is large enough for several people to interact with it simultaneously.

The *LDRs* become “motion” sensors with the use of theatrical lighting. Beams of light are strategically positioned perpendicular to the Blanket’s top surface so that when it is put into motion, the sensors are exposed to variable amounts of light as they are brought nearer and further from the center of the beams. The voltage output of the sensor system is sampled by an Arduino A/D converter [1]. Preprocessing, mapping, and synthesis take place in Max/MSP.

III. GESTURE TRACKING

Gesture has been defined as an intentionally expressive bodily motion recognized in a particular cultural context



(a) The interface at rest



(b) Interaction

Fig. 1. The Blanket

[6]. However the Blanket system does not acquire the gestures of the human interactors. Rather it maps cloth movement to sound parameters that promote precognitive interaction. In this sense, what is tracked is the “gesture” of the fabric rather than that of the humans, using the word gesture to refer to a single contour within the sensor system’s total response arising through the application of some data extraction method. Because intentionality is purposefully omitted from our software model, “noise” is also defined by the data analysis approach, and consists of all of the confounding factors when attempting to isolate a specific cloth movement. Differentiating signal from noise thus necessitates a specification of what constitutes a unitary contour in the data stream. A cloth gesture could be said to be the human interpretation of a contour in feature space arising from motion due to human(s).

Because the goal of this research is to determine how to generate meaningful sound derived from the “raw” physicality of cloth, feature extraction was limited to three functions: absolute activity (the values of all sensors added), sensor velocity, and activity “spike” (a sudden change in value that exceeds a preset threshold). Each of these functions make available a dynamic data stream for mapping to audio parameters.

The sensor system is by its nature a two-dimensional array that moves in three-dimensional space, which prompts the question of how to sample the array. A sampling technique consists of a spatiotemporal sequencing of individual sensor outputs defining a unique “domain”. Domains are orderings of the sensor array that include all discrete sensing points once and only once, but may segment sensor readings into groups. These constraints are inspired by measure theory, in the hopes of obtaining a holistic “snapshot” of the state of the entire interface.¹ Contours in the signal correlating to gesture are mapped to sound synthesis parameters depending on how the array

¹To test for dimensionality of a set of points N , the points inside a given radius N_r may be counted. Because the two-dimensional sensor array may be approximated as a “point cloud”, using all points once and only once within a ball of radius r will best represent its dimensionality.

is sampled. Three approaches were taken to sample the two-dimensional array, defining the domains of *string*, *sectors*, and *atoms*. Each method offers its own distinct approximation of the gestures traveling through the cloth.

A. String

The string model treats the data points as a space-filling curve. In this schema, the two-dimensional array is linearized by scanning all sensor values via a “walk” along the surface of the fabric, consolidating the total output of the sensor array to drive a single synthesis parameter. The domain’s differentiating characteristic is that it is a one-dimensional ordering of the set.

The path of the scanning sequence can be arbitrary, but physical properties of the cloth at each point on its surface have a profound effect on the string. For example, because the corners of the cloth are immobilized by support ropes, there is a damping effect that is most pronounced at the Blanket’s corners and edges, causing an increase in the relative kinetic motion of its center. When the motion of the interface resembles a vibrating membrane, modal physics affects sensor outputs. Two scanning paths were utilized to observe how physical factors influence the perceived meaning of the sound output: a spiral and a switchback.

The spiral walk begins at one of the outer corners of the interface, where motion is dampened by the support ropes, and spirals inward to the Blanket’s center. For the first mode, the points with the lowest kinetic energy are concentrated at the beginning of the walk and those with the highest are at the end. In contrast, the switchback introduces a periodic damping effect, as peripheral points are equally distributed among internal points. (Arbitrarily chosen walks will, of course, always be subject to the effect of damping, light source placement, and human factors arising from the embodiment of interaction behaviours.) Because the string domain was used to map each sensor value to a virtual mass on a string for scanned synthesis using the *scansynth*~ object [15] [4], the effects of any walk are evident in the timbre of the resulting

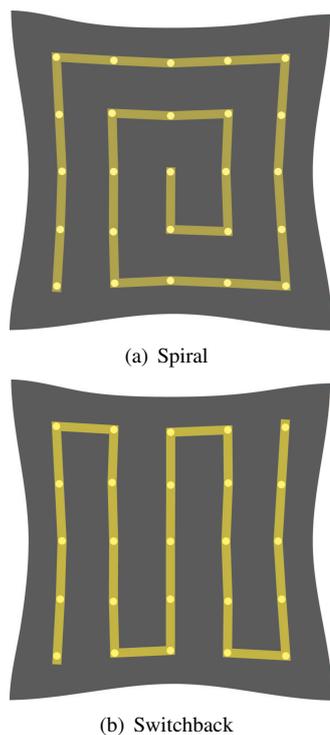


Fig. 2. Two possible paths for defining a “string” domain

sound. Out of the three domains, the string model is most responsive to the kinetic properties of the entire interface.

B. Sectors

The sector model is based on two principal functions, namely the dividing of the cloth into sonically-autonomous regions, and data smoothing within those regions by averaging all the points within them.

The *matrixctl* object in the patch allows a sound programmer to select multiple contiguous regions of the fabric and define them as sectors. The data associated with each sector can then be mapped to the control parameters of a sound instrument. Although no sensor ordering information is preserved within sectors, the membership clause of contiguity assures a relationship between the sectorized data points in two-dimensional space. Since sectorization downsamples the two-dimensional array to extract data values by area, the illusion of a unified response taking place over the entirety of each sector is created. As a result, gestures over the control surface take the form of an interaction between such distinct forces. Using each sector to control a separate instrument results in a polyphonic effect.

Two possible choices for boundary conditions include sectorization by surface geometry, or by physical properties such as damping conditions, proximity to light source, or performance-specific contingencies. Each choice implies an assumption about what data should be considered a unified force and so responds according to a different model of meaningful interaction. For example, sectorization by quadrant assumes that the most meaningful distinction between forces is absolute location of an

interactor, whereas a partition into concentric regions emphasizes the physics of the control surface. For the final implementation, the decision was made to group sensor data together by quadrant after observing the temporal scale at which a gesture would resolve over the entirety of the interface. Two results of this decision are that participants concentrated in one quadrant are prompted by sound to cooperate, and participants located in separate regions have primary influence over the voice(s) associated with their sector.

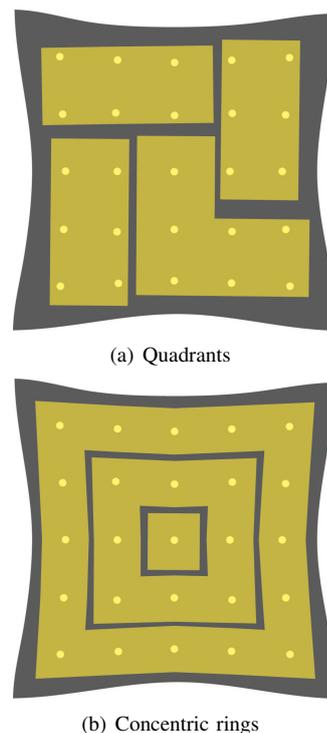


Fig. 3. Two “sector” domains

The sector approach is not without its weaknesses. Averaging several points runs the risk of including sensor data that is either redundant or not engaged in the interaction, as do the constraints imposed by contiguity and the inclusion of each data point exactly once. Additionally, sampling fewer than all sensors may give an equivalent result, while the activity of non-contiguous regions may be accurately represented by one voice due to modal properties of the interface. However these weaknesses can be seen as the price paid for choosing to preserve a holistic representation of the surface.

The quadrants were sonified by mixing four granular synthesizers. The intensity of each voice was controlled by the overall amount of activity in the associated quadrant, while textural properties such as amount of grains and pitch were controlled by rate-of-change parameters.

C. Atoms

Mapping each data point to its own individual sound generator utilizes the entirety of the interface’s output without making any assumptions about ordering or location. Treated as “atoms”, each sensor is given a unique

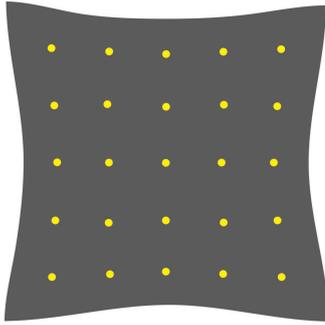


Fig. 4. Atoms

voice independent of the state of the entire interface. However, because the synthesizer responds to each input simultaneously, sound is very tightly coupled to gesture, which is expressed as temporal variation in the atoms. A relationship between points is manifest in the response of the sound instrument, but is not inherent in the definition of the domain. Without any ordering or grouping information whatsoever, this domain could be said to have zero dimensions; it makes no assertions about the dimensionality of the ambient object. Because assumptions about dimensionality are minimized, this approach may be said to avoid corruption of the data as a result of mapping choices. Gesture is represented in the sound, but not in the data.

The synthesis implementation consists of a wavetable “scrubbing” technique. Each atom controls the position of the playback head in its own *waveform*~ object. Velocity is mapped to the scrubbing speed and direction, so that kinetic energy and direction of motion are represented by pitch and timbre, respectively. A smoothing function has also been applied to act as a threshold, so that minute changes in sensor outputs while the interface is at rest do not cause low frequency noise.

IV. DISCUSSION

If the strict adherence to the definition of a domain is relaxed, many possibilities emerge. Using subsets of non-contiguous groupings of points might allow greater freedom to account for redundant or unnecessary information. Periodic motion or relative position may in fact be detectable without all of the sensor data.

The three dimension-based mapping strategies outlined here are similar in that they are ways of viewing the motion of a whole object. In practice sound instruments can adopt each paradigm in succession or can incorporate all three into their data extraction process. The three approaches may be used together in the same performance — either in parallel, to control different aspects of the sound feedback, or in sequence, to delineate game stages. Used in parallel, the domains could control multiple timbral features of the same sound instrument. In sequence, the shift from one domain to another could symbolize a state change arising either from interaction events or compositional choices. Further, audio output can

be produced by many sound instruments simultaneously, each enforcing a separate abstraction of the same data stream. The aptness of a given domain may be based on the sonic result of perceived gestures, or the ability of sonic feedback to direct gesture, indicate compositional choices, or influence play behaviour.

ACKNOWLEDGMENT

WYSIWYG is Freida Abtan, David Birnbaum, David Gauthier, Rodolphe Koehly, Doug van Nort, Elliot Sinyor, Sha Xin Wei, and Marcelo M. Wanderley. This project has been funded by Hexagram. The authors would like to thank Eric Conrad, Harry Smoak, Marguerite Bromley, and Joey Berzowska.

REFERENCES

- [1] Arduino. <http://www.arduino.cc/>. Accessed on June 11, 2007.
- [2] Claude Cadoz. Instrumental gesture and musical composition. In *Proceedings of the 14th International Computer Music Conference*, pages 1–12, Köln, 1988.
- [3] Lonny L. Chu. Musicloth: A design methodology for the development of a performance interface. In *Proceedings of the International Computer Music Conference*, pages 449–452, 1999.
- [4] Jean-Michel Couturier. A scanned synthesis virtual instrument. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression*, pages 176–178, 2002.
- [5] Jill Fantauzza, Joey Berzowska, Steven Dow, Giovanni Iachello, and Sha Xin Wei. Greeting dynamics using expressive software. In *Ubicomp 2003 Adjunct Proceedings*, 2003.
- [6] Adam Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [7] Daniel J. Levitin, Stephen McAdams, and Robert L. Adams. Control parameters for musical instruments: A foundation for new mappings of gesture to sound. *Organized Sound*, 7(2):171–189, 2002.
- [8] Joseph Malloch, David Birnbaum, Elliot Sinyor, and Marcelo M. Wanderley. Towards a new conceptual framework for digital musical instruments. In *Proceedings of the 9th International Conference on Digital Audio Effects*, pages 49–52, 2006.
- [9] Benoît Maubrey. Audio jackets and other electrasonic clothes. *Leonardo*, 28(2):93–97, 1995.
- [10] Eduardo R. Miranda and Marcelo M. Wanderley. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. A/R Editions, 2006.
- [11] Josie Moores. Sonic fabric. *Make: Technology On Your Time*, 5, 2006.
- [12] Maggie Orth, J. R. Smith, E. R. Post, J. A. Strickon, and E. B. Cooper. Musical jacket. In *International Conference on Computer Graphics and Interactive Techniques: ACM SIGGRAPH 98 Electronic art and animation catalog*, page 38, New York, NY, USA, 1998. ACM Press.
- [13] Dominic Robson. Play!: Sound toys for the non musical. In *Proceedings of the 1st Workshop on New Interfaces for Musical Expression, ACM SIGCHI*, 2001.
- [14] John Toth. *The space of creativity: Hypermediating the beautiful and the sublime*. PhD thesis, European Graduate School, 2005.
- [15] Bill Verplank, Max Mathews, and Rob Shaw. Scanned synthesis. In *Proceedings of the 2000 International Computer Music Conference*, pages 368–371, 2000.
- [16] Wolfgang Volz. *Christo and Jeanne-Claude: The Gates, Central Park, New York City*. Taschen America, LLC, 2005.
- [17] Marcelo M. Wanderley, Jean-Philippe Viollet, Fabrice Isart, and Xavier Rodet. On the choice of transducer technologies for specific musical functions. In *Proceedings of the 2000 International Computer Music Conference*, pages 244–247, 2000.
- [18] Sha Xin Wei. Resistance is fertile: Gesture and agency in the field of responsive media. *Configurations*, 10(3):439–472, 2002.
- [19] Sha Xin Wei, Yoichiro Serita, Jill Fantauzza, Steven Dow, Giovanni Iachello, Vincent Fiano, Joey Berzowska, Yvonne Caravia, Delphine Nain, Wolfgang Reitberger, and Julien Fistre. Demonstrations of expressive software and ambient media. In *Ubicomp 2003 Adjunct Proceedings*, pages 131–136, 2003.