A Multimodal Contact List to Enhance Remote Communication

Andrew Crossan[1], Grégoire Lefebvre[2], Sophie Zijp-Rouzier[2], and Roderick Murray-Smith[3]

[1] School of Engineering and Built Environment,
Glasgow Caledonian University, Glasgow, Scotland
andrew.crossan@gcal.ac.uk
[2] Orange Labs, Meylan, France
{gregoire.lefebvre,sophie.zijprouzier}@orange.com
[3] School of Computing Science
University of Glasgow, Glasgow, Scotland
roderick.murray-smith@glasgow.ac.uk

Abstract. The Multimodal Contact List provides a mechanism to browse context information and communicate with friends in a contact list both visually and through touch and sound. Each contact can share with their friend group selected information on their current context such as mood and availability. Users are able to gain a quick overview of the context provided over all of their contacts, allowing them to close the loop with them via touch, audio, visual feedback or a combination of all three. A user can then progressively probe the contact for more detailed information, eventually allowing the user to open a real-time multimodal voice and tactile communication channel to the contact for verbal or discreet tactile communication.

Keywords: Multimodal, Mobile, Remote Communication, Vibrotactile Feedback.

1 Introduction

Mobile devices provide the ability to stay connected and engaged with our friends, family and work colleagues in a wide range of contexts. They can now be seen as small, general-purpose, portable computers, but the ability to communicate with contacts has remained one of their most important and compelling features. Some of this connectedness now occurs through online social networking such as Facebook and Twitter, as well as the more traditional voice and SMS communication. The communication offered by these devices through voice or text however is in some ways limited in expressiveness due to its inherently remote and distant nature. Messaging systems attempt to get around this lack of expressiveness through the introduction of emoticons. Similarly, in remote voice communication the non-verbal cues from the context as well as contextual cues are not present. Humans are extremely good at reading these non-verbal cues in face-to-face conversations that can

R. Murray-Smith (Ed.): MSSP 2010, LNCS 8045, pp. 84–100, 2014.

guide the tone and quantity of communication as well as affecting the meaning. King [10] argues that context is fundamental to communication. In mobile communications, we lose a lot of the context information that we take for granted in everyday conversations. Partially, this relates to the loss of non-verbal cues during a conversation. However, King's definition of context includes factors such as psychological, environmental, cultural and situational contexts that can play an equally important role during communication. For example, the quantity and quality of communication might vary dramatically between situations where one of the parties is in a home or work environment, or when external sources of noise might interfere with the communication, or even when the remote person is engaged in some other activity that requires concentration. While these factors might be easy to discern in a shared communication space, important cues are lacking when the two parties are in different locations.

Sellen [15] demonstrates, however, that even in situations where both good quality video and audio are present, remote communication mediated through technology can result in communication that has a tendency to be more formal and less spontaneous suggesting that even with some context information, the remote nature of the communication can have an effect.

1.1 Scenarios

Here we propose a system where a user can choose to display certain context information (such as happiness or availability) to their friends through a multimodal contact list application. The ability to determine whether someone would be available for a phone call without calling is one immediate benefit of sharing context, helping prevent a contact's phone ringing at inappropriate times such as during a meeting. There are also for more subtle situations where privacy or discreetness might be an issue. Similarly, there are many situations where voice communication is impossible (e.g. on a loud subway), inappropriate (e.g. in a meeting or at cinema) or simply not desired (e.g.in public [13]). In these cases, non-verbal communication (touch-based device for example) could provide a private channel for communication. If a user sees a contact is unavailable, he may choose to tap out a message on the screen that the remote contact will feel as vibrotactile message. The length and rhythms of taps being sent might be used to convey some information to the remote contact, allowing expressive, discreet communication with the potential for people to develop simple personal tactile languages.

Shared context also allows users to filter the contact list according to some parameters, for example to find out about people's availability and general mood. So assuming the scenario of a night out, a user may look for friends to go out by filtering for contacts that are happy and not busy. Contacts are then filtered according to these input parameters, thus showing people who are most likely to be interested in going for a night out.

There is also the potential to multiplex the different channels of communication to provide either context or emphasis at different points in the communication. During a phone call, feedback presented through the standard vibrotactile actuator in a phone could be used to convey the emotional state of a caller or more contextual information for the remote contact in the communication displaying for example whether he or she is walking and how quickly, or whether they are communicating through a headset, or in a car, for example.

The multimodal nature of the communication also offers possibilities for providing an accessible means of communication. Touch-enhanced communication opens another potential remote communication channel for people with visual or hearing impairments or both. The addition of an expressive vibrotactile channel would provide obvious benefits over preset vibration patterns, allowing users to customise their messages and develop more complex communications.

2 Related Work

There are different mechanisms that have previously been used to attempt to convey more context information in communication (see [5]). One common mechanism is the social network ‘status update’ that allows a user to display a general purpose piece of information to a group of friends. Here, the user manually enters whatever information they choose into the message. With the availability of more sensors in phones, we are starting to see more systems that take advantage of automatically sensed context. Location aware systems would be the most common example where applications such as maps or navigation aids can use the user’s location to present context sensitive information. Google’s Latitude system now extends this location awareness to incorporate social networking features where friend groups share their locations. Many users are now willing to share more context information with selected friends or colleagues, leading to social features being incorporated into a growing number of applications, allowing users to maintain an awareness of their friends’ and families’ days, even if they do not get a chance to meet up with them.

The remote communication problem extends to issues of the availability of conversation partners. It is difficult to get an indication of whether someone is free to be contacted or whether current situational and environmental factors will make it difficult or socially inappropriate to respond. Researchers have started to address these issues in a number of ways. In instant messaging systems, they adopt the strategy of allowing users to explicitly state whether they are available or not. This concept has been extended to more general communication in systems such as the BuddyClock [9]. BuddyClock is an augmented alarm clock that allows people to know whether it is appropriate to make a call by alerting a friend group whether a person is awake or asleep. Using this friend group feature, it further incorporates social networking aspects to allow people to share sleeping behaviours with each other.

More recently, in a mobile setting, the Feelabuzz system [18] provides a mechanism for communicating continuous awareness signals between sensor-enabled mobile devices. An accelerometer is used to measure movements and these data are transmitted in real time to a receiver in a direct, non-abstract way, without the use of pattern recognition techniques, in order not to destroy the 'feel'. The goal is that over time, a user would learn to recognise common vibration patterns such as walking or on a bus. This enables direct communication as well as implicit context communication with display to the receiver through vibrotactile actuators. The real-time aspect of the communication allows the context information to enrich and augment voice communication between remote communicators.

In related work, Murray-Smith et al. [12] examined how to augment voice communication using the vibrotactile channel to indicate to each participant the walking behaviour of the other participant. They were able to demonstrate how this additional channel could affect synchronisation of the step rate between participants with the results being different for spontaneous and scripted communication. Brown and Williamson [3] allow users to send audio/tactile messages to each other using gestures. These gestures can be used to represent common events such as 'home safely' and are translated into preset audio/tactile messages that the remote user can learn to interpret. There have been a number of systems based on asynchronous haptic messaging [1, 4, 6, 14]. Chang et al. [4] have examined enhancing communication through vibration, converting hand pressure on the sender's phone directly to the receiver, mapping pressure to vibrational intensity. Again this real-time tactile channel is used to allow the user to add expressiveness to the communication. They further allowed symbolic communicative function through a tactile interface that is attached to a mobile device and utilizes encoded haptic patterns in communication. The first claimed advantage of adding the tactile channel was redundancy between voice and haptic, for emphasis. The second was the ability to incorporate mimicry of haptic patterns to indicate attention and camaraderie. At the minimum, it represents an additional communication channel to provide a means to incorporate turn taking signals into the conversation. This area is of particular importance for interfaces looking to support intimate interactions.

There have been many examples of systems that aim to support the communication of emotion. As so much is communicated non-verbally when face-to-face, research has looked to replace these non-verbal cues to both communicate more accurately the intention of the speaker, but also to support couples in a relationship. Wang, and Quek describe Touch & Talk where squeezing a customised mobile device causes a remote armband to squeeze the arm of a loved one [19]. This example demonstrates the potential of tactile feedback for affective communications in a remote context.

Similarly, recent work by Hemmert et al. [7] uses unconventional interaction mechanisms to support couples in a long distance relationship. Users interact by grasping, kissing or whispering into a custom device. The sensations are conveyed to the remote user through tactile sensations with a band that contacts around the user's hand communicating a squeeze, air jets communicating a whisper and a semi-permeable membrane and water conveying a kiss. These sensations are designed to be as close a representation of the sensation as technology allows, and as such requires specialist hardware.

Kontaris et al. [11] convey emotion through spatially distributed vibration patterns designed to supplement video communication. The patterns were generated by gesturing on a custom touch surface, with the vibrations being displayed through a 4x4 array of tactile actuators. The authors suggest that custom gestures could eventually form part of a personal tactile language between the couple.

Kaye [8] examines a different way of supporting couples in remote relationships without the need for custom technology. His minimalist approach uses a 1-bit communication channel initiated by clicking on a button. Although minimal information was conveyed, Kaye's study suggests that a rich channel of communication can be built over time by incorporating context. Alternatively, Bales et al. [2] use explicit codes for building tactile messages that support a couple's awareness of their partner's location. By defining recognisable vibrotactile patterns for arrive and depart, they could combine these cues with locations such as home or work to build a tactile message in a language that can be learned by both partners in the relationship.

One common factor in all these systems is that they use touch as an output channel. Touch provides a very personal communication channel that naturally lends itself to supporting affective means of communication.

Another relevant are of research is automated sensing of context. For example, Williamson et al. [20] demonstrate how accelerometers could be used to determine context of a user in a usability study during the morning commute. They were able to determine whether the participants were stationary, walking or using public transport.

We may also choose to automatically sense the mood of the content of a communication. Shirazi and Schmidt [16] suggest how audio alerts shaped by the content of a message can be used to provide non-visual feedback to the user. Their work identified four different types of common messages, happy messages, sad messages, questions and answers and responses. Through identification of these different types of messages using emoticons, punctuation and common words, the system can give the user an audio preview of the message content or mood.

3 The Multimodal Contact List

Here, we augment both the representation of a user in a contact list, as well as communication between the users to provide a mechanism for sharing context information. The goal is to provide a system that allows a user to initially get a brief overview, or glimpse of the status of users within their contact list. Further, the system attempts to provide a way to probe for more detailed information about a contact, and eventually open a multimodal communication channel with that contact. We choose to augment the contact list to display this information as it is an application already used for more traditional forms of communication and the user's contacts are already collected in one place.

The Multimodal Contact List has been implemented for Android Phones (2.0 and above). Figure 1 shows the main interface. The header describes the current situation: the last event status (here "New Mood Vector" message from Gregoire) and the current tactile communication status (i.e. communication enabled/disabled with voice on/off). Each entry in the contact list represents a user and provides some context information about that user visually through a texture and, when explored on screen, through the phone's internal vibrotactile actuator. This context information is stored as a texture with parameters controlled by some aspect of the users' context. These context parameters can be automatically inferred from an automatic classification of a combination of sensor readings, or could be controlled explicitly by the user. In this example we use the concept of a 'Mood Vector' associated with each user. This vector can be explicitly set by the user to indicate different aspects of their current mood with each parameter affecting the visual appearance and tactile representation of the contact. Currently the four parameters of the mood vector are 'Joy' (to indicate happiness or sadness), 'Aggression' (to indicate calm or stressed states), 'Mobility' (as an indication of their recent motion and mobility characteristics) and 'Busyness' (as an indication of their current workload). Similar parameterizations of mood have previously been used successfully in the widely used Moodagent music player that builds playlists based on the user's mood (moodagent.com). Each parameter can be set explicitly using the 10-point sliders shown in Figure 2. Eventually, the mood vector concept could be extended to inferred mood using sensor information such as phone location or acceleration activity, along with other contextual cues such as calendar entries, current music playlist mood content, or whether the contact is engaged in a call or not.

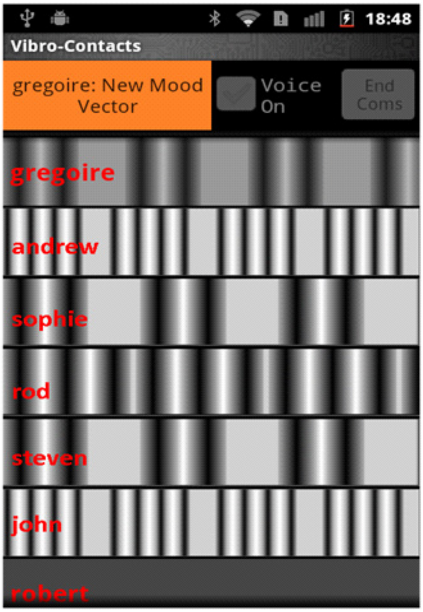

Fig. 1. The main Multimodal Contact List interface. The status bar at the top shows the most recent message received and the current communication status. Below is the list of contacts with mood vector indicated by their texture.

The parameters are then mapped to a visual and tactile texture displayed to the user for each contact when browsing the contact list. The visual texture initially provides a user with a low effort mechanism for gaining a fast overview of the general context information shared by their friends allowing the mood of each contact to be ascertained (for instance the contact named Robert is offline; consequently its visual texture is black). For more detailed interactions with a particular contact, we use a 'modality scheduling' mechanism [6]. When the user is interested in more detail about one particular contact, they can use a low-attention approach, where they interact through the phone's touchscreen, with increasing engagement with a contact leading to increasingly detailed multimodal feedback, and eventually tightly-coupled direct communication. Interactions with the contact list can be separated into vertical swipes (to move up and down the contact list), horizontal swipes (to feel the texture of a contact's Mood vector), and horizontal exploration (moving along a contact to find out more detailed information and eventually open a communication channel).

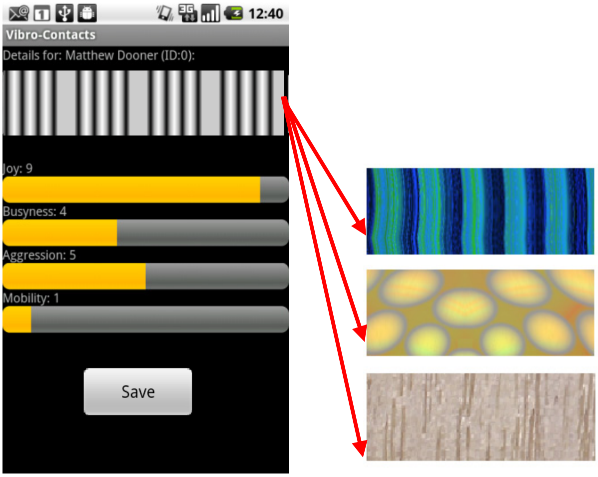

Fig. 2. Users set their Mood Vector parameters between 1 and 10 using the four sliders shown (Left). There are many ways this vector can be mapped to an arbitrary texture (right).

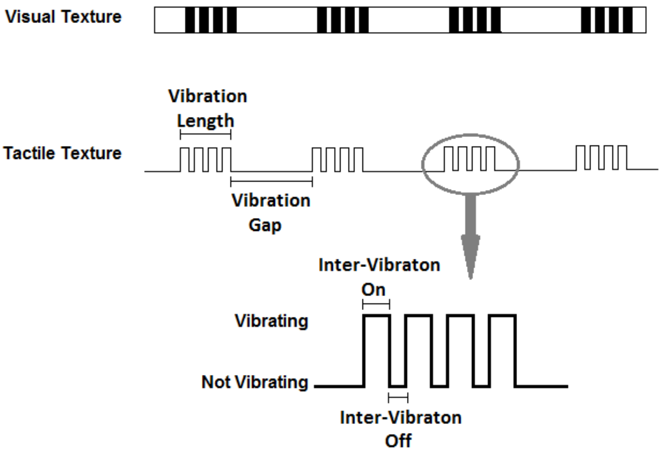

Fig. 3. An example of the visual and vibrotactile textures used. The four parameters of the texture are varied as the mood parameters change.

3.1 Mapping the Mood Vector to a Texture

There are many different ways we could map an arbitrary vector to visual and tactile texture. For example, using an approach such as physical modeling, rich representations of a vector can be built that vary with any number of parameters. The specific mapping should be tuned to best fit the technology available. In the current implementation, we have designed it for use with a phone's standard internal vibrotactile actuator.

Our current mood vector has four parameters that are mapped to a visual and tactile texture. Our texture is constructed through short regular pulses of vibration patterns. We define the four parameters of our texture as follows: the Vibration Length (VL), the Vibration Gap (VG), the Inter-Vibration On Length (IVOn), and the Inter-Vibration Off Length (IVOff). Figure 3 shows how each of these parameters are mapped to the vibrotactile signal. VL and VG are used to set the low frequency of the texture, with IVOn and IVOff contributing to the higher frequency components of the vibration that affect the perceived roughness of the vibration signal. Here we make simple mappings from these characteristics to the moods: IVOn length increases as 'Joy' increases, the VL increases as 'Aggression' increases, IVOff increases as 'Mobility' increases, and VG decreases as 'Busyness' increases. Consequently, when user is joyful and in urban transport, other contacts can feel a vibration intense with a high impulse rhythm. In opposition, when this user is sad and in his office, the vibration is smooth and slow. When the user aggression and busyness are high, pattern vibration are long and gaps between them are short revealing user is occupied and has few time for communication.

Likewise, the visual texture is built with ridges and grooves from a gradient pattern from the previous four parameters. The light grey corresponds to the vibration gap, and the grey levels reveal in black the grooves and in white the ridges.

3.2 Tactile Communication

The paper [11] proposes a method for creating tactile textures without force feedback by using a simple motion sensor and a single vibrotactile actuator. This proposal is based on wavetable synthesis driven by the user's hand movements. The results show envelope ridge length and spatial density were distinguishable design parameters and ridge length and spatial density influence perceived roughness and flatness similarly as with real textures.

We use here three types of tactile feedback automatically generated by user finger movements. Firstly, the tactile channel is used for list manipulation: the user can feel a short buzz for each contact moved over when scrolling up or down the contact list (vertical touchscreen movement). Secondly, the tactile channel is used to display the remote contact's mood vector providing context awareness (horizontal touchscreen movement along the contact item): when the user glides his finger along a contact item, he can feel a vibrotactile texture representing the mood vector of this contact, sharing their current mood. Finally, we use the tactile channel for inter personal communication augmenting voice communication (moving back and forward along a contact item): when the user strokes the contact item, the remote contact receives a (visual or tactile) notification. The remote contact can then open the tactile communication channel by stroking the user's contact list entry. Then they can start a tactile dialogue, playing with the duration of vibration and duration of silent: each time the user touches his screen, the device of the remote contact starts to vibrate. When the user stops touching his screen, the remote device stops vibrating.

The tactile communication can either take the form of preset patterns of vibration that might be customisable to a user, but have a specific meaning to that user (such as a notification for a particular contact being unique). Alternatively, real time communication allows us to provide a channel to support the more affective aspects of the communication. Initially, as one user explores the context of a contact, the remote contact may be made aware by feeling a tactile representation of the finger movements of the user across the touch screen. Mapping the vibration pattern to finger speed could provide a simple but potentially expressive method of interaction, with the potential to allow users to develop their own tactile languages.

3.3 Real Time Communication

With the advent of the *eXtensible Messaging and Presence Protocol* (XMPP – **www.xmpp.org**), the transfer of generic data in real time is now available on a wide range of devices. Traditionally, this protocol is used for instant messaging and VoIP applications, however we can take advantage of this infrastructure to start transmitting different types of data. Here we use XMPP to provide a mechanism for communication of generic data across a communication channel in real time. Currently, these messages are received either as Voice Over IP (VOIP) messages or in coded tactile form allowing simultaneous tactile and audio communication.

Figure 4 presents our real time communication architecture. The local phone uses sensor information as microphone or finger displacement on touchscreen to build a message. The message is processed and compressed to follow the XMMP recommendation. When receiving the message, the remote phone parses the input stream and decompresses the content to resituate the initial information through visual, audio or tactile feedback.

4 Evaluation

Prior to user testing, we evaluated the real time nature of the communication in different circumstances with particular focus on the latency of the communications. We do this over WiFi and 3G as well as in close proximity (both participants in the same city) and over a larger distance (where one participant is in the UK and one in the France). We also examine latency during low bandwidth (in this case tactile messages) and high bandwidth (when the audio channel is open) communication. Table 1 shows mean latencies in the different situations. Unsurprisingly, there is a larger latency when messages are being sent over longer distances and lower bandwidth communication channels. The round trip time of around 200ms in these instances translates to a network latency of approximately 100ms for a communication. This value might be acceptable for one-way communications and manageable for standard voice communication, but may cause issue in any more complex interactions where users try to synchronise their movements. One example of this would be a more complex version of the tactile feedback where supporting a remote relationship with both users simultaneously sending messages simultaneously where the timing of the messages may become important to any personal tactile language developed.

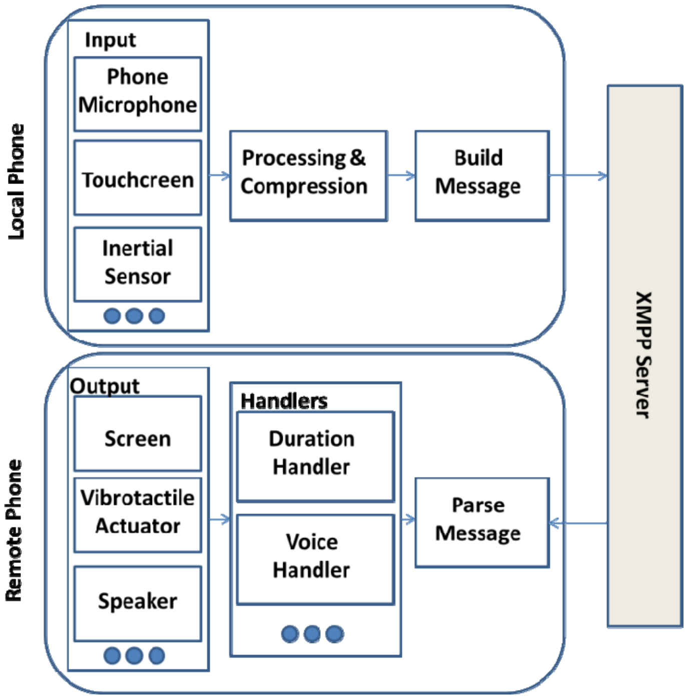

Fig. 4. The real time communication architecture

In the close proximity WiFi situations, a one-way latency of less than 20ms may lend itself better to more complex real time interactions.

We have completed an early user evaluation of the system. This took place in two stages; firstly to examine usability and user acceptance of the interactions and secondly to test the system with a small group of potential users.

4.1 Focus Groups

The first session took place as a focus group between 3 pairs of friends. The goal here was to test their acceptance and understanding of the application. The focus groups were used to generate discussion on potential usage and issues with the system. The two users were co-located but each had their own phone (Google Nexus One) with the app installed. The interactions were first described and demonstrated to the participants and then they were given time to explore the interactions and communicate with each other. A discussion of each of the features of the system and their potential uses was then held. Each focus group session took around 30 minutes.

While the reception of the technology was extremely positive, these sessions raised some potential issues that users will face using the multimodal contact list. These finding we feel will generalise to any multimodal communication system that attempt to use technology to communicate in unusual ways. The four main findings are described below.

Table 1. Round Trip Time (ms) data for different situational contexts and communication channels. Close refers to two users communicating in the same local area, where as remote refers to communications between two users in the UK and France. Low data load refers to situations where tactile message were being sent and high data load refers to RTT during audio communication.

	Round Trip Time (ms) WiFi - WiFi	Round Trip Time (ms) 3G - 3G
Close (low data load)	37.4 (std. dev. 19.6)	153.8 (std. dev. 38.8)
Close (high data load)	33.1 (std dev 22.6)	223.7 (std. dev. 59.6)
Remote (low data load)	96.6 9 (std.dev. 19.2)	N.A.
Remote (high data load)	163.2 (std. dev 34.9)	N.A.

4.1.1 Appropriate Metaphors Are Required

When developing novel methods of communication, we must be aware that users are not use to communicating in this manner and the emphasis is on the designer to provide methods that allow users learn new interaction techniques. One way of supporting this learning is by using interface metaphors. In this application the users stroked the appropriate contact onscreen to open a tactile communication channel. The contact then lit up on the screen indicating to the user that communication could begin. This stroke interaction and lighting up effect proved confusing. We could potentially take cues from face to face communication where a tap on the shoulder or nudge is sometimes used to alert the other participant discretely that you want their attention. Stroking an onscreen contact may be more appropriate for more intimate communication (between a husband and wife for example) and less appropriate when chatting to friends or work colleagues.

Additional feedback could support the metaphor when the channel is opened. For example, showing the user's face to indicate that you now have their attention.

4.1.2 Be Wary of Using Touch as a One Way Channel

Touch is naturally a bi-directional channel. We tap someone on the shoulder to get his or her attention, but we also received feedback through our haptic sense that we have performed the action. Here we use touch communication as two separate one-way channels. We open a tactile channel and send tactile feedback when the user presses the flat touchscreen. There was a strong sense in the focus groups that this one-way communication was not appropriate. Even co-located, participants were sending a signal and then asking “Did you feel that”. This lack of feedback adds a level of confusion to the communication that is rarely an issue with channels such as audio. One immediate change to the interface that was made was to include visual feedback to the user that their tactile message was being sent. A more complete system might also use phone sensors on the remote side (such as accelerometers or capacitive sensors) to determine whether the remote user is holding their device and communicate this back to the sender to close the loop and complete the communication channel.

4.1.3 Unusual Combinations of Technologies

Voice and vibration are rarely used together in phone communication. In this case the vibration interfered with the voice communication, particularly if the device was on a hard surface. An audible vibration could be heard through the audio communication channel. While this was an issue for this experiment, we do not believe this is a fundamental issue, as the problem of crossover and feedback has been dealt with in normal speakerphone design, and it should be possible to develop appropriate filters.

4.1.4 Overloading the Vibrotactile Channel

In the real world, the tactile channel is relatively high bandwidth and can detect subtle differences in the shape, texture, softness or temperature of objects. The standard vibrotactile device on phones severely limits this communication channel providing limited control and generally a single actuator. In this application we use the vibrotactile channel for different purposes such as list scroll events, receiving one or more tactile messages, and when exploring the mood vector. The key issue here is ‘how does a user distinguish one vibration from another?’ We can choose to use different vibration profiles for each, however the poor level of control for the actuator limits what can be achieved here. There is also the issue that two simultaneous events (such as two tactile messages sent by different users) will interfere. It is difficult to separate these events one tactile message from an ongoing communication with another from a new user without resorting to visual or auditory feedback. Care must be taken when rely on the tactile channel for multiple purposes that the user can distinguish between the separate events.

4.2 User Evaluation

We evaluated a modified version of the Multimodal Contact List app. The app was modified to better support feedback during communication, indicating more clearly when a communication channel was open and when tactile messages were being sent. A group of seven users took part in the study over a period of five days. Six users were work colleagues and friends while the seventh was in a remote relationship with one of the other participants. Each user had the same contact list containing all seven users of the system including them so that they could see their own mood vector and test the system by sending messages to themselves.

For this user study, we used a simplified version of the mood vector. We map two parameters onto the texture: Joy and Busyness. Joy is increased by increasing the 'Inter-Vibration On' parameter and Busyness is increased by increasing the 'Vibration Gap' parameter. Examples textures generated from this mapping are shown in Figure 5.

Training was provided to each of the users taking part in the study with a user manual describing how to perform each of the interactions being provided along with a training session where the experimenter demonstrated the different features of the application and then asked the user to try each of the interactions until they were satisfied that they could set their mood vectors, and send audio and tactile messages successfully. Results were collected at the end of the study through a questionnaire, which probes user acceptance of each of the features and their attitudes on how these ideas could be used in a wider context.

Here we discuss the main findings from the study for each of the functionality and an overall view of the goals of the app separately.

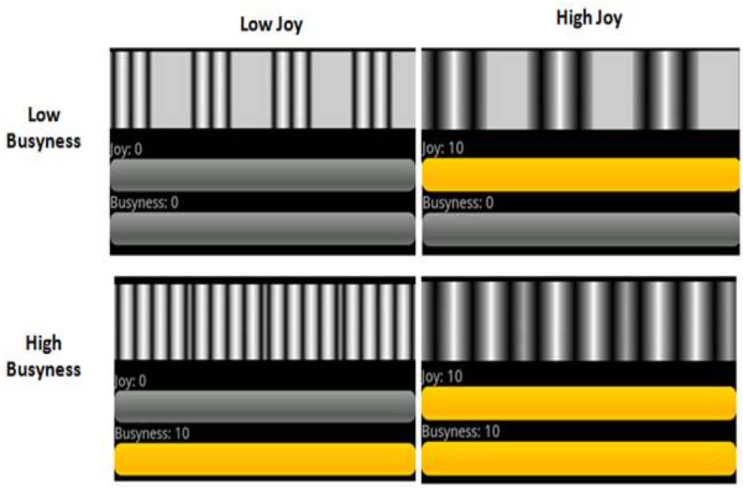

Fig. 5. Examples of the extremes of the mood vector used during the study. This application mapped 2 parameters onto the texture; Joy and Busyness.

4.2.1 The Mood Vector

The participants generally received the mood vector negatively. One participant "found the mood vector to be the least useful of the functions". The mapping chosen was described as confusing by more than one participant, which may have contributed to it being updated infrequently. This infrequent use again did not help reinforce the mapping with the users leading to problems remembering the mapping once learnt. Further to this, users seemed to use the visual mood vector but had little need for the tactile version.

The abstract nature of the mapping forced users into an extra ‘interpretation’ stage that could potentially have been avoided through an iconic representation that had a semantic link between the image and the user’s mood. The goal of this feature was to allow users to gain a lot of information about the general mood of their contacts with a visual glance. A better semantic link using visual properties quickly and easily identified with a glance may better support this goal.

4.2.2 Vibrotactile Messages

All users perceived Vibrotactile messages positively. There was a general feeling that it provided a discreet means of communication that would be appropriate in a number of different situations both in work contexts and when with family and friends. It was seen as a means to initiate a conversation in real time and potentially probe contacts about their availability to communicate.

There were reservations with the difficulty of communicating more complicated messages through vibration. With extended use, some “vibration code” may develop between participants but this was not evident in this short trial. Users drew attention to the fact that vibration was not suitable in all instances. Firstly, users felt that the vibrotactile feedback was a very personal way of communicating and not appropriate for unknown contacts.

Secondly, it was not considered appropriate for emergencies situations where it was important the other contact got the message. The discreet nature of tactile feedback is advantageous when communicating in situations where a lot of motion or noise may be inappropriate. However, this also leads to the fact that if the user is not touching the phone, the message is easily missed. Potentially, this could be resolved by using other sensors in the phone (such as accelerometers or capacitive sensors) to detect whether the user is holding the device and can therefore perceive the vibrotactile message.

4.2.3 Voice messages

Users were obviously more used to voice communication with phones. One innovation with this system however is that users could initiate a voice call without calling the contact first. Participants felt that some form of notification and acceptance was essential to avoid noise in inappropriate settings. The couple in a remote relationship particularly highlighted this. They chose to use the system in a playful manner “like two kids with walkie-talkies”. They used it to supplement Skype communication combining voice and vibrotactile messaging with a video feed.

There were also concerns about the bandwidth required for VOIP. This was to some extent mitigated by allowing users to restrict network communication in situations where no WiFi connection was available, however as this was a global setting on the app this also did not allow communication using the other modalities. Allowing more control to the user over the connection to allow low bandwidth forms of communication while blocking high bandwidth data would resolve this issue.

4.2.4 General Perspectives

There were a number of changes and additions to the available functionality suggested by the users. The major suggestions for change revolved around improvements to the mood vector to provide useful and easy to interpret feedback on the current context of contacts. This could be supplemented by text to support learning of the mapping and provide more detailed information when required. Availability for communication was a key factor here. This is a key feature missing from standard phone communication that is available and plays a useful role in many other forms of remote communication such as Instant Messaging and VOIP systems.

The mood vector concept could easily be adapted to support the user in their choice to communicate with a contact just now or later. The vector could be extended to show different information to different groups of users (e.g. Joy might be appropriate for friends but not work colleagues), however appropriate representation is key to allowing the user to browse quickly and easily extract the important information. This could also be enhanced by allowing the user to filter contacts on different parameters (e.g. Find me contacts that are not currently busy).

Participants also felt there were other types of information that could potentially be used to communicate. Sketches and handwriting were suggested as useful forms of communication that could support both useful and playful communications. Images were not seen as something that would regularly be shared, as there are other more traditional channels that allow easier sharing to a wider group of friends.

The vibrotactile messaging was seen as a useful feature, however, key to acceptance is providing feedback on whether the message was received or not. Even with the additional feedback added after the focus groups, there was still no indication of whether the contact received the message unless they replied. It was generally used to notify contacts, however there is the potential to include more information. The vibrotactile messages generated were simple vibrations of variable length (controlled by the user). With more complex control, users could potentially build up more powerful forms of vibrotactile communication over a longer time period.

One oft-mentioned issue was that users were acutely aware of their data limits. In an app such as this which relies of an always on data connection, it is often difficult to track data usage which could lead to expensive overruns of data over the 3G network. This is particularly true when using a higher bandwidth channel such as audio. Without any sort of visibility of the data usage, users were choosing to restrict the app to WiFi-only usage. While this allowed participants to use the app without worrying about cost, it meant that they were restricted to interacting only in fixed locations; usually at home or at work. This situation could have been somewhat mitigated by providing an option for tiered access to the 3G network for low bandwidth channels allowing tactile messages and mood vector messages to be sent without access to WiFi. This would have allowed users to maintain more of a presence in the system and encouraged more interaction between users.

Other improvements will look to support a history of communications such that any messages sent and missed will be available to be experienced later.

5 Conclusions and Future Work

The Multimodal Context List allows users to share context information and communicate both verbally and, discreetly, through touch. Visual and tactile textures are used initially to provide a quick overview of the context information for contacts in a contact list. By interacting further with the contact on the screen, we can 'drill down' deeper into the contact's context information, eventually opening a multimodal audio and tactile channel of communication. This information will allow a user better understanding of the psychological, situational, and environmental context that the remote contact is in.

The context represented can be based on explicitly set parameters (as described here with the Mood Vector which provide psychological context) or through some context inferred through a fusion of sensor values that maybe used to provide environmental and situational contexts. The application had been demonstrated on small numbers of Android 2.0 phones, which provide the basis for a larger field trial that will investigate the benefits in a longer-term usage scenario over a larger number of users.

References

1. Ahmaniemi, T., Marila, J., Lantz, V.: Design of Dynamic Vibrotactile Textures. IEEE Transactions on Haptics 3(4), 245–256 (2010)
2. Bales, E., Li, K.A., Griwsold, W.: CoupleVIBE: mobile implicit communication to improve awareness for (long-distance) couples. In: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (2011)
3. Brown, L.M., Williamson, J.: Shake-to-talk: Multimodal Messaging for Interpersonal Communication. In: Proceedings of the 2nd International Workshop on Haptic and Audio Interaction Design, Seoul, Korea (2007)
4. Chang, A., O'Modhrain, S., Jacob, R., Gunther, E., Ishii, H.: ComTouch: design of a vibrotactile communication device. In: Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, p. 320. ACM (2002)
5. Dey, A.K., Hakkila, J.: Context-Awareness and Mobile Computing. In: Lumdsen, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, ch. 13. Idea Group, Inc. (2008)
6. Heikkinen, J., Olsson, T., Vaananen-Vainio-Mattila, K.: Expectations for user experience in haptic communication with mobile devices. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2009). ACM, New York (2009)
7. Hemmert, F., Gollner, U., Löwe, M., Wohlauf, A., Joost, G.: Intimate mobiles: grasping, kissing and whispering as a means of telecommunication in mobile phones. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, Stockholm, Sweden, August 30-September 02 (2011)
8. Kaye, J.: I just clicked to say I love you: rich evaluations of minimal communication. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2006), pp. 363–368. ACM, New York (2006)

9. Kim, S., Kientz, J.A., Patel, S.N., Abowd, G.D.: Are you sleeping?: sharing portrayed sleeping status within a social network. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 619–628. ACM (2008)
10. King, D.: Four principles of interpersonal communication, `http://www.pstcc.edu/facstaff/dking/interpr.htm` (accessed June 10, 2010)
11. Kontaris, D., Harrison, D., Patsoule, E., Zhuang, S., Slade, A.: Feelybean: communicating touch over distance. In: Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts, CHI EA 2012 (2012)
12. Murray-Smith, R., Ramsay, A., Garrod, S., Jackson, M., Musizza, B.: Gait alignment in mobile phone conversations. In: Proceedings ACM MobileHCI, pp. 214–221. ACM, New York (2007)
13. Rico, J., Brewster, S.: Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2009, Bonn, Germany, September 15-18, pp. 1–2. ACM, New York (2009)
14. Rovers, A.F., van Essen, H.A.: HIM: A framework for haptic instant messaging. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2004), pp. 1313–1316. ACM, New York (2004)
15. Sellen, A.J.: Remote conversations: The effects of mediating talk with technology. Human-Computer Interaction 10, 401–444 (1995)
16. Shirazi, A.S., Schmidt, A.: Using Mobile Phones to Maintain Intimacy and Connectedness. In: Proceedings of MobileHCI (2011)
17. Strachan, S., Williamson, J., Murray-Smith, R.: Show me the way to Monte Carlo: density-based trajectory navigation. In: Proceedings of ACM SIG CHI Conference, San Jose, pp. 1245–1248 (2007)
18. Tünnermann, R., Mertes, C., Hermann, T.: Feelabuzz – Direct Tactile Communication with Mobile Phones. In: Proceedings of Workshop on Mobile Social Signal Processing (2012)
19. Wang, R., Quek, F.: Touch & talk: contextualizing remote touch for affective interaction. In: Proceedings of the Fourth International Conference on TEI, pp. 13–20. ACM (2010)
20. Williamson, J.R., Crossan, A., Brewster, S.: Multimodal Mobile Interactions: Usability Studies in Real World Settings. In: Proceedings of ICMI, Alicante, Spain, pp. 361–368. ACM Press (2011)