A Study I'd Live Streaming Practices in China

As early as 2005, live-streaming culture began in China when public video chat room services such as YY were repurposed by users to host public performances []. Continuous growth of live-streaming in China has resulted in more than 200 million viewers watching the streamers perform live each night on more than 200 live-streaming platforms, creating an estimated 15 billion dollar industry in China in 2016 [moshinsky_2016]. However, despite the early appearance of game live-streaming like Twitch.TV [hamilton2014streaming] in North America, it has only more recently expanded into mainstream culture with Facebook Live, Youtube Live and Periscope [].

Chinese live-streams differ greatly in content, style and form compared to those in North America. Previous studies of US and Canadian live-streaming have found that users in these jurisdictions use streaming almost exclusively for either live events or sharing among close-tie friends []. In contrast, we have found that these uses constitute a minority of Chinese use of such services. . there are various pan-entertainment live-streams like talent shows, variety shows, talk shows, e-commerce, personal knowledge sharing and personal experience sharing [36krreport]. Further, in contrast to North American corporate dominance of livestreams, Many Chinese streamers have more than 100,000 regular viewers, and in many cases treat their streams as a primary source of income.

Beside from the differences in their use, the financial incentive models services offer to their streamers are also differentiated in China. In North America, streamers are compensated primarily through a share of advertising revenue from their streams, or through pay-per-view [cite]. In contrast, live-streaming services in China use a reward-based system where viewers are enabled to purchase virtual "gifts" for the streamer during the stream. . When this happens, animations and other forms of notification are shown to the streamer and all other viewers. Once streamers receive these gifts, they can then exchange them for cash. Each streamer also has a leaderboard of the gift-senders. The community around Chinese live streamers is fostered by streamers use of additional social media. Frequent streamers also have fans group on the live-streaming platform they stream on, or even have chat groups of fans on instant messaging apps like WeChat and QQ. They keep interacting with the viewers in the fans group when they are off their live streams, and form a special community with their fans.

Although several research has explored how people are using mobile live-streaming apps like Periscope [J. Tang 2016] and how people currently engage with live events [Oliver L 2017], few research has explored how people are using live-streaming apps in China and what makes individual live-streaming, such as showroom live-streaming and personal experience sharing on live-streaming, engaging. Individual live streaming in China is more popular with larger number of professional streamers sharing their talents, experience or knowledge to broader audience compared to North America. The reward-based system in live-streaming apps in China further motivates the streamers to continue providing more engaging content and entertains viewers with gamification elements like leaderboard or achievement badges. To better understand this socio-technical phenomenon, we conducted a mixed methods study including a survey (N = 527) and interviews (N = 14). We explored what live streamed content is being viewed and is perceived engaging by viewers, the motivations of people watching and doing live streaming, and what is engaging about watching individual live streaming in China. We referred to several previous

research on user engagement [] for the scale in our survey, and focus on dimensions that are most relevant to viewing individual live streaming: entertainment, sociality, information and interactivity.

Through qualitative analysis of stories told by respondents in our survey, we find a broad array of practices enacting varied degrees and styles of engagement with live streaming in China. Our analysis also provides an account of the motivations and reasons for people engaging in live streaming. We contribute an empirical understanding of how people watch individual live streaming in China and what makes this experience engaging. <some more analysis results to be added>. providing an overview of the opportunities and challenges of conditional anonymous live streaming exposed by this research, to inform the design of future platforms and services that support this model.

This is My Hand: The Influence of Motion Paralysis and Appearance on Avatar Hand Embodiment

Eliciting the sense of presence, a state wherein users feel as though they are physically located in a another world [19], remains a fundamental challenge in virtual reality (VR). One way to create the illusion of presence is to induce feelings of virtual embodiment; that is, the feeling of self-representation by an avatar [28, 27]. Our work explores how various attributes of the avatar hand influences the user's perception of embodiment. According to Kilteni et al., virtual embodiment is "the sense that emerges when B's properties are processed as if they were the properties of one's own biological body" [16]. Kileni's conception of virtual embodiment consists of three components– namely, self-location, where one feels co-located with the avatar; ownership, which refers to a sense of possession for the avatar and self-attribution toward its experienced sensations; and finally, agency, the perception of being the source of the avatar's actions. Some of the more difficult aspects of replicating human hands are its motion and appearance. Comprised of twenty-seven degrees of freedom, hands are one of the most expressive parts of the body, making the task of tracking and animating them an arduous process. Presently, motion tracking technologies face impediments due to their cost, complexity and accuracy; and likewise, photo-realistic avatar hand rendering is stymied by the limitations of computer graphics capabilities. With the growing pervasion of VR, it is imperative to understand how appearance fidelity and mobility might interplay with one another, to guide designers in creating more immersive and engaging environments. This paper reports the results of a user study that sought to investigate the following questions: (1) Given limited tracking capabilities, should hand appearance actually be lower fidelity, to match the expectations set by low mobility? Or can improved visual fidelity compensate for poorer tracking? (2) Does appearance or mobility have a greater bearing on engendering the sense of embodiment?

LitSense: Supporting the Literature Review Workflow

Conducting literature reviews and reading papers is a key activity of academic research, and is a challenging and time-consuming task []. It involves searching for previous research in a field that one wants to contribute to, and understanding what has been done before [Citesense]. There are many challenges to performing a literature review. First, the growing scale and scattered nature of existing academic literature has become challenging to navigate [Papervis]; researchers have a difficult time finding the big picture of their field of research. Second, experience is needed to be able to make connections between different papers and properly manage all the information gathered during their literature review. As a new researcher, this is a difficult thing to learn to do [Papervis, Guided Tour]. Currently, there are tools that exist to assist researchers in performing literature review. Reference management tools such as Mendeley [] and EndNote [] assist in managing references and papers that are gathered while building a literature review. However, they are limited in that they assist only in document management; while they provide the ability to organize papers into folders, filter through them, and extract meta-data automatically from them, they do not provide much more than a way to manage a large collection of papers. Past research in data visualization literature has looked into ways of visually exploring a large collection of papers, and navigating through different relationships between them, such as citation connections [], where papers have been published [], etc. However, literature review is seldom performed this way. Literature review is an iterative process, that commonly involves slowly building a library of information from the bottom up [Paperquest], and continuously making semantic connections within that body of information and knowledge [Citesense]. What has not been fully explored is the literature review workflow – that is, the iterative knowledge-building process that involves the discovery of relevant papers, finding semantic connections, and synthesizing new information, which leads to a broad and thorough understanding of a research question [Citesense]. A few works have attempted to support aspects of the literature review workflow, including the discovery of new related papers [Papervis, Paperquest], and the task of establishing connections between documents and synthesizing new information [CiteSense, SenseMaker]. By focusing on only small parts of the process, however, these works enable only limited insight. Given the multifaceted nature of academic literature, researchers may benefit from a more integrated approach that supports and capitalizes on the various dimensions of literature review. An integrated approach can also support the process of building upon previous knowledge and exploring new ideas discovered through previous searches [SenseMaker, Browsing And BerryPicking]. In this work, we attempt to tackle the gap of performing both exploration and understanding of literature. We sought to understand the workflow that researchers currently follow when building an understanding of previous literature in their field, and to address the limitations of existing tools and research. To do so, we first conducted a formative study, in which we asked researchers about the process that they follow when performing literature review, and the struggles that they face. We then created a proof-ofconcept tool called LitSense, which offers a way for researchers to visually represent their library of papers and information, and assists in bottom-up curation of their library, and making semantic connections as new information is found. We performed an evaluation study to examine its feasibility in supporting this workflow. We then propose a set of design guidelines that should be considered when designing a tool to assist a researcher in the collective process of literature review.

We make the following contributions with this paper: 1. Conducted a formative study, asking researchers about the process that they follow when performing literature review 2. Designed, developed, and evaluated a proof-of-concept tool called LitSense, which offers a way to support

the collective process of literature review 3. Proposed a set of design guidelines for designing a tool to assist a researcher in the collective process of literature review

MedStory: exploration and validation of hybrid text visualization strategies for clinical overview

Clinical practice is a deeply interpretive activity, that seeks to apply the scientific medical knowledge to the needs and circumstances of the individual patient [hunter91,]: occupation, daily habits, religion and personal life history all contribute to determining how patients cope with illness and what treatment options may lead to better quality of life. Clinical text (in the form of medical narratives) plays a fundamental role at accommodating this qualitative complexity, providing the necessary flexibility and expressivity to effectively reason about and document such illness trajectories [greenhalgh1999, johnson2008, hobbs2003]. As such, and despite many attempts to further standardize clinical documentation input, free text remains a common practice for medical documentation [hayrinen2008] and also the component physicians typically spend most of their time on when studying a patient chart [reichert2010]. Text usage, however, presents a number of scalability challenges (larger records that become time-consuming to peruse) and computability challenges (it is difficult to automate tasks around unstructured text) that can severely impair physicians due to the time-sensitive nature of medical practice [christensen2008, ref]. As a result, summarization strategies have been proposed to facilitate access to longitudinal patient charts (and in particular medical text), including natural language processing (NLP) and data visualization solutions [pivovarov2015]. These strategies focus on extracting clinically relevant information from text and then presenting it in a simplified manner – be it a textual summary [hallet2006, hallet2008] or a graphical visualization summary [plaisant1996, plaisant1998, hsu2012]. While these works make important research contributions, there are a number related of technical, practical and conceptual shortcomings. First, there is a trend of employing increasingly complex summarization techniques which are potentially incomplete or error-prone; this raises issues of trust [jensen2016], introduces additional overhead for human verification (be it a data curator or the physician herself) [hsu2012], and hinders effective integration of such systems into real clinical practice [pivovarov2015]. Second, information extraction has typically focused on "strictly medical" details that does not encompass the qualitative richness present in medical records. Contextual aspects are often lost in the process, such as family situation, patient occupation, how the patient is coping with the disease, and hints on how physicians perceive those patients [ref]. We argue that a fundamentally different approach to clinical text needs to be considered, if we are to address the above issues in a practical manner. In this paper, we explore a more "text-centered" visualization approach for clinical overview, which preserves the more subjective facets of the medical text and supports user trust via extensive linking to original content. We establish the foundations for this research direction via a user centered process, including a qualitative formative study grounded by relevant literature in medicine and HCI, followed by the design and evaluation of MedStory, a prototype we built to validate the concept. Our contributions are as follows:

1. A deeper understanding of use of text for clinical overview, informed by qualitative studies and contextualized by existing research (background and related work);
2. Design and evaluation of a data visualization prototype for clinical summarization and overview3. A collection of design guidelines for clinical summarization visualization systems

Exploration of the Categorization and Audio Features of Think-Aloud Tests in Revealing Usability Problems

Think-aloud testing (TA) is a widely used usability evaluation method and valued by usability evaluators [7][13]. While using think-aloud, participants verbalize their thoughts out loud when working on the given task. Thus, TA provides an access to users' thoughts which cannot easily be found with observations alone [3] and thus can be useful in identifying potential usability problems.

The analysis of TA tests is necessary for usability evaluators to identify evidence to bolster their usability claims [8] and for the audience of usability reports to better understand the rationale behind a specific design recommendation [14]. Unfortunately, despite its popularity in usability testing, think-aloud tests are often not getting sufficiently analyzed [8].

Although there has been a body of research examining the validity of TA in practice (e.g., [1][5][6][9]) and the effect of TA in workloads (e.g., [12]), little has been focused on analyzing the content of TA. Lynn studied the verbalization of TA and identified four categories from the content of TA, which are reading, procedure, observation and explanation [2]. These four categories were further examined and confirmed in Elling et al.'s study [3]. With eye-tracking data collected during the TA sessions, both works suggested that silence and verbal fillers (e.g., "um", "let's see", "ah") occurred during participants' verbalization revealed useful information about participants' cognitive load [2][3]. Although they categorized the content of verbalization during TA, they did not study how these categorization, silence, verbal fillers might be related to potential usability problems. Our research builds upon their works and explores how these categories, silence and verbal fillers reveal useful information about potential usability problems.

We are also inspired by previous research that suggests the way in which people speak can also reveal a great of information, such as their feelings and mood (e.g., [10]). Thus, besides silence and verbal fillers, we also explore four more audio features (i.e., sentiment, speech rate, loudness, and pitch) that might be affected by usability problems. The sentiment is related to the users' feelings. For example, when users are confused or frustrated in using devices or following the instructions, their verbalization may exhibit negative sentiments. Speech rate is relevant because when users are engaging in cognitively demanding parts of the task, they may inevitably slow down their verbalization or even fall into silence to focus better on the immediate task [4][11]. Similarly, the loudness of their verbalization may also drop when they prioritize their cognitive energy on the tasks [2][3]. Pitch may become higher when users are excited or surprise.

In this research, we aim to explore two research questions related to the content of think aloud tests:

- How are the categories of the content of TA related to potential usability problems?
- How do audio features of the content of TA signal useful information about potential usability problems?

To explore the research questions, we first recruited participants to use commercial products by following their original product instructions. We asked them to think aloud while working on the given tasks and audio recorded the entire TA sessions. We followed Ericsson and Simon's guidelines to ensure the validity of their verbalization [4].

Next, we recruited participants who had usability testing knowledge as usability evaluators to explore the audio recordings of TA sessions. We asked usability evaluators to first segment and categorize the content of TA using the four categories, and then identified potential usability issues and explored different audio features to support their usability claims. To help evaluators to complete these tasks, we built tools for categorizing audio content and visualizing the categories and audio features (i.e., silence, verbal fillers, sentiment, speech rate, loudness, pitch) of the content of TA.

We analyzed their usability reports and found that: 1) not all categories are equally related to the usability problems. 2) observation reveals most of the usability problems followed by explanation; 3) the back and forth pattern between reading and observation may indicate usability problems in the instructions; 4) all audio features are used in supporting the usability claims with the sentiment information being rated the most useful one.

Categorizing the audio content into four categories is time-consuming. To alleviate the burden of categorization of audio content, we segmented audio recordings into smaller segments and used crowd-sourcing effort to complete the categorization. We recruited another group of evaluators, who were given the categories information along with the same set of audio features without needing to categorizing the audio recordings themselves. Results show that they discovered the roughly equal amount of usability problems and used all audio features to back up their identified usability problems.

In summary, we make the following three contributions in this research:

- Identification of the correlation between the category information of TA sessions and potential usability problems, which provides insights for usability evaluators to prioritize their attentions among different categories;
- Demonstration of the role of audio features (i.e., silence, verbal fillers, sentiment, speech rate, loudness, pitch) in signaling potential usability problems, which provides heuristics for usability evaluators to locate patterns for usability problems;
- Demonstration of the equal effectiveness in categorizing the content of TA audio recordings by crowd workers and usability evaluators, which suggests the possibility of shifting the workload of categorization to crowd workers who are not necessarily trained in usability testing.

Thor's Hammer: Ungrounded Force Feedback Device using Propeller Propulsion

Feeling the force plays an important role of perceiving physical properties of an object or the environment. To enable computers to create a more realistic virtual representation of the physical environment, there has been an intensive research on developing force feedback devices.

Majority of the force feedback devices require a grounded structure to move mechanical joints [mj, phantom], to pull strings attached to a handheld device [SPIDAR], or to place air jet actuators [airjet]. The need of a grounded structure, however, limits the interaction area to where close to the system is located.

Creating a large force without a grounding structure has been a challenging problem. Researchers sought ways to create ungrounded force feedback by making the reference frame to be worn on the back [HaptiGEAR, SPIDARW], by using a haptic illusion made by an asynchronous oscillation [oscs, traxion], or by using electrical muscle stimulation [Pfeiffer,

Lopez]. However, these methods have limitations that a user has to wear a very large device [HaptiGEAR], that the illusion could only create small virtual force (<0.3N) and the effect decreases when used for multidimensional forces [traxion], and that a user has to attach electrodes to the body and run a calibration prior to use.

In this paper, we present Thor's Hammer, a new ungrounded 3-DOF force feedback device that uses propeller propulsion. Thor's Hammer has six propellers installed on the six facets of a cubical structure (Figure 1b) and creates thrusts in 3-DOF with two propellers work as a pair creating a bi-directional force in a single axis. With an optical tracking system and an inertial sensor embedded in the handle, the device creates a 3-DOF force invariant to its orientation. Different from other ungrounded force feedback methods, propeller propulsion can create a large 3-DOF force almost anywhere without the need of attaching electrodes, per-user calibration, or wearing a large structure on the back.

Our technical evaluation showed that the Thor's Hammer could create a 4N force in three dimensions with a precise strength control (<0.2N SD), and could create a force in an arbitrary direction with less than 10 degrees of average orientation errors. We also conducted two user studies to determine 1) how well users recognize the force that Thor's Hammer generates in regards to its orientation and strength and 2) user experience of the force feedback when used in Virtual Reality (VR) environment. The results showed that the participants could recognize the force direction with 90% accuracy and the force strength with less than 10% of average error. Participants also reported that the force feedback makes the VR environment more realistic, immersive, and engaging (results to be updated after the study).

In summary, our paper makes the following contributions:

- Design and the implementation of Thor's Hammer, an ungrounded force feedback device using propeller propulsion;
- Force characteristics analysis of Thor's Hammer;
- User evaluation on the recognition of the force feedback generated by Thor's Hammer;
- Qualitative user feedback on Thor's Hammer used in VR environment.