Express it!: An Interactive System for Visualizing Expressiveness of Conductor's Gestures

Kyungho Lee, Donna J. Cox, Guy E. Garnett, Michael J. Junokas
Illinois Informatics Institute
University of Illinois at Urbana-Champaign
NCSA Building, 1205 West Clark, Urbana, IL 61801, USA
{klee141, donnacox, garnett, junokas}@illinois.edu

ABSTRACT
A conductor provides a single unified vision of how to interpret and perform music. However, perceiving a conductor's musical intention and expression is quite challenging as they convey information to performers with subtle, nuanced, and highly individualized gestures. This artwork visualizes the conductor's gestures in order to give the audience a better understanding of its expressivity. To represent the expressivity of the gestures, we created motion profiles over eight frames, at 30 frames per second, and compared them to previously modeled gestures using three motion factors, called Weight, Space and Time from related concepts in Laban Movement Analysis (LMA). Based on this, we have created a real-time, interactive visualization that is driven by the motion factor parameters. The visualization receives the input video stream, and it is transformed into a representation of the three motion factors extracted from the real-time conducting gestures.

Author Keywords
Visualization; expressivity; conductor's gestures; gesture recognition; music

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
"Arms carve the air. A hand closes as if to pull taffy. An index finger shoots out. The torso leans in, leans back. And somehow, music pours forth—precisely coordinated and emotionally expressive—in response to this mysterious podium dance." [27] This quote describes how a conductor's gestures look on the stage. Just as a theater director makes various interpretive decisions with a Shakespearean tragedy, conductors study the score to find the significance they want to bring out in a performance. While leading the orchestra, conductors use their body movement as a medium to deliver musical intentions and expressions to the ensemble, indicating dynamics, tempo, articulation, balance, and general qualities of sound and performance. The art of conducting is more than just synchronization and signaling gestures: it is a fascinating—and under explored—gestural interaction model for conveying high-bandwidth qualitative and quantitative information from the conductor to sometimes hundreds of performers.

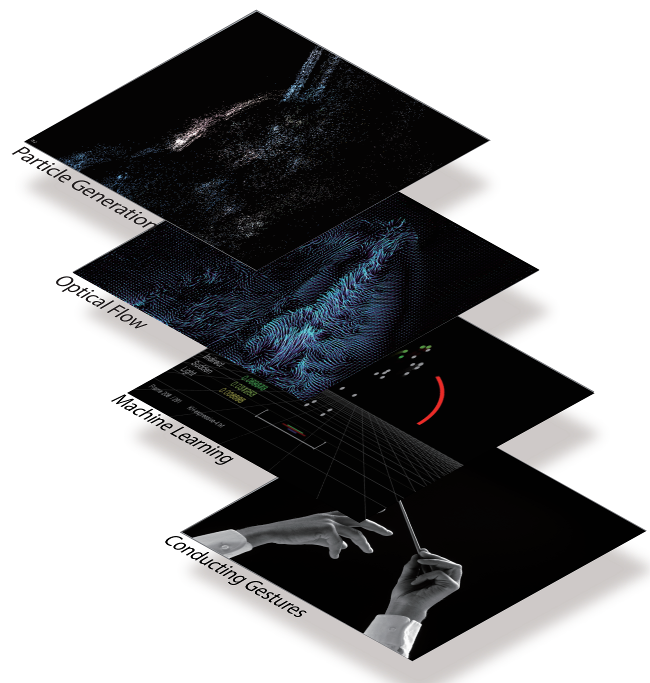

Figure 1. Express it!, a four-layer, machine-learning driven interactive visualization system.

Inspired by these characteristics of conducting gestures, many researchers have adopted the unique aspects of conducting gestures to develop interactive systems, driven by expressive, musical gestures. For instance, Mathew's Radio Baton is one pioneering attempt to control the musical tempo, dynamics, and expression of the music with expressive conductor-related gestures [18]. More directly, Nakra

C&C 2015, June 22–25, 2015, Glasgow, United Kingdom.
2015 ACM. ISBN 978-1-4503-3598-0/15/06$15.00
DOI: http://dx.doi.org/10.1145/2757226.2757243

[20] developed the Conductor's Jacket to capture and interpret physiological and motion information from musicians in order to better understand how they express affective and interpretive information while performing. Lee et al. [15] also created an interactive interface, iSymphony, that enables users to control the tempo, dynamics, and instrument emphasis of an electronic orchestra in a pre-rendered video clip by using three different baton techniques. However, little attention has been paid to quantifying, characterizing, and visualizing the expressiveness of conducting gestures themselves. Our work does exactly this: we use machine learning and other data analysis tools to drive a creative visualization of the data that is a personal yet informative reflection of the conductor's gestures and contributes to a better understanding and appreciation of the underlying meaning.

This uncharted territory is challenging due to several reasons: (i) there is no general framework for characterizing the expressive aspects of the conductor's gestures distinctly from the synchronization and signaling aspects; (ii) even given such a framework, it is not clear how to extract these expressive aspects automatically; (iii) once extracted, it is difficult to render them visually in a way that represents the key expressive features and does not merely duplicate the signaling content. In this paper, we contribute in each of these areas, but we focus on designing an interactive visualization system as an art form to represent the expressiveness of conducting gestures. We develop, inspired in part by LMA, a rudimentary framework to characterize the expressive gestures, then use a statistical template matching process to generate analytic data that is then visualized. Finally we show how it is apprehended by the audience at the exhibition.

EXPRESSIVITY IN CONDUCTORS' GESTURES

Even though conductors do not participate in producing a particular sound directly, they make a large number of diverse and often idiosyncratic physical signatures—by which we mean to include things such as facial expressions and body postures in addition to the usual hand and arm gestures—to deliver musical directions and expressions with their body movement at each moment. While some of these signatures are acquired by training in the conventional grammar of conducting—the latter including such elements as indicating beat placement and tempo through baton technique—the ways of communicating the expressive information tend to be more individual, *ad hoc*, and subjective.

Furthermore, conducting gestures, as a particular subset of musical gestures, contain intentions and the physical expressions of those intentions. Underlying these gestures are two tasks: First, the conductor maintains an inner-attitude that enables them to conceptualize their musical intention and translate it into gestures that are communicated to the ensemble musicians. Second, physical activity creates ongoing streams of information, enabling conductors to share their musical intention and to shape the playing of the orchestra. These nuanced gestures can include discrete events, such as an ictus to mark a temporal location, or continuously varying parameters, such as the moment-by-moment velocity or acceleration of the baton, etc. How does an observer, in the ensemble or the audience, interpret these gestures? How can we describe a conductor's signatures in a way that allows us to extrapolate expressivity from them? To begin to answer these questions, we consider the conductor's gestures as existing along a continuum, such as the "Kendon's Continuum" defined by McNeil, [19]. See also [11]. In order to describe the characteristics of the gesture, many researchers in the musical domain have adapted the Kendon's Continuum to consider movement in the perspective of linguistics.

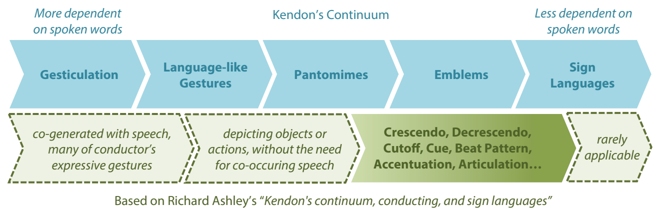

Figure 2. A diagram of the original Kendon's continuum and the musical analogue to Kendon's continuum.

Figure 2 illustrates how Kendon's Continuum can be reinterpreted in terms of expressiveness revealed in movement from the musical context [1].

This continuum proceeds from *Gesticulation* (motions that are merely related to and more-or-less synchronous with the semantic content—in McNeill's case speech, in our case music), to *Speech-framed Gestures*, to *Emblems* (which are signatures with a pre-existing conventional meaning–such as the beat patterns, accentuation, legato or staccato articulations, etc.),to *Pantomime*, to *Signs* (which are signatures with specific semantic content–such as, musically speaking, the sign for a cutoff at the end of a fermata, etc.). We focused mainly on the part of the continuum from gesticulation to emblems. While some progress has been made on very restrictive cases, we do not yet have a general theory of conducting gestures that will enable us to extract all the expressivity.

These *ad hoc* movements also incorporate a series of interrelated, yet separate streams of expression with the body. Referring to Kendon's Continuum, we can consider these types of gestures as conducting Emblems since these gestures create highly conventionalized, context-independent meanings that function as signs which can be 'read' by the musicians [7]. By using such emblemized gestures, conductors can efficiently (i.e., with little rehearsal or explaination) communicate some of their musical intentions to musicians. Simultaneously, conductors use body expressions as a medium to deliver expressive information, which is difficult to communicate solely through emblematic gestures. Revisiting Kendon's work, we can consider these gestures as having another dimension functioning as Gesticulation. According to McNeil, gesticulations are usually co-generated with speech, but having no conventional meanings and are highly dependent on context for their interpretation.

In addition to the above, we needed to keep in mind the spectrum of variability in individual conductor's approaches. For example, Lee [16] claimed that different conductors educated in different institutions and traditions, such as, Rudolf, Saito, and Green, will apply different gestural ideas, expressions, and technical principles, even for making simple 4/4 legato bean patterns. He also argued that conductors will greatly influence the orchestra and offer more nuanced and expressive performance as determined by his/her imaginative gestural interpretation. From a more practical, empirical perspective, Buck et al. pointed out that expressive musical gestures can be considered an affordance as well as a natural mediator between the mind and body movement [2]. They also believed that expressive gestures add extra-musical parameters to the audience's perception of music, creating a multi-modal experience wherein the visual affect of body movements are perceived as another means of communication. In addition, Jensenius et al. suggested the possibility of interpreting musical gestures as a body-mediated metaphor, meaning that gestures are able to function as conceptual objects that project physical movement, sound, or other types of perception to cultural topics [9].

Nevertheless, in order to extrapolate and visualize the expressiveness of conductors' gestures in a more generalized way, we needed a theory of meaning for movement that was related but distinct from the linguistic perspective of McNeil. For this, we turned to Laban Movement Analysis (LMA). For this work, we simplified and adapted elements of LMA, since it emphasizes an understanding of body movement as one aspect of a projection of an intentional process. Researchers in the field of LMA believe that the dynamics of the human body should be perceived as an outward expression of inner intent of the subject that eventually leads people to create phrases and relationships of movement that reveal personal, artistic, or cultural style.

It provides sophisticated languages and grammars that enabled us to observe and break down movement to the tiniest detail. Though LMA recognizes categories of Body, Effort, Shape, and Space (BESS), we focus here on Effort. Effort deals with the way the body moves: *"Every human movement is indissolubly linked with an effort, which is, indeed, it's origin and inner aspect. Effort and it's resulting action may be both unconscious and involuntary, but they are always present in any bodily movement."* [12]

The elements of Effort that we focus on are three of the fundamental motion factors: Space, Time, and Weight. These are each defined in terms of a continuum between two poles: Space goes from Direct at one extreme to Indirect motion at the other; Time goes between Sudden and Sustained; and Weight goes between Strong and Light. With this perspective, we created a simple low-level model of human movement that provided a baseline for representing the expressive content of the human body movement.

Charles Gambetta adopted the concept of Effort in the world of conducting in order to expose conductors to a more expansive range of movement possibilities for inspiration and exploration of their potential. In his dissertation, he used LMA to help four conductors in their training sessions. As a result, he reported that there were significant changes in movement choices [7]. With these promising results in mind, we prototyped a machine learning system to parameterize expressivity along the lines suggested by LMA, to see how elements of musical expression are represented in conductors' gestures and how they can be visualized to benefit conductors, musicians, and audiences.

DESIGNING AN INTERACTIVE SYSTEM

Motivation and Goal

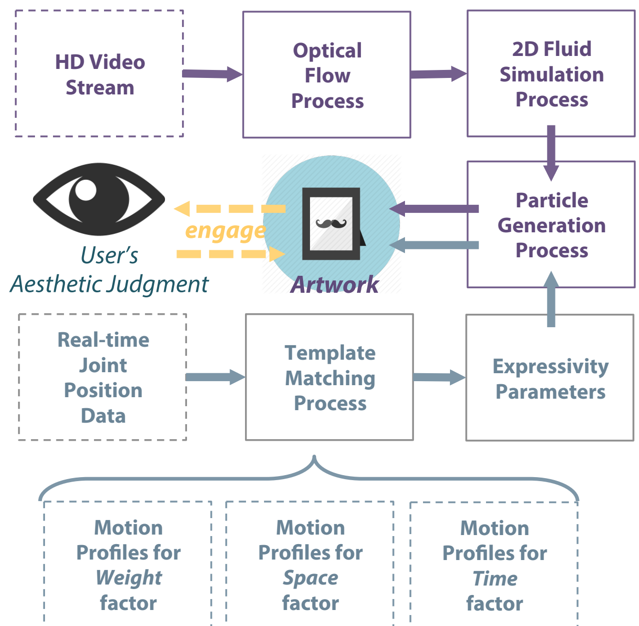

Figure 3. A flow diagram illustrating how the artwork is created by incorporating two processes, and how it is engaged with the user.

As we briefly described above, a conductor's gestures contain two different dimensions: intentions and expressions. It is difficult to discern the former from the physical manifestation of the latter. Yet, the wealth of expressive detail, combined with our intuition and personal experience of conducting, encouraged us to explore creative visual mappings of numerical data in order to evoke expressivity. We defined this interactive system as an artwork, rather than an interactive information visualization system, because our current focus is on aesthetic value rather than utility. By referring to the aesthetic value, we focused on visual style and experience as well as user input and feedback as Lau and Moere suggested in their domain model for information aesthetics [14]. We hope the artistic visualization techniques explored here facilitate the communication of meaning in a way that transcends analysis and affords the audience a deeper, intuitive appreciation of gestural interpretation. Figure 3 shows how our vision and technical specifications are intertwined in the artwork.

Design Process

We created a practical framework, as a design process, based on Camurri et al's conceptual frameworks [3], which enables capturing, processing, and visualizing expressivity in conducting gestures. Figure 4 illustrates how we built a framework on their concepts. The layers on the top show Camurri's four layered framework with our process highlighted in blue,

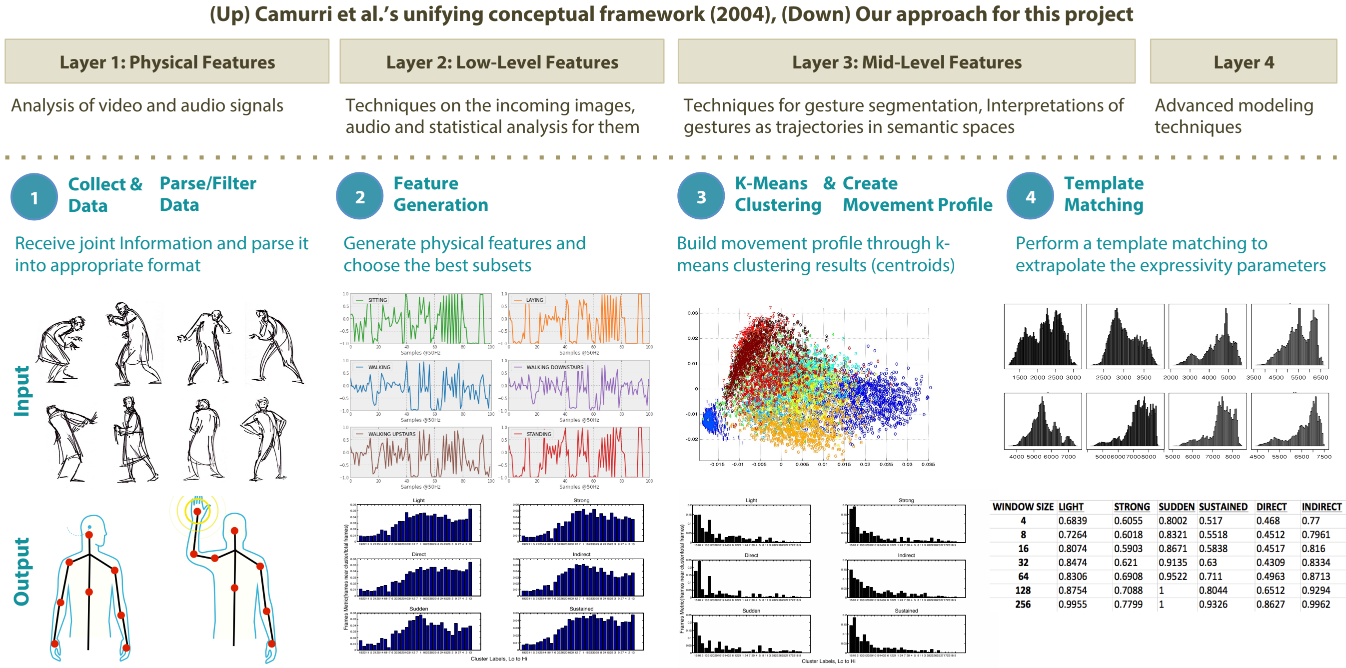

Figure 4. A four-layer framework for designing an interactive, machine-learning driven visualization system.

aligning each phase with Camurri's interpretation. In more detail, they proposed a conceptual, multi-layered framework for automatic expressive gesture analysis. In their work, expressive gestures are described using a set of motion features and the expressive contents from movement using advanced computational models. Inspired by Camurri's conceptual work, we built a four-phase design process: to collect data; generate features; perform k-means clustering to create a movement profile; perform template matching to extrapolate expressivity from the incoming movement data.

Building a Machine Learning Process

Winkler [28] pointed out that each part of the body has unique limitations in terms of direction, weight, range of motion, speed and force. The underlying physics of movement led us to the best representations of movement and its expressive capabilities. in this perspective, we computed basic physical features representing the conducting data by using positional coordinates extracted by the Kinect at 30 frames per second. The conductor's gesture was then denoted as an ordered sequence of these feature vectors, projected into classes by unsupervised clustering algorithm.

We clustered the data and generated characteristic templates for each motion factor from these clusters in order to design a machine learning process that enabled us to describe the characteristics of a conductor's movement. Clustering algorithms create a specified number of groups that maintain an internal coherence, yet are distinguishable from other groups. In other words, features within a cluster should be as similar as possible while features between clusters should be as dissimilar as possible. The initial clustering for our training datasets were based on velocity, the normalized dot product of successive velocity frames, and the combination of these two features. For generating clusters, we used the K-means algorithm, which generates centroids that minimizes the average distance between features, thus grouping features with maximum similarity around these centroids [10]. We determined the cardinality of the clustering (the number of clusters; the 'K' of K-means) using a heuristic method. Empirically, we determined that 32 clusters demonstrated a good tradeoff between distinctive grouping between classes and efficient computation. The latter is important to consider in order to run the final algorithm in real time, capturing conducting gestures during a performance. Once the clusters are computed by the K-means algorithm, we arranged them from low to high velocity and generated a set of histograms. We referred to these ordered histograms as movement profiles. We then used movement profiles to perform a template matching process between training and testing data. The template matching process determined the similarity of test data based on the Euclidean distance between the learned movement profiles and the input movement profile. The input data is then classified as the movement profile that it is 'nearest' to, with distance inversely proportional to the similarity. By comparing the similarity, the computer classified the chunk of movement into one of six different categories: three motion factors of LMA in binary pairs. The template matching process is briefly visualized in Figure 6.

Creating a Generative Visualization Process

In its broadest sense, visualization is the process of making the invisible visible [4]. In more detail, Cox claims that *"The specific characteristics used to visually represent data – such as color, shape, scale, and movement, perform a semiotic function. They don't just inform; they signify."* By reflecting on that insight, we created digital representations communicating the meaning of gestures, invoking personal reflections to intensify the audience experience. From this perspective, we designed a generative visualization process which is illustrated in Figure 7.

From a design perspective, we were inspired by the visual work of tienne-Jules Marey and Gjon Mili, pioneers who attempted to capture sentient beings moving through both time and space.

Marey considered the human body an animate machine, and he dedicated his life to analyze the laws that governed its movements in a visual way. Mili used stroboscopic light to capture the motion of everything from dancers to jugglers in a single exposure. Their works depicted how the movement is being unfolded in space in one single frame. In their artwork, the subject's movement shows delicate relationships from the beginning to the end.

From a more technical perspective, were inspired by interactive artworks that adopt a generative visualization process. The optical flow algorithm (Horn-Shunk) which was used in Sand and Teller's Video Matching [23] to generate optical markers that represented the differences between two scenes acted as a primary inspiration. Shiffman's Reactive [24] was a good example of how to design a particle systems. Forbes [6] showed how interactive fluid simulation and vector visualization techniques can be utilized for media arts projects. Based on these artworks, we created a generative visualization process that enabled the audience to view the expressivity of a conductor's movement.

Our generative approach adopted a 2D fluid simulation controlled in part by video image and expressivity parameters derived from the template matching. To put energy into the 2D fluid simulation process, we first computed motion vectors (velocity and density) from the input video through the optical flow (the first row in Figure 7). Then we delivered the result to the 2D fluid simulation process to manipulate factors of the fluid (the second row). The expressivity parameters, generated from the template matching process, drive the whole visualization. In Figure 5, we can see the influence of high effort (top) and low effort (bottom) on our visualization.

Implementation

Based on this machine learning process and inspirations, we designed an interactive system using two inputs and one output for visualization. One input was from the Microsoft Kinect, tracking the movements of the two arms in order to perform the template matching process. The other was receiving video from an HD camera, capturing a base scene that was transformed into particle images.

With the Kinect input, the template matching process continuously computes histograms and generates three motion factors as expressivity parameters. At the same time, the visualization process generates particle images based on a low-resolution replication of the original video input. The parameters extrapolated from the machine learning process drive the overall visualization system. The software was created using the Model-Viewer-Controller (MVC) design pattern [26]. This concept allowed us to separate modeling, presentation, and action easily. The model managed the behavior and data responding to requests for information about its state and instructions from the controller. The overall system was built in Max6, which was responsible for receiving data, extracting features, and performing machine learning. The visualization is composed in the Quartz Composer using OpenGL and is responsible for generating the particle system. Lastly, the controller, which interprets the conductor's movement, is designed in Visual C++ and informs the model to change appropriately using the Microsoft official Kinect SDK to receive/send the skeletal data via OSC (Open Sound Control) [29].

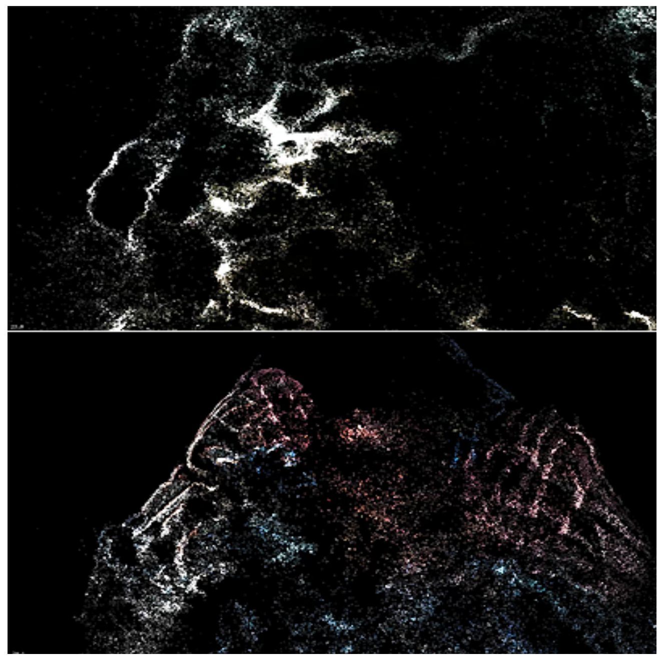

Figure 5. The screen capture of two sample visualization results: (top) the visualization with the efforts factor (strong, indirect, sustained). (bottom) the visualization with the efforts factor (light, direct, sustained).

INSTALLATION AND USER STUDIES

User Study Design

Prior attempts to observe how interactive visualizations or performances are perceived by the audience have been based on qualitative feedback from the audience, case studies, and analysis from the domain expert perspective [25][8] [13]. Starting from these evaluation methods, we evaluated the quality of our system's user experience in terms of engagement, defined by O'Brien et al. [21] as: *"desirable even essential human response to computer-mediated activities."* They suggested that engagement is heavily influenced by the user interface, its associated process flow, the user's context, value system, and incentives. In their work, they described the process of engagement as consisting of four distinct stages: point of engagement, period of sustained engagement, disengagement, and re-engagement. They explained

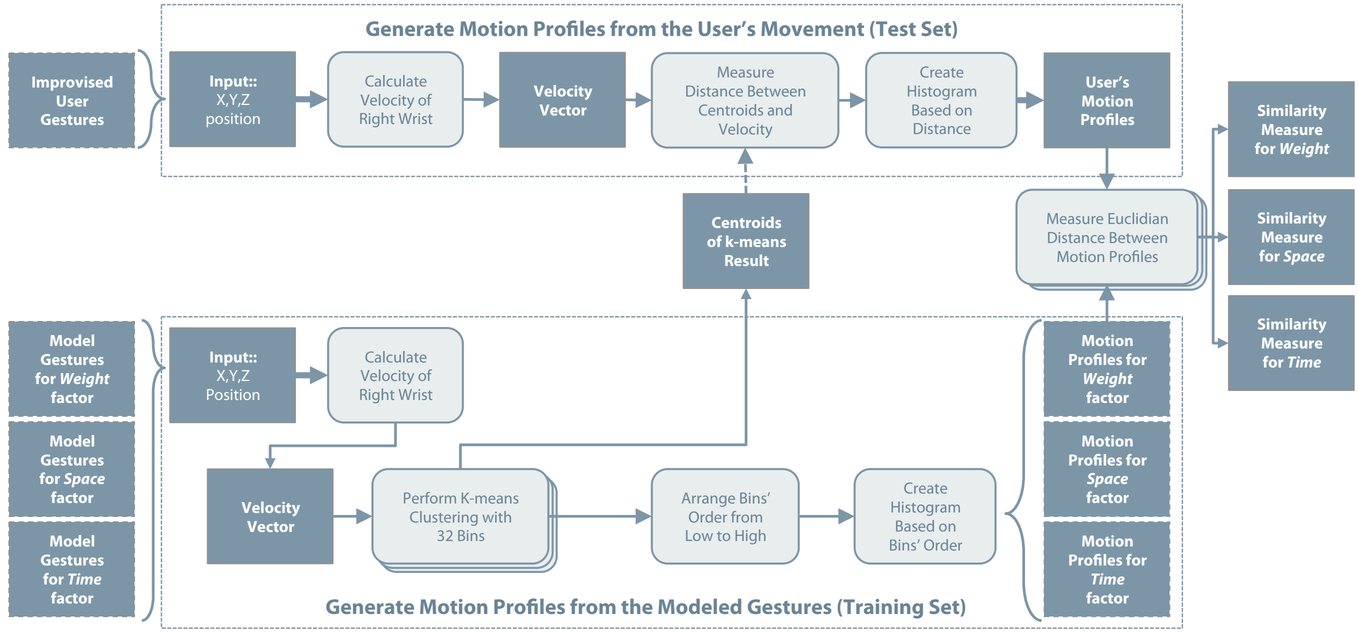

Figure 6. The flow diagram which is depicting how the template machine process is initiated, performed, and extrapolate, the expressivity as parameters to drive the visualization, from the body movement. The rounded rectangle illustrates each process while each filled rectangle represents an outcome.

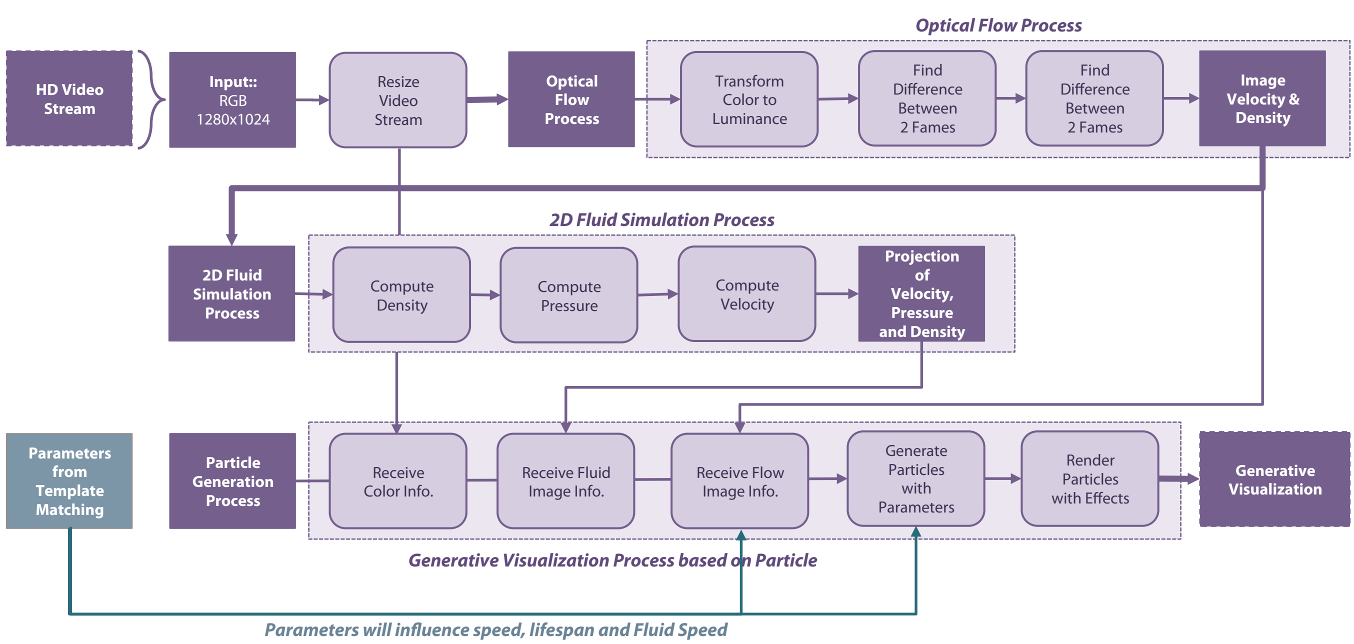

Figure 7. The flow diagram which is depicting how the generative visualization process is processed and rendered with the expressivity parameters extrapolated from the template matching process. The rounded rectangle illustrates each process while each filled rectangle represents an outcome. Note that rectangle and line colored in the dark cyan represents how the expressivity parameters intervene the generative visualization process.

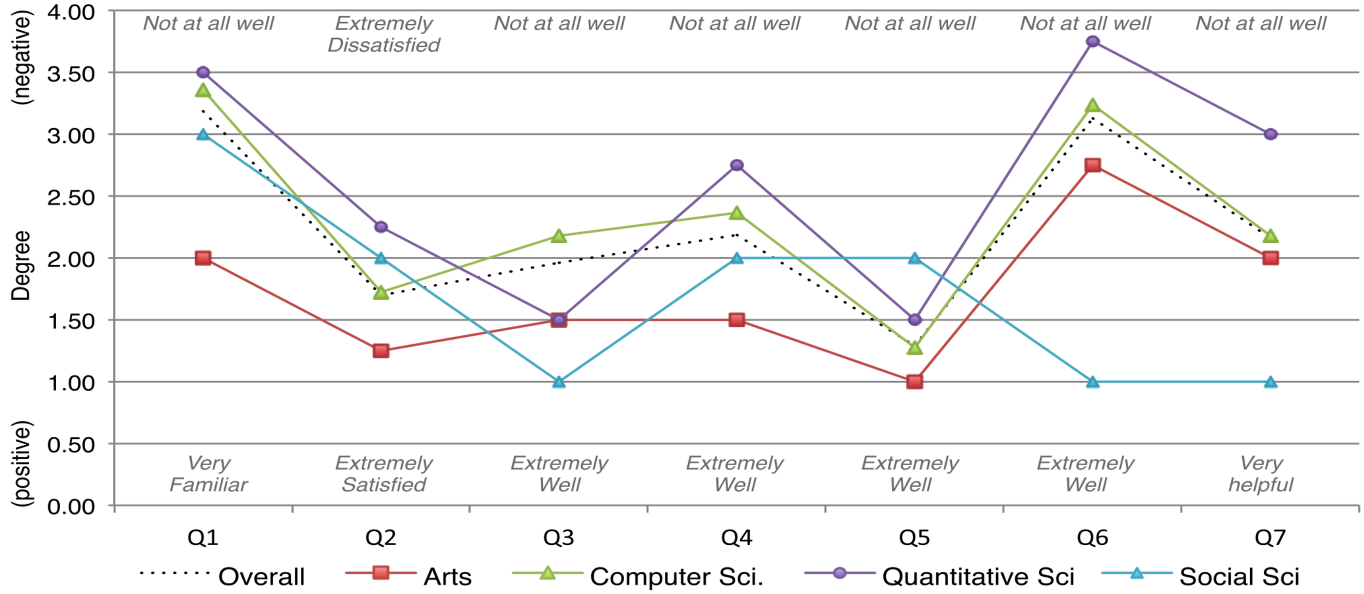

Figure 8. The responses from the different user groups. The x-axis represents the question number from 1 to 5. The y-axis represents the 5-point-scale. Each line shows how different user groups responded to each questions.

that the *point of engagement* described how participants' experiences began. They described the *period of engagement* as the participants' focus on their task and its interface. During this period, the participants' experiences the interactive system and continuously evaluates it based on the novelty of the experience, their level of interest, and their perceptions of challenge, feedback, and user control. The *disengagement* stage explains why participants stops the activity, or which factors in the participants' external environment caused them to stop or to proceed to a *re-engagement*. They suggested several attributes to evaluate these stages such as aesthetics, novelty, interest(curiosity), interactivity, or feedback. We adopted these three stages of engagement and its attributes to set up criteria for evaluating our artwork. With the notion of engagement, we designed a questionnaire of eleven questions using a 5-point scale. Q1, Q2, and Q3 were formed in order to evaluate the user's *point of engagement*, in terms of the category of the aesthetics, novelty, interests, and motivation. Q4, Q5, and Q6 were designed to address which types of engagement factors could arouse the user's engagement most. Q7 was designed to see whether the user could find any sense of utility or experiential goal while he or she was experiencing the artwork.

- Q1. How familiar are you with a conductor's role in an orchestra?
- Q2. What's your first impression to this visualization?
- Q3. Overall, are you satisfied with the visualization?
- Q4. How well, if at all, does the word "INNOVATIVE" describe this visualization?
- Q5. How well, if at all, does the word "AESTHETIC" describe this visualization?
- Q6. How well, if at all, does the word "INFORMATIVE" describe this visualization?
- Q7. How well do you feel the visualization helps you to evoke expressiveness in conducting gestures?

In the meantime, Q8, Q9, and Q10 were designed to address the attributes which might cause *disengagement* for the user. Finally, Q11 was also given to the users in order to have direct feedbacks or responses that were not covered through the survey just in case.

- Q8. If it DOES help you to evoke expressiveness in conducting gesture, what are the top two things you generally consider? Check two boxes among five selections: Abstraction of image; Fluid particle effects/motions; Interactivity (Gesture tracking via Kinect); Sound elements; The way of Installation in the venue.
- Q9. If it DOES NOT help you to evoke expressiveness in conducting gesture, what are the top two things you generally consider? Check two boxes among five selections: Abstraction of image; Complexity in fluid particle effects/motions; Lack of interactivity (need more interactivity); Lack of audio effect (need more audio); The way of Installation in the venue; Lack of explicit explanation
- Q10. If you're improving this visualization, what are the top two things you consider? Check top two things among five selections: Using more concrete/metaphorical objects (snow, fire, wind, rain, rock); Adding more concrete/metaphorical background; Adding sound/audio effects to indicate some aspect of gestures; Adding background music with the installation, Using more color and textures.
- Q11. In your own words, what are the things that you would most like to improve in this visualization?

User Studies at the Exhibition

In November, 2014, we exhibited our artwork at the IEEE VIS 2014 conference which was held in Paris, France to meet diverse audiences across disciplines and observe their response [17]. At the exhibition, we explained the concept and the way of experiencing the artwork briefly before the users interacted with it. After they experienced the artwork, we asked them to answer the questionnaire via Google survey anonymously. No incentives were given. We had more than 200 visitors at the exhibition. Among them, 32 participants voluntarily participated in the survey. We asked them to indicate their background in the beginning of the survey in order to analyze how different subject groups experienced the visualization. The result of the user study is shown in Figure 8.

Based on these questions, we had interesting responses from the users of different backgrounds or research interests. (Q1) the overall user group showed that they were neither familiar nor unfamiliar with the conductor's role (M=3.19, STD=1.37). Some of them were answered that they have seen the conductors on the stage or on the broad casting media. (Q2) Their first impression of the work was quite positive (M=1.70, STD=0.78). They felt extremely satisfied about the quality. (Q3) However, there was a slight disagreement from the computer science group in terms of the satisfaction measure. One of the major reasons that they showed disagreement was that they felt the visualization was too complex.

Figure 9. Users who are interacting with a relatively high level of engagement.

(Q4) Many of people showed excitement and evaluated the work as innovative approach (M=1.96, STD=1.08). However, people from the computer science group said that visualization is frequently used in their domain, so it did not seem innovative to them. (Q5) As the author hoped, all participants agreed that the work has its own aesthetic value,

Figure 10. Users who are interacting with a relatively low level of engagement.

illustrating expressivity of movement; this question showed the lowest standard deviation among all questions (M=1.29, STD=0.46). (Q6) Meanwhile, participants felt the the particle system did not provide enough information. (Q7) In general, the participants indicated that the visualization could help the audience better understanding the expressive contents in conducting gestures (M=2.16, STD=1.05). Since this question had one of the highest standard deviation, we inferred that participants experienced some conflict between the visualization's aesthetic quality and utility. Participants expressed that Fluid particle effects/motions and Interactivity (Gesture tracking via Kinect) are the most impactful factors to evoke expressiveness in conducting gestures.

(Q8) However, participants had difficulty understanding the visualization due to it being very abstract and without audio feedback. They suggested that we should use more concrete/metaphorical objects for visualization and sound/audio effects in order to indicate movement. Some asked for an explicit explanation besides the visualization.

DISCUSSION

It was our hope to design an interactive artwork that enabled the audience to witness the expressive contents in their movement through a generative visualization that used machine learning. Through the installation and the exhibition, we believe we showed how experimental interactive artworks can prompt an audience to actively investigate their movement and ponder its meaning. With the generative visualization, audiences tried to explore and make meaningful trajectories in a creative way. With the concept of user engagement and observations from the exhibition, we extracted three insights: *different levels of engagement, creative exploration with aesthetic interaction,* and *self-reflection behavior.*

As we can see in Figure 9 and 10, users from different disciplines showed different levels of engagement. According to our result, users who have an arts related background were actively engaged. They attempted to make movements that explored the limits of the visualization. These users expressed extreme satisfaction towards the artwork and showed interest in the abstract aspect of the visualization as well as its generative algorithm. Meanwhile, users who expressed a relatively low intensity were mostly focused in computer science/engineering or quantitative science backgrounds. In many cases, they seemed to try 'reading' the artwork, emphasizing its technical aspects. For this reason, we inferred that they marked relatively lower scores in the questions measuring innovativeness and aesthetic value due to the technologies not meeting their technical expectations.

Nevertheless, our audience generally expressed very high levels of satisfaction experiencing the artwork. They also showed interest in the aesthetic value of the visualization regarding creative exploration. The results confirmed that the audience was willing to engage the interactive artwork with high intensity even though the theme, concept, and utility were less clear. By exploring the artwork, the audience became more engaged with the concept, technical background, and the way of visualizing movements, asking diverse questions at the venue. Our observations and user study results support the claims of Petersen et al.'s [22] that the goal of aesthetic interaction can go beyond usability and usefulness, and it will promote curiosity, engagement, and imagination in the exploration of an interactive system. We also found that the O'Brien's model of engagement can be a good measure to observe and describe the quality of the user experience.

One last interesting implication concerns self-reflective behavior. Even though the artwork was motivated by the conductor's gestures, the audience did not use their gestures to mimic conducting. They engaged the artwork with quasi-conducting gestures but progressed to different gestures reflecting their thoughts and feelings in order to see how the visualization would reflect and transform their abstract ideas into a visual dimension. In the study of AnyType, Devendorf and Ryokai [5] suggested that their subjects tended to reflect their thoughts on capturing and creating types with the application. We observed similar behaviors at the exhibition. We tentatively conclude that the lack of control, interactivity (in terms of direct manipulation), and awareness could cause negative affects toward the work; if the user cannot create expected results within the space, they become disengaged.

CONCLUSION

Conductors put tremendous effort into gesturally communicating the imagined-ideal of a composer's intentions to an ensemble, much of which is difficult for the audience to see. In this project, we presented a creative way of visualizing the expressiveness of gestures in the form of an interactive artwork, inspired by conducting gestures. We synthesized two approaches for our work. The first approach used a generative visualization method to enable a robust amplification of expressive contents in gestures by using the optical flow and 2D fluid simulation with the particle generating system. The second approach used template matching, which enabled us to extrapolate the expressivity from the movement and classify them into the 6 categories from the perspective of Laban Movement Analysis (LMA) theory. With the interactive visualization, an audience was invited to perform their own conducting gestures in order to witness how the expressivity in movement can be revealed visually.

As a result, our project allowed us to observe how different audiences engage interactive artwork and how they explore a creative space to generate meaningful visualizations. Based on the observations, we developed three key insights into user interaction concerning different levels of engagement, creative exploration with aesthetic interaction, and self reflection. We could see how the artwork combined with cutting-edge technologies can support individual's creative pursuits. We expect that the impact of visualization will be increased if we could: (i) collect many more movement data from subjects so that we can create more generalized movement profiles; (ii) build a more robust machine learning process that is able to provide better classification results; (iii) create a simple yet impactful mapping that will reduce the complexity in visualization.

ACKNOWLEDGMENTS

We would like to thank the researchers that participated in the MovingStories project in Canada. We also appreciate for the 2014 IEEE VISAP program since it provided a wonderful venue for selected artists to meet the best audience from related fields. We would like to thank Mohammad Amanzadeh and Yishuo Liu for their discussions and critiques.

REFERENCES

1. Ashley, R. The pragmatics of conducting: Analyzing and interpreting conductors' expressive gestures. In *Sixth International Conference on Music Perception and Cognition*, ICMPC'00 (2000).

2. Buck, B., MacRitchie, J., and Bailey, N. J. The interpretive shaping of embodied musical structure in piano performance. *Empirical Musicology Review 8*, 2 (2013), 92–119.

3. Camurri, A., Volpe, G., Poli, G. D., and Leman, M. Communicating expressiveness and affect in multimodal interactive systems. *IEEE Multimedia 12*, 1 (2005), 43–53.

4. Cox, D. J. *Astral Projection: Theories of Metaphor, Philosophies of Science, and the Art of Scientific Visualization*. PhD thesis, University of Plymouth, 2008.

5. Devendorf, L., and Ryokai, K. AnyType: provoking reflection and exploration with aesthetic interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 1041–1050.

6. Forbes, A. G., Höllerer, T., and Legrady, G. Generative fluid profiles for interactive media arts projects. In *Proceedings of the Symposium on Computational Aesthetics*, CA'13, ACM (2013), 37–43.

7. Gambetta, C. L. *Conducting outside the box: Creating a fresh approach to conducting gesture through the principles of Laban Movement Analysis*. D.m.a., University of North Carolina at Greensboro, 2005.

8. James, J., Ingalls, T., Qian, G., Olsen, L., Whiteley, D., Wong, S., and Rikakis, T. Movement-based interactive dance performance. In *Proceedings of the 14th annual ACM international conference on Multimedia*, MM'06 (2006), 470–480.

9. Jensenius, A. R., Wanderley, M. M., God\øy, R. I., and Leman, M. *Musical gestures: Concepts and methods in research*. Routledge, 2010.

10. Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. An efficient k-means clustering algorithm: Analysis and implementation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 24*, 7 (2002), 881–892.

11. Kendon, A. Do gestures communicate? a review. *Research on Language & Social Interaction 27*, 3 (1994), 175–200.

12. Laban, R. v. *The Mastery of Movement on the Stage.* Macdonald and Evans, 1950.

13. Latulipe, C., Carroll, E. A., and Lottridge, D. Evaluating longitudinal projects combining technology with temporal arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, ACM (2011), 1835–1844.

14. Lau, A., and Moere, A. V. Towards a model of information aesthetics in information visualization. In *Information Visualization, 2007. IV'07. 11th International Conference of*, IV'07 (2007), 87–92.

15. Lee, E., Kiel, H., Dedenbach, S., Grüll, I., Karrer, T., Wolf, M., and Borchers, J. iSymphony: An adaptive interactive orchestral conducting system for digital audio and video streams. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, ACM (2006), 259–262.

16. Lee, K. *Towards an Improved Baton Technique: The Application and Modification of Conducting Gestures Drawn from the Methods of Rudolf, Green and Saito for Enhanced Performance of Orchestral Interpretations*. D.m.a dissertation, The University of Arizona, 2008.

17. Lee, K. Visualizing expressiveness in conducting gestures. *Proceedings of the IEEE VIS 2014 Arts Program* (2014).

18. Mathews, M. V. The radio baton and conductor program, or: Pitch, the most important and least expressive part of music. *Computer Music Journal* (1991), 37–46.

19. McNeill, D. *Gesture and thought*. University of Chicago Press, 2008.

20. Nakra, T. A. M. Toward an understanding of musical gesture: Mapping expressive intention with the digital baton. Master's thesis, Massachusetts Institute of Technology, Program in Media Arts & Sciences, 1996.

21. O'Brien, H. L., and Toms, E. G. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science & Technology 59*, 6 (2008), 938–955.

22. Petersen, M.G., I. O. K. P., and Ludvigsen, M. Aesthetic interaction: a pragmatist's aesthetics of interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, DIS'04 (2004).

23. Sand, P., and Teller, S. Video matching. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH '04, ACM (2004), 592–599.

24. Shiffman, D. Reactive. In *ACM SIGGRAPH 2004 Emerging Technologies*, SIGGRAPH '04, 22–.

25. Subyen, P., Maranan, D., Schiphorst, T., Pasquier, P., and Bartram, L. EMVIZ: the poetics of movement quality visualization. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging* (2011), 121–128.

26. Veit, M., and Herrmann, S. Model-view-controller and object teams: A perfect match of paradigms. In *Proceedings of the 2nd international conference on Aspect-oriented software development*, ACM (2003), 140–149.

27. Wakin, D. J. Breaking conductors' down by gesture and body part, 2012.

28. Winkler, T. Making motion musical: Gesture mapping strategies for interactive computer music. In *Proceedings of the International Computer Music Conference*, ICMC'95 (1995), 261–264.

29. Wright, M. Open sound control: an enabling technology for musical networking. *Organised Sound 10*, 3 (2005), 193–200.