The Use of Audio in a Stock Control Application

DIMITRIS RIGAS, IBRAHIM BAHADUR
Department of Computing
University of Bradford
Bradford BD7 1DP
UNITED KINGDOM

Abstract—This paper describes two main experiments that have been carried out on stock control system. The experiments examined how auditory stimuli can be used in stock control systems to communicate information. The inclusion of auditory stimuli can benefit many different types of users including the visually impaired. The experiments demonstrate the successful use of non-speech sound in stock control applications. In the first experiment short Earcons were used to communicate stock levels (in integer values). The pitch of the sound rose as the stock levels increased and two instruments were used to help users to understand information communicated. Organ notes communicated the first value in steps increasing by 10 and piano notes communicated the second value in steps increasing by 1. The two distinct instruments were used to aid disambiguation. The second experiment moved further to communicate several functions of the stock control system using earcons and auditory icons. The results of both experiments were very promising and they showed that the use of non-speech sound will be successful when used not just in stock control systems but many other types of applications.

Key-Words—Human-Computer Interaction, Multimedia Stock Control, Interactive Information Systems, Earcons, Sound, Auditory design.

1. Introduction

The use of audio in interactive information systems has attracted the interest of academic and industrial researchers in the last few years. Auditory stimuli can be used in many different ways. For example, it can be used on its own for special users such the visually impaired, complement multimedia systems or improve the usability of interactive information systems. Simple events can be communicated using simple sounds (e.g., a single pitch or a series of pitch) and more complex events can be communicated by introducing structure, repetition or iteration and higher structures of musical stimuli such as rhythms, tunes, harmony and timbre. These structures have a large number of parameters and characteristics and therefore many different types of combinations can be created.

A number of limitations also exist when using auditory stimuli. First, auditory messages are objects in time and thus once they are presented, there is no other reference apart from the memory of the users. Second, as they are limits in the speed that users can read text, there are also limits to the speed that users can process (i.e., identify, understand and interpret) auditory messages. Third, auditory stimuli do not provide as high resolution data as the visual stimuli. Fourthly, when non-speech sound is used, a change in one or more parameters often results in changing the quality or signature of the sound. This is often referred as low orthogonality. This paper describes two main experiments to at least partially address the possibility of maximising the usability of interactive information systems by introducing various types of auditory stimuli in addition to the visual stimuli. These experiments have also started to examine the volume and degree of complexity of the information that can be communicated in this problem domain using sound. In the following sections, we review relevant work in the literature and we describe four experiments and their results under two platforms that investigated auditory information processing in the domain of stock control applications.

2. Relevant Work

Software designers often use speech and non-speech sound in interactive information systems. Non-speech sound can be categorised into auditory icons (i.e., environmental sounds), earcons (i.e., short musical
The use of environmental sounds as communication metaphors are also known as auditory icons. Research in this particular area includes the SoniFinder interface [1], [2] in which various environmental sounds were used to communicate information in an interface. Further reading on everyday sounds or environmental sounds can be found in [3], [4],[5]. Earcons can communicate many different types of information in user interfaces of information systems. For example, an error message can be communicated using a digitised sound, a sound created by a synthesiser or a single note. These types of earcons are often referred to as one-element earcons. Many parameters can be used to characterise an earcon. These parameters include rhythm [6], [7] and pitch [8], [9]. Earcons could use repetition, variation and contrast in order to be easily remembered, recognised and understood. A more complex form of earcons involves combination, inheritance and transformation [7]. Earcons could be presented serially [10], [11] and in parallel [12]. Some guidelines for the creation of earcons exist [11], [13], [14], [15], [16].

When long structured musical stimuli are used, the physical characteristics of the sound (e.g., timbre, rhythm, and pause) suggest segmentation or breaking up of the stimuli into separate pieces. Tan, Aiello and Bever investigated the use of equal duration note sequences. Those sequences contained two melodic phrases and they ended with a cadence. Users were asked to determine if those sequences contained two particular probe notes. Three different forms of probe notes were used. The first pair of probe notes was ending the first phrase, the second pair was beginning the first phrase, and the third pair was straddling the phrase boundary. Users recognised the first two pairs of probe notes better than the third pair. This was because users created better accessible memory representations of the intervals within the probe notes, as they were determined by cadencies, than the ones created when notes were equally close in time but were deriving from different phrases [17]. Miller suggested that human memory decreases as more items are added and it is likely that the ability to remember previously stored items is lost. The memory problem can be solved by relating or associating different items together. When structural musical stimuli are used as a communication metaphor, these relations or associations are often found in the form of patterning and structuring. Other evidence suggests that musical extracts with strong representative titles were remembered better than other musical extracts with abstract conceptual titles [18]. The use of those representative titles enabled the creation of a story in the memory of users that can help users to remember parts of the musical stimuli. Other research work indicates that when two simultaneous tone sequences (one sequence was played to the left ear and the other to the right ear using headphones) were played to users, all high notes were reported that they were deriving from the right headphone and all the low tones from the left headphone. This in-fact is called the phenomenon of scale illusion, as users hear the tones as two scale passages rather than as two angular melodic contours that were actually present [19]. This pitch grouping phenomenon is present even when real instrumental sounds are used and the tones are presented with spatially separated speakers [20]. In this work it was found that most users grouped tones by pitch even when distinctive timbre was used. Experiments have also been performed to determine the memory of individual notes. In these experiments users were represented with two notes that were separated by 5 seconds interval. The pitch of the notes was the same for half of the trials and separated by a semitone for the other half. Users were asked to compare whether the two notes were of the same pitch. In one set of experiments the five second interval was silent. The results indicated that most of the users were able to judge the pitch of the note accurately [21], [22], [23], [24].

The use of auditory stimuli has demonstrated to be useful in a number of problem domains such as software engineering environments [25], [26], to help visually impaired users [13], [26], [27], and many others [28], [29], [30], [16]. In the light of the applicability of sound in all these different problem domain, we have developed two research platforms to investigate the suitability of speech and non-speech sound in the problem domain of stock control. Stock control systems often involve complicated visual displays in order to accommodate the functionality required by users. For instance, many users of stock control systems require complex graphical browsing of the stock available or the stock recently sold or arrived. Traditionally, most of the output in such systems is visual and often is complex as the volume of information to be communicated increases. An additional communication metaphor, such as the auditory, has the potential of increasing the channels of
communication to the user and therefore the information that can be communicated at any given instance of interaction can be larger in volume and yet maintain a user friendly presentation. Thus, the starting point of this research was to develop a basis under which experiments in auditory information processing can be performed. Both experimental platforms were developed using Visual Basic.

3. MUSCOS Experimental Platform

The MUSCOS (MUltimedia Stock COntrol System) platform was based on the results of the initial platform but offers more capabilities for long term experiments with general users as well as with special groups of users (e.g., visually impaired users). The platform has been designed in a way that it offers visual, auditory or audio-visual feedback of typical stock control functions in order to enable us to investigate the use of auditory stimuli in this problem domain. The interaction provides speech and non-speech feedback in order to enable users to perform various tasks.

3.1. Stock Levels: Structured Rising Pitch

In this experiment stock levels (range: 1 to 100) were communicated to users with rising pitch notes using organ and piano synthesised voices. One note using organ communicated 10 items in stock and one note using piano communicated one item in stock. For example, 11 items in stock were communicated with one note using organ and one note using piano and 9 items in stock were communicated with 9 rising notes using piano and so on for the rest of the stock levels. The results showed that all users successfully recognised the stock levels communicated.

3.2. Control Operation: Structured Tunes

In this experiment, five typical control applications were communicated using short tunes. These control operations tested were delete, add, search, view and refresh. The tunes were chosen using results from previous experiments [32], [33], [34], [35]. Results showed that all 75 users successfully recognised the delete and view control operations, 74 out of 75 users recognised the add operation ($X^2= 71.05$, df=1, critical value $10.83$, $p<0.0005$), 72 users recognised the search operation ($X^2= 63.48$, df=1, critical value $10.83$, $p<0.0005$), and 71 users recognised the refresh operation ($X^2=59.85$, df=1, critical value $10.83$, $p<0.0005$). Figure 1 shows these results in percentages. The control operations add, search and refresh were not immediately recognised by all users as the delete and view but they are still within acceptable levels. Statistical significance was reached by all operations and thus we can infer that the answers provided by the users were not due to chance. We believe that the users who failed to recognised the control operations (1 user for the add, 3 users for search and 4 users for refresh) would improve by offering some additional short training. The overall result however of this experiment demonstrates that virtually all users were able to use and understand these typical 5 control operations of a stock control application as they were communicated using structured tunes.

4. Discussion

Post-experimental research indicated that 100% of the users did identify what integer was being communicated. This figure was the same when two different timbres were combined to communicate larger integers. However for the second experiment the results were not perfect but very promising as the delete and the view operations indicated that 100% of users identified the stimuli correctly. Also 98.6% identified the add stimuli and 96% accurately identified the search stimuli.

The results of these experiments and other experimental findings [36], [37] demonstrate the addition of non-speech and speech messages to communicate information in stock control systems. The various types of auditory stimuli could be integrated with each other (e.g., earcon and speech or earcon, auditory icon and speech) or with the visual stimuli in order to deliver information in a multimedia...
fashion. The auditory stimuli would be used in most cases in addition to the visual stimuli to either highlight the information communicated visually or to communicate different types of information. For example, the visual channel will communicate one type of information and the auditory channel will communicate some other type. In this way the information communicated from the visual and auditory channels will complement each other and thus the volume of information communicated to users will be larger. In the design of the auditory stimuli, designers would need to take into account the experimental findings discussed below. Synchronisation of the various types of metaphors is another major part in the audio-visual design of a system. Our previous work [27], [32], [33], [34], [35], [36], [37] and these experiments have demonstrated that information communicated visually and aurally must be properly classified in order to perform initial trials to test the viability of the various metaphors used. For example, some information would be communicated visually, some other information would be communicated using synthesised or recorded speech and yet some other information would be communicated using earcons. The distribution of this information plays an important role for a successful communication.

First, a common case is that the design of the earcon is correct but it is not being recognised by users. This will result to confusion and often users will misinterpret the information communicated. Second, the earcon might not be designed in a way that could be perceptually combined with other communicated stimuli (e.g., visual). In this case, information in the problem domain has not been successfully communicated by either channel. Third, the earcon is successfully recognised either on its own or in relation to other stimuli (e.g., visual). Designers must be aware of these possibilities and cater early in the design stage to ensure that the whole design works well with the different types of metaphors used. The use of structured rising pitch was very useful. The grouping of pitch can communicate values with a very good degree of accuracy. The use of different timbre and stereophony is also quite important to aid disambiguation. We believe that the addition of other timbre could enable us to enlarge the range of the numbers that could be communicated. Users understood the earcons that communicated numerical values well because they had a simple structure and they were consistent in their design. They had a common design principle underpinning all earcons. Therefore, simplicity and consistency appear to be important issues when auditory messages are designed and certainly contribute to correct recognition by users. Short speech messages and simultaneous use of speech and non-speech sound was found to be useful. The use of synthesised and recorded speech messages either on their own or combined with other auditory stimuli demonstrated that it is possible to utilise the audio channel with different types of auditory stimuli in order to further increase the volume of information communicated aurally. When these combinations of auditory messages are communicated in the presence of visual stimuli, the information communicated will increase even further. Although users understood both synthesised and recorded speech, most of users expressed a preference to the natural recorded speech. Results have also demonstrated that the use of rhythms and tunes helped users to easily remember and recognise the earcons. Rhythm added structure to the earcons by dividing the earcon into smaller parts. These smaller parts showed to be easily understood by users. Tunes, on the other hand, enable users to remember the earcon more easily. It is very important that users easily remember earcons as often many Earcons are required in order to communicate a reasonable amount of information.

Finally, an on-line sound dictionary with examples of typical audio-visual interaction helps users. The on-line dictionary must contain all the pitch reference scales from which the non-speech sounds were taken and a library of the speech messages used. The on-line dictionary will also serve as a short user training facility and for further reference and consultation during the interaction process. This facility would therefore aid user comprehension of the auditory messages.

5. Conclusion

The experiments described in this paper have demonstrated that both speech and non-speech audio can be successfully used to communicate several types of information in stock control applications. The MUSCOS experimental platform developed have demonstrated that the use of audio can be used in different interface circumstances in stock control system interaction. For example, the MUSCOS has demonstrated the use of audio in more complex instances of interaction. The results of these
experiments and the lessons learned regarding the design of the audio will be used under the MUSCOS platform to continue this research.

We need to establish the degree of complexity that we can reach in communicating stock control type information. The simultaneous use of speech and sound has demonstrated that we can increase the volume of information communicated to users at any given time of the interaction process. These results could also apply in other problems domains and can contribute in the audiovisual interaction design of information systems. The next step is a larger scale investigation on synchronisation and design issues of interactive information systems in which speech messages communicate one set of events, earcons and other non-speech sounds could communicate another set of events, environmental sounds communicate some other set of events and special effect are also used to communicate a further set of events. In this way the auditory communication could be used to maximise the volume of information communicated.

References:


