Emotion Expression for Affective Social Communication

Prabath WEERASINGHE*, Ashu MARASINGHE*, Rasika RANAWEERA**, Senaka AMARAKEERTHI*** and Michael COHEN**

* Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
** University of Aizu, Kamiiawase 90, Tsuruga, Ikki-machi, Aizu Wakamatsu, Fukushima 965-8580, Japan
*** Sabaragamuwa University of Sri Lanka, Belihuleya, Sri Lanka

Abstract: Human interaction with social networking services (SNS) is currently a very active research area. SNS posts, such as tweets, allow users to broadcast their ideas in short form of text, voice, or images, using mobile devices and computers. Text and speech enriched with emotions is one of the major ways of exchanging ideas, especially via telephony and SNS. By analyzing a voice stream using a Hidden Markov Model (HMM) and Log Frequency Cepstral Coefficients (LFPC) based system, different emotions can be recognized. Using a simple Java client, recognized emotions can be delivered to a server as an index. A mobile client can then retrieve the emotion and display it through colored icons. Each emotion is mapped to a particular color, as it is natural to use colors to represent various expressions. Not only colors, we also use avatar animation models in different environments for the expression of different emotions.

Keywords: Emotion, Kansei Engineering, Virtual Environment

1. INTRODUCTION

NTT DoCoMo, Japan’s largest mobile network provider, features a service called iαppl, an integrated platform for Java-based applications. We developed a program using this platform to display emotions of human voice using colors. Colors have a strong impact on emotions and feelings [2]. Although perceptions of color are somewhat subjective, some color effects have global associations [1]. To display various kinds of emotions on a mobile screen, they were mapped to carefully selected colors. Interaction between a mobile phone and a MATLAB program, which recognizes emotions from voice, is performed through a collaborative virtual environment (CVE) server, a Java client–server architecture.

Our application can also interact with Open Wonderland, a Java open source toolkit for creating collaborative 3D virtual worlds. Using the “Wonderland–CVE Bridge” [9], we can display the location of a Wonderland avatar on a mobile screen.

Modern mobile devices — such as feature phones, smart-phones, tablets, netbooks, etc. — are internet-capable and feature extended hardware capabilities. With such advanced capabilities, people can perform many daily life activities (such as shopping or dining arrangements), business activities (such as banking, auctions, and correspondence), and social activities (such as events, chatting, meeting, and dating). Today’s mobile devices not only have larger screens, high color depths (Table 1), high speed processors, and huge memory, but they also support a wide variety of programming languages, including Java, Python, C/C++, and Objective-C. Vendor companies also encourage developers to create new applications by opening the marketplace. NTT DoCoMo provides an integrated platform for Java-based applications, a JME profile-compliant framework called DoJa (DoCoMo Java), to develop such applications using the iαppl framework [12].

2. HUMAN EMOTIONS

Every day humans experience lots of emotions. Feelings or emotions cannot be isolated with clearly defined limits. In some situations, emotions mix with each other. It is a challenge to model such emotions in virtual environments, and because of the complexity of emotions, selected discrete emotions are considered. A textual characterization of a writer’s mood or a facial expression is called an emoticon. Emoticons can improve interpretation of

<table>
<thead>
<tr>
<th>Handset Model</th>
<th>Color Depth</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galaxy P3210</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>Sony HD SO112</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>HTC Desire U</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>iPhone5</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>BlackBerry Q10</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>LG Lucid2 vs870</td>
<td>16 M colors</td>
<td>24</td>
</tr>
<tr>
<td>Motorola XT760</td>
<td>16 M colors</td>
<td>18</td>
</tr>
</tbody>
</table>
plain text. Character usage for emoticons varies across communities or application conventions. Emoticons can be classified by style, western or eastern. Cross cultural studies have revealed that as a group East Asians are more likely than Americans to show a dialectical emotional style [3,10]. Anyway, the existing literature targets at a group level and does not consider how individuals experience different kinds of dialectical emotional styles.

3. EMOTION REPRESENTATION USING COLORS

In the real world, colors are used to represent various emotions and feelings. For instance, in most cultures white denotes purity [11]. Red, and red denotes love or anger, evokes emotions ranging from feelings of warmth to feelings of anger and hostility [1].

The English vernacular is full of associations between color and mood. When people are sad they are blue, they are green with envy, and anger is often associated with red [4]. Colors can also be classified on warm—cool, modern—classical, clean—dirty, active—passive, hard—soft, tense—relaxed, fresh—stale, masculine—feminine, and like—dislike bipolar continua. Experimental results show no significant differences between male and female data for color preferences, whereas different results can be found between different cultures [6–8].

- Black represents unhappiness, sadness and death. In some of the societies people wear black dresses in funerals to show their grief [4].
- Red is associated with emotions like love and danger. It also represents life and vitality [4]. In Japanese society the combination of red and white denotes joy or cheerfulness [11].
- Green is for calming and healing, and because of that the staff in some operating theaters of hospitals wear green color scrubs [4].
- Blue is associated with confidence, security, order or loyalty. In some armed forces blue uniforms convey such emotions or feelings [4].
- Purple represents arrogance and mourning, as well as royalty [4].
- Yellow is also a warm color [1], denoting joy and happiness [4].

In their research, Claudia Cortes, Naz Kaya, Shirley Willett, Johann Wolfgang von Goethe, and Robert Plutchik (Figure 3) have suggested different models for emotion–color mappings [5]. Perceptions of color are somewhat subjective [1], and classifications of colors differ across different cultures [6]. Even though emotion–color models do not follow a standard set of emotions, we used emotions classified by Paul Ekman, as shown in Table 2.

3.1 Emo Sim Architecture

An emotion detector MATLAB program, developed by a fellow student, analyzes voice streams and outputs six weighted emotions. The highest value is considered the current emotion and sent to the CVE server (Figure 6) as an extra parameter. To simulate this process we developed a simple CVE client named “Emo(tion) Sim(ulator)” [13]. Once this parameter is captured from the CVE server, it multicasts to other clients subscribed to the same channel. Another program, called “CVE servant,” a servlet running on top of Apache Tomcat server, which is also a CVE client, caches this value. The extended mobile client cannot connect continuously to the CVE like workstation clients, but instead it can retrieve/send data from/to the CVE servant through HTTP. Once an HTTP request is sent to the CVE servant, it responds with a query string, which consists of position data and extra parameters. These values get updated when the CVE server receives new values. Once a query string is received it is parsed into key (name)–value pairs, in which one of the parameters will be emotion. Emotion value is then mapped to a particular color and displayed on the screen (Figure 1 & Figure 2).

As each user/avatar connects via different channels, this process is repeated for all subscribed channels.

3.2 Emotion Map as Avatar Gesture

In virtual worlds applications an avatar is a projection of a user and as such should have the capability to reflect one’s emotional state to other participants in a virtual

Table 2: Mapping between colors and emotions of different models

<table>
<thead>
<tr>
<th>Plutchik’s Model</th>
<th>Anger</th>
<th>Dislike</th>
<th>Fear</th>
<th>Happiness</th>
<th>Surprise</th>
<th>Sadness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>Purple</td>
<td>Green</td>
<td>Yellow</td>
<td>Cyan</td>
<td>Blue</td>
<td></td>
</tr>
<tr>
<td>Cortes’ Model</td>
<td>Red</td>
<td>Purple</td>
<td>Yellow</td>
<td>Yellow</td>
<td>—</td>
<td>Blue</td>
</tr>
<tr>
<td>Kaya’s Model</td>
<td>Red</td>
<td>Green-Yellow</td>
<td>—</td>
<td>Yellow</td>
<td>Yellow-Red</td>
<td>Purple</td>
</tr>
<tr>
<td>Goethe’s Model</td>
<td>—</td>
<td>Red-Blue</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Blue</td>
</tr>
<tr>
<td>Willett’s Model</td>
<td>Red</td>
<td>—</td>
<td>Yellow</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Summary</td>
<td>Red</td>
<td>Purple</td>
<td>Green</td>
<td>Yellow</td>
<td>Blue</td>
<td>Black</td>
</tr>
</tbody>
</table>
space. Understanding how emotional expressions and “body language” relate to and reflect emotional state, and developing avatars capable of projecting that state, is an extremely important aspect of enabling rich social interaction in virtual environments. Human beings are, by nature, sensitive to facial expressions, postures, and gestures that reflect wide ranges of emotion along a large number of dimensions. On any encounter we can discern whether an individual standing before us is happy or sad, angry or calm, bored, inquisitive, frightened, curious, exasperated. The question is what can be projected through an avatar to convey such a variety of emotional states. We focus on body language, posture and gesture, in addition to facial expression to convey emotional state. Avatars can be developed to project distinct human emotions and show clearly what a user is feeling (or wants to indicate they are feeling), irrespective of the presence of facial features. Wonderland has an embedded text chat system besides its voice chat system. Our technique is to introduce emotional communication to Wonderland to make it more realistic and natural. In our approach, rather than introducing new gestures to avatars of the virtual world, existing gestures were mapped. The basic primitive emotions currently supported are anger, dislike, fear, happiness, sad-ness, and surprise. Table 3 & Figure 4 show the emoticon gesture map implemented in Wonderland.

Our intention was to allow dynamic interaction, so that we extended the text chat to express emotions via animations of the avatar. The emoticons, which a user enters through the chat window, are mapped to predetermined animations and displayed at runtime. The avatar API provides accessory to get the available animations for the avatar depending on gender, by triggering animation with animation name pro-vided the avatar can be animated. Even though the animation of avatars is straight forward, accessing the current avatar is not. We used Presence Manager, another Wonderland module, to find the cells, and mapped the session to the avatar. Basic primitive emotions can also be displayed in Alice, as shown in Figure 5.
Table 3: Animation of selected classified emotions in Alice, Wonderland and Second Life

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Alice</th>
<th>Wonderland</th>
<th>Second Life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Strange</td>
<td>Take Damage</td>
<td>Anger</td>
</tr>
<tr>
<td>Dislike</td>
<td>Demonstrate</td>
<td>No</td>
<td>Dislike</td>
</tr>
<tr>
<td>Fear</td>
<td>Cover Ears</td>
<td>Take Damage</td>
<td>Fear</td>
</tr>
<tr>
<td>Happiness</td>
<td>Give Thumbs Up</td>
<td>Laugh</td>
<td>Happiness</td>
</tr>
<tr>
<td>Sadness</td>
<td>Sigh Disgustedly</td>
<td>Go Home</td>
<td>Sadness</td>
</tr>
<tr>
<td>Surprise</td>
<td>Scratch Head</td>
<td>Cheer</td>
<td>Surprise</td>
</tr>
</tbody>
</table>

(a) Anger  
(b) Dislike
(c) Fear   
(d) Happiness
(e) Sadness 
(f) Surprise

Figure 4: Emotions in Open Wonderland

Figure 5: Emotions in Alice

Figure 6: System Schematic
4. OPEN WONDERLAND & ALICE

Open Wonderland is a 100% pure Java framework originally developed by Sun Microsystems for creating collaborative 3D virtual worlds like “Second Life.” Wonderland is completely extensible: its functionalities can be extended to create entirely new worlds or add new features to existing worlds. Within those worlds, users can share live desktop applications, and collaborate in an education, business, or government context. Organizations can use Wonderland to create a virtual presence to better communicate with students, customers, partners, or friends. Individuals can do their real work within a virtual world, eliminating the need for a separate collaboration tool when they wish to work together with others. The types of collaborations that can happen within the space include audio communication, live desktop applications, and collaborative creation of world content (both graphical and procedural). Open Wonderland has high-quality, lightweight 3D rendering and high-fidelity audio, which immerse users in synthetic worlds, and its extensibility inspires developers to extend its functionality.

Alice is an innovative 3D rapid prototyping-programming environment that makes it easy to create an animation for telling a story, playing an interactive game, or a video to share on the web. Alice is a freely available teaching tool designed to be a first exposure to object-oriented programming. It allows one to learn fundamental programming concepts in the context of creating animated movies and simple video games. In Alice, 3D objects (e.g., people, animals, and vehicles) populate a virtual world and users create a program to animate the objects. In Alice’s interactive interface, drag and drop graphic tiles create a program, where the instructions correspond to standard statements in a production-oriented programming language, such as Java, C++, and C. Alice allows users to immediately see how their animation programs run, enabling them to easily understand the relationship between the programming statements and the behavior of objects in their animation.

5. SECOND LIFE

Second Life (SL) is a well-known virtual world, developed by Linden Lab, which simulates real (‘first’) life by providing a space for rich interactions and social events. A free client program, called Second Life Viewer, enables its users, called Residents, to interact with each other through avatars, delegates of the users logged into SL. These avatars are highly customizable. SL allows residents to have real world-like experiences, including strength-ening interpersonal relations, participating in various individual and group activities, and creating and trading virtual property and services. SL even has its own economy, based on the currency “Linden Dollars.” SL has a considerable collection of built-in emotional animations, which can be triggered by sending Internet Relay Chat (IRC)-like commands, pressing a HUD (head-up display) button, invoking via a Linden script, or triggering an action. Some built-in animations, such as “express fear,” directly express the respective emotion, but most animations are not well categorized, although many can indirectly express emotions. For example, kiss blowing can suggest affection and laughing can suggest happiness. Anyway, one can easily assemble a set of animations to represent whatever emotion is needed. In Second Life terminology, a “gesture” refers to an animation or collection of animations. Second Life residents can create gestures with a few clicks in the SL viewer interface, as shown in Figure 8 & Figure 9.

Also, one can wear emotion HUDs, which can be downloaded. Many emotion-based animations are available from the Second Life official website and other third-party websites. Those files, rich with gestures, postures, and facial expressions, are in the Biovision Hierarchical motion capture data (BVH) format. A .bvh file is a text file that has data captured from a moving skeletal system. (Another name for this type of data capture is “Motion Capture,” abbreviated “mo-cap.”) A .bvh file has basically two main parts: the hierarchy of the skeletal system, and the coordinates of the individual movements. When highly customized animations are needed, software-based or motion capture-based .bvh files can be uploaded.

The current version of Alice models body language, but not facial expression. Because of sponsorship of game publisher Electronic Arts, Alice galleries include assets from the popular “Sims” video games. Through the Collada digital assets exchange (.dae) interchange file format, such objects can also be imported into Wonderland scenes, including via runtime drag-and-drop.

6. PUPPETRY

Puppetry involves giving life-like characteristics to inanimate objects. Puppeteers’ artists create movements, voices, and personalities for them. There are many different varieties of puppets, made of a wide range of materials, depending on their form and intended use. They can be extremely complex or very simple in their construction. Puppetry is used in almost all human societies both as an
entertainment in performance and ceremonially in rituals and celebrations such as carnivals. Our approach is use computer keyboard to animate objects. In traditional puppetry, people use string or sticks to control each component of object (doll). In our system avatars are dance according to rhythmic data. We can make some additional movement-using keyboard. It is kind of new puppetry dance, as shown in Figure 7.

7. CONCLUSION

By extracting emotions from voice, better immersiveness can be conveyed, especially if such operations are performed in real-time. We developed a system, which can relay emotions to a CVE server so that each avatar’s emotion can be displayed along with location on a mobile screen. We would like to improve our system by handling mixed emotions. In this case we could display emotions as pie charts, which seem natural as signal indicators. We also consider porting our program to run on different phones, such as Google Android or Apple iOS iPhone/iPad. We have developed a mobile application, which can display human emotions using colors. Basically we recognize six different emotions, getting the highest probability emotion, mapping the emotion to a carefully selected color, and displaying it on a DoCoMo mobile phone. Accuracy of recognition can be improved by incorporating natural language syntax and semantics cues, other than purely acoustic data. An extension of this system could also be used with mobile applications whose user-represented avatars can communicate with each other, displaying avatar emotions on mobile screens as colorful icons.

Words are only a shallow representation of ideas, and facial expressions and body language are indirect
expressions of internal emotional state. Nevertheless, we can augment a primary, linguistic channel, such as voice chat or text chat, with a secondary, paralinguistic emotional channel featuring affect-triggered body language for avatars. Our experiment on emotion expressing of the three exercised systems described here (Table 3), SL is the most mature, and has built-in facial expressions and body language. Existing (bundled) animation can change one’s face, which suffices for expressing simple expressions. There is also software to compose animation, “AVImeter,” but it lacks parameterizability, so, for example, to represent quantized low, medium, & high intensities, it would be necessary to precompose three separate gestures. Crossing such articulation with other, independent controllers would combinatorily explode, making such an approach impractical for all but sparsely sampled combinations SL can be Animated with motion capture (“mo-cap”) data — self-captured (Figure 10), from a gallery such as OpenMoCap.com, or composed in a cad application such as Maya, 3DStudio/Max, or Motion Builder.

Open Wonderland is distributed at the lowest level of specification, and in some sense is the most flexible, but it is also the most demanding development-wise. We hope to use it to experiment with adjustment of an “idling” gesture. Whenever an avatar is not actively moving, via locomotion such as walking, swimming, flying, or being conveyed or transported via virtual vehicle, it can express a subtle “heartbeat,” a quiet postural sway that serves as a background or default “sign of life.”

“Pad,” acronymic for pleasure/arousal/dominance, is a well-known and oft-cited taxonomy for emotions [14], as shown in Figure 11. Emotions can be characterized by fitting them into the appropriate octant, with respective dimensional intensity scaled by displacement from the origin. These mappings are admittedly naïve or simplistic, but perhaps representative of articulatory control. The basic idea is analysis of signal, in particular realtime voice, extraction of features, inference of voice quality, estimation of emotion, and appropriate expression of such affect through figurative avatars in an online environment.

REFERENCES
empowering collaborative virtual environments with emotions. His current research interests include mobile Computing, HCI and Information Systems, the University of Aizu, Fukushima, Japan. He has worked as a Teaching Assistant, the Graduate Department of Information Systems, the University of Aizu, Fukushima, Japan. He received a master in Computer Science & Engineering from the University of Aizu, Fukushima, Japan. He has worked as a system engineer at Ridgecrest Financial Inc. in mobile interfaces, In Proc. HC-2010: 13th Int. Conf. on Humans and Computers, pp.28-31, Aizu-Wakamatsu, Japan, Dec. 2010.


Prabath WEERASINGHE
Prabath Weerasinghe is a Doctoral Student in the Department of Management and Information Systems Engineering, Nagaoka University of Technology, Nagaoka, Japan. He received a master in Computer Science & Engineering from the University of Aizu, Fukushima, Japan. He has worked as a system engineer at NS Service CO., LTD. in JAPAN from 2007-2010. He was a Research Assistant, the Graduate Department of Information Systems, the University of Aizu, Fukushima, Japan. His current research interests include mobile Computing, HCI and empowering collaborative virtual environments with emotions.

Ashu MARASINGHE
Ashu Marasinghe is an Associate Professor in the Department of Management and Information Systems Engineering, Nagaoka University of Technology, Nagaoka, Japan and visiting Professor at International University of Japan. Prof Ashu obtained his B.Sc. degree in Physics and Mathematics from the University of Colombo, Sri Lanka, M.Sc. in Computer Science & Engineering from the University of Aizu, Japan and Ph.D. in Computer Science & Engineering from the University of Aizu in 1997, 2001, and 2004 respectively. He has worked in the University of Aizu, Japan. Prof Ashu’s research interests are: Engineering Management, Construction Management, Intelligent Control System, Multi Agent System, Ubiquitous System, Social Informatics, and Kansei Engineering. Prof. Ashu served as a General Chair in several International Conferences include ICBAKE 2009, HC-2008, KEAS’08, CDG’08, KEIS’06

Rasika RANAWEERA
Rasika Ranaweera is a Doctoral Student in the Graduate School of Computer Science and Engineering, University of Aizu, Fukushima, Japan. He received a master in Computer Science & Engineering from the University of Aizu, Fukushima, Japan. He has worked as a system engineer at Ridgecrest Financial Inc. in mobile interfaces, In Proc. HC-2010: 13th Int. Conf. on Humans and Computers, pp.28-31, Aizu-Wakamatsu, Japan, Dec. 2010.

Senaka AMARAKEERTHI
Senaka Amarakeerthi received his B.Sc. degree in Computer Science from Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka in 2000 and M.Sc. degree in Computer Science from the University of Peradeniya, Sri Lanka in 2006. He has obtained his PhD degree in Computer Science and Engineering from University of Aizu, Japan in 2012. He has been serving as a lecturer at Sabaragamuwa University of Sri Lanka since 2006. His current research interests include mobile Computing, HCI and empowering collaborative virtual environments with narrowcasting.

Michael COHEN
Michael Cohen received a Sc.B. in EE (electrical engineering) from Brown University (Providence, Rhode Island) in 1980, M.S. in CC (computer science) from the University of Washington (Seattle) in 1988, and Ph.D. in EECS (Electrical Engineering/Computer Science) from North- western University (Evanston, Illinois) in 1991. He had post-doctoral appointments at the University of Washington and at NTT. He is a visiting professor at the Human Interface Technology Lab, University of Washington (Seattle). Currently, he is a Professor of Computer Arts Lab at University of Aizu, where he is also Director of the Information Systems and Technology Center (ISTC). He has interests in telecommunications semiotics, groupware and CSCW (computer-supported collaborative work), audio windows and stereotelephonics, computer music, digital typography and electronic publishing, hypermedia, and virtual reality.