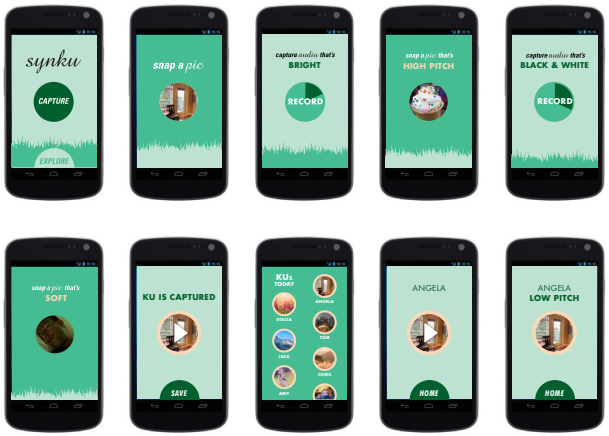
synku

SYNESTHETIC HAIKU AS A DESIGN PROBE

HOW DO YOU HEAR VISUALS?



HOW DO YOU SEE SOUNDS?

WHAT IS SYNKU

SynKu first processes the recorded images and audio to produce categorizations along with various spectra (saturation, lightness, and hue for images; volume and pitch for sound). Once processed, SynKu prompts the user to collect a different type of media according to that new category (e.g., "capture a sound that is bright" or "snap a pic that's slow"). SynKu interweaves the algorithmic processing of five such media recording activities using ten media attributes (e.g. light / dark, loud / soft) with users' cross-sensory interpretation of those attributes. People can then share the resulting media assemblage as a short video narrative (or "Ku"). By creating and sharing Kus, they produce new sensory objects interpretable by people and software. Here we discuss how this research enables us to examine the digital expression of sensory perception.

SCENARIO OF USE

Angela woke up to a sunny morning in Seattle. The light streamed through her windows and illuminated her room, inspiring her to open the SynKu app and snap a photo of the morning sunshine. The app immediately processed the image, categorized it as bright, and prompted Angela to capture a 'bright' audio clip. Angela decided to walk to the park. Thrilled to play on the swings at the playground, the children at the park struck Angela as particularly bright, and she opened SynKu to record the sound. Processing the sound and classifying the recording as high-pitched, SynKu then prompted Angela to capture a 'high pitch' image. After spending a long day out, Angela decided to stop for some ice cream before heading home. She chose chocolate and vanilla, squealing as she finished her cones. Finding this an appropriate 'high pitch' moment, Angela took a photo of the ice cream. After processing the image, SynKu generated a new prompt: "record audio that's 'black and white'." Angela waited until her children were in bed before reopening SynKu. They like to get their dad to read a bedtime story to them, and 'The Three Little Pigs' is their favorite. Angela decided to record the storytelling session on SynKu. The recording was then processed to provide a final prompt: "snap a pic that's 'soft'." After a long day, Angela ends her Ku with a photo of her partner falling asleep, feeling thankful to have her family by her side. Angela built a sensory object in a Ku, ready to share and discuss with her family and friends.

HOW IT WORKS

The SynKu application is built on the Android platform (supporting Android OS 4.0-5.0). The application uses Parse API for backend operations such as login, storing and retrieving audio and image files. The application has image and audio processing modules for generating prompts.

The image-processing module generates prompts for recording audio using a histogram based heuristic - light or dark based on lightness, bright or dull based on saturation and colorful or black and white based on hue. To generate the prompt, the algorithm first down-samples the captured image and calculates the weighted average of three color channels, hue, saturation and value, using separate histograms. Next, the deviation is calculated for each of the channels from a threshold set for each of them, 5 for Hue and 127 for Saturation and Value. The deviation is then normalized to a 0-1 scale for comparison. The system then selects the channel with the maximum deviation and generates a corresponding prompt.

The audio processing module generates prompts for capturing images: soft or loud based on maximum amplitude and high-pitch or low-pitch based on frequency. The algorithm first processes the PCM (pulse code modulation) audio data to find the maximum amplitude in dB. Then the PCM data is processed for average frequency using a Fast Fourier Transform based YIN pitch tracking algorithm (a frequency estimator based on autocorrelation) from the Tarsos DSP library. The thresholds used for these classifications are 500Hz for pitch and 90 dB for loudness.

Our software saves the audio and image files along with the prompts on the Parse servers for subsequent engagement by other people using the SynKu app.

FUTURE RESEARCH

Research in Science and Technology Studies (STS) has begun to probe the limitations of representing bodily knowledge through computing. In her ethnographic examination of surgical simulation, STS scholar Rachel Prentice [12] describes the process of 'mutual articulation' or the shaping of bodies by machines and machines by bodies. Both the bodies of the surgeons and those of the patients on which they practiced became present and knowable through the simulator. Building on Latour's [11] notion of 'articulation' through which bodies and knowledge get mutually enacted, she writes:

[…] *researchers pooled various disciplinary knowledges of anatomy, surgery, computation, education, cognition and engineering to develop an object (a model, a software program, a device) that has a particular relationship to the user's body. At each point, then, researchers are working to create interpretations of what human bodies are in relation to these objects; that is, to articulate the body in new ways* [12:862].

Body objects, according to Prentice, emerge as representations of human bodies that inhabit computers for use by people and computers. Drawing on this category, our design team will use SynKu to investigate what it would mean to build sensory objects, or algorithmic representations of sensory experiences for interpretation and use by people and software. For interaction design, sensory objects exemplify a type of 'strong concept' [9], an intermediate level of knowledge between theories and particular instances of design. Given that self-tracking activities have become increasingly shared [2], we wonder what form such sensory objects may take within everyday contexts of communication. How might we use sensory objects to articulate the body in new ways? What are the consequences of these new forms of articulation for understanding and engagement? We use design research to speculate on these questions: exploring and intermingling how people see through the eyes of a vision algorithm and hear through the ears of sound processing.

In the months ahead, we plan to expand our study of SynKu in two ways. Empirically, we plan to launch a study of our application within a local art community in Seattle, Washington in partnership with Seattle's Design Festival "Design in Public," a two-week affair that features public demonstration and celebration of the Seattle design community. Theoretically, we will use the notion of sensory objects - or the processing, interpretation and use of sensory experience - to investigate new forms of articulation in HCI.

HCDE Human Centered Design & Engineering
University of Washington

ANKUR AGRAWAL ANKURAG@UW.EDU
WENVI HIDAYAT WENVI@UW.EDU
MARK STAMNES MSTAMNES@UW.EDU
ARAVIND RAVI ARAVI@UW.EDU
MEISHEN YIN MEISHENY@UW.EDU
DANIELA ROSNER DKROSNER@UW.EDU

depts.washington.edu/tatlab/