Improv Remix: Mixed-Reality Video Manipulation Using Whole-Body Interaction to Extend Improvised Theatre

Dustin Freeman
dustin@cs.toronto.edu
Occipital
San Francisco, CA, USA

Ravin Balakrishnan
ravin@dgp.toronto.edu
Dynamic Graphics Project
University of Toronto
Toronto, Canada

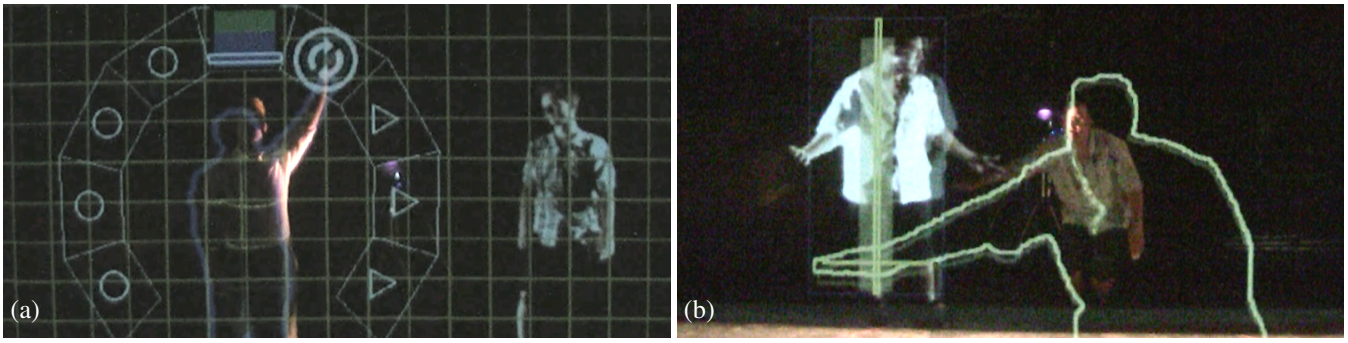

Figure 1. The view of Improv Remix from the audience. In the first frame, a live performer (left) accessing scenes using our novel Vitruvian Menu, and a video of a playback performer (right) is paused before playback. In the second frame, a live performer (right) scrubs his previous self (left).

ABSTRACT

Improv Remix is a mixed-reality system for live onstage video editing, using whole-body interaction, as an extension of improvised theatre. This work documents the process of analyzing an art form, then building technology that supports its extension. We tested the potential for video integration in improv in a exploratory workshop, then determined features and interaction techniques through iterative development with improvisors. We demonstrated the final iteration in a public showcase. Our contributions are (1) the documentation of the process, (2) a basic set of always-on interaction techniques tailored for performers standing adjacent to a large display, (3) methods to remix stage performance video, and, (4) a collection of creative use cases for the system, with an analysis of how the system extends improvisation.

Author Keywords

Whole-Body Interaction; Video; Improvised Theatre; Art; Mixed Reality;

ACM Classification Keywords

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

DIS 2016, June 04 - 08, 2016, Brisbane, QLD, Australia.

ACM 978-1-4503-4031-1/16/06$15.00
DOI: **http://dx.doi.org/10.1145/2901790.2901894**

INTRODUCTION

The usage of advanced technology in theatrical improvisation is rare, and this observation prompted the authors' investigation. In this work we sought to extend, and thus be faithful to, modern theatrical improvisation, as observed through one of the authors' 14 years of personal experience.

It is not mandatory or even recommended that we use technology in the creation of art. However, many traditional art forms now have advanced digital tools developed for them (e.g. painting, sculpture, film). Even scripted (non-improvised) theatre now benefits from digital tools for complex lighting changes and projection mapping (e.g. Troikatronix's Isadora[1]) — why not modern theatrical improvisation?

With Improv Remix, we *extend* the features of improv, allowing performers to create content not possible before, while still feeling in the spirit of improv, and taking advantage of pre-existing skills of improv performers. Our approach is to enable recontextualization of previous scenes by providing techniques to record, re-project and manipulate video of the stage (Figure 1). Thus, the live, present performers appear with their video recordings on a synthesized *mixed reality* stage. The system allows improvisors to construct complex, multi-layered scenes. The interface is designed specifically to be used spontaneously from the stage itself, reducing iteration time between ideation and instantiation of new theatrical ideas. Often, the reuse of a scene's video out of its original context is unexpected, surprising, and enjoyable. While other stage performance contexts could benefit from a mixed-reality stage system, this work examines the specific case of modern improvised theatre.

[1] **http://troikatronix.com/isadora/about/**

To validate that use of mixed-reality video on stage was compelling, and to prototype the features for Improv Remix, we ran a set of workshops: informal drop-in session for performers to experiment, unstructured, with a prototype. The system as described in this paper was evaluated by presenting it at a set of showcases to an audience. Our showcases were in theatre spaces, with scheduled showtimes, where public audience members were invited to watch performers use Improv Remix, and then use it themselves.

FEATURES OF MODERN IMPROVISATION

Modern improvised theatre is popularly seen on the TV show *Whose Line Is It Anyway?* Improv scenes progress via unpredictable, exploratory content generation, to the joy of performers and audience members. Through this process, ideas are discovered that would have been difficult to do so otherwise. Sets under 5 minutes are called "shortform", and longer sets are called longform. Extending the genre of longform improvisation, which can explore a wide variety of themes, was the primary motivation for this work.

Here are some definitions for the purposes of this document: An *action* is anything a performer does, while an *interaction* is an action meant to be interpreted by a digital system. Actions that are *onstage* are intended to be part of the show, while actions that are *offstage* are not. A *live performer* is performing in the present, whereas a *playback performer* is a performance that was recorded previously, playing back now.

To extend an art form, we must deeply examine its fundamentals. We shall describe three inspiring features of improv, derived from the personal experience of one of the authors, who has spent 14 years as an improv performer, and 5 as a director.

Minimalism

Jerry Grotowski argued in *Towards a Poor Theatre* (1967) [9], that to compete with other medias such as film, theatre should strip away superfluous elements. Others have argued that theatre's liveness is what makes it unique, and the inclusion of technology can be problematic to keeping it so [5].

Improv intentionally minimizes influences on the spontaneous generation of content; most groups use neutral clothing and few props. Most special effects in theatre, such as sound effects, lighting or projections, are controlled by technicians from offstage, using a pre-arranged list of cues. This would inhibit improvisation's ability to customize for the present audience, or recent events, as occurred centuries past in *commedia dell'arte*. Thus, we argue that any system meant to be used during improvisation must have as few constraints and distractions as possible.

In the last century, when control of technology *onstage* has been increasingly experimented with, it is the opinion of the authors that the interactivity has not been fully realized. Either the technology intrudes so much on the show that it becomes cyborgized — that the show becomes *about* technology, or, the technology is merely responds to performers in an aesthetic way, as often appears in dance. In this paper, the goal of the technology is not be aesthetic, but rather support the pre-existing task at hand: improvised theatre, though admittedly a very novel version of it. A similar tool for dance would support dancers planning the steps in a performance or controlling lighting changes from the stage itself. The point of an Improv Remix performance is not the tool, in the sense that the point of a film is rarely to show off the video editor used in its creation.

Recontextualization

Recontextualization, or a callback, is taking previous events or themes in a performance and juxtaposing them against others in the present. The most prevalent longform (>5 minutes) improv structure is called the *The Harold* [10]. In a Harold performance, improvisors will do several short, and initially disconnected, scenes, and as the performance progresses, themes, locations and characters will be revisited, so that the performance has a high density of self-reference. Callbacks are frequent in stand-up and sketch comedy, where an innocuous joke or event at the beginning appears again at the end.

The term *Chekhov's Gun* originates from Russian playwright Anton Chekhov, who observed that "If you say in the first chapter that there is a rifle hanging on the wall, in the second or third chapter it absolutely must go off"[3]. In the beginning of a longform improvisational set, performers produce a series of initially disconnected ideas, but during the set they weave them together in amusing and unexpected ways. This feature of improvisation is amenable to additional techniques to call back previous ideas — in our approach, to enable playback and manipulation of video of performers themselves, in new contexts.

Coordinating Gestures

Improvisors use coordinating gestures to manage group content generation. This is performer-performer interaction as a layer external to in-character actions. Coordinating gestures are an existence proof that interaction and performance may co-exist. A secondary stakeholder in the design of coordinating gestures is the audience, who must be able to understand if an onstage action is part of a performer's acting, or the work of planning the scene. Here are three examples:

Sweep: A performer runs from one side of the stage to the other at the front. Similar to a film wipe, this indicates the group will transition to the next scene.

Tag-out: A performer from offstage taps an onstage performer. This indicates the offstage performer will replace the onstage performer as a new character.

"Cut to that!": A performer calls "cut to that!" to tell the group to transition to depicting an event just mentioned in the current scene.

Note that we cannot reuse any coordinating gestures already in common use in improvisation for our system, as performers will instinctively use them during performance. The nature of these gestures merely serve as inspiration for gestures we may use in Improv Remix, which must co-exist safely alongside currently-used gestures.

RELATED WORK

As a whole, this work explores a technological extension of an existing art form, via an iterative process, creating new interaction techniques to support an extended feature set. This approach is similar to similar to Kazi et al., who extended the art form of sand painting to a digital touch surface [16]. To contextualize the work in this paper, we will discuss the topics *Liveness*, *Mixed Reality*, *Interfaces for Live Video* and *Whole-Body Interaction*. While some work has explored theatrical improvisation as an interesting AI problem, this is out of scope for our current work [19].

One interesting prior work that covers a few of the topics is a scene in Blast Theory's *10 Backwards* where the actress uses a remote control to record and play back video of herself eating a bowl of cereal [28]. A camera on a tripod is facing her, and behind the camera is a large projection screen, where she and the audience can see the output from the camera. She records herself using the camera, and then plays herself back, trying to imitate her actions with slight exaggeration. She then records this exaggerated performance, and then repeats the same self-imitating procedure, eventually using the remote to go forwards and backwards frame-by-frame to imitate exact facial expressions.

Liveness

It is widely believed that liveness, the dramatic sensation that a performance is happening in the present, is important for compelling art, but it is unclear how to foster liveness with non-live (mediatized) components, such as pre-recorded video [1, 13, 22]. There has been much analysis of the relationship between live and projected images of performers [5]. *Gertie The Dinosaur* (1914) is an early example of combining live and pre-recorded interaction, where a physical prop transitioned from a live performer to an animated character [20]. In the Blast Theory example, the performer is not using a version of themselves that is exactly live, but is in the very recent past, similar to Improv Remix. *Sampling* content from another source is common in art, and an interesting examination of liveness in digital art [5, 21, 27]. Creative workers have observed that it appears to be human nature to ascribe meaning to chunks of content after it has been randomly assembled, as if a large part of the origin of the story is the observer themselves [18, 33].

Mixed Reality

For creating *Mixed Reality* in a theatre context, it is difficult technical problem to project images in-place on the stage without requiring any augmentation of audience members. Some works have had performers wear completely white clothing, so that the image of a human figure can be projected on top of them [5]. Tsuchida et al. created a system for rehearsing multi-person dances by placing a projection screen on top of a robot; the robot can move around the space arbitrarily and a calibrated projector can project human figures to that position [29]. The technique we use for Improv Remix is a variant on the extensively-used stage technique *Pepper's Ghost*, where a transparent projection screen is placed in front of a real, conventional scene, allowing virtual objects to appear (and disappear), apparently from the real scene [26].

Interfaces for Live Video

The interface used in the Blast Theory example is a basic TV remote, recording to or playing from a VHS tape. Even with this setup, a highly interesting performance can occur. However, there are many other instances of interfaces designed for live video output, particularly VJ software [8, 12]. Outside live network TV, we could find few cases that include live video input as well [7]. One feature in Improv Remix — polite playback — is automatic non-linear playback of video in response to live performers. There has been work on combining several mobile novice-recorded videos together — in one case, a human "director" is designated to manage all the input streams from various people [6]. There has also been much work on automatic editing of video [32, 31], and even semi-automatic editing with some user input [14]. For Improv Remix, we must parse a recorded performance into chunks based on audio; a similar technique appears in DemoCut to reduce content size for instructional video [4].

Whole-Body Interaction

In Improv Remix, interaction is possible with the limbs and the coarse position of the body. Early examples of work on unencumbered Whole-Body Interaction are CHARADE [2] and VIDEOPLACE [17]. Vogel explored the notion of *proxemics* with a large display, where different functionality could be present at different distances [30]. Shadow Reaching exploits perspective projection so the user can operate at different scales depending on their position [25].

Even outside theatre practice, there is concern on how interaction is perceived by observers. The interaction design values for a generic user (such as minimal effort, or avoiding awkwardness) [11, 24] are different from theatre performers' values [23]. Audience members are explicitly present to watch performers, and the nature of interaction can affect their perception of the performer. We argue that for theatre, we should choose interaction techniques so that they influence the semantics of the show as little as possible; by Reeves' analysis, this suggests that both manipulations and effects should be made clearly visible to the audience [23]. Juhlin and Önnevall's work on collaborative gestures in front of TV screens also shows support for gestural visibility between users as being important [15]. In the Blast Theory example, using an interface that is not novel to the audience (a TV remote) means that interactions with it are well understood, even though they may not see specifically which buttons the performer is pressing.

EXPLORATORY STUDY: WORKSHOPS

We ran informal exploratory workshops with a prototype enabling capture and playback of stage video. The purpose of the workshops was to (1) validate that using playback performers was compelling and (2) discover design issues and opportunities that Improv Remix would face. The backend is a bespoke colour and depth media storage container designed for simultaneous read and write of multiple video streams [2]. This container was later improved and used in the final system.

[2] Source `http://github.com/dustinfreeman/riffer`

Participants
We recruited participants by word-of-mouth (21 total, ages 20 to 60, all had at least 2 years experience with longform improvisation). No compensation was given.

Prototype
We used a Kinect to record colour and depth video, and projected it onto a wall at the rear of the stage area. Performers could play the last scene with a "clap" gesture, detected with the Kinect skeleton tracker. A new scene started recording when someone entered the stage area, and stopped once the stage was empty again. Scenes could be played back once or looped, controlled by the researcher from a laptop.

Procedure
Performers could drop in any time during the workshops' 6-hour duration. The researcher explained the prototype functionality if necessary. However, they preferred that experienced performers explain the prototype and its use cases to new arrivals. Occasionally the researcher would have to complete descriptions, or correct misunderstandings.

Results
Improvising with playback performers was found to be definitely compelling. Use cases were discovered during the workshop that we iteratively expanded on and explored in Improv Remix — these are described in a later section. Below we discuss issues identified during the workshops

Sightlines
When humans typically speak to each other, their torsos are directly facing. In a performance, performers will "cheat out", i.e. rotate their torsos unnaturally towards the audience so they are more visible. In the prototype, the playback performers were projected behind the live performers, while the audience was in front. This made it difficult for the live performers to present the front of their torso to the audience, while still being able to see the playback performers.

Distinguishing Interface Actions from Performance
Performers' interaction techniques were not robust. The clap gesture to play the last scene frequently caused false positives and often had to be disabled. The gesture to stop and start recording; walking on or off stage, appeared at the beginning and end of every recorded scene and was obtrusive.

The Importance of Spontaneity
To execute ideas, performers would have to describe them to the researcher operating the prototype. New ideas are vague and difficult to express, and if the performer was uncertain, they would abandon them. We feel that operation of the system from the stage itself, by performers, is very important.

Live and Playback Scene Responsiveness
A common use case was to record a *Template Scene*, consisting of a performer on one half of the stage, speaking to the empty half. Performers would play back this scene several times, responding creatively to the playback performer. However, the playback performer would tend to speak over the live performer, who would naturally react as if they were offended, constraining and derailing a scene.

IMPROV REMIX IMPLEMENTATION
Improv Remix represents the final iteration in a collaborative design process with experienced improv performers, intended to be well-suited to spontaneous interaction directly from the stage. Our resulting interaction techniques and feature set were selected to be simple, easy to teach indirectly, and frequently used. During the iteration process, we dropped (1) interaction techniques that were too finicky or difficult to teach and (2) features that did not create interesting performances. As such, the resulting system is not intended to represent a thorough, rich, expressive exploration of the interaction design space, but a polished, lightweight, pragmatically usable instance of a system that creates compelling performances.

During the process, we found a particularly important property — *exposure* — when a performer did an interaction, the degree to which an observer (audience members or other performers) could discern the intent and the effect of the interaction. To support coordinated improvised scene-making, it was beneficial to design interactions so that they were highly exposed. This is one reason why we kept all interactions whole-body, instead of pursuing other interaction techniques.

Example Usage
Here is an abstract description of an improv set using some features in Improv Remix. Interaction techniques will be described in later sections.

Alice starts behind the performance area, and Bob and Carlos start off to the side.

Alice steps forward, and records a performance of her monologuing about a topic (cats). When finished, she steps off the stage and the recording stops.

Bob walks into the performance area, and instantiates a playback performer of Alice's monologue on cats. Bob acts as a supportive listener, nodding in agreement with Alice's playback performer, and sometimes lightly clapping.

Before Alice's playback performer is finished, Carlos indicates to Bob to get off the stage. Bob leaves, and Carlos jumps on stage.

Carlos clears the stage of Alice's playback performer, and restarts it from the beginning, this time setting the interface to record his own performance as well.

While listening to Alice's playback performer, Carlos acts as an unsupportive audience member, loudly laughing, expressing disbelief, being offended and booing.

When Alice's playback performer finishes, Carlos steps off the stage. Improv Remix now has two performances in its library: Alice's cat speech, and Carlos being a rude listener.

Bob steps on stage and instantiates Carlos' playback performer. Bob begins a speech on another topic, acting as a mayor who is opening a new kids hospital. During the speech, Carlos' playback performer keeps interjecting rudely in apparent opposition to what Bob is proposing. Bob becomes increasingly comically frustrated.

Hilarity ensues.

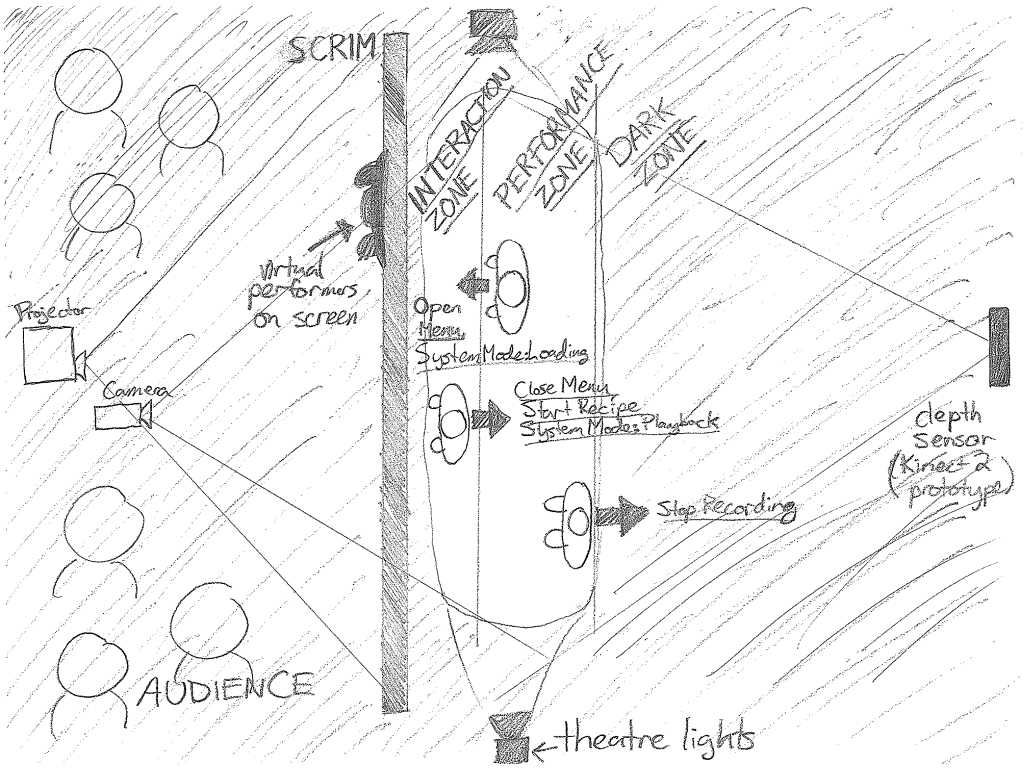

Figure 2. The final physical setup of Improv Remix. Middle: scrim. Left: projector, camera and audience. Right: lit performance space. Far Right: depth sensor (Kinect 2). Proxemic zones affect interface features: the Interaction Zone closest to the scrim, the Performance Zone in the lit area, and the Dark Zone past the light.

System Interface Overview

Improv Remix has three global modes: *Loading*, *Performance*, and *Library*. In Loading, performers instantiate playback performers using our novel *Vitruvian Menu* (Figure 3). In Performance, the interface is clean and any playback performers may be directly manipulated by live performers. In Library, a performer may browse all recorded performers, and slot them into the Vitruvian Menu for quick access.

We designed interactions to be robust in our inherently noisy and ambiguous environment. The scrim is invisible unless lit, so playback performers and UI elements appear to hover in mid-air. Having the scrim between live performers and the audience improves sightlines over our workshop prototype. We discovered the bright lighting inhibits visibility of UI elements that were farther away or used colour; thus, we used audio cues when possible.

Proxemic Zones

The stage has 3 proxemic zones (Figure 2): the *Interaction*, *Performance* and the *Dark*. A performer's zone is determined by distance from the rear depth camera, with hysteresis applied to prevent debouncing[3]. Audio feedback informs users of zone transitions. When a performer transitions zones, it affects system mode and recording: Stepping forward from the Performance Zone to the Interaction Zone sets the global mode to Loading, and the performer's Vitruvian Menu (described later) opens. The performer can choose to record a scene, or instantiate playback performers from previous scenes; we call this the *scene recipe*. The scene recipe does not execute until the performer steps back to the Performance Zone, which starts a 2.5 countdown until execution. To cancel a scene recipe, the performer steps back from the Performance Zone to the Dark Zone.

We use proxemic zones to enable performer access to the menu — this was found to be more robust than other approaches: (1) having the widgets always available, but in hard-to-reach places to prevent accidental activation, and (2) having the menu appear only if the user stood still for a period of time.

The Vitruvian Menu

Our novel *Vitruvian Menu* (Figure 3 a&b) is designed for a standing human; its name inspired by Leonardo Da Vinci's *Vitruvian Man*, as it is usable by arms and legs. We chose this user-relative design after finding that widgets with absolute position were difficult to acquire.

The Vitruvian Menu supports quick, spontaneous access of scenes, yet is also robust to the noisy nature of whole-body interaction. When the user's silhouette is overlapping with a button in the Vitruvian Menu, it gradually fills over 300 ms, and when untouched, it slowly un-fills; this mechanic, an extended form of dwell, ensures the button does not activate without a degree of certainty, but also makes it possible to

[3]Debouncing is an electronics term for strategies to prevent a switch rapidly alternating between states, by making a state "sticky", tending to preserve itself.

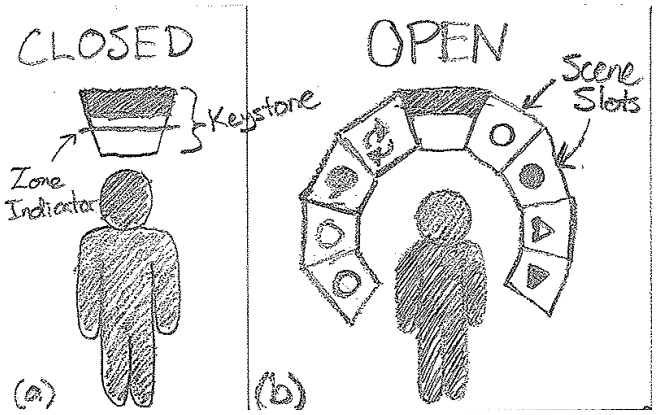

Figure 3. The Vitruvian Menu. (a) Closed, with only the Keystone visible. In the Keystone, we provide zone feedback. (b) Open, showing all scene slots, accessible by arms and legs. Icons in each slot indicate whether it has a scene, and its playback type in the Scene Recipe.

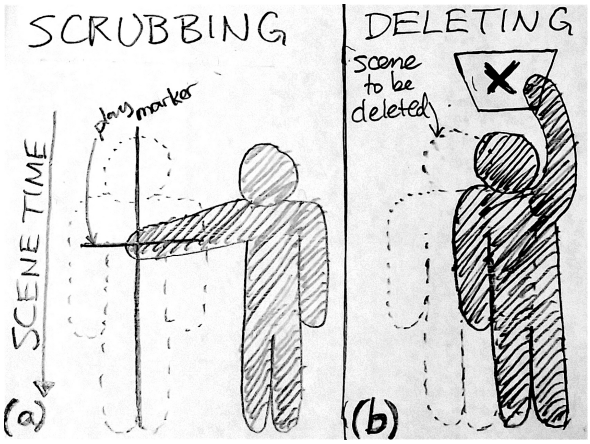

Figure 4. Interaction with Playback Performers: (a) Scrubbing a playback performer by reaching into its bounding box. The play marker is set to the centroid of the user's overlap with the timeline. (b) Deleting a playback performer by standing in its bounding box turning the Keystone into a delete button.

activate if the user has trouble maintaining overlap with the button. While activation by dwell is slow, other approaches were tried, such as (1) a specific gestural direction motion, and (2) instant activation if any part of the body touched a widget. However, these were found to not be robust enough for a performance setting.

When the performer is in the Performance Zone, their menu is closed, with just the Keystone visible; when the performer is in the Interaction Zone, the menu is open, with all buttons visible and Improv Remix in Loading mode. We use the Keystone to display feedback of the performer's current proxemic zone, while every other button is slot for a scene. Slots start empty, with Empty Circle icons resembling record buttons, whereas slots with scenes have Empty Triangle icons, resembling play buttons. When the user activates an icon it becomes solid, indicating that it is part of the scene recipe. An activated play icon means the system will instantiate a playback performer from that scene, and an activated record icon means that the system will record a new scene, and save it in that slot. When activated, a slot with a scene will cycle through possible playback behaviours: *none*, *play once*, *looping* and *polite* (described below).

Interaction with Playback Performers

Live performers may *scrub* or *delete* playback performers (Figure 4), by intersecting with a bounding box around the playback performer. We tuned this size of this so it was closer than natural acting speaking distance, to prevent accidental interaction.

Scrubbing

We define scrubbing as the ability to freely move a video backwards or forwards in time, where a release resumes normal play. Our scrubbing interface is composed of a vertical timeline, as well as a horizontal line representing the current play marker. To scrub the playback performer, the live performer touches the timeline, which sets the playback performer's play marker to that time. To avoid accidental scrubs when a live performer walks past a playback performer, we ignore silhouette overlaps with the bounding box above a threshold.

Scrubbing enables playful re-imagining of scenes, a powerful control which has not been possible before onstage. Scrubbing was found to be very intuitive by almost all who used it. In informal workshops, scrubbing was used innovatively by performers outside our primary target group. Dancers would finely control previous versions of themselves co-dancing with their arms, legs, elbows, knees, and even head. Vocal musicians set previous tracks to play at points that were easy to memorize, exploiting proprioception.

We found performers had fine control over the scrub position. Scrubbing is possible both from the Performance Zone and the Dark Zone. When performing with an audience, control from either zone takes on different characteristics; where hidden interaction from the Dark Zone feels like puppeteering. Due to perspective projection, scrubbing from the Performance Zone is best done with an ankle or a wrist, while from the Dark Zone, scrubbing may be done with the finger.

Deletion

Playback performers in *looping* or *polite* playback persist onstage until deleted manually. When a live performer and a playback performer overlap, a delete icon appears in the live performer's Keystone. Tapping the Keystone deletes all playback performers they overlap with. The stage can be quickly cleared by holding one's arm in the keystone while walking across, much like the *sweep* coordinating gesture in modern improv.

Playback with Manners: Polite Playback Performers

When a playback performer's mode is *polite*, its behaviour is to play random excepts from the original scene so as to create the appearance of the playback performer delivering new, often surprising, material. We accomplish this by using a naïve algorithm to parse the scene into suitable utterances and controlling their behaviour carefully so they appear to have manners with respect to the other stage occupants.

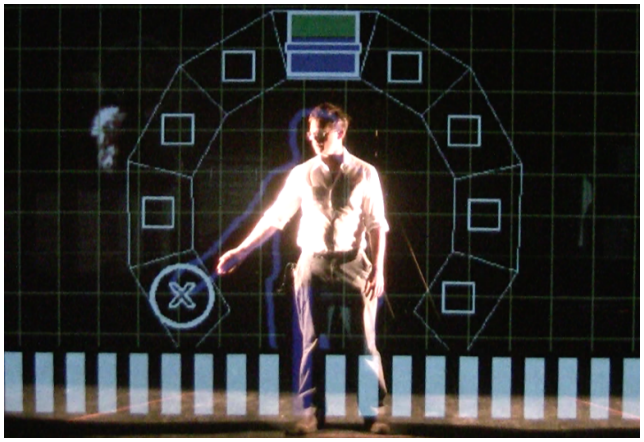

Figure 5. A performer using the Scene Library. To his left is the projected image of a puppet from his currently selected scene.

The default behaviour of a polite playback performer is to loop the longest non-utterance or "idle" section of the scene. This creates the appearance that the playback performer is listening, and avoids the jarring effect of freezing playback. Our system is careful when to trigger playback of utterances so that playback performers do not to talk over each other or live performers on stage. The *Manners Module* keeps a record of the last time the volume of the stage exceeded a threshold representing speech. If the onstage volume is silent for a given duration (we used 1.8 seconds) the manners module plays an utterance from a polite playback performer. If there are multiple onstage polite playback performers, the manners module maintains a circular queue, so that they take turns. After a playback performer finishes an utterance, it returns to looping the idle section.

The Scene Library

The Vitruvian Menu is meant for quick access, and thus only has 8 scene slots. Library mode is used to fetch scenes from the large quantity that may be generated over a typical 2-hour session (50+). To access the Library, the performer taps the Keystone when in the Interaction Zone (Figure 5). In Library Mode, a slot with a scene in it displays an X — if the performer activates this slot, they empty it. An empty slot has a square icon, indicating it can be filled with a scene. In the Library interface, scenes are represented by white rectangles on stage. The user selects a scene by walking left or right to stand over it, while the first frame of the selected scene is projected onstage. To fetch that scene into a specific empty slot, the user activates it.

SHOWCASE OBSERVATIONS AND USE CASES

We installed Improv Remix in a theatre space and ran three 2-hour public showcases. These were advertised over social media, and had approximately 50 attendees in total. We recruited 3 performers to create a demonstration of system features and known use cases. The structured demonstration of the system lasted an hour. In the second hour, audience members could come up and use Improv Remix, which including playback performers from earlier in the show.

The purpose of the showcases was to test the entertainment value of the system in front of a live audience. While a use case could be academically interesting in a workshop, we could not assume it would be in front of a live audience. The showcases also allowed us to observe if performers and audience members could invent use cases spontaneously

Physical Interaction

A performer created a chain of videos where he high-fived previous iterations of himself, with the audience clapping at each virtual impact. In a more complex example, a performer captured video of a pillar prop at three positions, and then instantiated them as looping, effectively static, objects. He made a puppet jump between the pillars (Figure 6a).

Collages

Performers layered several looping playback performers enacting a sound and/or physical action. In one example, performers recorded themselves individually dancing silently, intending to create a "dance party" when combined (Figure 6b). One of the performers played a song [4] on his smartphone next to his lapel mic. The resulting collage has him standing on the side of the stage, awkwardly swaying to the music, in contrast to the dancers. While the end result of a collage may be interesting, the audience can feel bored watching the build-up. It may be possible to make this payoff more exciting with performer practice, or it may be useful to bore the audience, lowering their expectations, before the payoff.

Music

One of our performers was an accomplished beatboxer and singer. He recorded himself beatboxing while dancing to create a layer, and then would begin to rap on top of that (Figure 6c). He constructed another example by singing the same song 3 times, and coordinated pointing to himself between the different video tracks (Figure 6d).

In many of these musical examples, the performer would construct the composition on the performer side of the scrim, then create a scene recipe of the prepared playback performers, and then excitedly run around to the audience side of the scrim to observe, with everyone else, the quality of their result. Musical samples from highly-skilled performers would often be brought back to accompany performers with less skill, and they would lip sync or react to the music. An interesting observation is that a short chunk of performance can become musical over the course of many (usually 10+) repetitions, regardless of orginial tone.

Constructed Scenes

The ability to record a scene while playing a previous one allows individual performers to construct complex scenes, exploiting timing and alignment. We have previously noted that the build-up of separate scenes intended to be combined is often not interesting. Here is one example where a performer recorded a series of scenes to that were interesting to watch under construction and satisfying in combination.

[4] with intentional irony, Robyn's *Dancing On My Own*

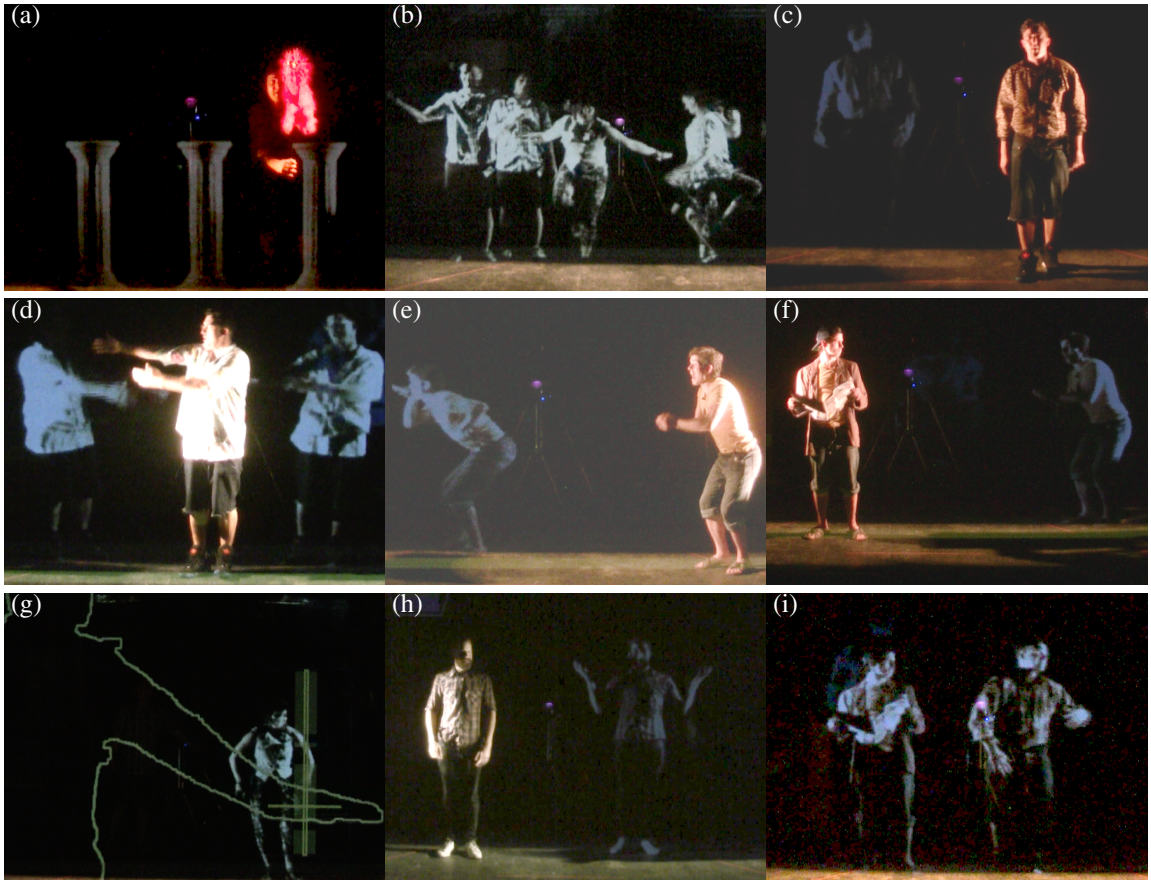

Figure 6. Example Use Cases: a) Physical interaction: A live puppet climbing static playback columns. The puppeteer hides his body in the Dark Zone; b) Collage: A dance party of several playback performers; c) Music: A beatboxer/dancer layering another instance of himself; d) Music, recontextualized: A performer replays two singing tracks so he may harmonize and gesture with himself; e) Constructed Scene: Part 1, an exuberant man beckons a duck; f) Constructed Scene: Part 2, a dishevelled (live) man responding to the beckons of a (playback) exuberant man; g) Responsive Scene via Scrubbing: a performer controls their own video from the Dark Zone; h) Polite Playback: A live performer records himself saying "true" or "false", then instantiates a polite playback performer and monologues, with his playback self informing him if he is lying; i) Failed Dissonance: The beatboxing playback performer re-appropriated alongside the dishevelled man. The dishevelled man's angry looks, originally intended for the exuberant man, appear to be directed at the beatboxer.

Scene 1: *Acting like a duck, the performer waddles from stage right to stage left, occasionally looking behind. Finally, it turns around and waddles back slightly faster, stands up and kisses an empty spot in the air.*

Scene 2: (Figure 6e) [With Scene 1 playing] *The performer stands on far stage right repeating exuberantly "Come here duck!" and beckoning as the duck, from Scene 1, waddles away. Finally, the performer loses his patience and loudly yells "Hey duck!", at which point the duck turns around and comes back. When it gets back to the performer, he picks it up and kisses it, saying "You're so cute!"*

Scene 3: (Figure 6f) [With Scene 2 playing] *A dishevelled man stands on stage left reading a newspaper. The exuberant man from Scene 2 repeats "Come here duck!", annoying and distracting the dishevelled man. Initially, the dishevelled man assumes that the exuberant man must be not be speaking to him. As the exuberant man continues, the dishevelled man becomes angry. Finally, the exuberant man from Scene 2 yells "Hey duck!" and the dishevelled man drops to the ground, ducking from possible danger. Seeing there is none, he charges the exuberant man, saying "Hey, buddy, what's the big idea!?". The exuberant man kisses him and the dishevelled man slaps him in response.*

Responsive Scenes

Both scrubbing and polite playback made for novel content which live performers had to respond to spontaneously.

Performer A asked the audience for 3 simple phrases, getting "It's raining", "These pretzels are making me thirsty." and "That's what she said". Performer A recorded himself saying these from stage right, facing (empty) stage left. Performer B, who had been sent out of the room, was called back and stood opposite playback performer A, while live performer A scrubbed his performer from the dark (Figure 6g).

One audience member created a polite playback performer saying "True" or "False" (Figure 6h). He then began a lecture and let his playback self choose whether the statement he just made was true. If a lie, he had to change his statement.

Failed Dissonance

Two seemingly unrelated scenes would often create an unexpected synthesis. A meaningless combination would be expected, but, in fact, often the dissonance *fails*. This effect is likely due to the unavoidable sense-making activities of human beings [18, 33]. We had two examples of failed dissonance: a) the dishevelled man from our Constructed Scene alongside our beatboxer, where his anger now appears directed at the beatboxer (Figure 6i) and b) An accidental recording of the beatboxer discussing and idea offstage, including the phrase "bastardized the equilibrium of harmonized core sounds". This was brought back amusingly against a multitude of other scenes, appearing to be overwrought and over-serious academic commentary on what is a bunch of silly messing around on a stage.

DISCUSSION: PERFORMERS AS SPECIAL USERS

Improv theatre performers are an unusual user group to design for, and here we will discuss observations of designing technology for them in a performance setting.

Instructing Interaction

We have made the specific choice in this project to not encumber the performers with hardware, and instead use coarse whole-body gestures, which are naturally noisy. Earlier in the project, we tried to search for gestures that would reliably not appear in theatrical performance, and could be reserved for system interaction. However, performance gestures are far too unpredictable, and if we banned certain, specific movements for performers, then performers became self-conscious. Fortunately, performers have a high degree of control over their bodies; when they are told they must pose or move in a certain way, they can consistently reproduce it. Performers are practiced at being cognizant of lighting on themselves and their visibility to the audience, and many theatre performers also have formal movement or dance training. The lesson is that "Do it exactly this way" works well, while "Don't do this" is received poorly.

Importance of Direct Control

A consistent observation during this work is that novel theatrical ideas are difficult, even tedious to explain in the middle of creative work. One goal of Improv Remix was to empower performers to have control of technical effects from the stage, thus, performers could execute half-formed ideas without having to articulate them, i.e. "brevity is the soul of wit". In the workshops, performers would occasionally abort ideas that were too difficult to articulate. It is not that performers are less persistent than the average user; just that they are used to generating a large amount of novel content quickly. If they are in a fast, creative mental state, as opposed to a seated formal discussion, then an impediment preventing them from prototyping an idea is in danger of making them abandon that idea for an "easier" one.

We noted that the coordinating gestures used in modern improvisation are fast and unambiguous to all performers on stage and the audience. The goal for our system was for its interaction to have the same degree of immediacy and clarity, which we termed *exposure*. We feel we achieved this, and performers were able to execute complex ideas without having to describe them. The visibility of interaction in our system also meant that there was no mystery as to how they achieved the result they intended, and others could replicate and build on their work.

Feature Elicitation

Improv is known for its "don't say no, say yes" rule. In truth, this is more nuanced — while rejecting an idea is often negative, improv teaching focuses on what to do with new information during a scene. Longform improv focuses on the "game of the scene" [10], a discovered pattern of behaviour that the performers may generatively explore and heighten or intensify. From a system developer perspective, this can run counter to the goal of finding a finite set of features to implement. Achieving a consistent and predictable map between user intention and result is not a good measure of success if our desire is to produce new tools for performers. Perhaps a better goal would be if a tool has properties that are initially surprising and novel.

Performers frequently asked questions of the form "Is it possible to...[X]" and the answer instinctively given back is whether X is possible with state-of-the-art equipment and algorithms, not whether it is possible to implement within the attention span of the asker. However, the system developer was cautious to say "no", as there was a worry that they will stop asking for interesting ideas. In practice, the best answer was "Yes, but not during this session".

Part of software development is finding and removing bugs. Performers' reactions to bugs was frustratingly positive — surprise and joy — as if the system was talking to them. Performers are trained to make every situation interesting — this means they are highly flexible in what others would consider a stressful situation (unexpected behaviour). It became difficult to work with this mindset during the early stages, when the goal of the developer was to end each workshop with a precisely-prioritized list of action-items.

CONCLUSION AND FUTURE WORK

We have presented *Improv Remix*: a design exercise in extending an art form through applied technology. Our primary contribution is our documentation of this process, from an analysis of the art form, to proposing an extension approach (performer control of video from earlier in the improv set), to designing a system, and observing use cases that would not have been possible before. This work has inspired the construction of novel interaction techniques, such as the *Vitruvian Menu*. We have also demonstrated a way to make verbal video content recombinable in a way that is satisfying to watch, with *polite playback*.

This work was inspired by existing practices; the goal was not to replicate them in digital form, but to extend and extrapolate from them. The extension work was specific to improvised theatre, though we feel there is general value in what we discovered to other theatre and art-making forms. We propose that extensions to other well-defined art forms and genres are possible using a process similar to that used in this work.

ACKNOWLEDGMENTS

This project would not be possible without Montgomery C. Martin, as the theatrical projection technical director. We also thank members of the DGP lab for putting up with a long-term physical setup, especially Haijun Xia, Jota, and John Hancock. We would also like to thank the members of the Toronto theatre community, especially Xander Williams, Alex James, Justine Cargo, Joe Law and Oliver Georgiou.

REFERENCES

1. Auslander, P. Live performance in a mediatized culture, 1999.
2. Baudel, T., and Beaudouin-Lafon, M. Charade: Remote control of objects using free-hand gestures. *Commun. ACM 36*, 7 (July 1993), 28–35.
3. Bill, V. T. *Chekhov–the Silent Voice of Freedom*. Philosophical Library, 1987.
4. Chi, P.-Y., Liu, J., Linder, J., Dontcheva, M., Li, W., and Hartmann, B. Democut: Generating concise instructional videos for physical demonstrations. In *Proc. of ACM UIST* (2013), 141–150.
5. Dixon, S. *Digital Performance: A History of New Media in Theatre, Dance, Performance Art, and Installation*. The MIT Press, 2007.
6. Engström, A., Esbjörnsson, M., and Juhlin, O. Mobile collaborative live video mixing. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, MobileHCI '08, ACM (New York, NY, USA, 2008), 157–166.
7. Freeman, D. E., Santosa, S., Chevalier, F., Balakrishnan, R., and Singh, K. Laces: Live authoring through compositing and editing of streaming video. In *Proc. of ACM CHI* (New York, NY, USA, 2014), 1207–1216.
8. Graphics, L. D. M. Resolume VJ software.
9. Grotowski, J., Wiewiorowski, T., and Morris, K. Towards the poor theatre. *The Tulane Drama Review 11*, 3 (1967), 60–65.
10. Halpern, C., Close, and Johnson, K. H. *Truth in Comedy: The Manual for Improvisation*. Meriwether Publishing, 1994.
11. Hansen, L. K., Rico, J., Jacucci, G., Brewster, S., and Ashbrook, D. Performative interaction in public space. In *Proc. of ACM CHI EA* (2011), 49–52.
12. Hook, J., and Olivier, P. Waves: multi-touch vj interface. In *Proc. of ACM ITS* (2010), 305–305.
13. Hook, J., Schofield, G., Taylor, R., Bartindale, T., McCarthy, J., and Wright, P. Exploring hci's relationship with liveness. In *Proc. of ACM CHI* (2012), 2771–2774.
14. Hua, X.-S., Wang, Z., and Li, S. Lazycut: content-aware template-based video authoring. In *Proc. of ACM MULTIMEDIA* (2005), 792–793.
15. Juhlin, O., and Önnevall, E. On the relation of ordinary gestures to tv screens: General lessons for the design of collaborative interactive techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 919–930.
16. Kazi, R. H., Chua, K. C., Zhao, S., Davis, R., and Low, K.-L. Sandcanvas: A multi-touch art medium inspired by sand animation. In *Proc. of ACM CHI* (2011), 1283–1292.
17. Krueger, M. W., Gionfriddo, T., and Hinrichsen, K. Videoplace an artificial reality. In *Proc. of ACM CHI* (1985), 35–40.
18. Kulešov, L. V. *Kuleshov on film: writings by Lev Kuleshov*. Univ of California Press, 1974.
19. Magerko, B., Manzoul, W., Riedl, M., Baumer, A., Fuller, D., Luther, K., and Pearce, C. An empirical study of cognition and theatrical improvisation. In *Proceedings of the seventh ACM conference on Creativity and cognition*, ACM (2009), 117–126.
20. McCay, W. *Gertie the dinosaur*. Museum of Modern Art, 1909.
21. Murray, J. H. *Hamlet on the holodeck: The future of narrative in cyberspace*. Simon and Schuster, 1997.
22. Phelan, P. The politics of performance. *London and New York: Routledge 4* (1993).
23. Reeves, S., Benford, S., O'Malley, C., and Fraser, M. Designing the spectator experience. In *Proc. of ACM CHI* (New York, NY, USA, 2005), 741–750.
24. Rico, J. Evaluating the social acceptability of multimodal mobile interactions. In *Proc. of ACM CHI EA* (New York, NY, USA, 2010), 2887–2890.
25. Shoemaker, G., Tang, A., and Booth, K. S. Shadow reaching: a new perspective on interaction for large displays. In *Proc. of ACM UIST* (2007), 53–56.
26. Steinmeyer, J. The science behind the ghost: A brief history of pepper's ghost, 1999.
27. The Builders Association. Xtravaganza, 2002.
28. Theory, B. Ten backwards. DVD, 1999.
29. Tsuchida, S., Terada, T., and Tsukamoto, M. A system for practicing formations in dance performance supported by self-propelled screen. In *Proc. of ACM AH* (2013), 178–185.
30. Vogel, D., and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proc. of ACM UIST* (2004), 137–146.
31. Wang, J., Bhat, P., Colburn, R. A., Agrawala, M., and Cohen, M. F. Interactive video cutout. *ACM Trans. Graph. 24*, 3 (July 2005), 585–594.
32. Yip, S., Leu, E., and Howe, H. The automatic video editor. In *Proc. of ACM MULTIMEDIA* (New York, NY, USA, 2003), 596–597.
33. Zimmerman, E. Creating a meaning-machine: The deck of stories called life in the garden. *Second Person* (2007).