Using Comics as a Visual Metaphor for Enriching SMS Messages with Contextual and Social Media

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

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

ABSTRACT

Comics can serve as visual artifacts, representing numerous qualities in a phone text message, including its participants, genres, and community, all tied to the context of the text. In this paper, we introduce a system called SMSComics that explores the usage of comics as a visual metaphor for SMS conversations sent by users on their mobile handsets for creating greater user engagement. The resulting imagery is delivered in the form of a comic strip stylization that improves context-awareness and relevance in a mobile social networking setting. Firstly, SMSComics takes images from a user's social communities such as Facebook to display the picture of the sender and receiver. The service then includes the detected location with geo-tagged images from Flickr. Lastly, through semantic analysis, the topic of the message can be represented by an image. For example, if a message is about meeting for a cup of coffee, a stock photo of a cup of coffee might be chosen. With SMSComics, we can study the semantics, usage and serendipitous nature of combining contextual and social media with text messages.

Categories and Subject Descriptors

H.5.2 [**UserInterfaces**]: Graphical user interfaces (GUI)

Keywords

SMS, text messages, comics, context, mobile interface

1. INTRODUCTION

The amount of personal digital information is increasing. Consumer devices such as camera phones greatly facilitate the daily creation of not just digital media such as, images, audio, and video, but also other types of data such as Short Message Service (SMS), voice calls, and contact information. In particular, SMS text messaging is the most widely used data application on the planet, with 2.4 billion active users, or 74% of all mobile phone subscribers sending and receiving text messages on their phones [2].

This data is most often created, exchanged, then stored in logs that are rarely used thereafter. In addition, such text communication can be perceived as lacking some components of communication, such as, intonation, pitch, gesture, facial expression, and visual cues that are essential in making these textual logs useful and engaging for post consumption.

Figure 1: SMSComics viewed on a phone browser.

Further, the use of context in mobile devices such as location, is receiving increasing attention in mobile and ubiquitous computing research. This continuous creation of SMS content gives us an interesting challenge to explore: how to leverage large amounts of personal mobile content, how to augment such data with awareness of their environment and situation as context, and how to offer users playful representations that will let them enjoy more, remember better personal and social moments captured in SMS dialogues.

In order to further address this design challenge, we present and discuss a visual interface to bridge the gap between the "conversation" metaphor of textual exchange in SMS, and the prevalent "model world" metaphor that users are most familiar with. Thus, the interface ideally provides an abstract but visually accessible world of objects and locations. Mobile computing aids in this bridging this gap, by being able to investigate ways of linking situational context of SMS dialogues into a richer visual interface.
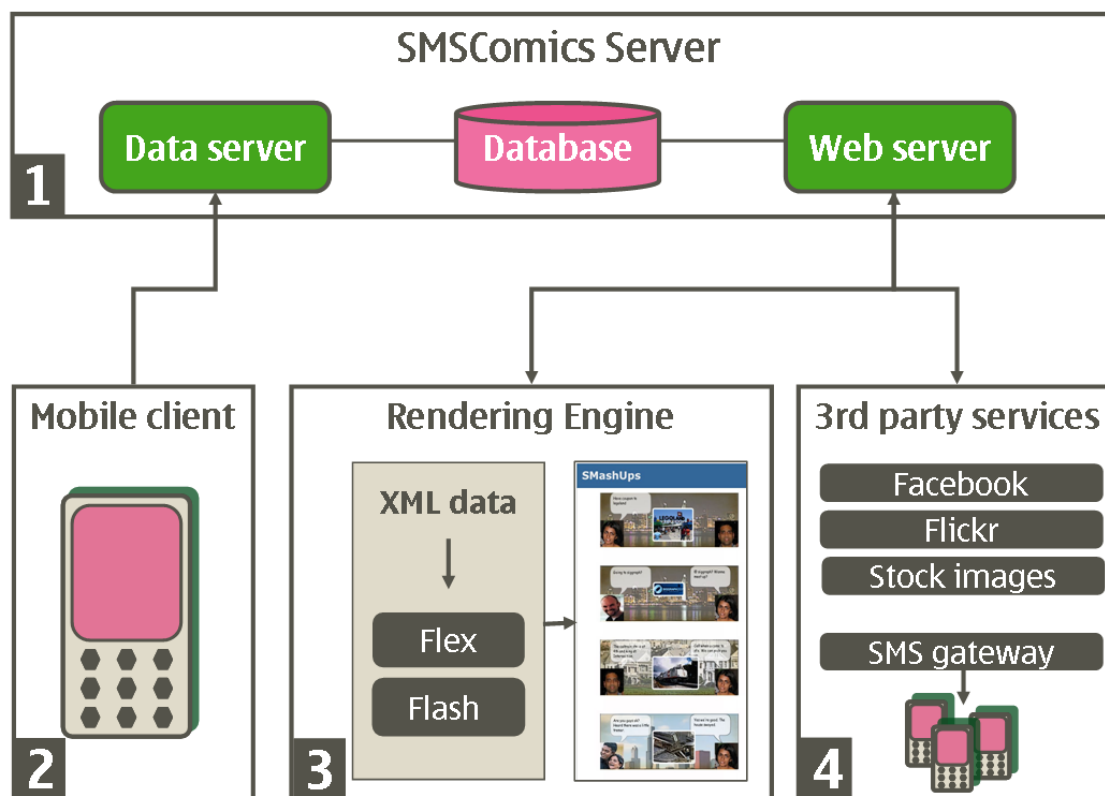
Figure 2: The overall SMSComics architecture: The mobile client application (2) sends the SMS and context data to the data server (1) which performs preprocessing on the data and stores it in a database ; the web server (1) enables the visualization of the user's SMSComics through interactive Flex/Flash rendering (3) ; third-party services APIs (3) are used to retrieve user's profile, spatial and topical images, as well as access an SMS gateway to let users share their comics with their friends.

2. RELATED WORK

Numerous studies have explored the use of text messages, primarily among teenagers [13],[15],[23],[9]. These studies all indicate the enjoyment and usefulness of text messaging. However, one of the lacking elements with text messages is the inability to enrich these messages. 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 human embodiment provides a richer experience for both the sender and receiver [4]. This provides some of the basis for exploring the use of comics for SMS representations.

Several related systems, such as ExMS, Comeks, Comic Chat, and Comic Diary, have touched upon similar designs around comic and avatar representations of text-based communications, though SMSComics stands alone as a unique contribution in terms of its use of existing text-based social conversations that are automatically converted to comic form based upon the semantics of the SMS data. We now discuss each of these related systems.

Su and Yee proposed a mobile chat message interface that could be used to help a user graphically see their messages on a mobile device [21]. Their interface uses colors and a timeline representation for easy navigation and display of chat messages. Messaging can be further enhanced by incorporating other forms of graphical representation. ExMS [18]

is a system that provides a rich experience through the use of avatar animations. The tool provides a library of expressions, animations and text input fields for creating avatars and sharing with others. ExMS is primarily an authoring tool for creating animated avatars for communication. In contrast, our work utilizes existing data to automatically convert one's messages and contextual information into comical form.

Unlike chat interfaces and avatars, other systems have explored the use of comics to embody emotion and represent context. Comeks [20] allows users to augment and annotate images in an MMS editor with speech bubbles and other comic-related accessories. The tool is used to author MMS messages and does not automatically convert existing phone messages into comic form. Our system automatically converts an existing SMS message into comic form using several rendering algorithms.

Comic Chat [14] includes balloon construction and layout, placement and orientation of the comic characters. Similar to our work, Comic Chat can adjust the background or scene elements to reflect the topic of conversation. 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 leverages processing algorithms to better identify the

semantics of SMS data for a richer and more relevant comic experience.

Comic Diary is a personal guidance system that visually represent one's experience as comics [22]. The system provides users with personal guidance to navigate places such as museums, trade shows, and conferences. Unlike SMSComics, their system does not capture social conversations that occur.

3. DESIGN MOTIVATION
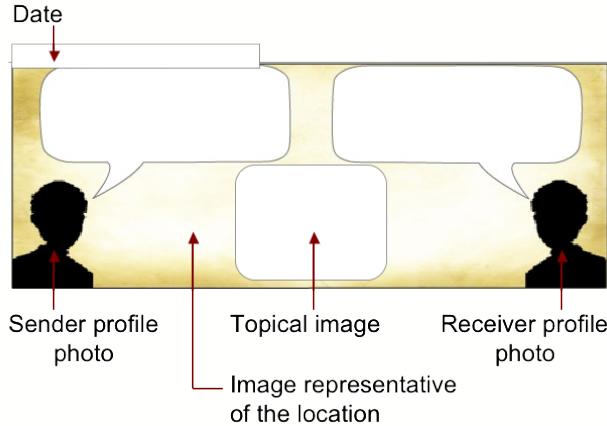
Figure 3: A sample design template used in the SMSComics system for visually depicting an SMS conversation between two people.

The increasing availability of computer systems with high-quality graphics and fonts has created an opportunity for using graphical design practices to design more effective user interfaces. "A picture worth a thousand words must first be a good picture" [6]. The goal of effective user interface design is to reduce the gulf between users and applications [7] and graphical techniques are a good way to achieve this goal [11]. Well designed graphical techniques can improve the presentation of application or task information, as well as improve recall of application or task information and make computer systems more satisfying and easier to learn [16]. Unfortunately, it is not easy to design graphical user interfaces. The problem is that technological advances, such as hardware support for high resolution color and real time animation, have created an overwhelming space of design choices.

A fundamental assumption behind the theory for designing effective interface is that graphical presentations are sentences of graphical languages, which are similar to other formal languages in that they have precise syntactic and semantic definitions [5]. The effectiveness of this vocabulary depends on how well the visual language can serve as a natural metaphor to the underlying information needed to be presented. In the context of SMS conversations, we went through various visual concepts to ascertain how well the visual interface could bridge the paradigm between the conversation metaphor of SMS, and the model world metaphor that users perceive around them.

Scott McCloud's work discusses the usage of comics as a form of creative expression that effectively deals with the arrangement of images and text to narrate a story or dramatize and idea [17]. Comics are a popular yet versatile narrative form that is familiar with readers for exercising both visual and verbal interpretive skills [8]. They are immediately recognizable and they appeal to a large population across ages, leading to a commonality of experience. Comics have also a rich visual vocabulary, and are graphically very flexible (and forgiving) in terms of layout, coloring, mix of textual and pictorial content, flatness or 3D layering, and levels of detail. Comics can provide visual representations of the different facets that take place in a SMS conversation, using symbols, imagery and color. In addition, comics can serve as social artifacts, referring to numerous qualities in a conversation, including its participants, genres, and community, all tied to the context of the text.

4. SYSTEM

The SMSComics service is comprised of a phone client application (Figure 1) for uploading SMS and viewing the comics, a web server, and a backend infrastructure as shown in Figure 2. The web server interacts with other social network service APIs such as Flickr and Facebook to leverage the user's friend connections. The functional flow of SMSComics generation involves the following processes: (1) Uploading SMS Message, (2) Linking contextual and social data such as location, images of contacts, topic of the conversation with the SMS messages, and finally (3) Rendering SMSComics. We now describe each of these processes in more detail.

4.1 Uploading SMS Message

Each SMSComic is rendered based on information and media objects collected. The user's personal phone is the primary source of information as it is the device where text messages (SMS) are sent and received between the user and his phone contacts. The collected phone data contains the text message information, along with other contextual attributes such as contact names, time, and location (using GPS coordinates and/or GSM cell tower ids when the GPS is unavailable). The phone client then transmits the phone data to the web server. The uploading of SMS data to the server from the phone client can be done in manual or automatic mode. In the manual mode, the user explicitly chooses the messages that he would like to upload, while in the automatic mode, all SMS messages are automatically uploaded after a user set time interval.

4.2 Linking Context and Social Data

The first linkage is to images of the people involved in the SMS conversations. At the time of registration with the service, the user is asked to link his newly created account to his Facebook and Flickr accounts. This information is used to obtain the user's profile photo along with photos of his contacts who are in the SMS conversation. In order to determine the profile photo images for the SMSComic, the system needs to resolve the "people" entities from the phone's contacts and the social network available online from Facebook and Flickr. However, there are some inherent challenges in performing this entity resolution. For example, if 'Joe' is in the phone's contacts, but there are two 'Joe's in the user's online social network, the system would need to resolve this conflict before generating the SMSComic. Our entity resolution algorithm creates a matrix of the metric distances based on Levenshtein distances between the phone
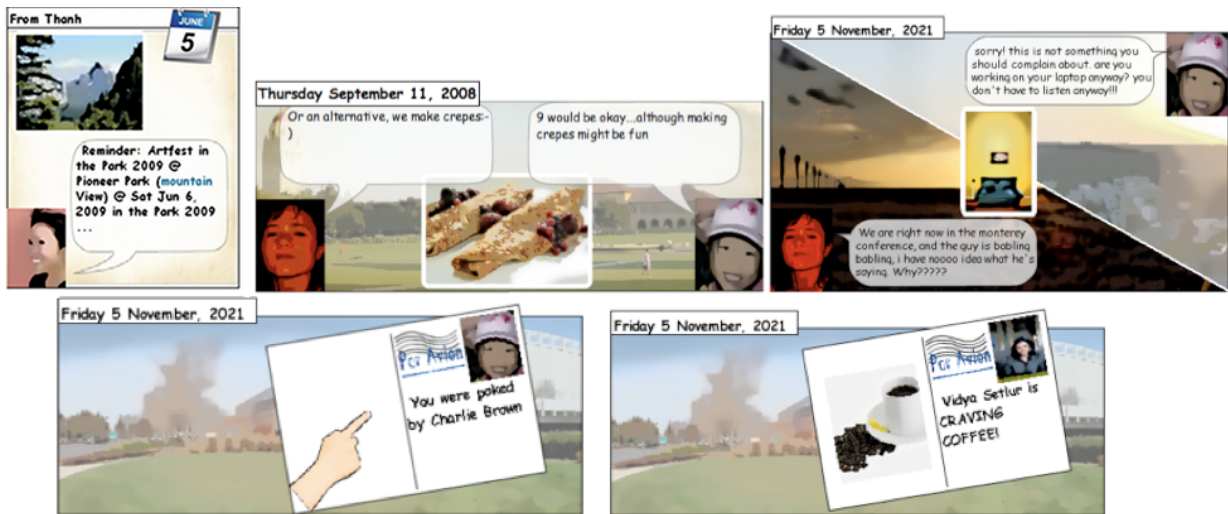
Figure 4: Examples of comics generated using various templates. Top row, from left to right: Depicting a single SMS message with an image depicting the key topic 'mountain.' There is no background image in this example, as the location from where the SMS was sent is unknown.; Depicting a conversation between two people with the background image of the sender's location only as the recipient's location is unknown.; Depicting an SMS exchange between two people from two different locations. Bottom row, from left to right: Depicting a Facebook 'Poke'. The poke feature can be used for a variety of things on Facebook. For instance, you can poke your friends to say hello; Depicting a Facebook status update from a friend on Facebook that is received as an SMS. In both the Facebook 'Poke' and status update messages, the background image denotes the current city of the sender in his Facebook profile.

book contact names and the online social contact names. The Levenshtein distance [1] is a metric for measuring the amount of difference between two string sequences (i.e., the so called edit distance)and is given by the minimum number of operations needed to transform one string into the other, where an operation is an insertion, deletion, or substitution of a single character. We shuffle the phonebook contact names in order to generate 'first name last name' and 'last name first name' combinations, and increase the number of perfect matching instances between the phonebook contact names and the social contact names. We then create a matrix containing several metric distances between all the combinations of phonebook contact and the online social friend names. We use 3 possible metric distance values ranging between 0 and 1, where 0 indicates no match and 1 indicates a perfect match:

- The first distance value is computed between the full names of the phonebook contact and online social friend.

- The second distance value is the best matching value between the phonebook contact's first name and either the online social friend's first name or last name.

- The third distance value is the best matching value between phonebook contact's last name and either the online social friend's first name or last name.

If the first distance value is 1, then it indicates that the phonebook contact name matches perfectly with that of the online social friend. If the first distance value is less that 1, we consider the second and third distance values. In this case, the system provides a visual affordance in the interface, indicating to the user that the match is only partial.

In order to extract the most important keyword in the SMS text, we adopt a simple approach of first filtering out stop words (e.g. 'and', 'the', 'or') , common SMS abbreviations (e.g. 'brb' for be right back) and perform word stemming (e.g. 'running' becomes 'run', 'eaten' becomes 'eat') to this text. For the comic representation, displaying a person, place or thing, *i.e.* a noun, tends to better conceptualize the subject, rather than action words which are often more ambiguous. Thus, by prioritizing our search on terms that are predominantly used as nouns as opposed to terms which have an uncommon noun sense, we avoid spurious associations such as an image of musical notes for the phrase 'beat Rockies.' We use the concept of polysemy to determine whether a term is a noun or not. A term is said to be polysemous if it has more than one meaning. For example, the term 'beat' has at least two meanings - a verb sense that means 'to strike' and a noun sense that means 'a metric for music.' In order to determine word senses and polysemy, we use Wordnet, a semantic lexicon for the English language [3]. A term is deemed a noun if WordNet ranks the noun senses of the term as more popular than any other part of speech. If there are multiple nouns identified in the text, we leverage Wordnet's ranking scheme for word familiarity to choose the most familiar noun in the text, as users tend to associate better with more familiar words.

Finally, based on the location coordinates associated with the corresponding uploaded SMS log, we apply reverse geocoding to compute a coarsely granular location (city, state) where the SMS was sent or received.

4.3 Rendering SMSComics

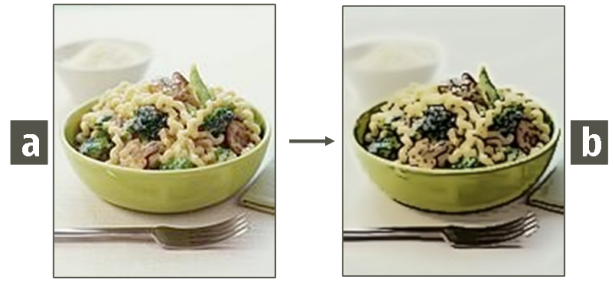An SMSComic is composed of at least one of six visual elements depending on the type of template used for the
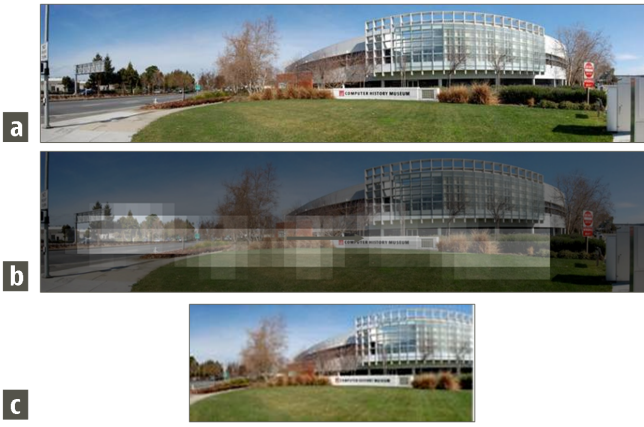
Figure 5: (a) Representational panoramic image retrieved from Flickr using the GPS information from the phone. (b) Using image saliency for image cropping. (c) Cropped image.

Figure 6: (a) Original image. (b) Tooned image.

compositing: the profile photo of the sender and the profile photo of the receiver, a topical image that relates to the keywords extracted from the conversation, an image representative of the location where the message was sent or received, and textual blobs containing the content of the SMS. Figure 3 shows an example template used whenever an SMS conversation needs to be depicted between two people. Figure 4 shows examples of various comic styles based on the type of data.

Once the profile photos are retrieved, the image is automatically cropped around the face after applying a standard face detection algorithm [19], and resized according to the template dimensions. For obtaining an image of the location, the server queries Flickr for panoramic images based on the city and state obtained by performing reverse geocoding. As the location image is rendered as the background of the SMSComic, it is desirable that the image is of a landscape, devoid of central objects or people that would divert the user's attention. In order to obtain such characteristic images, we apply an image saliency algorithm to determine the distribution of salient points on an image [12]. The saliency algorithm is used to extract attended locations in complex scenes based on low level image features such as color, intensity, and orientation, which represent the visual scene. The algorithm returns a gray scale map depicting regions of importance in an image, where higher the intensity, more important that given region is. Figure 5 shows how saliency is applied to location images. The location image is then centered and cropped around the central salient points, and resized. Similar heuristics are used for topical images, but since the topical image is composited in the center of the SMSComic, a desirable image would require a central object of attention. The system chooses those images that have a very focalized distribution of salient points.

Finally, all the images of the people, location and topics are tooned to conform with the abstract look-and-feel of a comic, as show in Figure 6. Tooning is accomplished through a quantizing technique in HSV space [10]. The hue H of each pixel is constrained to the nearest of twelve primary and secondary colors, and the saturation S and value V are clamped to 15% and 25%, respectively. We find that by rendering the outer contour of the extracted object with a black outline, the resulting cartoon image is aesthetically more pleasing. Figure 4 shows how the final SMSComic renders.

5. PRELIMINARY USER STUDIES

SMSComics was deployed in the field for a period of several months among 8 participants. With participant consent, user behaviors were logged. Upon completion of the field deployment of SMSComics, participants were interviewed regarding their experiences with SMSComics, frequency and usage patterns, utility of its features, and thoughts on improving the system. The participants experienced both text-only versions and comic-enhanced versions of SMSComics. For equal time periods (20 days with text-only version and 20 days with comic-enhanced versions), participant behaviors were logged and observed.Each interview lasted approximately 1.5 hours.

Participants said that they mostly used SMSComics for communicating with family and friends and that they wanted more populated comics (i.e., more images of locations or images from the phone). One participant mentioned that his teenage daughter and her friends would need to be able to post their comics on social networking sites if they were to use SMSComics. Some thought that SMSComics would be useful as a creative storytelling tool or as a digest of their activities in the past.

The follow-up interviews with participants resulted in several ideas for improving the design of SMSComics and rethinking SMS communication, more broadly. Rather than using SMS for short, asynchronous, text-based communications, this technology can be re-appropriated for more conversational, more synchronous, and visually enhanced communications. Furthermore, it may hold the potential for new usage scenarios such as archiving and storytelling around past communications.

6. DISCUSSION

SMSComics is a system created with the goal to explore augmentation of SMS data with contextual and social information to create comics whose visual elements relate directly to the persons conversing, and the topic and context of their conversation. SMSComics thus, provides a platform for discussing several important questions related to mobile social media: is the large adoption of SMS as communication mode tied to its overall simplicity, or can it be the basis for emerging concepts using more complex social and contextual media? What specific characteristics of SMSComics

provide additional value to the user, for instance in terms of enjoyment, social connectivity, or improved performance? Is the use of context within the practice of texting valuable? Do specific privacy issues emerge from the creation of comics? Do communication patterns change with the use of SMSComics?

7. CONCLUSION

Inspired by the comic form of communication through both images and text, we designed and prototyped a system 'SMSComics', which explores the usage of comics as a visual metaphor for SMS conversations sent by users on their mobile phones. While SMS communications have historically been used for short text-based communications, SMSComics a step in the direction of enriching such communications with social and contextual media.

8. REFERENCES

[1] Levenshtein distance. http://en.wikipedia.org/wiki/Levenshtein_distance.

[2] Wikipedia: The free encyclopedia. http://aquaryus.wordpress.com.

[3] Wordnet: Semantic lexicon for the english language.

[4] 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: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 161–166, New York, NY, USA, 2005. ACM.

[5] J. Bertin. *Semiology of Graphics*. The University of Wisconsin Press., Wisconsin, 1983.

[6] W. Bowman. *Graphic Communication*. John Wiley and Sons, New York, 1969.

[7] S. W. Draper, D. A. Norman, and C. Lewis. *User Centered System Design: New Perspectives on Human-Computer Interaction*. Erlbaum, Hillsdale, NJ, 1986.

[8] W. Eisner. *Comics and Sequential Art*. W. W. Norton, 2008.

[9] R. Grinter and M. Eldridge. Wan2tlk?: everyday text messaging. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 441–448, New York, NY, USA, 2003. ACM.

[10] C. G. Healey and J. T. Enns. A perceptual colour segmentation algorithm. Technical report, 1996.

[11] E. L. Hutchins, J. D. Hollan, and D. A. Norman. Direct manipulation interfaces. *User Centered System Design: New Perspectives on Human-Computer Interaction*, pages 87–124, 1986.

[12] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. volume 20, pages 1254–1259, 1998.

[13] E.-L. Kasesniemi and P. Rautiainen. Mobile culture of children and teenagers in finland. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 170–192, New York, NY, USA, 2002. Cambridge University Press.

[14] 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.

[15] R. Ling and B. Yttri. Hyper-coordination via mobile phones in norway. pages 139–169, 2002.

[16] D. Markoff and D. Cartel. *Total Recall*. Grosset and Dunlap, New York, 1976.

[17] S. McCloud. *Understanding Comics: The Invisible Art*. Harper Paperbacks, 1994.

[18] 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.

[19] M.-T. Pham and T.-J. Cham. Fast training and selection of haar features using statistics in boosting-based face detection. *ICCV 2007. IEEE 11th International Conference on Computer Vision*, pages 1–7, 2007.

[20] 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.

[21] D. K. S. Su and V. S. Y. Yee. Designing usable interface for navigating mobile chat messages. In *OZCHI '07: Proceedings of the 2007 conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artifacts and environments*, pages 291–294, New York, NY, USA, 2007. ACM.

[22] Y. Sumi, R. Sakamoto, K. Nakao, and K. Mase. Comicdiary: Representing individual experiences in a comics style. In *UbiComp '02: Proceedings of the 4th international conference on Ubiquitous Computing*, pages 16–32, London, UK, 2002. Springer-Verlag.

[23] A. S. Taylor and R. Harper. Age-old practices in the 'new world': a study of gift-giving between teenage mobile phone users. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 439–446, New York, NY, USA, 2002. ACM.