DISEMBODIED PERFORMANCE

Abstraction of Representation in Live Theater and Expressive Storytelling

Peter Torpey
Opera of the Future, MIT Media Lab

Disembodied Performance presents a new way to think about expressive presence. At its core, it allows for the abstraction of a person's emotive gesture away from the human body, allowing for the transmission of a performance across distances and into completely new forms of representation.

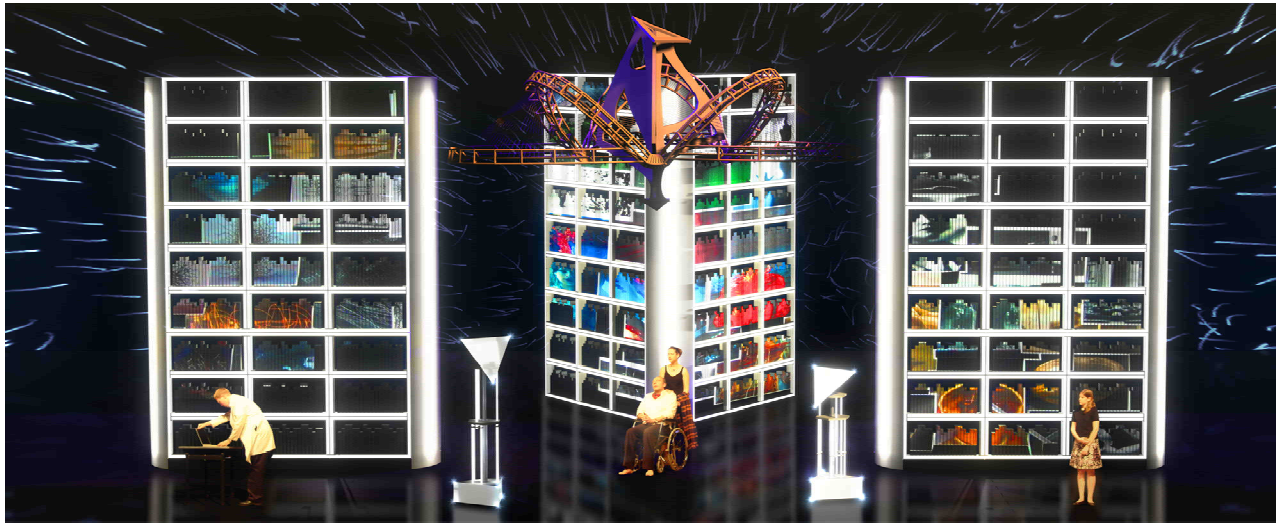
Adapting ideas from affective psychology, cognitive science, and the theatrical tradition, I propose a framework for thinking about the translation of stage presence. In order to distill the essence of a character, we recover performance parameters in real-time from gesture and physiological sensors, as well as voice and vision systems. This system allows an offstage actor to express emotion and interact with others onstage. The Disembodied Performance approach takes a new direction in augmented performance by employing a non-representational abstraction of a human presence that fully translates a character into an environment.

This paper briefly summarizes some of my research on this topic. For additional discussion of this work, please visit the links at the end of this document.

Death and the Powers

Early in Tod Machover's new opera *Death and the Powers*, the main character, Simon Powers, is subsumed into a technological environment of his own creation. The theatrical set comes alive in the form of robotic, visual, and sonic elements that allow the actor to extend his range and expression across the stage in unique and dynamic ways. The environment must compellingly assume the behavior and expression of the absent Simon.

Disembodied Performance was first conceived as the mechanism to realize this theatrical effect. Rather than present representational imagery through flat video projections or play back pre-recorded animations of set pieces and lighting, I sought a very real and ephemeral link between the behavior of the environment and the character that, in the world of the story, is embodied by it. An actor's performance onstage conveys a great deal of information about the character's internal state. The sense of presence alone carries weight. How can the same be accomplished through the theatrical set when the actor is not onstage?

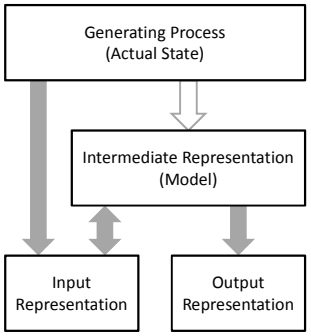


Conceptual rendering of the set for Death and the Powers

An Underlying Theory

Our world is highly ordered and the human brain and perceptual systems have evolved to quickly identify and interpret the patterns and relationships that exist around us. It has been asserted that perception is not

simply a process of recognition, identifying known features that we experience through our senses, but actually a process of explanation. That is to say our minds arrive at an understanding of what we see or hear or touch by making informed guesses about what would cause the qualities of the sensations. When we recognize features that are clearly not random or accidental, we determine their meaning by identifying the process that would generate this configuration of features. A vivid and communicative representation presents regularities in its features that readily impart information about the process that generated the representation.



Reified inference

In a typical mapping process, input parameters directly drive output parameters via some transformation. This type of mapping is especially common in augmented performance, where a sensor value is proportionally linked to a visual or sonic property. However, the individual features of a representation do not themselves contain meaning. It is the explanation of what caused those properties to exist that is of value.

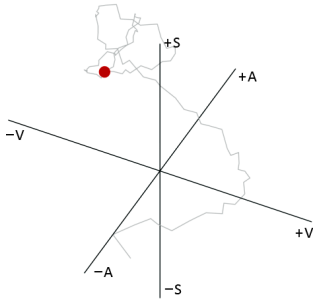
I introduce the method of *reified inference* as a structured approach to mapping representations across media and modalities. Let us consider the input to the system as a representation itself that is created by the generating process in which we are interested. Using inference, we can create a model from the input representation that approximates the state of the generating process. This model can then be used to generate new output representations, as if it were the generating process. As a measure of validity of the model to capture the intended salient information, the model can be used to recreate the informational impact of the input representation in its original modality.

Character Modeling

The work of Manfred Clynes and others has demonstrated that such features as I described in the previous section exist in our expressive behaviors. We move in certain ways, with certain qualities, in response to our emotional state and what we are either consciously or unconsciously trying to convey to those around us. This should come as no surprise to us. We very clearly understand gestures and body language when communicating with each other and this means of broadcasting one's inner state is at the very heart of acting. Thus, to capture a performance, we need to look at the gestural cues and signs that the performer makes. The output representation must then preserve the meaning of these gestures by reflecting their natural timing and quality, regardless of the form.

What then is the currency of this communication from the actor to audience? It is, in part, the affective state at any given moment of the character the actor is portraying. Applying reified inference, if we treat the actor's performance as a representation of the character's inner emotional state, then we can infer a model of that state from the performance. From that model, we can generate other novel representations in any form or modality that preserve the meaning of the expression (the state of the model).

In the current implementation of Disembodied Performance, I use a three-dimensional model of affect, based on the traditional circumplex model. This *affect space* has orthogonal bases of *arousal* (degree of alertness or heightened response to a stimulus), *valence* (whether a stimulus is favorable or unfavorable), and *stance* (level of engagement with the stimulus). Inferring the state of the model, the current locus of the character's psyche in affect space, from the input representation of the actor's performance is a problem of dimensionality reduction. The input data, such as from a multitude of sensors, is reduced to a point in the three-dimensional space that represents the character's current emotional state. Over time, the modulation of this state carves out a trajectory through affect space. This inferred state is then used to generate an output representation.



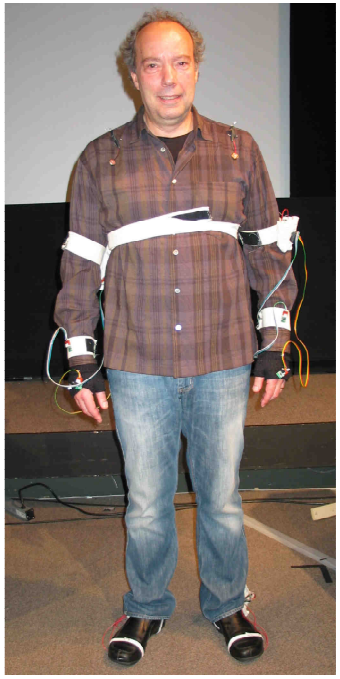
Trajectory through affect space

Performance System

The Disembodied Performance System is a hardware and software implementation of a show control system for *Death and the Powers* that facilitates the abstraction of performance from the offstage actor and allows for an expressive onstage representation through the set, robotics, projected imagery, sound, and stage lighting.

Wearable sensors, along with computer vision and vocal analysis, provide the input representation of the actor's offstage performance. Gesture sensors on each of the opera singer's arms, pressure sensors on the sole

of each shoe, and a respiration sensor all wirelessly transmit information to a central computer. These units capture behaviors over which the actor has control while performing in order to communicate his performance. They also measure gestures and posture that would be visible to the audience, as if the actor was performing onstage.



Opera singer James Maddalena wearing Disembodied Performance sensors

Software on the central system processes the data to infer the current affective state of the character. A distributed system of software renderers receives this model state information in order to compute graphical elements or robotic motion on stage. This Disembodied Performance System has been flexibly engineered to coexist alongside other designed elements and interface with existing theatrical control systems and technologies.

Additional Applications

The work discussed so far is readily generalizable to a broad range of performance applications. Though the system was designed for stage acting, with modifications to the model and to the methods of data capture, the system can be applied to other sorts of performances. For instance, a musician could play his instrument while generating a visualization or other representation of his experience of performing, rather than a visualization of the sound or music he produces.

Disembodied Performance has positive implications for the future of theater and storytelling media. Beyond remote and augmented performances, Disembodied Performance can readily bring affective and expressive representations to many computing domains. I believe that Disembodied Performance can be particularly useful in computer-mediated interpersonal communications, where the output representations can be viscerally understood. Informal online correspondence in instant messaging, e-mails, and social networks has demonstrated a need for affective augmentation of textual communication, as evidenced by the introduction of modern-day emoticons in 1982 and their widespread adoption and continued use. Disembodied Performance can fill such a niche by providing a simultaneous channel of affective communication to accompany text. Extending from the augmentation of text communications, Disembodied Performance techniques can contribute to telepresence. Since one of the fundamental goals of Disembodied Performance is to convey a sense of presence, it can supply alternatives to or enhance video communications in several ways.

Presence in virtual worlds and games can also benefit from the ideas I've laid out. Presently, avatars take several forms, but are generally humanoid representations in appearance and often gesture. In many cases, a user has the opportunity to configure their virtual self as a means of expression. However, the movement of the avatars is generally functional and related to navigating space or occasionally the state of activity, such as walking, standing and waiting, gesturing while talking (typically as the user is typing text to "speak" to another individual in the virtual space). Rarely, if ever, can these avatars convey the affective content of the communicated words or the emotional state of the user. In some cases, a few basic commands are provided to change the behavior or appearance of the avatar, but the result is much like the use of emoticons. Applying Disembodied Performance approaches, the avatar can emote continuously, informed by the affective model. The intermediate representation can be efficiently transmitted in distributed environments. The modes of sensing into the model would be different from those for performing in an opera. In the case of console games, for example, as input data, the system can sense properties of how the game controller is being held. At a higher level, how the avatar is being moved in space (erratically, slowly, quickly toward an object or other avatar) can contribute additional information to the model. Affective avatars could greatly enhance the immersive qualities of these virtual environments and interactions. The emotive humanoid form is only one possible output representation. Having expressive dynamic non-anthropomorphic avatars greatly broadens the palette of representations of an individual's identity.

Future Directions

An area of subsequent research to which I intend to contribute builds on the process of representation mapping using reified inference. Looking beyond modeling affect for performance contexts, I believe this approach can shed light on methods for translating representations in an informed way,

preserving important characteristics while changing the presentation of information for artistic, scientific, and communicative applications. The ultimate goal of continued research would be to extend the concept of reified inference in order to formulate a general theory and practice of representation mapping. The purpose of representation mapping can be to illustrate things that we normally have difficulty expressing or to observe higher-level structures than what we usually can see; to find the emergent patterns in a system or to simply feel them.

Disembodied Performance can play an innovative role in asynchronous interpersonal communications and memory-sharing. Technology is increasingly common in greeting cards and keepsakes to play music, record voices or, in the case of digital picture frames, to display changing images. Now, imagine an emotion-capturing keepsake that presented the recipient with a completely personalized and genuine experience of the giver's unique affective signature. A thank you card could capture some expression of just how grateful the sender was when she signed it. A holiday ornament could replay the awe and enthusiasm of your grandson on Christmas morning year after year. The simple model of affective state I have used can be readily stored in such devices and represented as sound and image.

With the ability to freely map between arbitrary representations, new types of human-computer interactions become possible. Another general implementation could be a computer user interface that is a proper superset of current interface technologies. Such an interface can present any information in one or more modalities, fully customizable to the user's needs, aesthetic preferences, and accessibility requirements. In this environment, a user can see, feel, or hear information as well as manipulate it in multiple ways. I believe that an interface capable of providing users and developers with this degree of configurability to choose the substance and manner of representation and change it at whim could revolutionize how we use computers.

Let us conclude with a brief look at how representation mapping can influence storytelling applications. An ancillary project to *Death and the Powers* called Personal Opera seeks to create a platform for individuals to tell their own story—to share their legacy—in a fun and intuitive way. Leveraging the power of music as a central narrative thread, Personal Opera will provide an easy-to-use interface to assemble images, video, sound, and text in a compelling manner, much like the innovative design of our research group's software Hyperscore facilitates music composition by anyone using a language of gesture and color to represent high-level musical properties. Personal Opera will leverage similar abstractions of shape and contour to control parameters of the story being told, the arcs of tension and resolution, and how media will be combined. Personal Opera may appear in a number of forms that allow for different types of creation and storytelling experiences. It will have an architecture well-suited to a computer application, web application, or even a mobile application where it can draw on a user's personal media collection and additional sources of information. Personal Opera can also appear as a large-scale interactive installation piece, using a gestural vocabulary for manipulation of media.

For more information, please visit these web pages:

Disembodied Performance
`http://web.media.mit.edu/~patorpey/projects/powers/system.php`

Death and the Powers
`http://powers.media.mit.edu`

Opera of the Future
`http://opera.media.mit.edu`

Peter Torpey may be reached at
`patorpey@media.mit.edu`

Conceptual rendering of Personal Opera