EXPLOITING INNATE RHYTHMIC SENSE IN A RINGTONE COMPOSER

Daniel Lock & Paul Cairns
UCL Interaction Centre
University College London
26 Bedford Way, London WC1H 0AP, UK
www.uclic.ucl.ac.uk/paul/taptone/

ABSTRACT

Mobile phone ringtones help their owners identify their own phone ringing amongst the many other possible phones that could ring in a given place. However, existing interfaces for personalising ringtones have real usability problems. This paper proposes a new design that exploits humans' innate rhythmic abilities. The design is implemented in a PC prototype and compared to the existing interface of a popular brand of phone. It is found to be more satisfying and suggests further development.

Keywords

Music Interface, Rhythm, Natural Interaction, TapTone, Mobile Phone Ringtones

1. MUSIC FOR THE MASSES

At one level, the ringtones of mobile phones need only communicate a single bit of information – whether someone is trying to call you or not. However, with the proliferation of mobile phones, that is not sufficient. Mobile phones are not like landline phones, but rather are distributed and often hidden around a person and their belongings. In a given place there may be many such phones. Thus, people do not need just to know that *a* phone is ringing but rather that *their* phone is ringing. Once into the realm of personalising ringtones, it is natural that manufacturers give users the ability to have not just a tone but a tune. Add in the fashion and status statements made by mobile phones, and it is not surprising that ringtone download business was worth $1.5 billion in 2001 in Europe alone [7].

The cost to the phone owner of downloading a ringtone into their phone was about $2.10 per ringtone on average in 2001 [7]. Phone owners of course could enter tunes for themselves but, as discussed in the next section, existing ringtone composer interfaces are far from usable. They are so far from usable that third party developers are even producing downloadable software to make ringtone composition easier [8]. Some web-sites also offer instructions on how to enter various tunes into different phones [2] but the phone owner must still struggle with the original interfaces even to do this.

There is clearly a demand, then, for ringtone composition. In this paper, we consider a new method for entering ringtones into a mobile phone. Of course, mobile phones present interesting constraints, having limited screen space and only push-button interfaces. They also present interesting opportunities because unlike the majority of desktop computers, the microphone is an essential feature of the device.

A guiding principle to the resulting design was the exploitation of natural abilities of humans, that is, music seems to be a universal human cultural trait so the interface should in someway cash in on music's "naturalness". The design was implemented in a working PC prototype so as to realistically evaluate its usability.

2. ENTERING TUNES INTO A PHONE

At its simplest, a tune is a *sequence* of notes each of which has *pitch* and *duration*. Of course for more expressive music, the notes may also have timbre, attack, volume and so on but for a ringtone, the simplest form suffices and is what is provided for on most models of mobile phone. Thus, any ringtone composer must allow the user to enter a sequence of pitches and durations.

Existing interfaces to phones exploit a wide variety of notations to represent ringtones. The ringtone composers are then editors for these particular notations. Some phones, such as the Samsung SGHr225, use Common Music Notation (CMN) that is, the familiar representation of notes on a stave. Nokia phones use a textual notation so that, for example, "4#c2" represents a quarternote (crotchet) at pitch C# in the second octave. Ericssons use a similar notation but there the same note would be "C#4" where the capital letter denotes the second octave.

Each of these notations is not without its problems. CMN is widely criticised even amongst musicians to the point that there is an association dedicated to finding better notations [15]. From the point of view of the untrained user, it has many problems – pitch is represented three ways (vertical position, key signature and accidentals), symbols for duration have no obvious mapping to durations. However, CMN is firmly embedded in the musical tradition so at least it is some concession to the musically trained.

The textual notations do not even achieve this. Pitch is represented using a system of capitals or octaves that is not standard. Duration is represented by values that increase with decreasing duration. To increase a sharpened note by a semitone requires deleting the note entirely and entering a new one – a significant editing task for a minor change. Thus, in many instances, the textual notations confound expectations of the musical and non-musical alike.

There are of course other musical notations that remove some of these problems. The piano roll notation represents pitch in vertical steps, each step representing a semi-tone, and duration through the length of a line. Though it lacks the expressiveness of CMN [18], it does at least consistently map attributes of notes to intuitive spatial features "taking advantage of physical analogies" [12]. Indeed, for Deutsch, a Professor of Composing, "The graphical [piano roll] window is far preferable [to the CMN display] in that notes can be positioned precisely in space without the use of arbitrary crotchet-quaver-minim coding" [4].

There is also the Klavarskribo notation favoured by the Music Notation Modernisation Association [15]. This is like a rotated version of the piano roll so that pitch is represented by horizontal position. The added feature of Klavarskribo is that white circles between lines represent white notes on a keyboard and black circles on lines represent black notes on a keyboard. Thus, there is a further explicit mapping between the representation and a particular musical instrument.

As far as we know, there are no commercial ringtone composers that use either of these notations but either would offer potential usability advantages over the textual methods, particularly for the musically untrained.

3. EXPLOITING NATURAL SKILLS

3.1 Why not voice?

Given that our design philosophy was to exploit natural skills, it seems that we could dispense entirely with complicated notations and use the phone's microphone to rely entirely on voice. However, this is not without problems. A feasibility study was performed using voice to MIDI software, Akoff Composer [1]. The user could sing a tune into the system and the recorded tune was digitised for play back. Many problems arose from two main sources: context and human ability.

Mobile phones can be used in a huge number of contexts. Recording a tune in all of these contexts could result in the device picking up unwanted background noise. Even in our study, which was conducted in a quiet, focused setting, the user on one trial did not notice a plane going overhead during recording until it was played back as a low bass tone in the final tune. Also, quiet notes tended not to be picked up at all.

Other limitations come from how humans sing. They do not arrive automatically on the correct note but reach the right note after a short feedback period [6,9]. Even a study of professional singers showed that the first 200 milliseconds of a sung note was generally flat [16]. Also, humans have a tendency to compress wide leaps in pitch and expand sequences of smaller intervals [11]. These factors mean that when singing a tune, humans do not automatically sing the correct pitch and, when they do, even the initial part of the note is not correct.

Supplement these inherent problems with current technical limitations and voice is ruled out as a useful input method.

3.2 Natural Rhythm

Though human reproduction of pitch seems flawed, there is good evidence that the human sense of rhythm is more reliable. When asked to tap a finger, most humans will fall into a beat fairly close to one tap every 0.6 seconds and there are many related phenomena [5]. For instance, a conductor can beat a large range of beats per minute with very good accuracy. Related to this, when asked to sing back a well known song, subjects were able to match the tempo of the original recording very precisely [10].

For entering tunes then, it seems not only do humans have a good sense of rhythm so that we could reliably enter durations but that we can also remember tempi and hence enter the *correct* durations for the tune that we want as the ringtone. Thus, our final design separates out entry of pitch and duration and implements a way of tapping out the rhythm for entering duration in a more 'natural' way.

4. THE TAPTONE DESIGN

We have dubbed the new design for a ringtone composer the TapTone design. The design has many features that directly address several usability issues. These include those raised here as well as others that are based on more traditional usability concerns such as easy correction of errors. Rather than give a full design description here, the reader is referred to our web-site where an executable can be downloaded, or to the presentation given during the talk. Instead, this section highlights key features related to our current concerns.

For evaluation purposes, the design is produced as if implemented on a Nokia 3210. This gives a realistic display area, resolution and set of buttons. The main composer display uses a piano roll notation to represent the tune with the edges of the display giving navigational cues. This is illustrated in Figure 1.

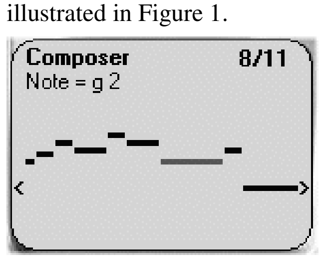

Figure 1: TapTone screen

To enter the duration, the user can use ‘8’ to halve the duration of the currently selected note and ‘9’ to double it (within limits). However, the novel method introduced is that the ‘#’ and ‘*’ keys (in the bottom right and bottom left of the Nokia phone numeric pad) can be used to tap out the rhythm. Both keys work exactly the same way. When the user taps the key repeatedly, notes on the piano roll display change durations (length) equal to the intervals between successive taps. Obviously, there is some rounding as arbitrary note lengths are not accepted but this simply helps to even out slight variations in the user rhythm. If the user has reached the end of the notes already entered, new notes are added as they continue to tap.

Two keys perform this same function because it was felt that when tapping out fast rhythms, it might be easier to use two fingers (thumbs) to do the tapping rather than just one. Also, the position of the keys were chosen to make it easy for a person to hold the phone with both thumbs over the tap keys.

The navigation operations allow the user to adjust the pitch or duration of each note. As each note is entered or changed, it is played giving the user immediate feedback on their actions. Also, the user can choose to playback the tune from the current point or from the beginning.

5. EVALUATING TAPTONE

To evaluate the TapTone design, its usability was compared qualitatively with the ringtone composer of an existing phone, namely the Nokia 3210. However, being a new design, the TapTone was implemented in Visual Basic on a PC and not implemented on a real phone. To remove the possible confounds that this may cause, the Nokia phone was simulated in Visual Basic as well. To keep the controls close to an actual phone, the simulation could only be operated using the numeric keypad of a standard keyboard. Both simulations are available from our website.

5.1 Experimental subjects

16 subjects were used in the experiments. Eight used the TapTone design and eight the Nokia simulation. Ages ranged between 22 and 51 years with the average of 27.2 (SD=6.85). Eleven subjects were male and five female. Five had used a ringtone composer before but only two had used them more than once or twice. 13 users claimed to have had some musical experience in the past but only three claimed to have reached more than an intermediate standard. Only one subject could sight read CMN and knew more than basic musical theory.

5.2 Experimental Method

Having been interviewed about their backgrounds, users were assigned to one of the simulations. The experimenter explained how to use the interface and demonstrated how to compose the well-known tune *Happy birthday.* Subjects were given five minutes to get to know the buttons, the buttons being labelled identically to those on a Nokia. Subjects were also provided with a crib-sheet of what each button did which they could refer to throughout the tasks.

After five minutes of familiarisation, subjects were given twenty minutes in which to compose a melody from the list of *Three blind mice, Jingle bells, God save the Queen, Ode to joy, Yankee doodle.* Subjects were asked to talk aloud as they did this and, as the experiment was not intended to test the users’ musical abilities, the experimenter provided any musical guidance as requested.

After the task was completed, the users filled in the “quick and dirty” SUS questionnaire to get a rough indication of satisfaction with the interface [3]. Also they were interviewed in a semi-structured way to elicit comments and views on the interfaces.

5.3 Results

The SUS questionnaire gave scores of 39.7 (SD=16.3) to the Nokia and 77.8 (SD=10.4) to the TapTone. This suggests that users found the TapTone on the whole more satisfying to use though there is no indication of why this might be. Indeed, it could simply be due to novelty of the interface.

The implementation on PC still caused some confusion as subjects tried to use the mouse or arrow keys rather than the keypad. Nokia users generally reported more trouble remembering the function of each key though TapTone users were also confused by the keys to change octave as well as those to change pitch by semi-tones.

Nokia users had considerable trouble entering pitch as they did not remember about the sharp key and that they needed to change octaves explicitly. Also, notes could only be changed by entering a new one and deleting the incorrect note. Users were surprised to find lots of extra notes as a result of editing. The only pitch problem encountered on the TapTone was that one user did not realise that notes had to be selected before the pitch could be altered.

For entering duration, Nokia users were confused (as expected) by the increasing number representing decreasing duration though some ignored the length of notes entirely. The TapTone caused problems due to ending the tapping – a final tap is needed to finish the last note. Also, entering rests caused problems as they did not seem to fit into the tapping style of entry.

In summary, the TapTone seemed to cause fewer problems in general except for when entering the duration of notes. In particular, pitch entry was much simpler than on the Nokia.

6. DISCUSSION AND FURTHER WORK

The piano roll notation of the TapTone seems to have been successful avoiding many of the problems of the textual notation of the Nokia. Unexpectedly the duration entry method caused more problems than that of the Nokia. However, this seems in part due to users actually trying to enter duration as opposed to avoiding it on the Nokia. Also, users tried more ambitious things such as rests. Tapping then seems straightforward in the main but rests and the end of the tune cause problems. Further design implementations and more formal quantitative evaluation should distinguish better between the usability of the different designs.

Note durations are currently constrained to powers of two. In principle, there is no particular bar to having arbitrary note durations but this would mean that the TapTone would not be able to even out rhythms. Some tunes require more complex rhythms but some rhythms can be successfully "approximated" by simpler ones. This, then, is a trade off which could only be resolved with specific testing of different implementations.

The TapTone design used Western tonality, so 8 notes in the well-tempered octave as opposed to any of the many oriental tonalities with 5, 7 or 13 notes in the octave in a variety of tunings. The Piano Roll notation does not represent absolute pitch so could easily be adapted for these other tunings.

Certainly the TapTone seems more pleasurable to use. It may be that for the less musical, any form of music entry that can match the user's intuitive sense of pitch and duration is more desirable than the current, confusing and complex notations. With the mobile phone market being so competitive, further development of ringtone composers could lead to an important competitive edge which might truly bring music composition to the masses.

7. REFERENCES

[1] Akoff Sound Labs, *Akoff Music Composers v2.0.* www.akoff.com

[2] Bernatchez E. (2002). Truly free ringtone composers and free melodies. cellphones.about.com/cs/ringtonesmelodies/

[3] Brooke J. (2002) *SUS – A quick and dirty usability scale* www.usability.serco.com/trump/methods/satisfaction

[4] Deutsch S. (1996). Music and technology: the composer in the age of the internet. *RSA Journal,* July media.bournemouth.ac.uk/smd/articles/article1.html

[5] Gura T. (2001). Rhythm of Life. *New Scientist,* **171,** 32-35

[6] Harvey N. (1985). Voice control in singing: a cognitive approach. In Howell P., Cross I., and West R. (eds) *Musical Structure and cognition.* Academic Press.

[7] Hayes K. (2001). Ringtones are music to content providers' ears. *M-Business,* **11,** 61 www.mbusinessdaily.com/magazine/story/MBZ20011 1003S0028

[8] Hirvonen J. (2000), *Melodia 0.7* www.geocities.com/toytools

[9] Howell P. (1985). Auditory feedback of voice in singing. In Howell P., Cross I., and West R. (eds) *Musical Structure and cognition.* Academic Press.

[10] Levitin D. J. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute.

[11] McNab R. J., Smith L. A., Witten I. J., Henderson C. L. and Cunningham S. J. (1996). Towards the digitial music library: tune retrieval from the acoustic input. *Proceedings of ACM DL'96,* 11-18

[12] Norman D. A. (1990). *The design of everyday things.* MIT Press.

[13] Picard R. (2000), *Affective Computing,* MIT Press.

[14] Pinker S. (1994). *The Language Instinct,* Penguin Books.

[15] Reed T. S. (1997). *Directory of music notation proposals.* Notation Research Press.

[16] Seashore C. E. (1967). *Psychology of Music*. Dover

[17] Spence R. (2001), *Information Visualisation,* ACM Press

[18] Wiggins G. and Smaill A. (2000). Musical knowledge: what can AI bring to the musician? in Miranda (ed) *Readings in music and AI,* Harwood Academic Publishers