Viewpoints AI: Procedurally Representing and Reasoning about Gestures

Mikhail Jacob
Georgia Institute of Technology
mikhail.jacob@gatech.edu

Alexander Zook, Brian Magerko
Georgia Institute of Technology
a.zook@gatech.edu, magerko@gatech.edu

ABSTRACT
Viewpoints is a contemporary theatrical composition technique for understanding the expressive powers of gesture used to formally describe a dance performance or theatrical movement (Bogart 2005). We describe a computational system that integrates a gesture-based interface (Kinect), theatrical aesthetics framework (Viewpoints), AI reasoning architecture (Soar), and visualized embodiment of the AI participant (Processing) to explore novel forms of meaningful co-creative theatrical interaction in an interactive installation piece. Providing this ability to reason about a gesture's meaning enables game designers to explore novel ways for players to communicate with intelligent game agents. Toward this end, we describe our prototype for live interaction with a projected virtual agent in an interactive installation piece.

Keywords
gesture, theatre, Viewpoints, artificial intelligence, procedural aesthetics, computational creativity

INTRODUCTION
Expressive artificial intelligence (EAI) strives to explore the affordances of AI architectures for the human creation of meaning (Mateas 2001). Expressive AI approaches tend to either involve the human interactor as a piece of a larger artistic system (e.g. *Terminal Time* (Mateas et al. 1999) or *DARCI* (Norton et al. 2011)) or with the AI agent in a dominant creative role compared to the interactor (e.g. the use of drama managers in interactive narrative systems (Roberts and Isbell 2008)). Rarely do systems approach human / AI creative practice with both in equal roles. This omission is largely due to difficulties in semantic understanding by computers; making meaning with a machine as an equal partner requires clear communication between both entities.

Our current work, called *Viewpoints AI* represents a co-creative human/AI experience where neither entity has privileged knowledge nor a privileged position in the process of creation. *Viewpoints AI* focuses on interpreting a continuous space of gesture meaning making described by a small set of procedures for extracting aesthetics derived from the

Proceedings of DiGRA 2013: DeFragging Game Studies.

Viewpoints theatrical framework. Viewpoints is a theatre composition technique that provides an aesthetic framework for understanding human motion and gesture while training actors (Landau and Bogart 2005). We chose to use Viewpoints for this breadth of applications that maps well onto creating expressive artificially intelligent co-participants for a live theatre performance.

In this paper we present *Viewpoints AI*'s integration of a gesture-based interface (Microsoft Kinect), an aesthetic theatrical framework (Viewpoints), an AI reasoning architecture (Soar), and a visualized embodiment of the AI's decision making (Processing) to explore novel forms of meaningful co-creative theatrical interaction. We describe a computational representation of a subset of the Viewpoints system developed with a trained Viewpoints actor and a computational system for procedurally reasoning on Viewpoints input from a human to produce expressive visual output. We envision this system enabling new kinds of video game interactions using gestures as expressive (and often ambiguous) components of play to be interpreted and replied to, rather than simply a novel input modality. Providing the ability to reason about gestures' meaning enables videogames to explore new avenues for players to communicate with game agents in novel ways. Toward this end we describe our working prototype for live on-stage interaction with a projected AI agent co-performer (called *VAI*).

VIEWPOINTS

The Viewpoints technique provides an aesthetic framework for understanding motion, training theatre actors in expressive action, and theatrical composition. Overlie formulated the six Viewpoints of *Time*, *Space*, *Shape*, *Emotion*, *Movement*, and *Story* to structure dance improvisation (2006). Landau and Bogart have since expanded and refined these to a subset of nine Physical Viewpoints and Vocal Viewpoints (2005). We employ the Physical Viewpoints—relating to gesture—rather than Vocal Viewpoints—relating to sound—due to the challenges of processing vocal signals and natural language.

Physical Viewpoints are the two dimensions of space and time. The Viewpoints of time are:

(1) tempo: how fast a movement is.

(2) duration: how long a movement lasts.

(3) kinesthetic response: the timing of a movement in response to external events.

(4) repetition: repeating something internal—within one's own body—or external—from outside one's body (e.g. another actor's motion).

Viewpoints of space are:

(1) shape: the outline of the body in terms of lines and/or curve.

(2) gesture: a movement of part of the body including beginning, middle, and end.

(3) architecture: the physical environment of the performance.

(4) spatial relationship: the distances among things onstage, particularly between individual bodies, individual bodies to a group, and bodies to the architecture.

(5) topography: patterns of movement through space (e.g. repeating a movement motif or treating areas of space as prone to more or less rapid movement).

Viewpoints serve as an external source of inspiration for actions when training actors with this aesthetic framework. Expressive actions must attend and respond to these aspects from outside the individual actor, placing a premium on awareness of one's environment. *Viewpoints AI* adopts this view through an AI system that uses these procedural aesthetics of expressive motion to guide an AI agent to interact with a human actor on stage.

The *Viewpoints AI* project as a whole explores co-creativity through procedurality, differentiating it from prior EAI systems emphasizing authorial control and instantial knowledge. Complementing interpretive and authorial affordances, co-participants require expressive affordances that allow them to convey meaning to an AI system. Complementing instantial assets, co-creative systems require procedural assets that enable meaning to be inferred from the range of actions expressive affordances provide. *Viewpoints AI* thus contributes to the goals of EAI by examining the space of computational expression in terms of procedurality and co-creation.

At the time of writing, we have implemented a subset of the above Viewpoints in *Viewpoints AI* targeting those central to Viewpoints practice and amenable to computational operationalization. Below we describe our implementation of these elements after first discussing related approaches to expressive AI systems.

RELATED WORK

We use a theatre aesthetic framework for understanding movement in a system for procedural co-creation of meaning with an AI agent. Developing such an expressive artificially intelligent system can enable new forms of game play and a better understanding of the affordances of AI techniques for computational expression. Below we situate our work in the space of computational systems leveraging narrative theories and discuss the relationship of our work to other approach to co-creation of narratives with computational systems. Unlike previous approaches, we have developed a system for mixed reality interaction based on gestures, focusing on procedural generation of proto-narratives. Proto-narratives are an abstract space of temporally and causally linked events not grounded in particular semantic expressions or actions. Viewpoints supports proto-narratives through expressive gestures that build a system for sharing meaningful actions (excitement, anger, connectedness, etc.) co-created by a group of actors, without linking these actions to particular characters or dramatic content.

Stanislavsky's System

Several theatrical frameworks—including Stanislavskian, Laban, and Improv theatre—have been applied to the creation of interactive narratives or the interpretation of gesture. Below we briefly discuss these approaches to contextualize our work and motivate our choice of aesthetic framework.

Stanislavsky developed a system to train actors to draw from emotions and create scenes with meaningful motivation. El-Nasr (2007) developed the Mirage system using these dramatic principles in an interactive narrative system. Mirage highlighted the use of "user character arc" to create engaging interactive narratives. User character arcs allowed for a gradual progression in the user's character over the narrative to give a sense of development and growth.

Morgenstern (2008) developed a commonsense logical formulation of Stanislavskian scene analysis. Rather than create interactive stories, this work emphasized the use of formal logic and planning to analyze existing scenes for coherence and appropriate motivation. Our use of the Viewpoint system shares the intent of understanding an existing scene and guiding it toward meaningful interactions. Unlike Stanislavskian approaches we focus on the interpretation of gesture without verbal narrative content. We note that Viewpoints stands as a reaction to the American misappropriation of Stanislavskian. American Stanislavskian theatre sought to induce actors to follow a particular emotional state (which Stanislavsky rejected in work after visiting America). Viewpoints instead emphasizes the actions taken on stage being grounded in the current setting (Landau and Bogart 2005).

Laban

Laban Movement Analysis is used to understand human movement both for analysis and composition. Schiphorst, et al. (1990) describe the COMPOSE tool supporting the creative process of Laban dance composition. Authors use the tool to create key frames of movement, with a rule-based AI system that complements the user by using constraint propagation to fill in missing frames in the movement. In contrast, *Viewpoints AI* puts human and AI on equal grounding, with neither performing a low-level task for the other.

Subyen et al. (2011) describe the EMVIZ artistic visualization based on the Basic-Efforts from Laban Movement Analysis. EMVIZ uses a machine-learning technique to render perceived motions from a wearable computing system into a vector representation of 8 Basic-Efforts. These efforts are transformed through a 2D line representation into an L-system model. The L-system generates a set of lines that are then colored drawing from Kandinsky's theory of colors. *Viewpoints AI* also employs both processing of human motion and a procedural visualization, although viewing these through the lens of Viewpoints theatre. Our choice of Viewpoints was motivated by local expertise of a performer and the adoption of Viewpoints in theatrical staging with actors. *Viewpoints AI* extends approaches like EMVIZ through an additional layer of AI reasoning on movement to select a meaningful response, before then rendering that response. *Viewpoints AI* is intended to support meaningful performance, rather than artistic visualization.

Improv

Several efforts have drawn from improv theatre to develop EAI systems. Magerko et al. (2011) developed an interactive AI system for the *Party Quirks* improv game. Human participants query an AI avatar about its character traits and the AI takes actions to suggest possible characters. Crowdsourcing information on prototypical representations of characters informs the AI system of appropriate ways to represent a given character. In contrast, *Viewpoints AI* avoids heavy use of such pre-created instantial knowledge and focuses on procedural expression via gestural interaction.

O'Neill et al. (2011) describe a knowledge-based framework for humans and AI agents to collaborate to create a scene introduction. Based on the *Three Line Scene* improv game, individuals take turns presenting actions in order to create an initial scene (including characters and activity) based on mimed motions. Piplica et al. (2012) describe a gesture interpretation system for improv theatre. Human motions are perceived through a Kinect and interpreted into basic components that are then further composed into more complex gestures with meaning. In contrast to these approaches, *Viewpoints AI* examines proto-narratives rather than improvised dramatic stories and aims for a full-length narrative, rather than only scene establishment. Viewpoints provides a means for *Viewpoints AI* to procedurally understand aesthetics, rather than required a pre-authored set of motions tied to a particular narrative.

Interactive Narrative

Interactive narrative (IN) systems use AI techniques for controlling narrative experiences while balancing between authorial control and user agency. Conceptually, most IN systems take the role of a real-time director guiding an interactive story, but do not necessarily draw from any single theatre aesthetic framework. AI drama managers employed in IN systems control the story world in response to player actions in order to convey an author's intended narrative to the player. Effectively fulfilling this role requires carefully balancing authorial intent against the freedom of the player to experience the narrative, while also preventing the AI systems from becoming the focus of user interaction.

Mateas (1999) reviews several IN systems focusing on the particular problem of balancing characters against story. Roberts (2008) reviews drama managers in terms of their computational techniques, finding statistical machine learning and AI planning approaches to be predominant recent approaches. Riedl and Bulitko (2013) expand on the authorial intent vs. user agency perspective to note three dimensions of IN systems: authorial intent, character autonomy, and player modeling. Authorial intent captures the extent to which an AI system is allowed free rein in changing a narrative or generating new narrative. Character autonomy describes how much characters are free from the control of a drama manager, from fully autonomous to completely controlled. Player modeling entails learning about and responding to user differences, typically with the goal of enforcing authorial intent. *Viewpoints AI* provides weak authorial intent through a generative AI system; currently employs a single character enacting the proto-narrative in cooperation with the user; and models and responds to user differences through tracking a shared history of patterns of actions and responses in terms of Viewpoints representations.

Horswill (2009) describes a procedural approach to abstract story creation in interactive narrative domains using light-weight procedural animation in the TWIG system. However, TWIG requires the developer of the interactive narrative to pre-author the necessary gestures, control-loops and other related instantial assets to accommodate the narrative domain. The *Viewpoints AI* system bypasses this limitation by basing all of *VAI*'s response gestures on the current gestural inputs (either the most recent gesture or gestures it has experienced in the past) in order to make the gestural interaction truly open-ended. These human-originated gestures are then transformed according to a library of functional transforms that are domain independent.

THE VIEWPOINTS AI SYSTEM

The *Viewpoints AI* system enables procedural interpretation of movement and gesture aesthetics for the improvisation of a proto-narrative. *Viewpoints AI* does this aesthetic interpretation using an agent-based model of the perception of a human's gestures, reasoning on the gestures' meaning, and action in response to those gestures. Three modules are responsible for this process of *perception*, *reasoning*, and *action*. *Viewpoints AI perceives* human gestures and renders them into aesthetically meaningful components using the Viewpoints framework. *Reasoning* decides on a gesture for *VAI* to perform by connecting the perceived gesture to previous gestures with multiple possible modes of response to choose among. Finally, *VAI acts* by procedurally visualizing its gesture back to the human interactor. Figure 1 presents the key components of *Viewpoints AI* discussed below.

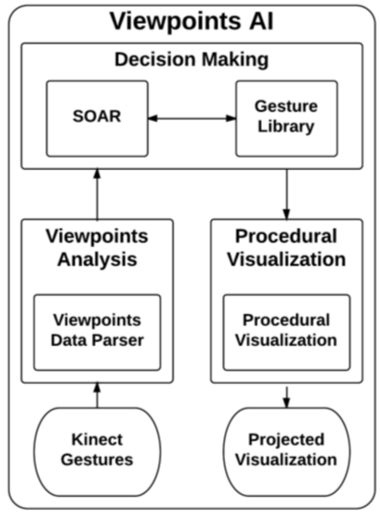

Figure 1: System architecture for the *Viewpoints AI* system.

Perception

The *perception module* (Viewpoints Analysis from Figure 1) of *Viewpoints AI* reads in data from a Microsoft Kinect and derives an aesthetically meaningful description of a perceived gesture using the Viewpoints of *time* (such as tempo or duration) and *space* (such as shape or spatial relationship). The raw gesture data from the Kinect consists of absolute positions of joints of the body, in relation to the Kinect sensor, over time. A gesture is represented by this positional data over a period of time delimited by a sequence of movement between two periods of stillness. *Viewpoints AI* has been implemented as a turn-based interactive experience to focus on the problems of improvisation and procedural aesthetics interpretation, rather than allow spontaneous turn-taking.

The perception module derives symbolic viewpoints information from the raw, positional Kinect data (see Figure 1 for a summary). Viewpoints information is represented as a set of Viewpoints predicates that are derived from Physical Viewpoints elements. Predicates are symbols describing a discrete state of some part of the world. For example, HEIGHT(TALL) describes an interactor (human or AI) standing at full height, with related predicates HEIGHT(SHORT) or HEIGHT(MEDIUM). The *Viewpoints AI* system's *reasoning module* (Decision Making from Figure 1) uses this symbolic data to operate its decision-making algorithms, converting continuous Kinect data into symbolic Viewpoints information.

The *Viewpoints AI* system's *perception module* provides a procedural rendering of gesture aesthetics based on Viewpoints techniques. This rendering enables game applications to reason about gestures in an aesthetic framework derived for orchestrating movement and reactions among individuals, such as in a dance game where interaction could move beyond learning pre-authored dance steps to open-ended expressive movements with a responsive AI partner. Below we discuss further elements of *Viewpoints AI* used in a human / AI interactive experience.

Viewpoints Predicate	**Viewpoint**	**Viewpoint**
Frame & Average Tempo	Tempo	The instantaneous (and average over a gesture) speed of movement.
Frame & Average Energy	Tempo	The instantaneous (and average over a gesture) amount of movement.
Frame & Maximum Smoothness	Tempo	The instantaneous (and maximum over a gesture) smoothness or flow of movement.
Total Duration	Duration	Duration of gesture or duration of movement between two periods of stillness (two poses).
Kinesthetic Response	Kinesthetic Response	Not yet implemented as current turn-taking approach prevents natural timing of kinesthetic response.
Repetition	Repetition	Limited implementation as output to *action module* for repeating output gesture.
Frame Height	Shape	Instantaneous height of the actor from knee to head.
Frame Size	Shape	Instantaneous size of gesture bounded by arms and body.
Limb Curve	Shape	Whether limb is bent or straight.
Body Symmetric	Shape	Whether body is symmetric or not.
Arm Position & Height	Gesture	Position and height of arm.
Hands Together	Gesture	Whether hands are together or not.
Average Limb Stillness	Gesture	Whether limb is still or not.
Average Limb Transversal, Longitudinal & Vertical Movement	Gesture	Whether or not limb is moving transversally, longitudinally or vertically.
Birth / Life / Death of Gesture	Gesture	Not yet implemented due to complexity of sensing and learning required for real time analysis.
Architecture	Architecture	Not yet implemented.
Frame Distance To Center	Spatial Relationship	Instantaneous Euclidean distance from center of stage.
Frame Distance To Other Actor	Spatial Relationship	Not yet implemented because current version of *Viewpoints AI* does not feed *VAI*'s position back to *perception module* from *action module.*
Frame & Average Facing	Spatial Relationship	The instantaneous (and most common over the gesture) stage orientation of the person during the performance.
Frame Quadrants	Topography	The instantaneous top down position (forming a path over time) of a performer in a stage quadrant system.

Table 1: Formalization of Viewpoints of *space* and *time* into Viewpoints predicates

The *Viewpoints AI* system's perceptual system can augment existing game designs that use Kinect data by providing an additional layer of aesthetic information relating human users to one another and/or the Kinect itself. Designers may use this information for many ends, ranging from guiding AI agent responses in games to exploring a space of game designs for best choreographing character actions in Viewpoints space (mirroring the uses of Viewpoints in theater composition). For example, in the first case Viewpoints can provide additional aesthetics information to determine if a free-form gesture by the user is 'aggressive' (say a *fast* tempo gesture, rapidly *reducing* distance between agents) towards a non-player character resulting in a believable reaction of self-defense or fear depending on the character and context. In the latter case while designing the behaviors or response gestures of non-player characters, exploring the entire space of Viewpoints for those responses can help improve the believability of the response.

Reasoning

Viewpoints AI sends the derived Viewpoints predicates to the *reasoning module* (Decision Making from Figure 1) in order to select an appropriate improvisational response to that gesture. The *reasoning module* consists of a rule-based system called Soar (that decides how to respond appropriately to a user's gesture) and a gesture library (used for storing and matching raw gestures from the Kinect for use by the *action module*). This decision-making uses background knowledge such as rules for selection context and aesthetic appropriateness derived from the expertise of an expert local theatre practitioner. The *reasoning module* uses this background knowledge in combination with architectural capabilities for experiential learning, memory and planning.

Soar

The Soar rule-based system (Laird 2012) is an agent-based model of cognition relying on procedural knowledge in the form of rules to operate on other knowledge stored in long term memory and working memory in order to execute goal-oriented behavior. Soar consists of internal states (analogous to mental states) and operators (analogous to actions) that modify those states in order to achieve some goal. Soar uses a constant decision cycle in order to decide what operator to execute. This decision cycle consists of reading in input, proposing possible operations to execute, elaborating knowledge in working memory based on the new inputs, selecting an operation to execute, and the final execution of that operator.

Soar is used in the *Viewpoints AI* system as an architecture for selecting and applying *response modes* (different modes of responding to a gesture), *improvisational strategies* (rules for choosing a *response mode*) and *response gestures* (the output response itself). Soar uses rules based on aesthetics and improvisation developed in conjunction with an expert Viewpoints practitioner. Soar's *response modes* correspond to different ways to respond to perceived gestures; *response gestures* are the gestures shown to the human interactor.

Soar starts by randomly choosing a period of history to consider when choosing how to respond. Soar then chooses a particular mode of responding to the user's gesture. At the time of writing, Soar may respond by: doing nothing, mimicking the user's gesture, transforming the user's gesture and then performing it, repeating a gesture it has learned during its lifetime of experience, or executing certain kinds of interaction patterns.

Functional Transform	Intended Effect
Reflect Limb Motion	Vertically, Transversally or Longitudinally reflect motion of a limb.
Switch Limb Motions	Switch the movements of two limbs.
Copy Limb Motion	Copy the motion of a limb to one or more other limbs.
Repeat Gesture	Repeat the *response gesture* multiple times.
Viewpoints Transformations	Transform any Viewpoints Predicate to another allowed value. Eg. Transform Tempo changes the tempo of the *response gesture*.

Table 2: Library of functional transforms used to modify user gesture into *response gesture*

Doing nothing and repeating a human's gesture are self-explanatory *response modes*. Functionally transforming a human's gesture occurs through selecting from a library of domain-independent functional transforms that the agent is aware of. These transforms operate on the Viewpoints predicates calculated by the *perception module*. The functional transforms change these Viewpoints predicates and modify the human participant's gesture (see Table 2 for examples of functional transforms).

The *response mode* of repeating a gesture from past experience allows the agent to extend its repertoire of movements beyond repetition and modification of the user's current input gesture. Currently, the system chooses a gesture from the past at random. Future work will bias this selection based on the human interactor's last gesture according to a measure of perceived similarity between the duration, energy, tempo or other Viewpoints predicates of the input and candidate gestures. The incorporation of these seemingly new 'riffs' into the performance permits the agent to take on a more equal role in the creation of the performance by providing the human a creative offer to build off of.

In the final *response mode*, *Viewpoints AI* has a limited capacity to analyze and utilize patterns of interaction between the user and the agent in order to decide how to respond. These interactional patterns can be of different types. The pattern could either be a pattern of gestures done in the past, a pattern of functional transformations done to gestures in the past, or a complex mixture of the two. An example of following a pattern of functional transforms would be to carry out the theatrical *rule of threes*, which states that comedic actions are generally done the same way twice, but are transformed or modified in some interesting way the third time. Soar operationalizes this by doing a response gesture, then exaggerating that gesture (using a set of functional transformations to make that gesture more prominent) and finally transforming it. This establishes an expectation for the pattern of the interaction from the audience, which is reinforced by the second exaggeration of that response and finally broken by the third transformation so as to create novelty and interest. *Viewpoints AI* currently uses patterns from its knowledge base of theatrical techniques, rather than learning them through interaction—acting as an encoding of Viewpoints technique rather than an open-ended system for any (potentially meaningless) interaction.

A future extension of this work is to extend the current limited pattern following into a more full-featured pattern learning and analysis system. Three types of patterns were formalized for extending this system. Gestural patterns are sequences of literal gestures that *VAI* and the human have executed in sequence, corresponding to rote learning of interaction sequences. Transformative patterns are sequences of functional transforms done by *VAI* learnt through analyzing the favorable functional transforms executed by *VAI* based on user feedback. Complex patterns are combinations of gestural and transformative patterns.

The *gesture library* (see Decision Making in Figure 1) is used as a store for raw gestures (from the Kinect) that the *action module* requires further down the pipeline for visualizing *VAI*'s responses. In addition it can do fast matching against the existing gestures stored in it to detect a historically repeated gesture from the human interactor. This gesture matching is important for interactional pattern usage. The raw gesture (from Kinect) sent to the *action module* from the *gesture library* and the corresponding Viewpoints predicates together determine the final expressive response that *VAI* performs.

The *Viewpoints AI* system's *reasoning module* provides a general framework for organizing responses to perceived gestures. This framework enables new game designs with agents and environments that respond to the underlying aesthetics and ambiguous meaning of human motion. Putting human and AI on equal footing provides a new perspective on user-generated content as co-created performances with AI systems. Performative games can explore alternative motivations for playing games, such as user's desires to have expressive motion. Proceduralizing a space of motion makes these perception and response techniques amenable to broad audiences and reusable across many designs. Proceduralizing the reasoning process enables designers to consider how to construct characteristic styles of interaction at the level of proto-narrative meaning, rather than being limited to discrete rules that are tethered to highly specified game states. The *Viewpoints AI* system's reasoning thus enables a new form of game mechanics built around the aesthetics of motion, pursuing ends similar to *Prom Week*'s (McCoy et al. 2011) proceduralization of social interaction knowledge to enable social game mechanics.

Action

The *action module* (Procedural Visualization from Figure 1) converts a selected response gesture from a set of Viewpoints and gestural predicates into a procedurally generated visualization. The visualization maps the predicates into visualization operations and functional transforms to perform on the positional gesture data sensed from the Kinect. A transformed output gesture is finally rendered as a human silhouette composed of a swarm of fireflies (see Figure 2).

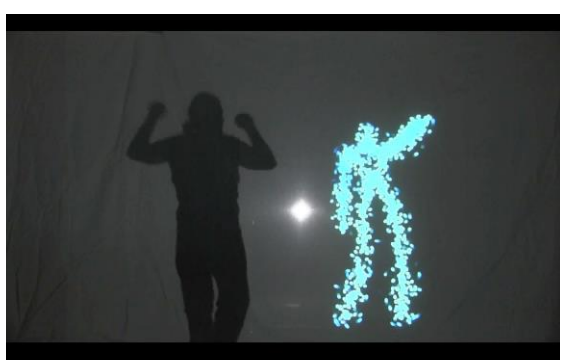

Figure 2: The human and *VAI* interacting in a movement-based expressive piece based on theatrical Viewpoints technique.

Viewpoints Predicate	**Visualization Effect**
Reflect Limb Motion	Reflect motions of one limb vertically, longitudinally or transversally.
Switch Limb Motions	Switch motions of two limbs.
Copy Limb Motion	Copy motion of one limb to one or more other limbs
Transform Tempo	Transform current speed of fireflies and current playback speed of response gesture to make it faster or slower.
Transform Duration	Transform duration of response gesture to make it longer or shorter by either repeating or truncating gesture playback.
Transform Energy	Transform energy of response gesture by changing colour of the fireflies from red to orange to white to blue (smoothly) in order of increasing energy. In addition, areas of higher energy reflect the energy gradient.
Transform Smoothness	Transform smoothness of response gesture by changing the length and duration of fireflies' trails creating flowing movements.
Repeat gesture	Transforms the total duration of the response resulting in repeated playback of the original response backwards and then forwards alternately.

Table 3: Mappings between Viewpoints predicates and Visualization changes to *VAI*

Together *Viewpoints AI* encompasses the above flow of perceiving a gesture, reasoning to choose an appropriate response based on Viewpoints technique, and finally acting to procedurally render that response in a visualization. Below we discuss the *Viewpoints AI* installation and an example human interaction with the AI performer *VAI* illustrating the full architecture.

Interactive Installation

The *Viewpoints AI* system is an interactive installation where a human and AI participant can take turns co-creating a movement-based proto-narrative. The installation creates a liminal virtual / real performance space for the human and AI to interact in using techniques from shadow play and digitally augmented theatrical performances. Spectators view the installation from the front, watching the human and *VAI* interact turn by turn, performing gestures and expressive movements.

The installation was designed to enhance the presence of the user experience by using the human's shadow as their avatar for the interaction. The use of shadow play has the desirable analog property of being hyper-responsive at light speed displaying nuanced expression of user movement, while being sufficiently abstract to focus attention on both interactors simultaneously. The human's shadow is rear-projected onto a semi-opaque muslin screen to allow simultaneous front and back projection. The digital rendering of *VAI* is front-projected onto the muslin screen to serve as the second participant in the interactive experience. The audience views the installation from the front (see Figure 3).

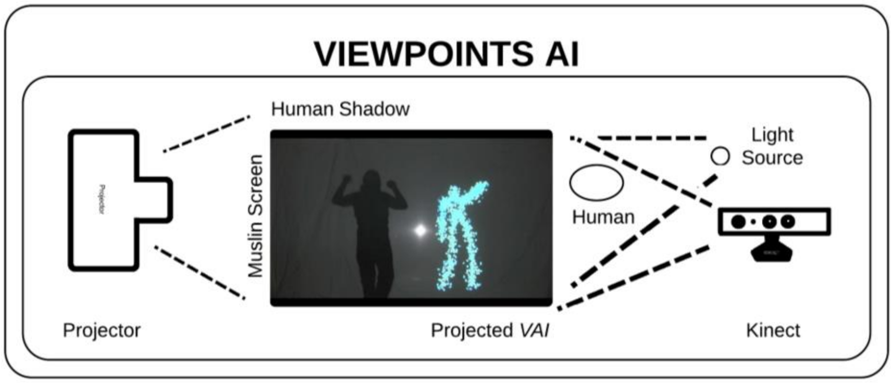

Figure 3: The liminal virtual / real interaction space of the *Viewpoints AI* installation, created through human shadow play and digital projection of the virtual interactor *VAI*.

Example

Below we describe a simple example of the interaction and internal processing involved in the *Viewpoints AI* system. The user starts the interaction by offering *VAI* a gesture consisting of walking from right to left in an exaggerated manner with long, purposeful strides (see Figure 4). The gesture is internally perceived by the *perception module* as having (a salient subset of perceived Viewpoints predicates) *long* duration, *medium* tempo, *high* energy, with only *longitudinal* limb motion and with average facing (stage orientation) *left of center stage*. The raw Kinect gesture and Viewpoints predicates are sent to the *reasoning module*.

Figure 4: The user walks from right to left in an exaggerated manner

Soar randomly chooses to look at just the last gesture by the human in order to decide its *response mode* (how to respond to the human's gesture). Soar decides to *transform* the user's last gesture as its current *response mode*. The functional transform *reflect gesture* is chosen for this response since there is expert aesthetics knowledge that promotes selection of reflection longitudinally when longitudinal reflection is perceived to be a highly noticeable transformation (*i.e.* when limb motion is longitudinal and when average facing or stage orientation is not facing stage center, presenting the interactor's profile to the audience).

The *action module* receives the transformed Viewpoints predicates and raw Kinect gesture from the *reasoning module*. The *action module* first maps the Viewpoints predicates it receives to parameters of the procedural visualization such as *high* energy to bright blue color and medium playback and firefly speeds to medium tempo. The *action module* then proceeds to playback the resulting *response gesture* sent to it by the *reasoning module*. It also carries out the functional transforms on *VAI* such as reflecting *VAI*'s leg movements in the longitudinal direction. The final result is *VAI*'s response of walking backwards from left to right in an exaggerated manner (see Figure 5). The reflected backwards walk would be physically impossible for a human, showcasing the benefits of augmenting analog reality with digital fantasy in an expressive installation.

Figure 5: *VAI* transforms the user's gesture by reflecting it longitudinally and walking backwards in an exaggerated manner from right to left.

CONCLUSION

Viewpoints AI is an exploration of a procedural rendering of the Viewpoints theatre technique to enable human/AI co-creation of proto-narratives. Unlike previous approaches to theatrical performance with AI, *Viewpoints AI* puts human and AI on equal ground in driving the meaning behind a performance. *Viewpoints AI* contributes models of perceiving gesture aesthetics in the Viewpoints framework; reasoning on how to respond to an interactor's gesture given the context of a history of interactions in real-time; and a procedural visualization method to render gesture responses. *Viewpoints AI*'s perception module can advance natural interfaces based on the Kinect to provide new forms of games based on a large space of gesture aesthetics, rather than relying on pre-coded gestures recognized for particular purposes. *Viewpoints AI*'s reasoning modules opens new avenues for interactions with game agents that understand an aesthetic history of interactions with a player and use this to guide intelligent responses that develop a meaningful interaction.

ACKNOWLEDGMENTS

The authors would like to thank Adam Fristoe for introducing them to the Viewpoints technique and for providing invaluable Viewpoints expertise and feedback in developing the *Viewpoints AI* system. The authors would also like to thank Gaëtan Coisne, Jihan Feng, Akshay Gupta, Rania Hodhod, Fengbo Lee, Paul O'Neill, Ivan Sysoev and Gaurav Gav Verma for their respective contributions in creating *Viewpoints AI*.

BIBLIOGRAPHY

Bogart, A. and Landau, T. “The viewpoints book: a practical guide to viewpoints and composition.” Theatre Communications Group, New York NY, 2005.

El-Nasr, M. S. (2007). “Interaction, narrative, and drama: Creating an adaptive interactive narrative using performance arts theories,” in Interaction Studies vol. 8 no. 2, pp. 209-240.

Horswill, I. D. (2009). "Lightweight procedural animation with believable physical interactions," in Computational Intelligence and AI in Games, IEEE Transactions on 1, no. 1, pp. 39-49.

Laird, J. “The Soar cognitive architecture.” MIT Press, MA, 2012.

Magerko, B., DeLeon, C., and Dohogne, P. (2011). “Digital Improvisational Theatre: Party Quirks,” in Proceedings of the 11th International Conference on Intelligent Virtual Agents, pp. 42-47.

Mateas, M. (1999). “An Oz-centric review of interactive drama and believable agents,” in Artificial Intelligence Today, Lecture Notes in Computer Science vol. 1600, pp. 297-328. Springer Berlin Heidelberg.

Mateas, M. (2001) “Expressive AI: A hybrid art and science practice,” in Leonardo vol. 34 no. 2, pp. 147-153.

Mateas, M., Domike, S., and Vanouse, P. (1999) “Terminal time: An ideologically-biased history machine,” in AISB Quarterly, Special Issue on Creativity in the Arts and Sciences vol. 102 (Summer/Autumn 1999), pp. 36-43.

Mateas, M. and Stern, A. (2003) “Integrating plot, character and natural language processing in the interactive drama Façade,” in Proceedings of the 1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment.

McCoy, J., M. Treanor, B. Samuel, N. Wardrip-Fruin and M. Mateas (2011) “Comme il Faut: A System for Authoring Playable Social Models,” in Proceedings of the 7th AI and Interactive Digital Entertainment Conference (AIIDE'11).

Morgenstern, L. (2008). “A first-order theory of Stanislavskian scene analysis,” in Proceedings of the 23rd AAAI Conference on Artificial Intelligence, pp. 496-503.

Norton, D., Heath, D., and Ventura, D. (2011). “An artistic dialogue with the artificial,” in Proceedings of the 8th ACM conference on Creativity and Cognition, pp. 31-40.

O’Neill, B., Piplica, A., Fuller, D., and Magerko, B. (2011). “A Knowledge-Based Framework for the Collaborative Improvisation of Scene Introductions,” in Proceedings of the 4th International Conference on Interactive Digital Storytelling, pp. 85-96.

Overlie, M. (2006). "The six viewpoints," in Training of the American Actor, pp. 187-221.

Piplica, A., DeLeon, C., and Magerko, B. (2012). “Full-Body Gesture Interaction with Improvisational Narrative Agents,” in Proceedings of the Twelfth Annual Conference on Intelligent Virtual Agents, pp. 514-516.

Riedl, Mark O., and Vadim Bulitko. (2013) “Interactive Narrative: An Intelligent Systems Approach,” in Artificial Intelligence Magazine vol. 34 no. 1 (Spring 2013), pp. 67-77.

Roberts, D. L., and Isbell, C. L. (2008). “A survey and qualitative analysis of recent advances in drama management,” International Transactions on Systems Science and Applications, Special Issue on Agent Based Systems for Human Learning vol 4 no. 2, pp. 61-75.

Schiphorst, T., Calvert, T., Lee, C., Welman, C., and Gaudet, S. (1990) “Tools for interaction with the creative process of composition,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 167-174.

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