PIXEE: Pictures, Interaction and Emotional Expression

Margaret E. Morris
Intel Labs
Intel Corporation
Hillsboro, OR 97124 USA
margaret.morris@intel.com

Murad Al Haj
Centre de Visió per Computador
Universitat Autònoma de Barcelona
Bellaterra, Barcelona, 08193 Spain
malhaj@cvc.uab.es

Carl S. Marshall
Intel Labs
Intel Corporation
Hillsboro, OR 97124 USA
carl.s.marshall@intel.com

James S. MacDougall
University of Victoria
3800 Finnerty Road
Victoria, BC V8N 1M5, Canada
jamiemac@uvic.ca

Mira Calix
Saxmundham, Suffolk
P.O. Box 50
IP171YZ UK
miracalix@mac.com

Douglas M. Carmean
Intel Labs
Intel Corporation
Hillsboro, OR 97124 USA
douglas.m.carmean@intel.com

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.

Abstract
An interactive system, PIXEE, was developed to promote greater emotional expression in image-based social media. Images shared on social media were projected onto a large interactive display at public events. A multimodal interface displayed the sentiment analysis of images and invited viewers to express their emotional responses. Viewers could adjust the emotional classification and thereby change the color and sound associated with a picture, and experiment with emotion-based composition. An interdisciplinary team deployed this system around the world to explore new ways for technology to catalyze emotional connectedness. This paper describes the system, design iterations, and observations about how people used it for self-expression and connection.

Author Keywords
Affect; social media; image sharing; sentiment analysis; emotion; interpersonal connectedness; interactive displays

ACM Classification Keywords
H.5.m. Information interfaces and presentation, miscellaneous.

General Terms
Design, Experimentation.

Introduction

Today's social media offers radical online networking but few tools have been explicitly designed to promote emotional connectedness. Many have argued that the demands of documenting and sharing experiences in real time can dilute engagement in face-to-face interaction, and contribute to isolation. Online dialogue about a shared image or status update is often constrained to an approval or a quip rather than an exploration of thoughts and feelings. As Turkle notes in Alone Together [14], the simple, often competitive communication practices on Facebook and other social media may preclude intimate communication.

Social media, by retaining social ties and offering channels of lightweight communication, addresses relational needs that are critical for physical and emotional wellbeing [6]. Interpersonal connectedness depends on communication nuances, however, that only some are capable of exercising within the constraints of today's social media systems. Specifically, connectedness requires emotional intelligence – an awareness of one's own states and the ability to empathize with others [4, 12]. Indeed, those who share content and respond substantively to others' updates report greater wellbeing and social capital than those who merely view or "like" content [3]. It may be that these individuals exercise emotional intelligence by sharing emotionally rich content and responding empathically to others. Another important element of emotional intelligence, not yet explored in social media capabilities or research on their use, is the capacity for experiencing positive and negative emotions at once.

In this exploratory project, a system to enable emotional rich expression and interpersonal connectedness in social media was developed. One of the fundamental means of promoting interpersonal connectedness was shifting attention from the mobile phone on which images were captured to a large public display surface where people could interact with the images collectively. In addition, the system included an emotional layer for image sharing. This emotional layer consisted of sentiment analysis of images and self-expression by viewers. The emotional classification from the sentiment analysis and self-expression was reflected in the colors and sounds associated with each picture on the display. The ultimate goal of this classification was to challenge the participants to change the existing classification according to their own emotional responses to the images. In many cases, participants re-classified images that were originally shared at a different event, often in a different country.

This project advanced previous work in affective computing, social media and interactive displays in several ways. First, sentiment analysis was applied to enrich the user experience rather than to analyze population trends. Second, the sentiment classification was intended as a probe for emotional exchange. This effort extends research by Sengers and colleagues [2], in which affect is measured and represented in order to invite reflection. In addition, the system was capable of tracking the ongoing reclassification of images by viewers. This forum for emotional exchange extends the experience offered by the popular tool "We feel fine" [5] which offers a compelling, playful visualization of the emotional context of images. Lastly, this work merges affective computing with large interactive displays, suggesting a new direction for personalizing public computing and creating true social settings for social media exchange.

The system components

Below is a description of capabilities and components of the interactive system.

Image sharing

This system enabled participants to submit, browse, and express emotional responses to media images on a large interactive display surface. The display was created with multiple, high power projectors, projecting on to one or more walls. The display size varied by venue, but was typically 5m x 2m. The interactive surface displayed content that had been shared on social media in a historical timeline. Participants could share new content by submitting images to applications including Twitter, Instagram and Weibo (Chinese social media) using a hashtag to reference the event. The system was notified as new social media content was posted and scraped the content to be displayed on the interactive surface. Images that were submitted, typically via participants' mobile phones during an event, appeared instantaneously on the display. At any one time the display showed about 70 photos, with several thousand photos available in the image timeline. The display retained the caption text and user name in each picture. The display is shown in Figure 1.

Interaction overview

The interaction design was modeled after modern smartphones to provide intuitive and a natural extension of the participant's personal device. In addition to submitting photos, participants could interact with the display in a variety of ways. The surface responded to three general types of gestures. First, swiping across the display allowed perusal of archived content. Participants could browse through images submitted over the course of seven months at a variety of cultural events in cities around the world. Second, briefly touching a photo enlarged that photo and others with similar affect classification. The highlighting of images with similar emotion, shown in Figure 2, was intended to visualize the association that is sometimes called emotional contagion. Viewers could opt to see images of a particular mood, and thus experiment with exposing themselves to different types of emotional influence. Third, by pressing an image for several seconds, participants could change its emotional classification and express their emotional response. This editing of emotional classification is elaborated in the Emotional Interface section below.

Figure 1. Interactive display of Instagram photos. The color each image's frame indicates its emotional classification.

Arbitrary interactive surfaces

A key feature of this system was its ability to transform any surface into an immersive, interactive computing system. The system was fully scalable, allowing multiple depth cameras to support participant input on large, irregular surfaces. In addition, the system used

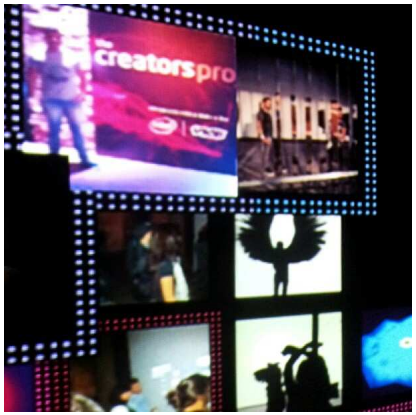

Figure 2. Emotional association among images.

projection mapping with support for multiple projectors, providing high resolution displays on arbitrary surfaces. The specific installation varied according to the physical constraints of the venue, with a typical interactive surface covering a wall approximately 5m by 2m. The venues were typically large warehouses repurposed for art and music events. The projection surfaces at these venues were concrete, brick or wooden walls. In many cases, the walls were neither flat, nor completely smooth, challenging the technology for interactivity.

The system consisted of pre-production hardware supplemented with commodity components when available. The software framework was fully custom, written in C++, Go, Python, C# and Objective-C. There were many custom computer vision algorithms that were developed to provide a natural, intuitive interface using multiple depth cameras. Beyond the simple challenges of distinguishing fingers, limbs and resulting gestures, the ability to disambiguate multiple depth cameras with irregular interference patterns is one of the many novel aspects of the system. This system advanced previous approaches for enabling touch interaction with depth cameras [15].

Sentiment analysis

As images were scraped from the various social media sites, the sentiment of captions was analyzed. The sentiment analysis software developed for this project classified images according to the Circumplex Model of Emotion [11]. That is, images were classified along the two dimensions of arousal (low to high) and valence (negative to positive). New categories of high and low arousal were added to the LIWC [9] software, in particular the positive/negative emotion categories. Second, the system screened for exact matches between caption text and the sixteen terms mapped out in Russell's Circumplex Model and validated in the development of Pollak's Photographic Affect Meter (PAM) [10]. Third, synonyms of these 16 terms were added: a set of contemporary English, Chinese, Korean and Brazilian terms frequently used in social media. The last addition was a set of commonly used emoticons, a corpus which expanded rapidly in Korea due to the extensive and diverse emoticons popular among different cultural niches. Lastly, the system incorporated expressive colloquialisms that were unique to each culture with a corresponding mapping on to Russell's Circumplex Model.

Emotional Interface

A visual frame around each of the 70 photos on the display was colored to reflect the emotion of the picture. The frame color indicated either the preliminary sentiment analysis of the caption text or reclassification by a viewer. Colors were based on common associations in Western culture such as red with anger,

pink with excitement, blue with calmness and grey with sadness. These color selections are aligned with cardiovascular models, in which red is associated with autonomic heating and blue with cooling, and research on emotional responses to color [7, 1]. Subsequent installations in China and Korea allowed the team to explore Eastern cultural differences in emotion-color associations. These cultural and individual variations may be addressed in the next phase of this work on expressive interfaces.

An emotion expression interface allowed people to adjust the classification from sentiment analysis and describe how pictures made them feel. This interface was based on the two dimensional Circumplex Model of Emotion [11] that was previously developed into a touch screen mobile application for experience sampling of mood [8]. There were 16 cells in this two dimensional grid, each associated with a color and an emotion term. As viewers moved an icon around the space, an emotion term appeared in the center and the associated color filled the interface and photo frame.

Music: Emotional Composition

Musical capabilities were added to the system to invite emotional engagement and experimentation with new forms of emotion based expression. The musical design was done in collaboration with a professional composer. Musical motifs were associated with each of the 16 areas of the affect grid, i.e. the Circumplex model. The musical experience varied depending on whether one was interacting with the timeline of 70 images or expressing a response to a particular image.

As one touched a photo on the timeline of images, a brief percussive sound played that corresponded with the image's emotional classification and color frame. The classification reflected either sentiment analysis or the adjustment to the classification by a viewer. Those images that were not classified by the software or by a viewer were tagged with a neutral sound. By touching a series of photos, one could compose a tune that reflected the emotional pattern of those images.

As one moved across the various areas of the emotion interface associated with a particular image, different melodies played. The 16 melodies worked in four part harmonies, with one four-part harmony per quadrant. This allowed participants to hear a seamless musical flow while moving a finger across the map. The intent was to heighten the emotional response to the image and re-enforce viewers' intuitive understanding of the Circumplex Model of Emotion.

The motifs were based on a combination of early music theory, composer intuition and iterative testing. Music theory by Shubart [13] and others associates major keys with joy, hope, excitement and moon keys to sadness, and anxiety. As an example of how this theory was applied in the current system, F Major, was associated with the calm (low arousal, positive affect) region of the emotional interface. Other rough design guidelines came from the principles that factors such as speed, rhythm, texture and timbre affect emotional response such that quick, loud, bouncy rhythms sound happier than slow quiet ones. Due to the inconclusive nature of research on music-mood relationships, these principles were used as starting points. Ultimately, the composer's intuition, understanding of the interface, and testing with other members of the team were critical for refining the set of motifs and incorporating them into the system.

An overview of the system described in this section can be seen in Figure 3.

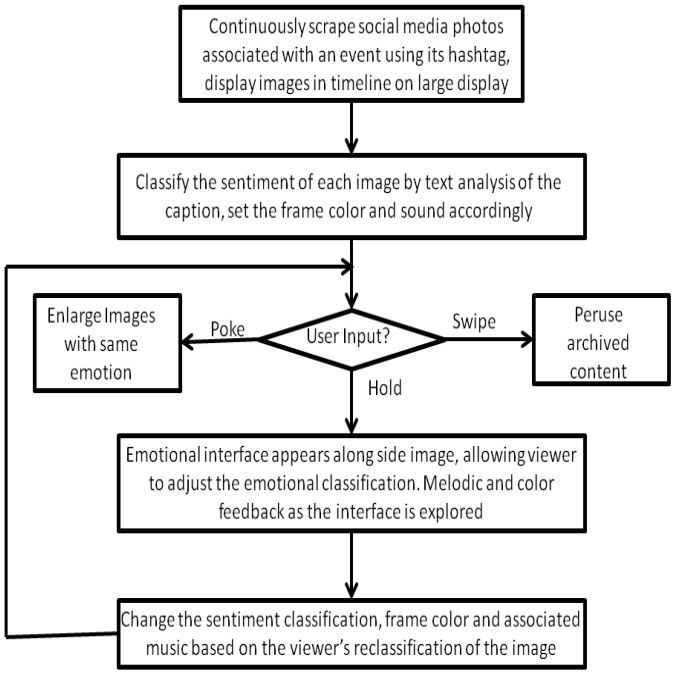

Figure 3. Flow chart of system

Design iteration

The system was tested and refined over the course of seven months as it was installed at nine events in six countries. Five of these events were art-music-technology festivals, and four were technology showcases. Observations of usage during these events drove design iterations.

A major challenge addressed through design iteration was conveying the emotional capabilities of the system without creating the experience of a psychological survey. Following is a short summary of the iterations related to this challenge. The first interface for emotional expression was completely abstract, without axis labels or text within the quadrants. In this initial interface, the selection of an emotional state occurred by touching different areas of an unlabeled affect grid. The grid and picture frame changed colors according to the selection. There was considerable experimentation with color pallets before settling on ones that aligned with the aesthetics of the system and with past research. Prior to deploying this system in different countries, particularly China, experts were consulted on cultural variations in color-emotion associations.. The complexities of individual and cultural differences led the team to retain the initial palette, but to observe how people responded to the system in different countries. Although viewers actively explored the color dynamics of the initial text-free installation, it was not obvious to all users that the colors represented emotional states until the team implemented the additional design changes below.

To convey the emotional capabilities of the system more clearly than was possible to do through color alone, emotional terms that corresponded with the affect grid coordinates were added. As viewers touched different areas of the grid, the color of the entire interface changed and specific terms appeared in English and local languages, as shown in Figure 4.

Another major iteration was the inclusion of music in the emotion interface. As described above, music was introduced to heighten emotional engagement with

particular images, invite expression, and enable experimentation with emotional composition across images. Different melodies played as one explored the areas of the emotion interface to express feelings about a particular image. And, as one touched a series of photos, percussive sounds illustrated the pattern of moods across the photos. These musical capabilities appeared to greatly enhance engagement and play among individuals and groups using the system.

In an effort to simplify the user experience, the system included no buttons for input. To differentiate the gesture types (swipe, poke and hold), varying time thresholds were applied. Optimal time thresholds were tested, revealing that even increments of 20 ms significantly affected user experience.

The display and interactivity also evolved in accordance with observations of people using the system in unexpected ways. For example, children often played with the bottom edge of photos as if looking for something underneath. To support this "peeking" behavior, an underlying layer of images (iconic cityscapes associated with the primary images) was added, appearing when one lifted the top layer of images. This particular design resembled raising a curtain to reveal the city of origin for the images.

Another iteration followed the observation of viewers inspecting computers that ran the installation and displayed the depth camera images. Noticing the strong interest across locations in this material, the team projected this "behind-the-scenes" content onto a second wall. This display of the inner workings of the system (shown in Figure 5) fostered a sense of participation.

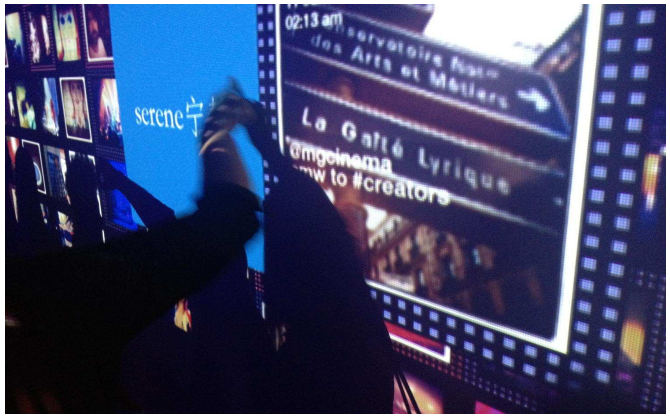

Figure 4. An image that was initially submitted at an event in Paris was emotionally reclassified by a viewer in Beijing.

Observed Emotional Self-Expression

As anticipated, people used the system individually, in groups and dyads. Some people intently explored the sentiment analysis and emotional expression interfaces and explicitly discussed their feelings about images with a companion. Some played with the color changes and grouping of images without apparent understanding of the emotional element of the technology.

In addition to expressing emotional responses to images, participants used the system in creative, unexpected ways to express identity, to participate socially and to represent mixed emotions. One common form of identity marking was sharing a picture of oneself or one's initials, then photographing oneself in front of that projected picture, sharing that second picture, and repeating the process to varying degrees. Many participants actively posed for photographs such that an array of images were projected onto their skin

Figure 5. Display of depth camera images.

and clothing. Frequently these photographs of a person covered with images of other people and art installations were themselves uploaded to the display. In this recursive process, people marked themselves as social screens in the recording of the event.

More complex forms of emotional expression were also observed. One of the ways people expressed complex feelings was by labeling seemingly positive pictures with negative emotions and vice versa. One woman laughed as she labeled a very attractive picture of herself as "miserable". Prior to deciding on this term, she explored the full range emotions on the interface, reading each term aloud while looking at her picture and asking the question "Am I happy?" . . ."Am I sad?" and so on. Smiling as she described herself and her glamorous photo as miserable, she demonstrates comfort with emotional complexity.

Another animated participant labeled as "happy" an image of a skeleton (Figure 6). She explained that she had taken the picture in a museum several weeks earlier, and upon seeing the interactive display selected it as a way to way to represent her recent break up. She thereby demonstrated a capacity to experience a number of emotions – sorrow, confidence, optimism – simultaneously. Such explorations suggest a desire for technology that allows for complex, playful expression of emotion.

Conclusion

Social media currently permits vast networking but could do far more to foster emotional connectedness. This project catalyzed connectedness by shifting interaction from personal devices onto a shared display and adding an emotional layer to image sharing. The

Figure 6. The skeleton image was reclassified as happy by the participant to express her mixed feelings about a breakup.

emotional layer captured emotional context of images and invited participants to express their emotional responses to images. Participants engaged with the system in creative ways to express themselves and connecting with others. This work explored new applications for sentiment analysis and interactive displays, suggesting a range of new directions for personalizing public computing.

Acknowledgments

We thank colleagues at The Creators Project and Intel particularly Ciel Hunter, Hosi Simon, Annie Dietz, Gerhard Stochl, Ryan Franzmann, Dave Haroldsen, Miro Bojic and Janet Tseng.

References

[1] D'Andrade & Egan. (1974). The Colors of Emotion. American Ethnologist, 1: 49–63.

[2] Boehner, K., Sengers, P., & Warner., S. "Interfaces with the ineffable: Meeting aesthetic experience on its own terms." *ACM Trans. Comput.-Hum. Interact.* Vol 15, no. 3. Nov. 2008. pp. 1-29.

[3] Burke, M., Kraut, R., & Marlow, C. Social capital on Facebook: Differentiating uses and users. *In ACM CHI: Conference on human factors in computing.* Vancouver, BC, 2011.

[4] Goleman D. Emotional Intelligence. Bantam Books. New York (NY), 1995.

[5] Harris, J. & Kamvar, S. An exploration of human emotion, in six movements. http://wefeelfine.org/

[6] Hawkley, L. C. & Cacioppo, J. T. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40*(2), 218-227, 2010.

[7] Mayer et al. (1990). Perceiving Affective Content in Ambiguous Visual Stimuli. Journal of Personality Assessment, 54: 772–781.

[8] Morris, M.E., Kathawala, Q., Leen, T.K., Gorenstein, E.E., Guilak, F., Labhard, M., Deleeuw, W. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *Journal of Medical Internet Research, 12,2* (2010), e10.

[9] Pennebaker, J.W., Francis, M.E., & Booth, R.J. Linguistic Inquiry and Word Count: LIWC 2001. Mahwah, NJ: Erlbaum Publishers, 2001. http://www.liwc.net/.

[10] Pollack, J.P., Adams, P. Gay, G. PAM: A Photographic Affect Meter for frequent, in situ measurement of affect. *CHI 2011,* Vancouver, BC, Canada, May 7–12, 2011.

[11] Russell, J. A Circumplex Model of Affect. *J. Personality and Social Psych, 39,* pp 1161-1178, 1980.

[12] Salovey, P. & Mayer JD. Emotional intelligence. Imagination, Cognition, and Personality, 9, pp 185-211, 1990.

[13] Steblin, R. A History of Key Characteristics in the 18th and Early 19th Centuries. University of Rochester Press. 2005.

[14] Turkle, S. *Alone together*. Basic Books. New York (NY), 2011.

[15] Wilson, A. D. Using a depth camera as a touch sensor. *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10),* pp 69-72. New York, 2010.