Using Gestures on Mobile Phones to Create SMS Comics

Vidya Setlur vidya.setlur@nokia.com
Agathe Battestini agathe.battestini@nokia.com
Timothy Sohn tim.sohn@nokia.com
Hiroshi Horii hiroshi.horii@nokia.com

Nokia Research Center
955 Page Mill Rd.
Palo Alto, CA 94304 USA

ABSTRACT
SMS messages provide an easy and simple method to communicate with others. These short messages are useful, but can sometimes feel restricted due to the limitations of textual communication. The ability to express subtle nuances and contexts around the message could help add enjoyment and amplify the emotions being expressed by the mobile user. We present SensorComix, a new way of creating comics using SMS messages combined with gestures on mobile phones. Comics are automatically generated from users' SMS messages, and augmented with visual icons based on the performed gestures. We demonstrate that gestures mapped to comics can help influence the expressiveness of messages sent by mobile users.

ACM Classification Keywords
H.5.2 Information interfaces and presentation: User Interfaces - Graphical user interfaces

General Terms
Human Factors

INTRODUCTION
Short Message Service (SMS) communication has become a popular alternative to mobile voice communication. In 2008 nearly 75 billion SMS messages were sent every month worldwide [3]. The popularity of SMS messages can be attributed to a variety of factors. SMS is an easy way to communicate when voice might be inappropriate, difficult, or too heavyweight. SMS is also an easy way to send quick messages to a loved one, such as "see you soon! i love you".

Although SMS affords simple lightweight messaging, communicating within a 160 character limit can restrict the rich nuances and meanings that people often want to express. One way to help individuals express richer content with their SMS messages is through comics. Comics are a popular versatile narrative form that are immediately recognizable and appeal to a large population across generations. Comics also have a rich visual vocabulary, and are graphically very flexible in terms of layout, coloring, mix of textual and pictorial content, flatness or 3D layering, and levels of detail. Using these graphical elements, comics can provide a medium to help mobile users express emotion through SMS messages.

Figure 1. Examples of SMS Comics created with our interface. Top: A message augmented with hearts on the comic using an affectionate gesture of rubbing the phone. Bottom: A message with the rub gesture used as shortcut for augmenting a smiley emoticon.

The key to enabling expressiveness in a comic is to provide flexibility in comic representation with minimal effort. Creating messages while being mobile can be cumbersome, thus extensive interactions tend to be expensive. Gestures on mobile phones provide an easy way to complement SMS message creation by enabling the user to amplify emotions already being expressed in the text. Several systems have explored the use of gestures through additional sensors or camera-based interaction [10]. Previous work has found that gestures can create a feeling of flow and emotion [2] and add to the spontaneity of the mobile communication. SensorComix uses gestures to complement text messages with this additional dimension of expression. A quick "i love you :)" message could be augmented with hearts on the comic by a short affectionate gesture after creating the message (Figure 1- top). Gestures can also function as shortcuts for inserting

TEI 2010, January 25 - 27, 2010, Cambridge, Massachusetts, USA.

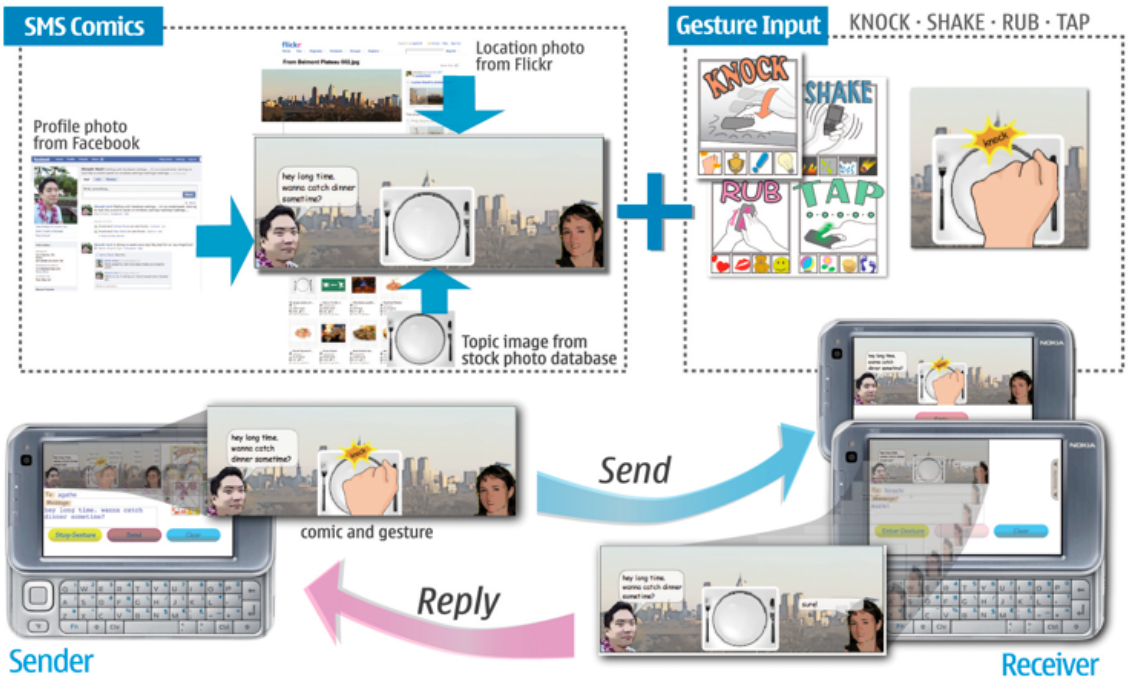

Figure 2. System overview. The interface resides on a Nokia N810 equipped with a sensor box. On the sender side, the user types an SMS message that gets converted into a comic. The people in the comic are retrieved from their respective Facebook profile pages. The location image in the background of the comic is based on the location from where the SMS is sent from the phone, and is retrieved from Flickr's geo-tagged photo collection. The topic image is retrieved from a stock photo database based on the keywords from the SMS text. The user then gestures with the phone, and an associated visual is augmented with the comic and sent to the receiver's phone.

emoticons into the text messages. For example, a gesture can alleviate the burden of typing ':' and ')' and directly insert a graphic instead (Figure 1 - bottom). These gestures provide a multimodal way to enhance SMS messages.

This paper describes SensorComix, a system to create SMS comics using gesture inputs on mobile phones. As the user is composing a text message, he can make an initial gesture to determine the visual added to the comic. Continuous gesturing progressively add more visuals or animate the visuals already composited on the comic. The comics vary by text bubble layout, colors, emoticons, and imagery. Since the gesture is performed while the message is being created, the interface is natural for the user to amplify the emotion expressed in the SMS message.

RELATED WORK

One of the lacking elements with SMS is the inherent inability to enrich these messages with additional content. Amin *et. al.* found that teenagers who text frequently, enjoy SMS, but miss the expressiveness of the application. They found that enhancing text messages with contextual information and graphical representation, provides a richer experience for both the sender and receiver [1]. This provides some of the basis for exploring the use of mobile phone gestures and comics for SMS representations.

Several related systems, such as ExMS and Comeks have touched upon similar designs around the comic metaphor and avatar representations of text-based communications. ExMS is a system that provides a rich experience through the use of avatar animations, and is primarily an authoring tool for creating animated avatars for communication [6]. In contrast, our work utilizes existing data to automatically convert one's messages and their associated contextual information into comical form. Comeks allows users to augment and annotate images in an MMS editor with speech bubbles and other comic-related accessories [8]. The tool is used to author MMS messages and does not automatically convert SMS messages into comic form.

Comic Chat [5] includes balloon construction and layout, placement and orientation of the comic characters. Several key differences are that the comic representation is scripted with distinct locations that the comic characters can enter (e.g., house, balcony, or fantasy world). Comic characters also have a palette of emotions that a user can choose from. Our work however, automatically utilizes the semantics and context of SMS data for a more relevant comic experience.

SENSORCOMIX

SensorComix is a system that lets users augment their text messages with emotional expressiveness using gestures. As text is sometimes insufficient to express communication, we employ phone-based gestures that a user performs while creating SMS messages to visually augment context to the messages. The system is comprised of a phone client application for uploading SMS and viewing the comics and back-end infrastructure that handles the comic generation (Figure 2). We equipped a Nokia N810 with the Shake SK-7 sensor

Figure 3. The SK-7 sensor box is attached to the back of the N810 to create a smooth handheld experience. The sensor box has 12 capacitive touch sensors on the outer shell.

box. The SK-7 augments a device with inertial, magnetic field and capacitive sensing in a sub-matchbox size enclosure. The SK-7 can transfer the sensor data to the mobile device via Bluetooth wireless protocols. The added sensor box fits easily on the back of the phone without significantly changing the form factor (Figure 3).

Building an SMS Comic

An SMS comic is composed of one or more of the following six visual elements:

- Profile photo of the sender
- Profile photo of the receiver
- Topic image that relates to the keywords extracted from the SMS conversation
- Image representing the location from where the message was sent
- Text bubbles containing the content of the SMS
- Visual effects created through gestures

Figure 2 shows an overview of how comics are generated. The system retrieves profile photos for both the sender and receiver by resolving the "people" entities from the phone's contacts and online social networks such as Facebook and Flickr using a Levenshtein distance metric [11]. The profile image is then automatically cropped around the face after applying a standard face detection algorithm [7], and resized according to the comic template dimensions.The sender's photo (who initiated the conversation) is on the left and the receiver's photo is on the right. SensorComix creates the background image by querying Flickr for panoramic images based on the sent message's city and state inferred from the onboard GPS unit. The image must not contain distracting objects or people that could detract from the generated comic. In order to obtain such a characteristic image, we apply an image saliency algorithm on the Flickr image to determine the distribution of salient points on the image [4]. The saliency algorithm is used to identify salient points in an image based on low-level image features such as color, intensity, and orientation. The location image is centered and cropped to contain the most salient regions, and resized to fit the background. Similar heuristics are used for topical images, but since the topical image is composited in the center of the SMS Comic, a desirable image would require a central object of attention. The system chooses topical images that have a focalized distribution of salient points.

If the user performs a gesture while creating an SMS, the icons represented by that gesture are shown in the comic. The sender in Figure 1 performed a rub gesture, so heart icons are created in the comic. As a final step, all the images of the people, location, topics, and gesture-based icons are tooned to conform to the abstract look-and-feel of a comic.

User Interface

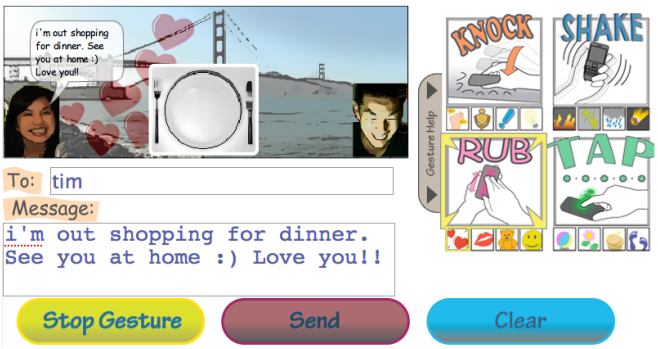

Figure 4. The Mobile Interface. The sender types a message and rubs the phone. The appropriate gesture is highlighted [in yellow] in the menu along with the corresponding visual, which are hearts.

The user interface provides text boxes to enter the SMS recipient's name (or phone number) and the text message. The phone contact and the sender's name are used to populate the faces in the comic. While the sender types the message, the text simultaneously appears in the text bubble of the comic (Figure 4). The sender then presses the 'Enter Gesture' button that also opens up a menu of possible gestures and their corresponding default visuals below (Figure 4). When the sender performs a certain gesture with the phone, it is highlighted (in yellow) in the menu for visual feedback. Each gesture maps to a set of icons representing a type of expression. For example, rubbing gestures are more affectionate, while shake gestures are more action driven.

The sender can change the visual for the gesture by clicking on another icon below the gesture, in the menu. When the sender is satisfied with his gesture and visual, he clicks

on the 'Stop Gesture' button and can press 'Send' to send the message to the recipient. The 'Clear' button resets the interface for the sender to start again. When the recipient receives the message, he can press the 'Reply' button which opens up the interface in 'Sender' mode, and any text entered will appear in a second text bubble near the recipient's face in the comic. This interaction helps create an ongoing conversation between the sender and receiver.

Gestures to Comics

SensorComix maps gestures to a family of visuals representing four types of expressions: affection (rubbing), poking (tapping), getting attention (knocking), and showing impatience or anger (shaking). We chose these gestures based on common hand motions people use to express these emotions. After entering the SMS message, the user starts with an initial gesture, and the corresponding visual for that gesture appears based on the menu (Figure 4). This first recognized gesture defines what type of visual appears in the comic. To add to the playfulness of the interface and let the user specify a range of intensity for each emotion, the corresponding visual is added to the comic progressively as the user continuously inputs the same gesture. For instance, if the user rubs the sensor, one heart will show up. Each subsequent rubbing gesture would add an additional heart to the comic until the maximum allowable icons are shown (Figure 1). The interface also supports continuous gesturing to make the interface more natural to use. For example, when the user performs another gesture such as shaking after the initial rub gesture, it results in the hearts already present in the comic, moving in interesting ways.

The gesture algorithm is implemented using Support Vector Machines (SVM). SVM is a supervised learning method used for classification and regression that is widely used for performing gesture recognition [9]. After the user clicks 'Enter Gesture', the sensor box continuously generates touch and accelerometer data at 10 Hz each. The accelerometer data contains x, y, and z values and the touch data contains one value for each of the 12 capacitive sensors on the sensor box (Figure 3). Each input signal is segmented to detect the onset of the touch or move gesture. This is achieved by thresholding the rising front of the input energy signal, obtained with the sum of the absolute values of each dimension. The feature values are selected by aggregating the signal values over 0.5s time intervals.

The features from the two sensor signals are used as input to two different classifiers. One classifier uses the capacitive sensor input to recognize three classes: tapping, rubbing, and absence of touch. The accelerometer input is used to recognize four classes: stationary, knocking, moving in hand, and shaking. SensorComix uses the output from both classifiers to recognize the four gestures used in the system (tapping, rubbing, knocking, shaking). The other three classes are used to assert that the sensor input is inactive.

Conclusion

In this paper, we present SensorComix, a new way of communication through comics using SMS messages combined with gestures on mobile phones. In the future we would like to allow users to create their own custom visuals to share with others. We would also like to enable tactile feedback on the device while performing gestures to add another dimension of play to the SensorComix experience.

REFERENCES

1. A. K. Amin, B. T. A. Kersten, O. A. Kulyk, P. H. Pelgrim, C. M. Wang, and P. Markopoulos. Sensems: a user-centered approach to enrich the messaging experience for teens by non-verbal means. In *MobileHCI '05*, pages 161–166, New York, NY, USA, 2005. ACM.
2. S. Bhandari and Y.-K. Lim. Exploring gestural mode of interaction with mobile phones. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 2979–2984, New York, NY, USA, 2008. ACM.
3. CellSigns. http://www.cellsigns.com/industry.shtml.
4. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, 1998.
5. D. Kurlander, T. Skelly, and D. Salesin. Comic chat. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 225–236, New York, NY, USA, 1996. ACM.
6. P. Persson. Exms: an animated and avatar-based messaging system for expressive peer communication. In *GROUP '03: Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, pages 31–39, New York, NY, USA, 2003. ACM.
7. M.-T. Pham and T.-J. Chain. Fast training and selection of haar features using statistics in boosting-based face detection. *IEEE International Conference on Computer Vision*, 0:1–7, 2007.
8. A. Salovaara. Appropriation of a mms-based comic creator: from system functionalities to resources for action. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1117–1126, New York, NY, USA, 2007. ACM.
9. C. von Wrede, P. Laskov, and F. F. Ida. Using classification to determine the number of finger strokes on a multi-touch tactile device. In *In European Symposium on Artificial Neural Networks*, pages 549–554, 2004.
10. J. Wang, S. Zhai, and J. Canny. Camera phone based motion sensing: interaction techniques, applications and performance study. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 101–110, New York, NY, USA, 2006. ACM.
11. L. Yujian and L. Bo. A normalized levenshtein distance metric. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(6):1091–1095, 2007.