GENERAL ARTICLE

Capturing the Body Live

A Framework for Technological Recognition and Extension of Physical Expression in Performance

ELENA JESSOP

ABSTRACT

Performing artists have frequently used technology to sense and extend the body's natural expressivity through live control of multimedia. However, the sophistication, emotional content and variety of expression possible through the original physical channels are often not captured by these technologies and thus cannot be transferred from body to digital media. In this article the author brings together research from expressive performance analysis, machine learning and technological performance extension techniques to define a new framework for recognition and extension of expressive physical performance.

We are all equipped with two extremely expressive, subtle and powerful instruments for performance: the body and the voice. In the context of theatrical, choreographic and musical performances that incorporate digital media, artists have often used technology to sense, analyze and augment the body's natural expressivity. However, the sophistication, emotional content and variety of expression possible through original physical channels is often not captured by the technologies used for analyzing them and thus cannot be transferred from the body to digital media. With the introduction of devices such as Microsoft's Kinect, cheap and reliable movement capture technology is now readily available, but the computer's limited high-level analysis of movement or voice significantly restricts artists' ability to take advantage of this wealth of data to augment a performance work.

What if systems could let someone creating a performance, an installation or a game choose what qualities of movement he or she wanted it to recognize? What if an artist could define for the system, through examples, a "small, nervous movement" or a "sharp, direct movement" and instruct the system on how these qualities should shape music, visuals or other media? How might this system be designed? Three main areas help shape a framework for exploring these questions: (1) performing expression: describing qualities and expressive features of movement and voice; (2) capturing expression: training computers to understand expressive performance information by building on existing machine learning algorithms; (3) extending expression: creating tools that incorporate high-level qualitative descriptions of performance to be used intuitively and creatively in performance-making. In this article I bring these areas together and define a framework for systems that recognize and extend expressive physical performance.

Elena Jessop (researcher). Email: <ejessop@gmail.com>.

See <www.mitpressjournals.org/toc/leon/48/1> for supplemental files associated with this issue.

THEORETICAL FOUNDATIONS FOR EXPRESSION CAPTURE

Performing Expression: Defining Gesture, Quality and Expression

There have been many efforts to study and categorize bodily movements and their relationship to emotional effects or linguistic concepts, from Quintilian's advice on appropriate gestures for Ancient Roman orators [1] to the systematized set of gestures in Delsarte's acting technique, developed from observations of naturalistic movement [2]. Researchers in the psychology of gesture and speech have also created many gesture categorization taxonomies [3]. However, the definition of the term "gesture" is not precisely fixed. Kurtenbach and Hulteen define a gesture broadly as "a movement of the body that conveys information" [4]. Kendon uses "gesture" to describe movement that appears intentionally communicative and deliberately expressive [5] and is identified by sharp onsets and offsets and being a temporary excursion from a position. This implies that certain temporal and spatial characteristics indicate a meaningful gesture. The term "gesture" can also be extended past physical actions into musical and sonic domains [6]. For example, Hatten broadly describes gesture as "any energetic shaping through time that may be interpreted as significant" [7].

For the remainder of this article, I define a *gesture* as a vocal or physical action that conveys expressive information—a gesture is *what* a performer does. Similarly, I define *quality* as the elements of movement or voice that convey individual variation, such as dynamics, timbre and timing. As described

 doi:10.1162/LEON_a_00935

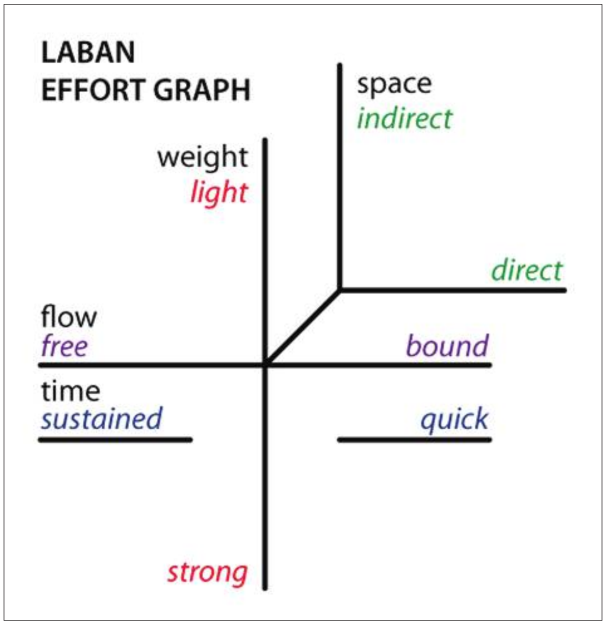

Fig. 1. The four-dimensional Laban Effort Space. (Graph by Raphaël Cottin, reprinted from Wikimedia Commons [40]. Creative Commons Attribution 3.0 Unported License [creativecommons.org/licenses/by/3.0/deed.en]).

by Kendon and Hatten, these temporal (and, where applicable, spatial) elements delineate an expressive gesture—this describes *how* a gesture is performed. I additionally define *expression* as emotional and evocative communication through movement or the voice. It is important to clarify, as Juslin highlighted [8], that expression is not a single-axis phenomenon of which a performance has "more" or "less," but a broad space outlined by multiple quality parameters.

Analysis of Expression in Performance

Contemporary dance performance is a valuable testbed for exploring expression in movement. Dance has a tremendous range of choreographic styles, methods of physical expression and movement vocabularies, ranging from formal ballet technique to pedestrian movements like walking, running and jumping. As various forms of dance focus on expression conveyed by the body, we can ask what qualities of movement have been identified in the study of dance performance and how they relate to the expressivity of the movement. Such descriptions of qualities and their identifying features will be important in creating frameworks for recognition of expression.

Martin divides the features of modern dance into four categories: *space*, *time*, *dynamism* and *metakinesis* [9]. Space includes the volumetric space of the dancer's body, of parts of the body and of bodies in relation to a performance space and to others. A dance can also be described by temporal elements: its patterns in time, speed, duration of parts of movement, rate of succession of movements and "regularity or irregularity of intervals between stresses" [10]. Stress and dynamism relate to the levels and variations of intensity in the movement. Finally, metakinesis refers to the "overtones" of movement that convey intention and emotional experiences.

Similar elements are found in Rudolf Laban's Theory of Effort, a major framework for analyzing movement dynamics including strength, control and timing (Fig. 1). Many researchers have used this framework to examine affective and qualitative movement information [11]. Laban describes the quality of any movement using a continuous "effort space" with the axes of *Weight*, *Time*, *Space* and *Flow* [12]. Weight denotes the amount of energy and intensity in the movement. Time denotes the speed of a movement. Space denotes the way that a movement travels through space around the body. Finally, Flow denotes a movement's smoothness of energy and tension.

Some efforts have also been made to describe vocal quality, although primarily for discussing vocal dysfunction rather than for exploring expression. Scherer has analyzed a variety of vocal features that convey expression, including vocal perturbation (short-term variation), voice quality (timbre), intensity, tempo and range of fundamental frequency over time [13]. As with movement qualities, temporal features and measurements of intensity and energy are particularly important in describing vocal qualities.

Capturing Expression through Pattern Recognition

For computer processes intended to identify information about qualities of movement and voice from a stream of data over time, we can apply techniques from machine learning. *Pattern recognition* is a machine learning technique wherein a computer is trained to discover a function that maps between input examples and output labels, with the goal of generalizing from known examples to appropriately handle new input. If the desired output of the system is discrete categories, such as identification of a specific gesture, the required process is *classification*. If the desired output is real-valued parameters, the appropriate recognition process is *regression*. Regardless of the algorithms used, these processes are similar: input sensor data undergoes *feature extraction* to produce a set of features that may be particularly descriptive. This processed data is used to *train* a model that represents a best guess at the relationships between the input feature vectors and the specified output values. Finally, that model is presented with new inputs and tested.

In the field of human-computer interaction, a significant amount of research has gone into gesture recognition [14]. However, traditional gesture recognition systems still lack aspects necessary for certain expressive recognition tasks. First, as Gillian discusses [15], systems need to be able to handle input and output ambiguity as well as user-specific vocabularies: Users may want to incorporate a variety of different sensors to detect performance input, choose their own vocabularies of recognition for particular pieces and have the results of the pattern recognition algorithms control many different kinds of output. In addition, most systems only recognize whether or not a specific gesture has been performed; we also want computers to be able to recognize time-varying expressive qualities of movement. No system yet exists that allows a user to train it to recognize specific continuous qualities rather than classifying movement into emotional categories or particular gestures.

Extending Expression

There is a long history of performance extension through technology, particularly in the field of digital musical instruments. These instruments often use the performer's movement as the primary control mechanism. Tod Machover's paradigm of Hyperinstruments provides virtuoso musicians with additional levels of expressivity through digitally enhanced traditional musical instruments [16]. Other digital musical instruments have novel interaction models, such as Waisvisz's "The Hands" or Julie and Mark Bokowiec's "Bodycoder," which have been used to shape and trigger sound in live performances through movements of the performer's arms and hands [17].

In the field of dance, there have been many additional examples of technological performance extension. In 1965, Merce Cunningham's *Variations V* incorporated photoelectric sensors and antennae to mark the positions of dancers; the data gathered then controlled electronic musical devices [18]. The dance company Troika Ranch has made use of dancers' movement detected by laser beams crisscrossing the stage, impact sensors on the floor or computer vision tracking points on a dancer's body to shape visual and sonic elements [19]. Other systems, including Sparacino's DanceSpace [20] and the Danish Institute of Electronic Music's Digital Dance Interface [21], have also been used for the real-time generation of music to accompany dancers onstage.

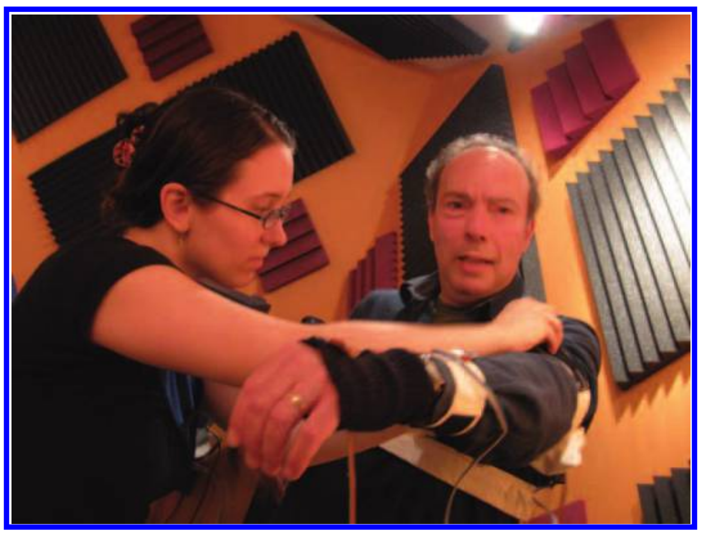

Fig. 2. Adjusting wearable expressive gesture sensors on baritone James Maddalena in rehearsal for Tod Machover's opera *Death and the Powers.* (© Tod Machover)

The underlying problem in a majority of these examples is one of *mapping*. How is the physical performance input (from sensors, microphones, video cameras, etc.) related (mapped) to the control parameters of output media? Many software tools exist to facilitate mapping tasks in performance, including Max/MSP, EyesWeb, Isadora and Field [22]. Some artists have made efforts to incorporate machine learning and concepts of expression into these tools [23]. For example, EyesWeb recognizes some pre-programmed qualitative features, and the SARC EyesWeb catalog incorporates simple gesture recognition [24]. However, there is still ample room to develop mapping tools that automatically simplify the process of working with higher-level expressive information. I believe that creating meaningful mappings becomes easier as the inputs to that mapping become more intuitive to the composer or choreographer. This is where pattern recognition can play a major role: by allowing the computer to learn how to transform sensor data into expressively meaningful inputs to the mapping process.

TOWARD NEW SYSTEMS FOR EXPRESSION RECOGNITION

A variety of my previous research projects have provided useful testbeds for the design of a framework for expressive movement capture and analysis. These include the Vocal Augmentation and Manipulation Prosthesis (VAMP), a gesture-based wearable controller for live vocal performance (Color Plate B). This controller allows a performer to capture and harmonize notes that he or she sings, extending his or her voice through free gesture [25]. Before developing the technology for VAMP, I designed a vocabulary of gestures based on choral conducting and ways to connect these gestures to sound manipulation: capturing a note with pinched fingers, extending the arm to control volume, raising the hand to add a harmony note and shaking the hand to add vibrato and overtones. This separation of the desired gestural vocabulary and control mappings from the technology needed to implement them resulted in a system that has intuitive and expressive interactions between movement and music.

Building on this importance of high-level movement analysis in mapping, I began developing systems to abstract raw sensor data into more meaningful movement descriptions. I developed the Gestural Media Framework to recognize gestures from continuous streams of movement data using Hidden Markov Models (HMMs). This system also incorporated concepts from a modified Laban Effort Space to characterize qualities of a performer's movement [26]. As part of the evaluation of the Gestural Media Framework, I choreographed a set of performance pieces that used this system throughout the performance and rehearsal process to map dancers' movements to manipulations of projected visualizations, sound and theatrical lighting. In this process, I discovered that the system's ability to recognize particular gestures and use them to trigger events was not particularly interesting; more important were the subtle details of how the movement was performed.

I took some of these lessons from my experience with the Gestural Media Framework into my work developing performance capture technologies for the Disembodied

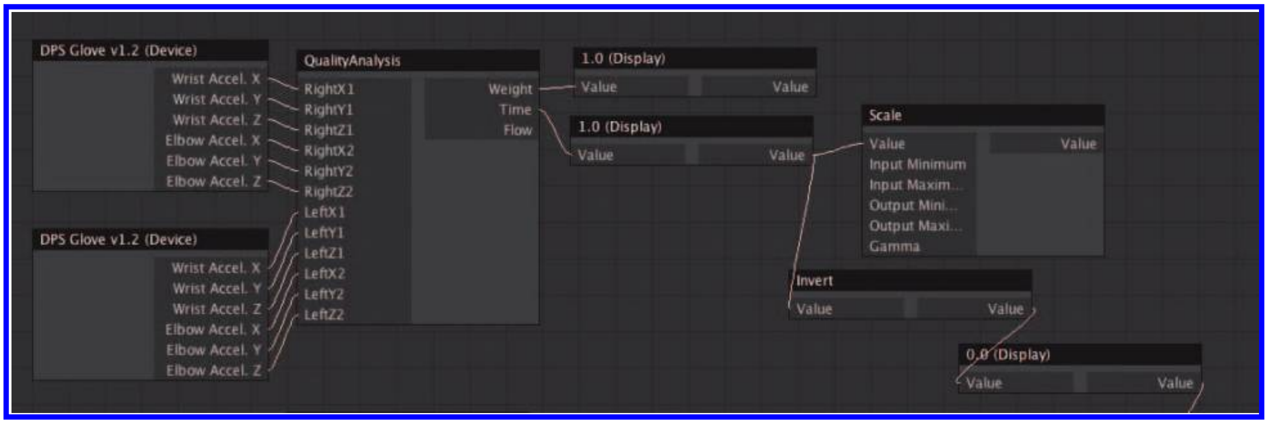

Fig. 3. Screenshot of the Disembodied Performance mapping system transforming raw sensor data into expressive movement quality parameters. (© Elena Jessop. System design by Peter Torpey.)

Performance System, developed with Peter Torpey [27] (Fig. 2). This system, created for Tod Machover's opera *Death and the Powers*, addressed how to map an expressive performance from the human body onto other modalities including non-anthropomorphic visuals and sound. In the opera, a wealthy man seeks to extend his life by uploading his consciousness into a computer system integrated throughout his house. Theatrically, this means that the character is first portrayed by a live opera singer and then embodied by the theatrical set and the entire performance space [28].

The performer leaves the stage after he is "uploaded" in the first scene but continues giving a live performance offstage. Wearable sensors and microphones measure the parameters of his movements and voice, which are then used to control the theatrical environment (Fig. 3). To do this, Torpey and I examined expressive qualities of the performer's voice, including intensity, frequency and timbre parameters. We also mapped movement input to qualities in a modified Laban Effort Space of Weight, Time and Flow, abstracting the movement away from specific sensor values into a higher-level quality framework. This process did not rely on a specific movement vocabulary but rather on the performer's natural expressive performance. The resulting qualities of movement and voice could then be mapped to control a variety of output. While the levels of abstraction in this system proved useful for creating mappings that closely connected performer and digital media, the system was still tied to a particular sensor set and performance scenario. A more generalized system is needed that can quickly learn the relationships between different sensor inputs and desired qualitative spaces.

A FRAMEWORK FOR EXTENDING EXPRESSION

My proposed framework for extension of physical expression has four layers of data abstraction (Fig. 4). The first level is the raw sensor data from inputs such as wearable sensors and video cameras. The second level consists of extracted features of the sensor data (particularly temporal features) that are related to expressive content. Pattern recognition techniques can then be used to associate the data at this level with the third layer of data abstraction: high-level parametric spaces of expression, such as Laban's Effort Space. Finally, those high-level spaces can be manually mapped to the fourth level of data: parameters for the control of output media. I agree with Camurri et al. that the majority of expressivity communicated through movement is conveyed by spatiotemporal characteristics rather than by syntactic meanings of particular gestures [29]. Thus, the goal of this framework is to analyze and recognize qualities of movement rather than to perform gesture recognition. This framework differs from existing expressive gesture analysis methodologies (such as the one proposed by Camurri and the Genova group [30]) in two main ways. First, my framework prioritizes the final representation of gestural expression as trajectories through continuous expressive spaces defined by semantically meaningful parameters rather than as discrete semantic categories. Second, this framework prioritizes creative thought at a higher level of abstraction than either raw sensor data or expressive features. The creator's mapping process between gestural inputs and output controls is designed to take place at the level of the continuous expressive space, while the machine learning algorithms and feature analysis tools built into the system handle the association between sensor data and that continuous space. The goal of generalization in this system is intrapersonal rather than interpersonal: The framework should learn different relationships between layers for individual artists, not force one general relationship between sensor data, features, expressive spaces and output control parameters.

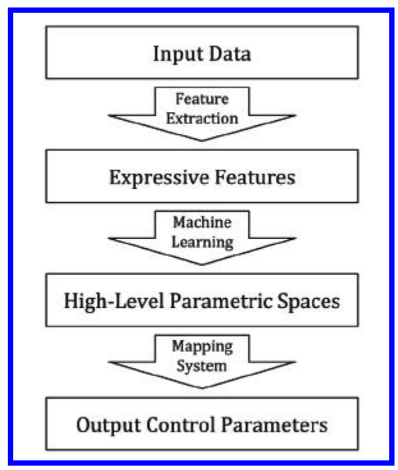

Fig. 4. The proposed Expression Recognition Framework. (© Elena Jessop)

Extraction of Expressive Features

In this process, the first step is to transform raw sensor data into features of interest that are likely to convey salient information. The selection of particular features is especially important in gesture and movement quality recognition, as some kinds of features may be more likely to contain expressive information. Meaningful features for both vocal and physical performances include the amount of energy in the input, the tempo and the changes in tempo, and the amount of variation in different input parameters. Temporal features are the focus of Manfred Clynes's research on "essentic forms": cross-cultural "spatio-temporal curves" of emotions determined by many users performing a simple button-press action while focusing on a particular emotional state [31]. Here, dynamic and time-dependent shapes, rather than static positions, describe an emotional space.

Algorithms for Expression Recognition

Below I examine a variety of pattern recognition algorithms for learning relationships between these selected features of interest and the desired expressive spaces. Such algorithms may also be useful at the feature recognition stage to relate particular sensor data streams to expressive features. It is important to consider several properties when selecting an algorithm to recognize expressive input. First, how well does the algorithm generalize from its training data? Can it handle the complexities of human movement and voice, given representative examples? Second, how quickly can the algorithm be trained, and how many examples does it need for training? Third, can it run its recognition process in real time? Fourth, can it handle temporal variability in its input, either by using algorithms that include a history of samples or by preprocessing input samples to obtain a time-normalized input?

The first pattern recognition algorithm I examine is the Hidden Markov Model (HMM), a probabilistic algorithm based on state machines and popularly used for recognition of sequential patterns in time-varying domains such as speech [32] and gesture [33]. A particular advantage of HMMs is that they can easily handle temporal inputs and inputs of varying lengths due to their internal incorporation of sequences and memory [34]. However, HMMs are generally only used for classification. In order to have HMMs learn the relationships between movement and a continuous space, such as Laban's Effort Space, we would need to adapt these models to produce continuous output. HMMs also require a large number of training examples and take a long time to train [35].

Another pattern recognition algorithm frequently used in gesture recognition is the Support Vector Machine (SVM) [36]. Some benefits of using SVMs for expression recognition include their strength at generalization and ability to handle small training sets. However, the algorithm generally performs classification rather than regression and becomes more complex if more than two output classes are necessary. Another drawback to using SVMs is that the input must be a fixed-length feature vector, so movement data streams of various lengths must be normalized to a fixed period of time.

A third algorithm that may be particularly beneficial for recognizing expressive parameters is the neural network. Neural networks (NNs), inspired by models of brain activity, consist of interlinked layers of nodes (or neurons), each of which activates if the sum of its inputs passes a given threshold. While NNs may take a long time to train, they quickly process new input data for testing and need few training examples [37]. Neural networks have been used frequently in the recognition of musical parameters [38], where their ability to perform regression is a major benefit. Another particularly strong point of NNs is their ability to produce meaningful outputs for inputs not included in the training set [39]. However, NNs present the interesting difficulty that the structure of the network does not reveal how it has learned to classify inputs or whether it has learned the "right thing."

High-Level Expressive Spaces

Regardless of the choice of pattern recognition algorithm, it is necessary to define the expressive space that the algorithm is to learn. A major goal of this framework is the transformation of sensor data into meaningful high-level parametric structures constructed from multiple expressive axes, each with a normalized range. The sensor input, ranges of parameters and desired expressive spaces will vary between different pieces and different performance-makers. Thus, an ideal system should let a user demonstrate examples of movements or vocal gestures that form particular points in an expressive space and then automatically model the relationships between input data parameters and that continuous expressive space. A user may also want to define high-level axes in addition to suggested parametric structures.

In determining the qualitative expressive axes to use for analysis of specific performance input, it is important to consider the relevance of temporal, spatial, dynamic and emotional parameters in describing that input. As an example, let us say that the input is sensor data from a solo ballet dancer, performing choreography that is inspired by the physical properties of fluids in motion. As this is a non-narrative piece without gestures intending to represent specific emotions, a set of expressive axes to describe this movement might include parameters like *pace* (slow to quick), *fluidity* (legato to staccato), *scale* (small to large), *continuity* (continuous to disjointed), *intensity* (gentle to intense) and *complexity* (simple to complex). These axes can be combined to form a high-level expressive space. Positions in and trajectories through that space can then be mapped to control parameters for multimedia. In all cases, the outputs of the pattern recognition algorithms should be continuous values, not classification; the goal is not to label a movement "staccato" or "angry" but to find a position on a set of continuous expressive axes that are meaningful for controlling output media. To extend the possibilities of expression recognition systems even further, these systems should analyze temporal behavior not only to determine the performer's current point in an expressive space but also to recognize particular trajectories or features of trajectories through expressive spaces.

Similarly, it is important to remember that expression comes both from the definition of a physical or vocal standard (the overall tempo information, the amount of energy, the specific gestural vocabulary, etc.) and through the ways in which a piece then varies moment to moment from its own established "norms." For example, a particular dance performance might have a quick, intense rhythm throughout; however, the moment-to-moment expressivity in the dance comes from how the movement deviates from that standard—through slower or faster subdivisions of tempo, the amount of fluidity versus rigidity in a motion or the amount of energy in a particular movement. A system to appropriately process expressive information must therefore recognize not only the baseline expressive parameters of a piece but also when and how the work varies from that baseline. Systems constructed according to this framework should also include these various aspects of expression in the parameters they include for defining a multi-dimensional, continuous expressive space.

CONCLUSIONS AND EVALUATION METRICS

This paper has outlined a framework for the development of performance expression recognition systems that have the potential to open new levels of artistry in live performance, interactive installations and even video games. Such systems would be able to recognize expressive feature spaces and could be taught, via machine learning techniques, relationships between those feature spaces and meaningful high-level quality spaces. Several metrics can be used to evaluate the success of these expression recognition systems. The first of these is technical, examining the accuracy of the system in the recognition process and the ease of training that system. Once the system has been trained, how well does it generalize to new examples? How many training examples are necessary to obtain sufficiently accurate predictions? Second, a system should be evaluated as an artistic tool, based on how well it represents sophisticated movement and gives sophisticated control. The more meaningful the expressive parameter spaces recognized by the system are to a performance-maker, the more meaningful the connections between the movement quality information and the resulting output media can be. Finally, systems should be evaluated on the extent to which they serve the interactive goals of performance, by examining whether the "liveness" of the system is observed by audiences. Is it clear that there is live control, affected by moment-to-moment human performances? Are the system's effects different every night? That, after all, is the fundamental goal: a system that can be a true counterpoint to and an extension of a live performance, recognizing and responding to subtleties of timing and expression that make every performance fundamentally unrepeatable and unique.

References and Notes

Unedited references as provided by the author.

1 Adam Kendon, *Gesture: Visible Action as Utterance* (Cambridge, UK: Cambridge University Press, 2004).

2 L'Abbé Delaumosne et al., *Delsarte System of Oratory* (New York: E.S. Werner, 1893).

3 David Efron, *Gesture, Race and Culture* (The Hague: Mouton, 1972); Kendon [1]; David McNeill, *Hand and Mind* (Chicago, IL: University of Chicago Press, 1992).

4 Gord Kurtenbach and Eric Hulteen, "Gestures in Human-Computer Communication," in Brenda Laurel, ed., *The Art and Science of Interface Design* (Reading, MA: Addison-Wesley Publishing Co., 1990) pp. 309–317.

5 Kendon [1].

6 Examples of research on music and gesture can be found in: Rolf Inge Godøy and Marc Leman, eds., *Musical Gestures: Sound, Movement, and Meaning* (New York, NY: Routledge, 2010); Anthony Gritten and Elaine King, eds., *Music and Gesture* (Hampshire, U.K.: Ashgate, 2006); and Marcelo M. Wanderley and Marc Battier, eds., *Trends in Gestural Control of Music* [CD-ROM] (Paris: IRCAM, 2000).

7 Robert Hatten, "A Theory of Musical Gesture and Its Application to Beethoven and Schubert," in Anthony Gritten and Elaine King, eds., *Music and Gesture* (Hampshire, U.K.: Ashgate, 2006), pp. 1–23.

8 Patrik N. Juslin, "Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance," *Psychology of Music*, Vol. 31(3) (2003), pp. 273–302.

9 John Martin, *The Modern Dance* (Princeton, NJ: Princeton Book Company, 1933).

10 Martin [9] p. 55.

11 These include Petra Fagerberg et al., "Designing Gestures for Affective Input: An Analysis of Shape, Effort and Valence," in *Proceedings of the 2nd International Conference on Mobile Ubiquitous Multimedia* (Norrköping, Sweden: 2003); Gualtiero Volpe, *Computational Models of Expressive Gesture in Multimedia Systems*, Ph.D. dissertation (University of Genova, 2003); Liwei Zhao and Norman Badler, *Synthesis and Acquisition of Laban Movement Analysis Qualitative Parameters for Communicative Gestures*, Ph.D. Dissertation (University of Pennsylvania, 2001).

12 Rudolf Laban, *Mastery of Movement*, 4th edition (Devon, U.K.: Northcote House, 1980).

13 Klaus Scherer, "Vocal Affect Expression: A Review and a Model for Future Research," *Psychological Bulletin*, Vol. 99 (1986) pp. 143–165.

14 Many gestures and gesture recognition technologies used in HCI are summarized in Dan Saffer, *Designing Gestural Interfaces: Touchscreens and Interactive Devices* (Sebastopol, CA: O'Reilly Media, Inc., 2009).

15 Nicholas Gillian, *Gesture Recognition for Musician Computer Interaction*, Ph.D. dissertation (Queen's University Belfast, 2011).

16 Tod Machover, "Hyperinstruments: A Progress Report 1987–1991," technical report (Cambridge, MA: MIT Media Laboratory, 1992).

17 Michel Waisvisz, "The Hands, a Set of Remote Midi-Controllers," in *Proceedings of the International Computer Music Conference* (Vancouver, Canada: 1985) pp. 313–318; Mark Bokowiec and Julie Bokowiec, "The Suicided Voice: The Mediated Voice," in *Proceedings of the International Computer Music Conference* (Barcelona: 2005) pp. 459–462.

18 Joseph Mazo, *Prime Movers: The Makers of Modern Dance in America*, 2nd ed. (Hightstown, NJ: Princeton Book Company, 1977).

19 Dawn Stoppiello and Marc Coniglio, "Fleshmotor," in Judy Malloy, ed. *Women, Art, and Technology* (Cambridge, MA: MIT Press, 2003) pp. 440–452.

20 Flavia Sparacino et al., "Augmented Performance in Dance and Theater," in *Proceedings of the International Dance and Technology Conference* (Tempe, AZ: Arizona State University, 1999).

21 Wayne Siegel and Jens Jacobsen, "The Challenges of Interactive Dance: An Overview and Case Study," *Computer Music Journal*, Vol. 22 (1998) pp. 29–43.

22 See <http://cycling74.com/>, <www.infomus.org/eyesweb_eng.php>, <http://troikatronix.com/isadora/about/>, and <http://openendedgroup.com/field/>.

23 Such as Javier Jaimovich and R. Benjamin Knapp, "Pattern Recognition of Emotional States During Musical Performance from Physiological Signals," in *Proceedings of the International Computer Music Conference* (2009), pp. 461–464.

24 Gillian [15].

25 Elena Jessop, "The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-Based Gestural Controller for Vocal Performance," in *Proceedings of New Interfaces for Musical Expression* (Pittsburgh, PA: 2009).

26 Elena Jessop, *A Gestural Media Framework: Tools for Expressive Gesture Recognition and Mapping in Rehearsal and Performance*, S.M. Thesis (Cambridge, MA: Massachusetts Institute of Technology, 2010).

27 Peter Torpey, *Disembodied Performance: Abstraction of Representation in Live Theater*, S.M. Thesis (Cambridge, MA: Massachusetts Institute of Technology, 2009).

28 See: <http://powers.media.mit.edu>; Tod Machover, *Death and the Powers* (New York: Boosey and Hawkes, 2010); Elena Jessop et al., "Music and Technology in *Death and the Powers*," in *Proceedings of NIME* (Oslo: 2011).

29 Antonio Camurri et al., "EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems," *Computer Music Journal*, Vol. 24, No. 1 (2000) pp. 57–69.

30 Camurri et al. [29]; Antonio Camurri et al., "A multi-layered conceptual framework for expressive gesture applications," in *Proceedings of the International MOSART Workshop* (Barcelona: 2001).

31 Manfred Clynes, *Sentics: The Touch of the Emotions* (New York: Doubleday and Co., 1977).

32 Lawrence Rabiner, "Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," in *Proceedings of the IEEE, 77* (1989), pp. 257–286.

33 Examples include Gillian [15] and Tracy Westeyn et al., "Georgia Tech Gesture Toolkit: Supporting Experiments in Gesture Recognition," in *Proceedings of ICMI* (2003).

34 Rabiner [32].

35 Gillian [15].

36 Shigeo Abe, *Support Vector Machines for Pattern Classification* (New York: Springer-Verlag, 2005).

37 Christopher Bishop, *Pattern Recognition and Machine Learning* (Science and Business Media, Springer, 2006).

38 Examples include: Rebecca Fiebrink, *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*, Ph.D. dissertation (Princeton University, 2011); Michael Lee, Adrian Freed and David Wessel, "Neural Networks for Simultaneous Classification and Parameter Estimation in Musical Instrument Control," in *Proceedings of Adaptive Learning Systems 1706* (Orlando, FL: 1992) pp. 244–255.

39 Fiebrink [38].

40 Raphaël Cottin, File:Laban-effort-graph.jpg, accessible at <http://commons.wikimedia.org/wiki/File:Laban-effort-graph.jpg>.

ELENA JESSOP*'s work explores the intersection of expressive movement, music and computer systems in the context of enhancing and expanding live performance. She is a choreographer, artist and technology designer who received her doctorate at the MIT Media Lab in the Opera of the Future research group.*

Manuscript received 23 August 2012.

COLOR PLATE B: **CAPTURING THE BODY LIVE**

Elena Jessop, the VAMP glove, shown here capturing a note from the singer's voice with a gesture. (© Elena Jessop. Photo: Peter Torpey.)

CALL FOR PAPERS

Pioneers and Pathbreakers

In anticipation of Leonardo's 50th anniversary, we are seeking papers dealing with the history of developments in the arts, sciences and technology. The aim of the project is to establish reliable, selected, online documentation about 20th-century artists, scholars and institution builders whose works and ideas are considered seminal in the development of technological art. Since its founding, *Leonardo* has accompanied and championed the work of the pioneers who were just beginning to use computers and other emerging technologies for artistic purposes.

We are interested in topics including the following:

- Memoirs written by pioneer artists working in new media (holography, computer and electronic arts, telecommunication arts, interactive arts, new materials, space arts, bio art, etc.). Texts should cover an extended body of work. Preference will be given to artists describing early work carried out prior to 1980.
- Memoirs of engineers and developers who collaborated with artists or whose engineering or computer science work in the 1960s, 1970s and/or 1980s proved to be important for developing the new art forms based on new and emerging technologies.
- Memoirs by curators who organized art-and-technology exhibitions in the 1960s, 1970s and/or 1980s.
- Memoirs by pioneering collectors who were early supporters of technological artists.
- Memoirs by pioneering institution founders of organizations, university programs or centers pre-1984.

Texts must be in English. Length may be up to 2,500 words, 6 illustrations.

Readers of Leonardo *are asked to encourage their colleagues to submit such memoirs, which will be invaluable primary documents for historians and scholars in the future.*

Interested authors should submit manuscript proposals or completed manuscripts to the Leonardo Editorial Office. Email to: <leonardomanuscripts@gmail.com>.