Conference papers

Digital Media Centre

2006-01-01

Harmonically Combined Contour Icons for Concurrent Auditory Display

Charlie Cullen
Dublin Institute of Technology, charlie.cullen@dmc.dit.ie

Eugene Coyle
Dublin Institute of Technology, Eugene.Coyle@dit.ie

Harmonically Combined Contour Icons for Concurrent Auditory Display

Charlie Cullen[ϕ] and Eugene Coyle*

[ϕ]*Digital Media Centre,,*
Dublin Institute of Technology,
IRELAND
E-mail:[ϕ]*charlie.cullen@dit.ie*

* *School of Control Systems and Electrical Engineering,*
Dublin Institute of Technology,
IRELAND
E-mail:[ϕ]*eugene.coyle@dit.ie*

This paper considers the harmonic combination of basic melodic shapes known as contour icons in concurrent auditory displays. Existing work in the field (such as that concerning earcons) has considered the combination of patterns designed using low level cognitive features, and so effective streaming is difficult. This work investigates means by which musical patterns with high level cognitive features (such as contour) representing data values can be rendered concurrently, so that multiple data sets can be effectively conveyed using an auditory display. The detection and comprehension of harmonically combined contour icons was tested in comparison to those combined uniquely (non-harmonically). Results suggest that significant improvement in pattern combination detection was made using harmonically combined contour icons, although limitations were observed due to the nature of the harmonic relations involved. Future work will investigate the most flexible methods of harmonic combination, to produce an effective method of auditory display.

Keywords – Contour Icons, Harmonic Combination, Auditory Display, Cognitive Science, Sonification.

I   INTRODUCTION

McGookin [1] defines an auditory display as "the use of sound to communicate information about the state of an application or computing device to a user". The roots of auditory displays can arguably be traced back to alarm and alert mechanisms, seeking to inform the user of an important or urgent condition. Such displays utilise many of the advantages of audio information delivery, notably the operation of an 'eyes-free [2]' interface. Focus independent systems of information delivery have been implemented in situations such as flying a plane [3] or driving a car [4] as an essential means of processing information, and in so doing highlight on of the main advantages of auditory display.

Auditory displays also possess the advantage of faster information delivery [5] than any visual mechanism can offer, and so are well suited to situations requiring alerts or alarms. Audio is also largely unavoidable [6] and so is again ideal for alert or alarm information, unlike the focus dependence of visual systems which may often be ignored by the user.

a) *Concurrent Auditory Displays*

Concurrent source auditory displays [7] and [8] offer advantages over other auditory displays due to the increase in data they convey. A concurrent display achieves a greater data bandwidth by using multiple sources and so gives the listener access to more data in the same length of time. This higher information rate is augmented by the potential for the user to focus on specific aspects of the data [9] and [10] at any time by using ASA (auditory scene analysis) [11]. In this manner, information rendered by a concurrent auditory display can be analysed either wholly or in part by the listener as they require.

Because of this, Concurrent auditory displays are potentially one of the most useful methods of delivering data using audio as they provide a means of delivering several related data sets in tandem. Brown [12] considered the potential of two sonified graphs presented in tandem as a means of detecting intersections of interest. Results showed that this

method of representation significantly reduced the time taken to determine intersection points, without reducing the accuracy of responses compared to sequential (individual) presentation. Such intersections can be better highlighted using more complex patterns (such as earcons) as shown by Hankinson and Edwards [13], who used compound earcons to indicate the validity of certain operating system tasks. The earcons used were designed with harmonious musical attributes based on valid operations, such that a copy earcon would sound harmonious with a file earcon but dissonant with a printer earcon. This method of concurrent representation suggests great potential when using musical patterns, as it affords the use of many of the traditional musical compositional techniques required of harmonic consonance [14].

II   CONTOUR ICONS

Contour icons are designed in a similar manner to earcon patterns [ref to design guidelines], although no motif information is conveyed within a contour icon. Instead, contour icons use the same features of detection as earcons, with the additional specification of a melodic contour based on a simple shape. In this manner, a more robust framework for pattern design can be considered which will ideally be transparent to all listeners (regardless of musicianship skills).

a) *Rhythm Pattern Matrix*

All contour icon patterns are designed using an overall rhythm pattern matrix (Table 1) which ensures that no two patterns have the same rhythmic signature.

Table 1: Rhythm pattern matrix for contour icon pattern design

The aim of the matrix is to provide means of distinguishing the rhythm of each pattern, removing the possibility of two patterns having the same (or similar) rhythm.

b) *Boundary Pitches*

Boundary pitches [15] allow the shape of a pattern to be accurately specified from point to point- a useful framework for contour design. This use of boundary pitches also suggests benefits when seeking to create a set of patterns as individual from each other as possible (Figure 1).
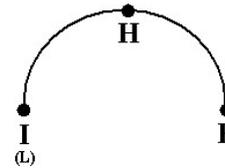
Figure 1: Example contour icon defined by boundary pitches

The boundary pitches used ensure that each pattern differs in overall pitch characteristics from its counterparts- alongside its unique melodic contour. In this manner, it is less likely that users will struggle to detect the beginning and end of each pattern used in a Sonification.
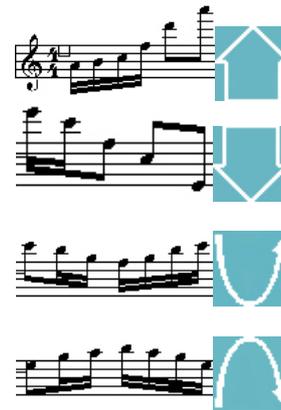
Figure 2: Example Contour icons

III   HARMONIC COMBINATION

When 2 musical pitches are sounded together they are defined as an interval [14], and the relation of these pitches determines the interval as consonant or dissonant [16]. The definition of consonance and dissonance in western music has changed with each period of compositional style [17], but relations of 3rd, 5th and octave [14] have remained sufficiently constant to allow implementation. The most recognisable interval is that of the octave [18], and so this relation was used as an initial basis for development.

a) *Register*

Earcon design guidelines [19] specify that register is a poor source of discrimination in isolation, but if used as part of a design template should ideally utilise gaps of 2 or 3 octaves. The trio format employed in this research could be considered in terms of bass middle and treble patterns (paraphrasing audio terms for frequency ranges [20]), and so the distance between the pitches in each of the bass, middle and treble patterns would be set at 2 octaves (Figure 3).

Figure 3: Register gap for contour icon (Up)

The bass patterns used are pitched in the range of A0 to A2, with middle patterns from A2 to A4 and treble patterns from A4 to A6. Although a strict gap of 2 octaves does not necessarily exist between all patterns, segregation is maintained by the use of boundary pitches, preventing any occurrence of the same pitch at the beginning or end of any pattern (regardless of register).

b) *Timbre*

In the harmonic tests, 3 distinct timbres were required that would define bass, middle and treble pattern streams in a Sonification. As a result, patterns were chosen based on the different timbre families defined by Rigas and Alty [21]. The bass patterns were allocated the 'picked bass' sound from the General Midi soundset [22], with middle patterns being allocated the 'drawbar organ' and treble patterns allocated the 'flute' sound.

c) *Spatialisation*

Investigation by McGookin [1] shows that the use of spatialisation has a significant effect in concurrent audio presentation. Standard locations within the azimuth [23] (left, centre and right) are used to maximise the potential for stream segregation. In 2 value Sonification (involving 2 concurrent pattern streams) the instruments used are panned left and right, while 3 instruments are allocated left, centre and right (Figure 4). This allows for the maximum difference in location possible within the stereo field, and so reduces the possibility of proximity effects [11].
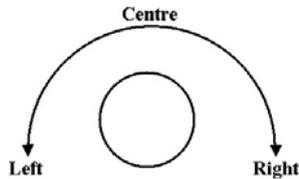

Figure 4: Example azimuth locations within the stereo field

IV   TESTING

20 participants took part in the experiment described in this section which was of a within groups design involving two conditions, the non-harmonic

combination condition and the harmonic combination condition. No participants were taken from formal music courses. Participants were randomly assigned to one of two groups to determine the order in which they would undertake the experiment. Each group contained the same number of participants, and both conditions consisted of training and testing phases (Table 2).

NH= non-harmonic		H= Harmonic	
Group	**1st session**		**2nd session**
	Training	Testing	Training	Testing
1	**NH**	**NH**	**H**	**H**
2	**H**	**H**	**NH**	**NH**

Table 2: Testing procedure for the non-harmonic combination vs. harmonic combination experiment

All tests were performed using Sonifications of 2 to 4 patterns, for 2 and 3 variable conditions. The independent variable in testing was the harmonic combination of patterns. The dependent variables were the number of patterns identified and the number of combinations identified. The workload placed on participants by each condition was also of interest, and so NASA TLX questionnaires [24] were filled in by participants after completing each condition.

a) *Training Phase*

Participants were first introduced to the musical patterns they would be using, followed by a brief period for questions about the testing. The use of contour icons was explained, alongside the means by which they would be combined harmonically. Each training phase also contained a tutorial on the Sonification method as it was employed during testing.

    All participants were played an example Sonification, with an accompanying visual listing of the contour icon patterns used in that Sonification. In the non-harmonic combination condition, participants were informed they were required to detect combinations of differing contour icons in 2 and 3 variable conditions. In the harmonic combination condition, participants were told they would be asked to detect a single contour icon played in harmony by 2 or 3 instruments. After the example Sonification, a further brief period was allowed for any other questions participants had about the experiment.

b) *Testing Phase*

In the testing phase participants were asked to listen to Sonifications of 2 to 4 patterns, for 2 and 3 variable conditions. If a participant could more accurately detect harmonic combinations of patterns, it would suggest that harmonically combined patterns were easier to detect. Using the TrioSon [25]

application (Figure 5), participants were played various predetermined pattern Sonifications.

Figure 5: TrioSon Sonification Application

All Sonifications were performed on a Compaq NX6100 laptop, using the onboard ADI AC97 soundcard. Participants were asked point estimation and pattern combination questions about the pattern Sonifications, to determine how effectively the data had been conveyed. Participants were allowed to listen to a Sonification once for each part of a question, ranging from 2 passes through to 6 for the last question of each test. After the 5 test questions were answered, participants were asked to answer post-test TLX questionnaires to determine how difficult they had found the process.

V   RESULTS

Overall results (Figure 6) show that performance improves from 58.25% in the non-harmonic combination condition to 75% in the harmonic combination condition. This shows significant improvement ($T(20) = -3.02$, $p=0.0043$) in performance between test conditions, suggesting that harmonic combination is more effective in multiple stream Sonification.
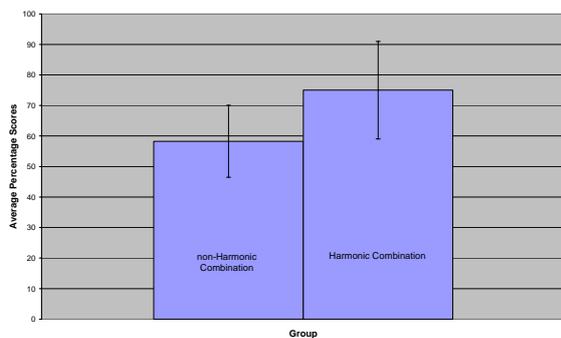

Figure 6: Graph showing overall average percentage scores (by test condition), showing standard deviations

a) *Effect of Harmonic Combination on Point Estimation*

Point estimation questions were asked to determine whether harmonic combination had any effect on individual pattern detection performance (Figure 7).
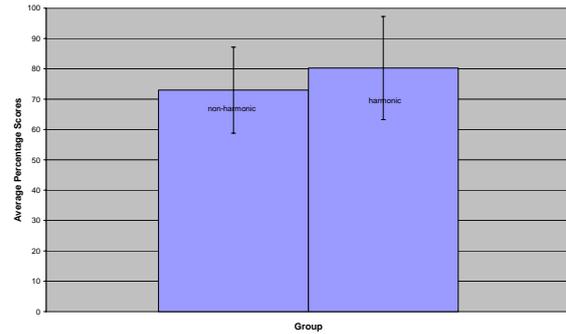

Figure 7: Graph showing average point estimation percentage scores (for each test condition), showing standard deviations

Results show that point estimation performance in the non-harmonic condition improves from 73% to 80.25% in the harmonic condition. This result shows that harmonic combination does not affect the point estimation of individual patterns in a multiple stream Sonification.

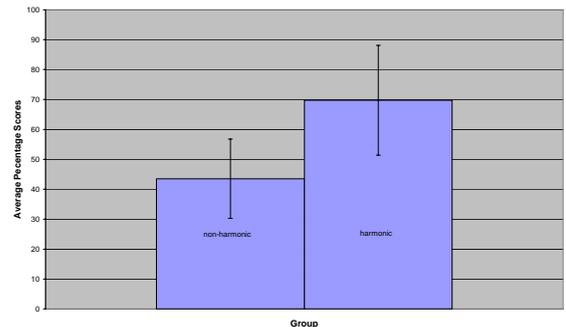b) *Effect of Harmonic Combination on Pattern Combination*

Figure 8: Graph showing average pattern combination percentage scores (for each test condition), showing standard deviations

Results for the pattern combination test questions (Figure 8) show significant improvement from 43.5% for the non-harmonic condition to 69.75% in the harmonic condition ($T(20) = -5.06$, $p<0.0001$). This result shows harmonically combined patterns are significantly easier to detect, and so the results for 2 and 3 pattern combinations were considered individually (Figure 9). Scores for 2 variable pattern combination questions show a significant increase from 50.42% for the non-harmonic condition to 79.17% for the harmonic condition ($T(20) = -4.36$, $p<0.0001$). A similarly significant increase is observed in the 3 variable pattern condition ($T(20) = -3.52$, $p=0.001$), with an average non-harmonic combination score of 33.12% rising to 55.62% in the harmonic combination condition.
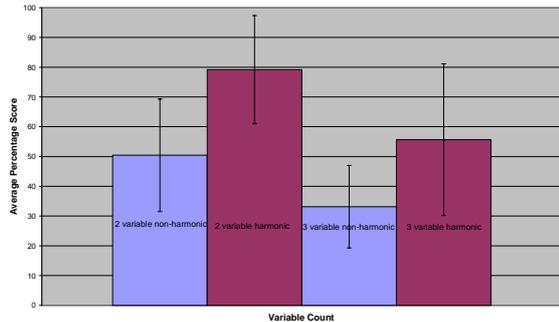
Figure 9: Graph showing average 2 and 3 variable pattern combination percentage scores (for each test condition), showing standard deviations

It is noted that performance in the 3 variable condition was not as effective as had been hoped, with some participants detecting false positive values in conditions where 2 instruments were in harmony (rather than the required 3). Future work will have to consider how harmonies of 2 instruments in 3 variable combinations could be made more dissonant, to highlight the presence of a conflicting value in the third instrument present.

c) *Post Test TLX Results*

Overall TLX results (Figure 10) show a significant reduction (T(20) =6.18, p<0.0001) in workload from 54.95 to 38.9 in the harmonic combination condition.
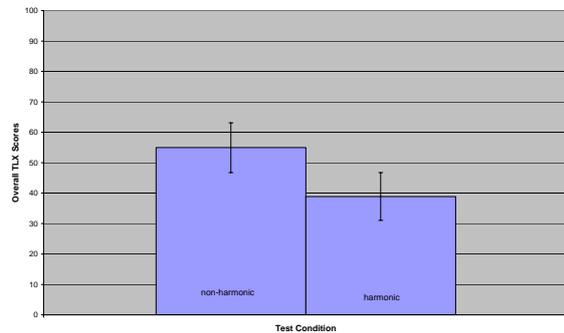

Figure 10: Graph showing overall average post-test TLX scores (by test condition), showing standard deviations
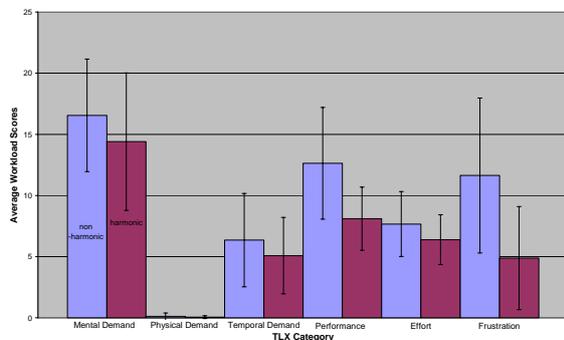

Figure 11: Graph showing average post-test TLX scores by category (for each test condition), showing standard deviations

Workload scores for each test category were examined individually (Figure 11), with mental demand scoring highest in both conditions (16.55 and 14.4 respectively). A significant reduction is found in the performance workload score from 12.63 to 8.1 (T(20) =3.77, p<0.001), suggesting that participants felt they had performed better when patterns were combined harmonically. This improvement is accompanied by a similar reduction in frustration from 11.63 to 4.88 (T(20) =3.87, p<0.001), which suggested that participants were more comfortable with harmonically combined patterns. No problems were encountered during explanation of the test schedule, and participants understood the principles of Sonification, contour icons and harmonic combination as they related to the tests.

VI   DISCUSSION

Tests show that significant improvement can be made when harmonic combination is employed during concurrent auditory display. Testing had employed non-musicians in a series of point estimation and pattern combination questions, which show that harmonically combined contour icons can achieve average recognition rates of 75%. Point estimation performance is not affected by harmonic combination, improving from 77.5% to 82.5% in the harmonic condition. Also, overall pattern combination results are significantly improved from 43.5% to 69.75% in the harmonic condition. Participants were comfortable with the use of harmonic combination and contour icons. This is considered an indicator of the potential of harmonically combined contour icons, in that no musical knowledge or training is required to recognise a simple harmony or melodic shape.

Post test TLX questionnaires show a significant reduction in overall workload due to harmonic combination, with significant reductions also being observed for performance and frustration scores. Several participants commented that harmonic combination was straightforward to understand, and improvements in performance observed during testing suggested that harmonic combinations were also easier to detect.

VII   CONCLUSIONS

Harmonic combination provides a more effective means of highlighting intersections between values during concurrent auditory display. Significant improvement is observed with harmonically combined patterns during testing, although limitations were observed due to combination restrictions (one to one mappings) and also due to the detection of false positives in 3 variable conditions. Future work will have to consider means by which other values could be defined in multiple combinations based on a single value from a certain

variable. Possible solutions may involve the use of other common harmonic intervals within the octave (such as the 3rd, 5th and 7th), and further investigation would be required to assess the potential of such combination. Harmonic combination also produces a significant reduction in workload during testing, and this indicates potential for the method.

VIII   FUTURE WORK

Harmonic combination was developed to improve pattern combination detection during concurrent representation. Although results show significant improvement, further work is required to develop it implementation fully:

a) *How many patterns can be combined effectively*

Higher counts may be possible if intervals within the octave are employed. Further work will consider whether such intervals allow for greater levels of combination, and if a limit of recognition exists (when proximity becomes a grouping factor).

b) *How can several patterns be combined distinctively*

Further investigation is needed into the role of consonance and dissonance in harmonic combination. Effects of instrument (timbre) and register must be considered, alongside means by which different contour icons can be made more dissonant during combination.

c) *How can one pattern be combined with more than one other in another given variable*

Further work will have to consider how to combine a pattern with several others in other registers, again observing gestalt grouping factors such a belongingness and similarity.

IX   REFERENCES

[1] (100) McGookin D, "Understanding and Improving the Identification of Concurrently Presented Earcons", PhD thesis, University of Glasgow, 2004.

[2] (101) Brewster SA, Lumsden J, Bell M, Hall M, Tasker, S, "Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices", In Proceedings of ACM CHI 2003 (Fort Lauderdale, FL), ACM Press, Addison-Wesley, pp 463-480.

[3] (102) Newell AF, "Extra-ordinary human-computer interaction", Edwards ADN ed., Extra-Ordinary Human-Computer Interaction: Interfaces for Users with Disabilites, Cambridge University Press, Cambridge, 1995, pp 3–18.

[4] (103) Vargas MLM, Anderson S, "Combining speech and earcons to assist menu navigation", International Conference on Auditory Display (ICAD), Boston, Massachusetts, 2003, pp 38–41.

[5] (8) R Kail, TA Salthouse, "Processing speed as a mental capacity," Acta Psychologica, vol. 86, pp. 199-255, 1994.

[6] (108) Kramer G, "An introduction to auditory display", Kramer G ed, Auditory Display: Sonification, Audification and Auditory Interfaces, Addison-Wesley, Reading, Massachusetts, 1994, pp 1–77.

[7] (110) Sawhney N, Schmandt C, "Nomadic radio: Scaleable and contextual notification for wearable audio messaging", ACM Transactions on Computer-Human Interaction (CHI), vol.1, Pittsburgh, Pennsylvania, ACM Press, 1999, pp 96–103.

[8] (111) Gaver WW, "The Sonicfinder: An Interface that Uses Auditory Icons", Human Computer Interaction, vol.4, no.1, 1989, pp 67–94.

[9] (112) Gaver WW, Smith RB, O'Shea T, "Effective Sounds in Complex Systems: The Arkola Simulation", ACM Transactions on Computer-Human Interaction (CHI), vol.1, New Orleans, ACM Press, 1991, pp 85–90.

[10] (105) Sawhney N, Schmandt C, "Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments", ACM Transactions on Computer-Human Interaction (CHI), vol.7, num.3, ACM Press, 2000, pp 353–383.

[11] (80) Bregman AS, "Auditory Scene Analysis: The Perceptual Organization of Sound", MIT Press, 1990, ISBN: 0262022974.

[12] (113) Brown L et al, "Browsing Modes for Exploring Sonified Line Graphs", Proceedings of British HCI 2002, vol.2, London, UK, 2002, pp 6-9.

[13] (114) Hankinson JCK, Edwards ADN, "Musical Phrase-Structured Audio Communication", International Conference on Auditory Display (ICAD), Atlanta, Georgia, 2000.

[14] (24) Warburton A O, "Harmony: New Edition", Cambridge University Press, 1995, ISBN 0521569028.

[15] (56) C Adams, "Melodic Contour Types," Ethnomusicology, vol. 20, no. 2, pp 179-215, 1976.

[16] (22) Taylor E, "The AB Guide to Music Theory, Part 1", Assoc. Board of the Royal Schools of Music, 1998. ISBN 1854724479.

[17] (178) Sadie S, "The New Grove Dictionary of Music and Musicians", vol.5, 2001, ISBN 0333608003.

[18] (179) Plomp R, "Pitch, timbre and hearing theory", International Audiology, vol.7, 1968, pp. 322-344.

[19] (57) SA Brewster, PC Wright, DN Edwards, "Guidelines for the Creation of Earcons," Proceedings of BCS-HCI 95, vol. 2, pp. 155-159, 1995.

[20] (180) Brice R, "Music Engineering, 2nd Edition", Newnes, 2001, ISBN 0750650400.

[21] (28) Rigas D, Alty J, "How Can Multimedia designers Utilise Timbre?", People and Computers XIII, Springer-Verlag 1998, ISBN 3540762612.

[22] (153) "The Complete Midi 1.0 Detailed Specification", Midi Manufacturers Association, 2001.

[23] (10) Moore BCJ, "An Introduction to the psychology of hearing", Academic Press, 1997, ISBN 0-12-505627-3.

[24] (164) Hart S, Staveland L, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research", Human Mental Workload, Hancock P, Meshkati N (Eds), North Holland B.V., Amsterdam, 1988, pp. 139-183.

[25] (25) Cullen C, Coyle E, "TrioSon: A Graphical User Interface for Pattern Sonification", International Conference on Auditory Display (ICAD), Limerick, 2005.