# A Material Computation Perspective on Audio Mosaicing and Gestural Conditioning

Navid Navab
Topological Media Lab
Concordia University
Montreal, Quebec, Canada
navid.nav@gmail.com

Doug Van Nort
Topological Media Lab
Concordia University
Montreal, Quebec, Canada
dvnt.sea@gmail.com

Sha Xin Wei
Director, School of Arts,
Media and Engineering
Arizona State University
Phoenix AZ, USA
xinwei.sha@asu.edu

## ABSTRACT

This paper discusses an approach to instrument conception that is based on a careful consideration of the coupling of tactile and sonic gestural action across the layers of physical and computational material in coordinated dynamical variation. To this end we propose a design approach that not only considers the materiality of the instrument, but leverages it as a central part of the conception of the sonic quality, the control structure, and what generally falls under the umbrella of "mapping". This extended *computational matter* perspective scaffolds a holistic approach to understanding an "instrument" as gestural engagement through physical material, sonic variation, and somatic activity. We present some concrete musical and installation performances that have benefited from this approach to instrument design.

## Keywords

Topological Media, Computational Matter, Mapping Control Structures, Sonic Gestures, Enchanted Objects

## 1. INTRODUCTION

In the case of digitally-based instruments, there has been much focus on defining new instrumental systems by mimesis of acoustic instruments (e.g. digital clarinets, zithers, guitars) and as instrument-inspired interfaces which seem themselves in direct evolution of acoustic performance tradition [5]. However, Magnusson[4] notes that often the tangible and immediately perceivable interface of digital instruments is merely a shell. He argues that the expressive potential of the instrument is (often) situated in the composed symbolic instructions of the designer, rather than in physical matter. From an interaction point of view, the expressivity is thus moderated by an elaborate sequence of musical events that are shaped through "higher-level" control of musical material by a performer. This point of view considers musical instruments as "cognitive extensions" of human musical thought, and expressiveness as the navigation of composed structures in the act of performance. This trajectory of instrument design has been articulated by Schnell and Battier [9] as a process of the dematerialization of the instrument – of the progressively increasing electromechanical and now computational mediation of the coupling between somatic-gestural movement and sonic result. In re-establishing this coupled link between action and sound, the nature of the representation that one is introducing needs to be carefully considered. This includes the musical nature of the representation as well as the level of immediacy and expressive intent that is being represented. One must consider whether they are designing for an interface wherein large musical structures are "steered" or triggered by performer (at one extreme), or the moment-by-moment actions defined at the lowest level by physical performer action (at another extreme).



In this paper, we propose a "material computation" ([12]) approach to rethinking the representations of musical thought and gestural engagement that are introduced in the process of coupling somatic action with sonic variation. We feel that it is productive to move away from a purely mimetic conception of digital performance systems, and from classic paradigms such as that of composer and interpreter [2]. In fact, we do not pre-suppose that musically-relevant excitations are only those that arise from the hands or mouth of a singular human performer. However, we do not suggest giving up on gestural immediacy and the physicality of the instrumental system in making this shift. Just as Van Nort [14] has proposed to consider instrument design through the lens of the sonic gestural affordance of a given system, we propose to consider gestural potential of matter as part of the design process. By this we do not mean simply the physical properties of sensing technologies, but the encoded gesturally and computationally modulatable potential of physical matter itself - the spatial and temporal encoding of gestural potential - and how this is coupled with environment, human interaction and sonic output in continuous and connected fashion.

These are a constellation of forces at work, ones which are surrounding but external to the black/white box systems view of the instrument. To this end we propose a design approach that not only considers the materiality of the instrument, but that leverages it as a computational substrate. Such computational matter then becomes a central part of the instrument's conception of the sonic quality, the control structuring and what generally falls under the umbrella of "mapping" design. As we will discuss, this extended computational matter-centric view is of benefit towards holistically understanding an "instrumental" gestural engagement, as it is realized through physical material, sonic gestural matter and felt human engagement.

## 2. CONTINUOUS MATTER/RESPONSE

The world around us is full of rich sounding matter, affording complex sonic experiences through our physical engagement. Naturally, this was articulated early on by Cage through his explorations of amplified objects [3]. This was also artfully expressed and systematized by Tudor, through his Rainforest series of works [1] which highlight the intelligence and beauty found in deep vibrational interactions between sonic and physical materials. In our work, we further augment, enrich and transcend the natural tendencies of matter to transcode gestural manipulations into audible sounds and tactile sensation. Simply placing a contact mic on a resonant material and skillfully manipulating it does already afford an extremely rich instrumental palette, which is the starting point for a countless number of experimental music performance practices. We extend this further by taking a material-computation approach to deterministic couplings with such manual engagement; in order to allow for the full thickness and boundlessly open set of experiences that are potentially realizable through interaction with computationally-enriched matter, we avoid strongly determined systems which recognize, learn or model specific sonic/haptic audio interaction or human experience.

Rather, to enable the continuous richness of potential computational response to non-schematized gesture we focus the design on the considered coupling of physical matter and sound synthesis/processing techniques. This includes the highlighting of natural affordances of vibrating matter, as we wish to transmute (augment/enrich) its given tendency to yield audible acoustic energy whenever manipulated. In addition to static qualities, there is an implicit temporality which arises through viewing the material as computational matter: through its encoding of gestural potential in a spatial form. This is written on its surface, in the folds, the density, etc. and this too comes into consideration in the course of designing control structures.

In practice, our design work has been realized through audio-mosaicing of haptic-sonic gestures using corpus-based concatenative synthesis (CBCS)[11]. In short, CBCS methods use a database of sound "snippets" to assemble a desired sound according to a target phrase. Recent developments ([11]) enable real-time sound generation by navigation through a multi-dimensional descriptor space informing unit selection within the *corpus* and giving access to specific sound characteristics. The corpus is a large database of *units* of either pre-recorded or live-recorded sounds, that have been segmented and descriptor-analysed in order to be placed within the descriptor space, according to the extracted features. Audio-mosaicing [13] can be seen as a special case of CBCS wherein the target is set from the real-time descriptor analysis of live audio input. This technique equally accommodates gestural engagement with physical objects, whole body interaction, as well as the augmentation of sonic material in live performances as well will highlight through out examples.

# 3. DESIGNING FOR MATERIALS TO COMPUTE

In the case of instruments based on contact microphones and manual engagement with physical objects, the textural and resonant nature of the physical material becomes a central component for consideration, along with the kinesthetic gestural interactions that are conditioned through the spatial and material structure of the object. For example, we have found that a vibration-isolated wooden surface with an evenly distributed textural roughness, enough acoustic conductivity, and a balanced impulse response is ideal for transcoding a wide range of gestural manipulations carried out via human skin, nails, and light objects. Such an object will transmute gestural interactions across a wider timbral spectrum, and thus provides an optimal platform for the continuous differentiation and distinct amplification of subtle changes in the process of haptic-sound feature extraction and sonification. Consider the subtle differences in measurable characteristics of the acoustic energy that results from a fingertip rubbing across a surface with varying degrees of applied pressure. In such a scenario, rough and sticky surfaces would provide a considerably improved acoustic response and a better signal to noise ratio in comparison with smooth and slippery surfaces.

Objects such as fruits, pine cones, combs, nails, the floor, and the human body can provide other computational operations such as band-limiting, resonance, convolution, smoothing, and spatial encoding that can be exploited to transform gestures from haptics into optimal acoustic energy. When talking of optimization, our focus is mainly on accessing, in the the acoustic response, the finest levels of intentionally nuanced gesture from noise. The question becomes: how do we extract useful and optimized information about the way (non)human performer interact with the objects' surface? How do we highlight important gestural nuances that yield minimal acoustic energy from the noise floor of an input system (instrument)? How can material thinking and material computation contribute to the design of haptic-acoustic instrument that distinguish the subtlest intentional change in the input sonic gestures?

## 3.1 Haptic-Acoustic Transcoding

Combining acoustic surface sensing with a signal-driven sonification strategy can take full advantage of the richness of the feeling of touch, and thus enable the performers to rely solely on felt engagement with real matter, discovering and inventing their own repertoire of meaningful gestures in the process. In previous work we have designed the continuous potential of computational response through an architecture that implements acoustic sensing coupled with physical modeling sound synthesis [12], which allowed participant to invent a wide range of gestural vocabulary and nuance with physical consistency between action and sound. With any one instance of a physical modeling synthesis however, the performer is restricted to a more or less uniform timbral universe specified by the sonic characteristics of the synthetic physical model and its couplings with natural matter and gesture. While the instrumental potential in this system is already vast, we are interested in control structures which allow us to adapt to differing qualities of timbre and output sonic gestures as well as to play with different action/sound gestural couplings in a poetic fashion.

## 3.2 Haptic-Acoustic Transcoding through CBCS

We approach a chosen synthesis technique as a software-domain computational processes which is always co-dependant with the computational properties of matter. Even if often our goal in the transparent coupling of action and sound is to construct perceptually singular morphologies, audio mosaicing has not necessarily been chosen for its tendency to perfectly mimic and imitate the target gesture. Rather, a considerably attractive quality of audio mosaicing arises from its ability to condition the potential degree of semblance of the resynthesized sounds to a given target phrase, while keeping the continuous morphologies of the target phrase intact. The target thereof could be thought of as an abstract gesture-template consisting of feature contours and their time-dependant variables [13]. When using the input haptic-sound as "target," one may preserve morphological continuities of the audio-encoded gesture in the sonification process and yet retain novel compositional control over the timbral qualities of the output.

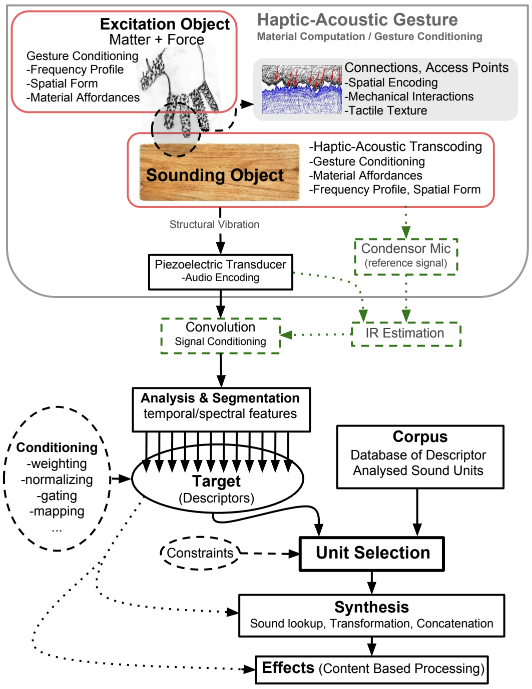


**Figure 1: Audio Mosaicing Haptic-Acoustic Gestures**

## 4. COMPOSING GESTURAL COUPLINGS

In order to condition target audio descriptors, a prioritized unit selection is parametrized by adjusting weights of each order-dependant descriptor. Then, in order to obtain perceptually differentiable results from fine gestural nuance, we normalize the target descriptors and map them to the full range within the corpus. In contexts where sonic resemblance and mapping transparency is not the main focus, one can freely scale, reverse, offset, swap and re-map the descriptors, without upsetting the consistency of co-dependanct descriptors. For example, when sonically augmenting objects with non-existent or static pitch features and yet needing to have some gestural control over pitch, one might want to explore mappings from arbitrary descriptors to pitch. Optimization and conditioning of the descriptor space for efficient exploitation of the corpus is, at the moment, time consuming and parametrically multi-layered and complex. In the future we intend to simplify this calibration process through a unified and visualized correlation space.

We are currently investigating the mosaicing of complex temporal patterns and sonic-tactile textures that result from mechanical interactions that lead to the generation of transient impact and friction impact micro-events: brushing or scratching an uneven surface with varying degrees of acceleration, ice cracking , drum rolls, etc. It is important to distinguish the analysis/synthesis of these dynamic and gesturally modulated impact, friction sounds from the current goals of sound texture synthesis [10] which deal with the resynthesis of more stable dynamic morphologies. Due to the desired immediacy of gesture to sound coupling, and the unpredictability of target texture features and morphologies, current statistical modeling approaches over longer time scales are not immediately applicable to sonic-tactile texture sonification with CBCS. We have concluded that instead an improved segmentation method as well as locally improved correlations between micro-events could lead to more promising results in the long run.

While a close coupling is easily achieved through audio-driven physical models, with audio-mosaicing we often find ambiguities within the micro event spaces. Using larger corpuses increase the likelihood of concatenated sounds which contain similar temporal profiles as the target units. However, the real achievable resolution of the morphological changes are often lower that one would expect for particular gestural expressions. Temporal and spectromorphological coupling ambiguities are particularly exaggerated in between the meso and micro event spaces where phenomenon such as tremolo and vibrato (5-8hz) occur: in between fastest repetitive gestures (12Hz) and the emergence of conscious time (600ms)[8]. We have found that small grain sizes (50-150 ms) improv the temporal coupling of mosaiced sounds with the gesture. Larger grain sizes (200-1300 ms) ambiguate the fineness of gesture to sound relationship but depending on the context and target material could lead to sonically more coherent results. In most cases, improved temporal couplings come at the cost of reduced spectral coherency in respect to the input sonic-gestures.

We have utilized several strategies to deal with this trade-off. Primary among them is the simultaneous analysis of multiple timescales (micro and meso) and incorporation of heterogeneous segmentation methods which are optionally applicable to individual micro-events for both the target and source sounds. Tremblay and Schwartz [13] also suggest a need for a binary descriptor for partially dealing with the complexity of the presence of transients within segmented units and to efficiently couple these transients with the target with low latency. On the synthesis side, we have found it beneficial to use two or more synthesis modules on the same target in order to mix different temporal scales to blend desirable responses on the fly. For example, short grain sizes(30-120) could be used to provide optimal coupling with the target sound's dynamic morphology while another synthesis module with longer grain sizes and stronger pitch variations could add a content driven soundscape in the background. We also add another layer composed from a different set of source sounds driven by the detection of significant onsets in order to provide independently determinable responses to high energy impact events. Finally, often one might want to use a limited collection of source sounds. For example: to mosaic the sounds resulting from gestural manipulations of a washboard as target using only limited samples of Mbira (thumb piano) source sounds. One main problem that often arises is that the target sounds are much more varied than what a limited corpus could offer. We have implemented attempts to systematically shape the synthesis and post-processing parameters based on input content. Mappings to the synthesis grain size and grain adsr envelope have been promising initiatives.

## 5. PERFORMING MATERIALITY

We maintain the position that matter can be a computational substrate. Such a perspective on new interfaces for musical expression can help us avoid some problematic divides in our design metaphors between performer/performed, instrument/score, intention/noise, software/hardware, digital/analog, speculation/action, and etc. Matter does not distinguish between performer intentions and material physics; we claim the same holds for computational matter. As designers of interfaces we can employ this inherent symmetry to design for arbitrary associations of agents doing arbitrary actions.

### 5.1 Gesture Bending

Gesture Bending, a generic term coined by the first author, refers to the poetic transformation, prolongation and enrichment of gestures through staged and unstaged technical mediation of movement – in this case through the incorporation of real-time sound instruments and computational matter. The goal of Gesture Bending is to continuously enact persuasive conditions for the transformation of the discursive networks of meaning production in the embodiment of movement. It can for example lead to the signification of an empty gesture or the abstraction of an inherent signifier (ie. within a beat gesture). Pervasive Gesture Bending can lead to the emergence of social experiments, multidimensional compositions and the creation of conditions that invite inhabitants to synergetically improvise with a hybrid expressive force.

In a recent workshop at the Topological Media Lab, the first author employed audio-mosaicing instruments from the Gesture Bending toolkit [6] to prepare the floor as an instrument; the event focused on populating this instrumental space with diverse set of activities and social events. The participants' gestures not only lead to unexpected musicality but to narratives about shaping relationships with the immediate world and recognizing daily life and the material world as a platform for play and for refined practice. Participants discovered that their everyday movement can create intricate sonic textures[1] and developed their own unique vocabulary of sound generation to sculpt musical events via engagement with the floor.[2] Others set objects such as tennis balls into motion, allowing objects to effectively "perform" music.[3] Such experiments illustrate how any physical gesture can effectively augment physical objects with gesturally-conditioned sound, augmenting said objects' material qualities. Through interactively varied augmentation of the object's natural acoustical response, an a priori distinction "synthetic" and the "natural" and the "performer" and "performed" becomes unnecessary. Performing a score or improvising music could turn into a hybrid mode of engagement and perception borrowing elements from gaming, playing, building, day to day living practices, puppetry, and performance art.

## 5.2 Practices of Everyday Life | Cooking

"Practices of Everyday Life | Cooking" [7] is the first part in a series of performances and installations exploring how everyday gestures could become charged with symbolic intensity and used for improvised play. A performance choreographed around a chef and sonified objects: fruit, vegetables, meat, knives, pots and pans, cutting board and table.

Cooking, the most ancient art of transmutation, has become over a quarter of a million years an unremarkable, domestic practice. But in this everyday practice, things perish, transform, and nourish other things. By augmenting the meats, wood and metal with sound and painterly light, we stage a performance made from the movements and gestures of cooking, both high cuisine and everyday. The performance features a dancer who is also a virtuosic chef who wields foods, knives, pans and spices transmuted gesturally into real-time sound instruments. Within our responsive scenography system, every cooking process is transformed into an environment thick with aroma, light, video, sound, movement, and objects. A knife sleeking against another knife, carrots vocalizing their unfolding mutation into a cacophonous a cappella, the sizzle of hot oil mosaiced into a downpour of Bartok pizzicati along with the aroma of onion and garlic immerses the audience in an ecology of remembrance and anticipation. At the end, the performer offers the audience a chance to taste the dish that is prepared.

**Figure 2: Practices of Everyday Life | Cooking**

The participants are given a chance to extract new and unbounded forms, meanings, affects and percepts from the otherwise familiar situations such as cooking. The emergent mental modalities are more likely to be in closer contact with ecological and mental complexities of socio-gestural behaviour than if they were left colonized by the subjugated tonality of standardized mentality.

# 6. CONCLUSION

Works such as "Practices of Everyday Life | Cooking" and other Gesture Bending experiments leverage material thinking and acoustic sensing techniques to symbolically charge everyday actions and objects in ways that combine the composer's design with the performer's contingent nuance. Our material computational design allows for any potential movement at all by the performer or the objects to turn into potentially musical gestures. This removes the burden of modeling the human experience and instead allows for such notions as gestural meaning, intentionality, expressivity, noise, musicality, and even performer, performed and speculator to freely arise from the context established in the moment of performance together with the theatrical apparatus of expectation.

# 7. REFERENCES


[1] J. Driscoll and M. Rogalsky. David tudor's rainforest: An evolving exploration of resonance. *Leonardo Music Journal*, 14:25–30, 2004.

[2] M. Gurevich and J. Treviño. Expression and its discontents: toward an ecology of musical creation. In *Proc. of the 7th international conference on New interfaces for musical expression*, pages 106–111. ACM, 2007.

[3] D. Kahn. John cage: silence and silencing. *The musical quarterly*, 81(4):556–598, 1997.

[4] T. Magnusson. Of epistemic tools: Musical instruments as cognitive extensions. *Organised Sound*, 14(02):168–176, 2009.

[5] E. R. Miranda and M. M. Wanderley. *New digital musical instruments: control and interaction beyond the keyboard*, volume 21. AR Editions, Inc., 2006.

[6] N. Navab. http://gesturebending.weebly.com/.

[7] N. Navab, M. Montanaro, and J. Delappiere. http://www.practicesofeverydaylife.com.

[8] C. Roads. *Microsound*. MIT press, 2004.

[9] N. Schnell and M. Battier. Introducing composed instruments, technical and musicological implications. In *Proc. of 2002 Conference on New Interfaces for Musical Expression*, 2002.

[10] D. Schwarz. State of the art in sound texture synthesis. In *Proc. Digital Audio Effects (DAFx)*, pages 221–231, 2011.

[11] D. Schwarz. The sound space as musical instrument: Playing corpus-based concatenative synthesis. *New Interfaces for Musical Expression (NIME)*, 2012.

[12] X. W. Sha, A. Freed, and N. Navab. Sound design as human matter interaction. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 2009–2018. ACM, 2013.

[13] P. A. Tremblay and D. Schwarz. Surfing the waves: Live audio mosaicing of an electric bass performance as a corpus browsing interface. NIME, 2010.

[14] D. Van Nort. Instrumental listening: sonic gesture as design principle. *Organised sound*, 14(02):177–187, 2009.


[1] http://vimeo.com/68112441
[2] http://vimeo.com/36977151
[3] http://vimeo.com/68105290