Research Statement

Anthony TANG
School of Computing and Information Systems, Singapore Management University
Tel: (65) 6826-4959; Email: tonyt@smu.edu.sg
1 (Day) 12 (Month) 2023 (Year)

Background

Computing is a pervasive part of our everyday lives. Yet, we still encounter challenges using these technologies, ranging from small usability annoyances such has having to pull out one's phone to complete a multi-factor authentication, to larger usability disruptions such as forgetting one's password, to complete failures such as encountering unhelpful failure messages that do not provide effective remedy to complete tasks. While some of these challenges can be attributed to poor programming (e.g. programming "bugs"), many of these challenges stem from a fundamental misunderstanding of people's needs, goals, and tasks.

My research in human-computer interaction (HCI) aims to address this shortcoming by focusing on people's needs, goals and tasks in their use of computing technologies. Developing this empathy helps us to envision, design, and refine computing interfaces that people use. With my students, we build new computing interfaces—usually to give people new capabilities to interact with their data or their world, or to interact with one another. We then evaluate our designs, by observing people's use of these technologies, characterizing successes and failures as descriptive theory. These theories are then used to inform the design of interfaces for future computing technologies.

This approach is critical in the face of new computing technologies—whether they are new interface capabilities (e.g. multi-touch interface technologies) or new underlying computing capabilities (e.g. large-language models)—which introduce new capabilities, and new challenges. Our research creates theoretical understanding of people's needs and goals. We express this understanding through functional prototypes of working systems or through toolkits that can be used by others. Finally, through our evaluations, we create descriptive theories that allow us to anticipate the kinds of mental models that people bring to their use of technologies. Together, this helps us to direct and shape the direction of computer interface design.

In practice, we use specific application and domain areas to develop these theories and ideas. Over the past two decades, we have explored a range of domains, including collaborative visual analytics tools, remote collaboration tools, augmented and virtual reality, electronic fashion, and most recently human-AI collaboration.

Research Areas

My current research focuses on three problem areas: collaborative systems; human-AI collaboration in the form of conversational agents and robots, and AR/VR systems. I briefly describe the work my research team and I have been doing in these spaces, and outline ongoing and exciting future research questions that we aim to address.

Collaborative Systems. How can we design tools to help people work together effectively? What characterizes effective work? And, how do new computing technologies create opportunities for new kinds of collaboration?

The COVID-19 pandemic created a situation that forced many of us to work and interact with one another at a distance. This allowed us to see how mobile video and AR/VR capabilities created new opportunities for everyday collaboration, as well as creating new challenges. For instance, how can I instruct my father how to cook a dish that we like [1], [2]? How can my son play with his friends over video [3], [4]? How can we have shared physical experiences when one or people cannot join [5], [6]? How can mobile video be used to help wilderness search-and-rescue workers collaborate with one another [7], [8]?

Our most recent work in this area explores how 360 videos can be consumed and explored by a group of people [9]. These videos are captured by low cost omnidirectional cameras that capture the entire environment around the camera, where a viewer can smoothly swivel around as if they were turning their heads. We designed a prototype system Tourgether360, which rethinks the interaction model with such videos, which typically rely on a time-based navigation model. In Tourgether360, we recover a trajectory through a 3D scene, and provide users with the ability to see one another in the video space. This creates a videogame-like experience that participants generally found more straightforward to navigate, and to collaborate with one another.

Our current work explores how we can effectively navigate through immersive videos captured through multi-camera systems (or using VR systems), where an avatar can help guide a person to learn about the environment, or coach them through the completion of a task in the environment.

Human-AI Collaboration. How can we use AI to superpower people's activities? How can we smooth interaction between humans and AI agents with an understanding of how people interact with one another? How can we make it easier for people to specify their intentions to computing agents? How can we make it easier for computing agents to clarify their understanding of people's intentions and needs?

Large-language models are creating new possibilities to create agents that are plausibly capable of responding to people's queries and needs. Chat-GPT, for instance, has shown us that general large language models are able to maintain a plausible multi-turn conversation with humans. In our work, we are interested in harnessing this type of capability for interaction—potentially conversational user interaction. At the same time, we are fundamentally interested in how people use body language in relation to spoken interaction to interact with one another. For instance, how can swarm robots use a user's body language to infer intent, and to smoothly navigate the environment? Similarly, how can body language be used in concert with speech to drive collaborative behaviours of a robot [10]–[12]?

Our most recent work in this space explores the design of a conversational programming agent that helps citizen programmers (people without formal training) to create interactive VR experiences. Here, the user can specify an interactive world through conversation, and the system responds by creating this world (under the hood, it relies on a large language model to create new computer code that can realize the user's intentions), which the user can then experience and further refine through speech. Our main contributions in this work are an understanding of how people specify and articulate their ideas, a fully realized system that democratizes the creation of virtual worlds.

Our current work furthers this exploration by exploring new techniques for resolving ambiguities, where we will explore both AI-based techniques and user interface techniques to help people to refine their intentions (and in practice, to help them to identify their intentions!). We are also exploring how these technologies can be applied to AR contexts. To this effect, we are exploring also conversational smart assistants that can aid people in performing tasks in the real world—we do this by supplying the agent an understanding of what the user sees, and what they are doing.

<u>Augmented/Virtual Reality Systems</u>. How can we enhance the tangible aspects of people's experiences in VR? How can we make it easier to create AR/VR and immersive experiences? How can we use AR/VR technologies to support collaborative activities?

AR technologies create visual experiences that are anchored and tracked atop the physical world, while VR technologies focus on an all-encompassing immersive experience. AR create new opportunities for how we can augment and interact with our real world—perhaps by supplying additional information, or supplying information about things that we otherwise could not see. VR also creates new possibilities for interacting with data and worlds that do not exist in reality. How can we ease the development of interaction techniques for such environments [13]? How can they create opportunities for exploring datasets with complex inter-relationships [14]? What kinds of interaction problems does wearing headsets create for people [15]? And, how can collaboration be viewed through the lens of AR technologies?

Our most recent series of works explores how we can create tangible (i.e. haptic) experiences for those wearing VR headsets. That is, how can we make it so people wearing headsets can physically touch and manipulate objects that they see in VR *without* having to give them an exact replica of those objects [16], [17]? We explored the use of object proxies, and the extent to which these need to mimic the physical properties of the object being represented. We have found that there can be significant discrepancies [18], where people can have high fidelity haptic experiences even though the proxy may not share the exact same properties [19], [20]. Further, we have found ways to create and use visuo-haptic illusions to create rich haptic experiences for VR users with comparatively impoverished physical proxies. This means that designers do not need to create the virtual world with physical objects order for people to have rich VR experiences with touch.

Selected Publications and Outputs

[1] O. Fakourfar, K. Ta, R. Tang, S. Bateman, and A. Tang, "Stabilized Annotations for Mobile Remote Assistance," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, in CHI '16. New York, NY, USA: Association for Computing Machinery, May 2016, pp. 1548–1560. doi: 10.1145/2858036.2858171.

[2] B. Jones, A. Witcraft, S. Bateman, C. Neustaedter, and A. Tang, “Mechanics of Camera Work in Mobile Video Collaboration,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, in CHI ’15. New York, NY, USA: Association for Computing Machinery, Apr. 2015, pp. 957–966. doi: 10.1145/2702123.2702345.

[3] M. Cohen, K. R. Dillman, H. MacLeod, S. Hunter, and A. Tang, “OneSpace: shared visual scenes for active freeplay,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI ’14. New York, NY, USA: Association for Computing Machinery, Apr. 2014, pp. 2177–2180. doi: 10.1145/2556288.2557117.

[4] S. E. Hunter, P. Maes, A. Tang, K. M. Inkpen, and S. M. Hessey, “WaaZam! supporting creative play at a distance in customized video environments,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI ’14. New York, NY, USA: Association for Computing Machinery, Apr. 2014, pp. 1197–1206. doi: 10.1145/2556288.2557382.

[5] J. Procyk, C. Neustaedter, C. Pang, A. Tang, and T. K. Judge, “Exploring video streaming in public settings: shared geocaching over distance using mobile video chat,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI ’14. New York, NY, USA: Association for Computing Machinery, Apr. 2014, pp. 2163–2172. doi: 10.1145/2556288.2557198.

[6] H. Shakeri, S. Singhal, R. Pan, C. Neustaedter, and A. Tang, “Escaping Together: The Design and Evaluation of a Distributed Real-Life Escape Room,” in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, in CHI PLAY ’17. New York, NY, USA: Association for Computing Machinery, Oct. 2017, pp. 115–128. doi: 10.1145/3116595.3116601.

[7] B. Jones, A. Tang, and C. Neustaedter, “Remote Communication in Wilderness Search and Rescue: Implications for the Design of Emergency Distributed-Collaboration Tools for Network-Sparse Environments,” *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. GROUP, p. 10:1-10:26, Jan. 2020, doi: 10.1145/3375190.

[8] B. Jones, A. Tang, and C. Neustaedter, “RescueCASTR: Exploring Photos and Live Streaming to Support Contextual Awareness in the Wilderness Search and Rescue Command Post,” *Proc. ACM Hum.-Comput. Interact.*, vol. 6, no. CSCW1, p. 113:1-113:32, Apr. 2022, doi: 10.1145/3512960.

[9] K. Kumar, L. Poretski, J. Li, and A. Tang, “Tourgether360: Collaborative Exploration of 360° Videos using Pseudo-Spatial Navigation,” *Proc. ACM Hum.-Comput. Interact.*, vol. 6, no. CSCW2, p. 546:1-546:27, Nov. 2022, doi: 10.1145/3555604.

[10] J. Li *et al.*, “Stargazer: An Interactive Camera Robot for Capturing How-To Videos Based on Subtle Instructor Cues,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–16. doi: 10.1145/3544548.3580896.

[11] K. Mahadevan, M. Sousa, A. Tang, and T. Grossman, “‘Grip-that-there’: An Investigation of Explicit and Implicit Task Allocation Techniques for Human-Robot Collaboration,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan: ACM, May 2021, pp. 1–14. doi: 10.1145/3411764.3445355.

[12] K. Mahadevan, Y. Chen, M. Cakmak, A. Tang, and T. Grossman, “Mimic: In-Situ Recording and Re-Use of Demonstrations to Support Robot

Teleoperation," in *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, in UIST '22. New York, NY, USA: Association for Computing Machinery, Oct. 2022, pp. 1–13. doi: 10.1145/3526113.3545639.
[13] M. Weigel, S. Boring, J. Steimle, N. Marquardt, S. Greenberg, and A. Tang, "ProjectorKit: easing rapid prototyping of interactive applications for mobile projectors," in *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, in MobileHCI '13. New York, NY, USA: Association for Computing Machinery, Aug. 2013, pp. 247–250. doi: 10.1145/2493190.2493242.
[14] S. Ichikawa, K. Takashima, A. Tang, and Y. Kitamura, "VR safari park: a concept-based world building interface using blocks and world tree," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, in VRST '18. New York, NY, USA: Association for Computing Machinery, Nov. 2018, pp. 1–5. doi: 10.1145/3281505.3281517.
[15] Y. Kudo, A. Tang, K. Fujita, I. Endo, K. Takashima, and Y. Kitamura, "Towards Balancing VR Immersion and Bystander Awareness," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. ISS, p. 484:1-484:22, Nov. 2021, doi: 10.1145/3486950.
[16] M. Feick, T. Mok, A. Tang, L. Oehlberg, and E. Sharlin, "Perspective on and Re-orientation of Physical Proxies in Object-Focused Remote Collaboration," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, in CHI '18. New York, NY, USA: Association for Computing Machinery, Apr. 2018, pp. 1–13. doi: 10.1145/3173574.3173855.
[17] M. Feick, S. Bateman, A. Tang, A. Miede, and N. Marquardt, "Tangi: Tangible Proxies For Embodied Object Exploration And Manipulation In Virtual Reality," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Nov. 2020, pp. 195–206. doi: 10.1109/ISMAR50242.2020.00042.
[18] M. Feick, K. P. Regitz, A. Tang, T. Jungbluth, M. Rekrut, and A. Krüger, "Investigating Noticeable Hand Redirection in Virtual Reality using Physiological and Interaction Data," in *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, Mar. 2023, pp. 194–204. doi: 10.1109/VR55154.2023.00035.
[19] M. Feick, C. Biyikli, K. Gani, A. Wittig, A. Tang, and A. Krüger, "VoxelHap: A Toolkit for Constructing Proxies Providing Tactile and Kinesthetic Haptic Feedback in Virtual Reality," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, in UIST '23. New York, NY, USA: Association for Computing Machinery, Oct. 2023, pp. 1–13. doi: 10.1145/3586183.3606722.
[20] M. Feick, A. Zenner, O. Ariza, A. Tang, C. Biyikli, and A. Krüger, "Turn-It-Up: Rendering Resistance for Knobs in Virtual Reality through Undetectable Pseudo-Haptics," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, in UIST '23. New York, NY, USA: Association for Computing Machinery, Oct. 2023, pp. 1–10. doi: 10.1145/3586183.3606787.