Musical Notation for Multi-Touch Interfaces

Warren Enström
University of Wisconsin-Milwaukee
3223 N. Downer Ave.
enstromw@uwm.edu

Joshua Dennis
University of Wisconsin-Milwaukee
3223 N. Downer Ave.
jndennis@uwm.edu

Brian Lynch
University of Wisconsin-Milwaukee
3223 N. Downer Ave.
bjlynch@uwm.edu

Kevin Schlei
University of Wisconsin-Milwaukee
3223 N. Downer Ave.
kevinschlei@gmail.com

ABSTRACT
This paper explores the creation and testing of a new system for notating physical touch actions on a surface. This notation is conceptualized through and initially tested on multi-touch interfaces. A brief examination of movement notation concepts is followed by an overview of design challenges considered. The notation symbols and formats are detailed. A testing procedure was carried out in order to evaluate how effective this notation system was. The results of the tests are analyzed, and results and criticisms are discussed.

Author Keywords
motion, gesture and music, notation, multi-touch interfaces, action-based music

1. INTRODUCTION
The need for notating multi-touch surface movement comes from the onslaught of new devices that rely on a multi-touch interaction. The different forms and shapes of these devices and controllers present a challenge if one wishes to reuse notation methods between instruments. Our aim is to solve the problem of notating music for these devices whether they may be on a tabletop [4], spheres [1], or arbitrary shapes like leaves on a plant [8].

The need for a universal touch-based notation has been noted by previous scholarship [2]. It would also provide composers a starting point for notating multi-touch interfaces. For example, Danny de Graan's concerto for iPad and orchestra *With(out) Words* uses Western standard notation with written instructions, such as "rub with fingers." The composer noted that he chose this approach in part because he knew of no tradition in of gestural notation.

It would be useful to have a system for notating actions to be taken upon a multi-touch interface of any variety, that is simple to understand and flexible enough for the idiomatic aspects of each individual application to be considered and notated through standard symbols.

2. NOTATIONS OF MOVEMENT
2.1 Notations of Physical Movement
Traditionally, a musical score uses notes and articulation markings to represent the sound that the ensemble is to produce. This approach assumes that the performer is already familiar with the physical interactions of their instrument, and can reproduce the desired sound using those movements.

NIME'15, May 31-June 3, 2015, Louisiana State University, Baton Rouge, LA.

Notes and articulation markings are less successful in the realm of multi-touch interfaces. Because of the vast number of synthesis and sampling possibilities that can be tied to multi-touch interfaces, creating a standardized sound-depiction is a daunting if not impossible task.

For this reason, the notation system conceived in this research aims to communicate physical touch movement on a surface, rather than the expected sonic result.

2.2 Related Works
Attempts have been made at notating human movement for centuries, often for the purpose of documenting dance choreography. Notation of dance in particular changes dramatically in style as emerging forms and movements became increasingly complex [3]. One can look at highly stylized engravings used in renaissance dance instruction as simply tracing a footpath, which suits the needs of its time perfectly. Contemporary choreography, on the other hand, requires the use of far more complex notational systems such as Labanotation [7] to map out three-dimensional movements of separate limbs and digits through time.

Underlying the thinking of composers such as Berio, Lachenmann, and Kagel was an unspoken engagement with enactive cognition, in which the act of doing is a model for creation and performance. There have been numerous musical compositions in recent decades that set about notating the physical space and actions of the performer rather than the sound per se [5].

The efficacy of these notations can be judged on the balance of elements such as intuitiveness, meaning how closely notation visually resembles the thing it is representing, and complexity, being the quantity of symbols a notation uses. The true metric for how different aspects of a notation are to be balanced is, ultimately, the use it is serving [7]. For this project, flexibility and simplicity were valued higher than absolute precision. Most of our symbolism is used commonly and is fairly recognizable. Elements such as location on a surface and rhythm can be written in a straightforward way that emphasizes gestural accuracy.

The notation was developed considering the parameters defined by Ganhör and Spreicer, with design modifications aimed at solving musical challenges [2]. Among other decisions, our system does not notate consequences of touch activity, such as application responses. Tap sequences were considered in the context of musical rhythm, which outgrew representation by concentric circles in [2]. Dual-direction arrows were eliminated for clarity, and acceleration during a drag was considered a significant musical requirement.

3. DESIGN CHALLENGES
A number of issues arose during early draft work. The first issue was conveying time. In a musical score, time is traditional notated from left to right. Because of the need to

demonstrate two-dimensional motion, this approach to time was deemed too impractical.

The next issue that arose was how to address temporal sequences of events. If there were several touches in a gesture, how would the performer know in what order to perform the touches? Further, there was a need to designate simultaneous touches.

Speed and acceleration of dragging a touch were also issues.

Early revisions showed that some complex gestures were visually overloading, and would benefit from simplification. For example, if a large number of touches were moving in a similar direction, many arrows would clutter the notation. Even worse, gestures within gestures were a dense cluster and very difficult to read.

4. NOTATIONAL DESIGN

The notation system devised uses a small collection of symbols drawn to represent touch movements on a two-dimensional plane. These symbols take place on the *canvas* (Figure 1L). The canvas itself represents the touch-sensitive area of a multi-touch interface. We defaulted to an approximately 4:3 rectangle, as we were testing with iPads, but the notation is designed to be stand-alone, not relying on any particular canvas shape (or edges at all) to be effective. It would be simple to design a canvas whose shape is not rectangular, i.e. drawing a leaf-shaped canvas if one wished to write a piece for the aforementioned leaf interface, or a circle if an interface for a circular tabletop was designed.

By designating the canvas as representational of the physical space used, we solved one part of our time problem. In order to notate time, we decided that a canvas would be followed by a timeline (Figure 1J), extending to the right of the canvas with a time marker (written in seconds in our tests) placed above the line. The gestures that appear on the canvas take place over the allotted time.

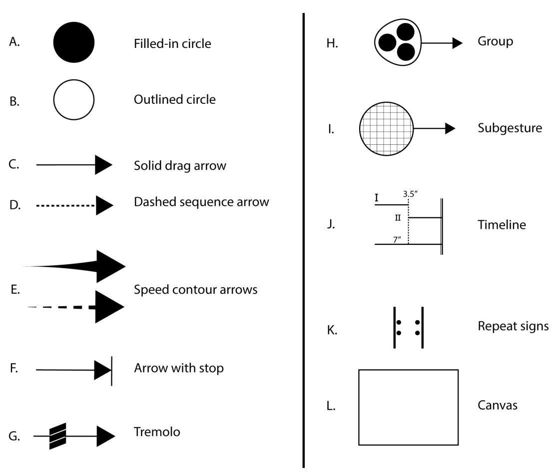

Figure 1 - Terms and Symbols

We use circles to represent touches as they represent a generic finger shape while still being legible at a variety of sizes. Filled-in circles (Figure 1A) represent touches that are held – the touch is initiated, and then remains present until the touch is removed. Outlined circles (Figure 1B) represent taps – a moment of short contact between fingertip and surface.

In regards to motion, we use arrows to designate motion. Solid drag arrows (Figure 1C) indicate a touch that moves in contact with the surface, while dashed sequence arrows (Figure 1D) indicate a sequential series of touches.

To show acceleration and deceleration, the thickness of the arrow can be varied. A thicker body indicates a faster speed, while a thinner body indicates a slower speed. Thus, an arrow with a thickness gradient from thin to thick (Figure 1E) indicates an accelerating motion. These gradients can also be applied to a dashed arrow to indicate increasing/decreasing time over a motion (Figure 1E). Finally, a solid drag arrow with a stop (Figure 1F) means the speed of the motion reaches zero where the head stops, whereas an unstopped arrowhead (Figure 1E) indicates the speed of the motion does not reach zero where the head stops (and represents a fling, or toss, or throw, where a touch has a non-zero speed when the fingertip loses contact with the surface).

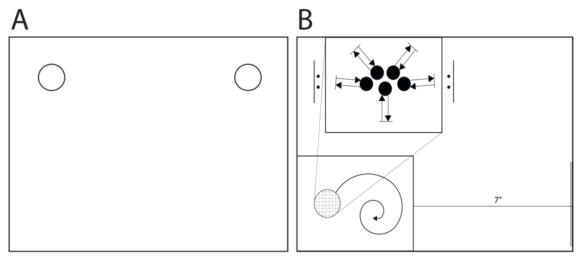

Figure 2 - Test Gesture 1 (A), Test Gesture 8 (B)

There are two data structures within our notation. First, if a number of touches are surrounded by an enclosure (Figure 1H) they are called a *group*. Groups are used to indicate when many touches move together as a unit. This is useful in cases where notating each touch individually causes clutter or confusion. Second, a shape filled in with a grid (Figure 1I) indicates a *subgesture*. Two thin lines extend from the subgesture symbol to a rectangular shape above the canvas, which contains touches and motion indicators. On the main canvas, these subgestures can then be given a larger motion scheme, which cleanly indicates complex motions that would otherwise require crossing or looping arrows.

Finally, a few aspects of our notation are pulled directly from Western standard notation. We designate repeated gestures with repeat signs (Figure 1K) placed around the canvas or subgesture that is to be repeated. We use tremolos to indicate dragged motions that are to be tapped quickly as the gesture progresses (Figure 1G). Finally, we used double bar lines to indicate the end of a piece (Figure 1J).

5. TESTING PROCEDURE

In order to get a sense of how well the notation communicated movement, we tested fifteen people of differing backgrounds with a series of gestures. Each of these participants was given a brief questionnaire asking if they had any musical experience or experience reading sheet music, and what their field of study or profession was. These few questions would later offer some insight on how different skill sets informed differing interpretations of the test examples. Thirteen of these fifteen people were musicians, ten of whom read traditional Western musical notation.

After participants filled out the survey, they were given a key similar to Figure 1 that catalogues all of the symbols. They were given as long as they liked to study this key and ask the test administrators questions for clarification. We gave each participant ten examples (Figure 2) engineered to convey different types of movement and physical interaction with an iPad surface. Participants were given the ten examples in sequence, which were to be studied and practiced for two minutes before ultimately being "performed" on the iPad three times. There was no structured form of practice; the participants practiced however they felt it was appropriate, often tracing the notation on the examples, or practicing directly on the iPad surface.

The examples given were meant to increase steadily in difficulty, introducing more complex ideas that require the use of multiple symbols and specific timing. For instance, initial examples were meant to test merely the amount of time to hold touches in place, location on the screen, or simple linear movement (Figure 2A). As they progressed in difficulty, concepts such as broken timelines, sub-gestures, speed gradients, and multiple groups were introduced and combined (Figure 2B).

The participants' hand movements were video recorded during their performance. These recordings were then observed and compared to see how each individual interpreted our notation, and specifically how they compared to what we had intended them to be. The iPad was loaded with a patch in TC-Data [6] that was sending MIDI data to a sequencer for recording. It was decided that the app should not produce sound, in order to prevent the participant being distracted. The visuals created by the app have no relation to the notation system we created.

The data sent included the total number of touches present at a given moment, the average distance each touch was from the center of a group, the x-coordinate of the center of a group of touches, the y-coordinate of the center of the group of touches and the speed of the center of the group of touches. We then catalogued and compared the trajectories of these continuous controllers, and the time taken in each example, in order to make normative comparisons of how certain ideas were understood. These are evidenced in Figures 4A, 4B, and 4C and are discussed further in Section 6.

After each participant finished performing all of the examples, they were then queried on how well they thought they adhered to the rules given to them, and they were asked for additional feedback that will be discussed in Section 6.

6. DATA ANALYSIS

6.1 Graphic Analysis

An important detail to consider is our use of qualitative, normative comparison for analysis. Our primary goal in creating this notation is to communicate musical gesture on a non-traditional playing surface. A precise replication of the trigonometric functions present in our design is of far less interest than adherence to general features of an idea, and the degree of precision to which one can interpret a notated gesture lies in the hands of composers, teachers, and performers.

We found that, overall, participants were able to replicate number of touches and location of touches on the screen. Further, the participants eventually properly replicated the drag and sequence arrow paths.

We recorded a large amount of data, both as video files and MIDI controller data, and to sort and analyze it we organized the MIDI data into graphs and compared them to a recorded archetypal performance created by one of us.

For example, Figure 2B contains a subgesture indicator, which shows five held touches expanding and contracting. This gesture was then repeated over seven seconds.

Figure 3A is the MIDI data graph of our "ideal" performance of Figure 2B. Compare this to Figure 3B, a "good" performance by one of the participants. Note how the overall curve shapes are generally replicated – number of touches stays relatively consistent at 4-5, while Group center X and Y values increase, then decrease, then increase again, closing in on the proper value, showing the spiral. We decided average distance to group center increases and decreases periodically as the touches expand and contract.

Figure 3C is a "poor" performance by one of the participants. This participant interpreted the repeat sign to apply to the bottom canvas's "looping" shape, taking roughly half of the time allotted to complete one of the subgesture shapes before repeating that process again to fill a seven second space. What we had intended in writing this particular example was for the expansion/contraction subgesture to be repeated indefinitely while the entire hand was to be traced along the path once over seven seconds. This interpretive error was not anomalous; many of the participants interpreted repeat signs in different ways, some repeating the entire length of a phrase, others only repeating once rather than indefinitely. Of particular importance here is that, despite aforementioned interpretive errors, the participant replicated many parameters correctly, namely the shape of the gesture and subgesture, here represented by the shape of the Average Distance to Group Center line. The participant's graph represents one period of our subgesture, while our "ideal" graph shows more periods over the same amount of time. Additionally, the participant's Group Center X and Y positions and number of touches are accurate.

6.2 Design Analysis

Most participants understood how the canvas and timeline worked. However, this often occurred after the initial two gestures, where most people erroneously assumed the canvas should be read from left to right. It was noted that participants who were familiar with proportional notation were more adept at keeping their gestures in time.

Similarly, most participants understood that the dashed sequence arrow indicated the order in which touches were to occur. There was some confusion between the dashed sequence arrow and the tremolo, as some participants interpreted the dashes to mean tapping.

In terms of speed, all participants understood that the solid drag arrows implied motion along the designated path. When introducing acceleration and deceleration, most participants attempted to emulate the speed gradient shown to them, but it was one of the hardest gestures for participants to perform. Additionally, the dashed sequence arrow, with a gradient applied, was ambiguous, with about half of the participants correctly understanding that the sequence of events is meant to accelerate, rather than each element individually.

Groups were one of the most successful notations used. Almost all participants performed the groups as intended, with only a few mistakes. Subgestures were similarly successful, with all participants understanding that the subgesture was to be performed while following the path shown on the canvas.

7. CRITICISM

7.1 Notation

As soon as the first tests had begun, we realized that there were a few things that would have to be changed. First, there was an ambiguity about what to do when two touches appear without grouping or a sequence arrow (as in Figure 2). We intended for non-grouped touches without a sequence arrow to be touched simultaneously. However, it was proven after testing that grouping touches was far more effective in showing simultaneity. In the easiest gesture that we designed (Figure 2) most of the participants touched the screen twice, with only a few participants touching the screen in the two locations simultaneously.

In situations when gestures were too clustered for proper detail, we used subgestures to enlarge and clarify. These subgestures cause problems because they alter the scale of the canvas between the full canvas and the subgesture. Once a certain expectation of the canvas was understood, participants had to adapt to a new scale as it relates to time and space on the page.

The repeat signs were ambiguous and received mixed interpretations. It would be better to design a scheme for repeats that explicitly communicates the desired repetition count or rate.

The repeat sign ambiguity lead to a significant realization about the notation system. When we used traditional music notation, performers who identified as musicians often had a difficult time contextualizing the old notation in a new way. The old meanings often created confusion over how the gesture was to be performed. However, the non-musicians in the study seemed to catch onto the use of these notations quickly, having no previous experience with them, and thus no pre-existing connotations. It may be better to rid the system of Western standard notation symbols entirely.

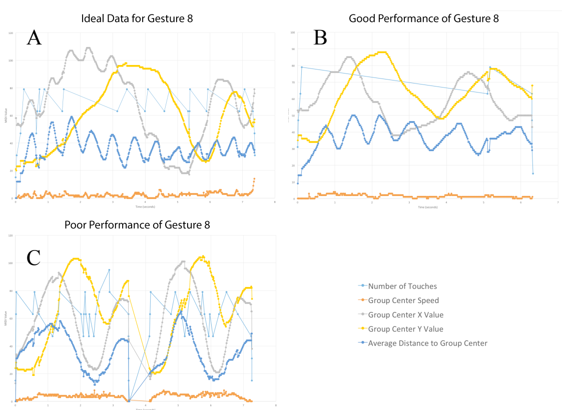

Figure 3 – Graphs of Performances of Gesture 8

7.2 Testing

During testing we had the participants test each gesture individually, that is to say, one gesture at time, with pauses, instead of as a fluid string. If these gestures were part of a larger piece it would be advantageous to show where each page fits into a larger timeline in order to coordinate events with other performers. In time-based notation the performers generally know where each page fits into the larger scheme, which is an idea absent from out tests.

Perhaps the most egregious problem with the testing was that we only tested the notation with one iPad app. The reason for this was the ability of the app to capture and record MIDI data, but it doesn't show us how the data would translate to another device, which was a main focus when we started the project.

8. CONCLUSION

8.1 General Conclusions

We designed a notation for multi-touch interfaces by representing touches and motion via graphics. The graphics, touches and motion can be structured into groups and subgestures to convey complex movements, all while still maintaining a simple notation that can be applied to many different sorts of multi-touch interfaces.

When designing the notation, we tried to use as intuitive symbols as possible in order to not confuse the performer, while still paying attention to overall messiness, legibility, and reproducibility. The notation allows for compatibility with a variety of interfaces and temporal possibilities.

Overall, the participants were able to read and understand simple information about touches and motion. Many participants quickly grasped the difference between a tap and a

hold, and also quickly understood how to perform our notation when it showed motion. As the complexity increased, participants had more interpretations of our notation, but groups and subgestures helped the participants to replicate certain phrases.

This notation has potential for many kinds of multi-touch interfaces, including non-flat multi-touch interfaces. Having a standardized notation for multi-touch interfaces will allow for greater communication between composer and performer.

8.2 Future Work

Our next steps forward will explore further methods of testing the effectiveness of our notation, including extended periods where participants could take passages home to practice. This method will offer more insight than our current results, which mainly gauge how well our system can be sight-read. It will also offer an opportunity to more accurately test how well complicated ideas can be communicated. Our participants were hampered in their ability to execute difficult passages due to the time constraint of the testing procedure.

In addition to giving people more time to study scores and practice them, we would like to test with a variety of multi-touch interfaces to see exactly how applicable our notation is to other systems. While our tests were successful for TC-Data, they might fall apart in other systems.

Finally, it would be interesting to see if people are able to transcribe existing gestures into our notation. In our tests up to this point, we tested the participant's understanding of our notation through their ability to read and perform it, but their understanding could also be measured through their ability to write in our notation as well.

9. REFERENCES

[1] R. Balakrishnan, H. Benko, and A.D. Wilson. Sphere: Multi-Touch Interactions on a Spherical Display. Monterey, CA. October 20, 2008. http://www.msr-waypoint.com/en-us/um/people/benko/publications/2008/Benko_Sphere_UIST2008.pdf

[2] R. Ganhör and W. Spreicer. Monox: Extensible Gesture Notation for Mobile Devices. MobileHCI. Toronto, ON, Canada. September, 2014.

[3] A. Guest. Dance Notation. *Perspecta* Vol. 26 (1990): 203- 214.

[4] J. Hochenbaum and O. Vallis. Bricktable: A Musical Tangible Multi-Touch Interface. Berlin Open. Berlin, Germany. June 22, 2009. http://www.berlin-open.org/uebersicht/116-brick-ein-projekt-des-california-institute-of-arts

[5] J. Kojs. Notating Action-Based Music. *Leonardo Music Journal* Vol. 21 (2011): 65-72.

[6] K. Schlei. "TC-Data: Extending Multi-Touch Interfaces for Generalized Relational Control." *ICMC – SMC*, Athens, Greece, 2014.

[7] S.H. Waters and A. Gibbons. Design Languages, Notation Systems, and Instructional Technology: A Case Study. *Educational Technology Research and Development* Vol. 52, no. 2 (2004): 57-68.

[8] F. Weil. Artistic Human Plant Interfaces. Master Thesis, Universität für künstlerische und industrielle Gestaltung – Kunstuniversität Linz.