2016-6

Human Pattern Recognition in Data Sonification

Charlie Cullen
Technological University Dublin, charlie.cullen@tudublin.ie

William Coleman
Technological University Dublin, d15126149@mytudublin.ie

HUMAN PATTERN RECOGNITION IN DATA SONIFICATION

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

William Coleman
Dublin Institute of Technology
d15126149@mydit.ie

ABSTRACT

Computational music analysis investigates the relevant features required for the detection and classification of musical content, features which do not always directly overlap with musical composition concepts. Human perception of music is also an active area of research, with existing work considering the role of perceptual schema in musical pattern recognition. Data sonification investigates the use of non-speech audio to convey information, and it is in this context that some potential guidelines for human pattern recognition are presented for discussion in this paper. Previous research into the role of musical contour (shape) in data sonification shows that it has a significant impact on pattern recognition performance, whilst investigation in the area of rhythmic parsing made a significant difference in performance when used to build structures in data sonifications. The paper presents these previous experimental results as the basis for a discussion around the potential for inclusion of schema- based classifiers in computational music analysis, considering where shape and rhythm classification may be employed at both the segmental and supra-segmental levels to better mimic the human process of perception.

1. INTRODUCTION

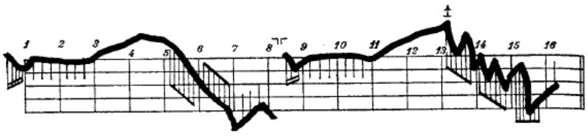
The innate audio processing capability of all humans (and indeed most animals (Kaas, Hackett, & Tramo, 1999)) is amply demonstrated by the ability of infants to discriminate between pitches (Olsho, Koch, & Halpin, 1987), melodic contour (Trehub, Bull, & Thorpe, 1984) and rhythm (Trehub & Thorpe, 1989) as well as an adult can. This ability even extends to the segmenting of melodies (Thorpe & Trehub, 1989) into smaller phrases, and the association of music with other events (Fagen et al., 1997) in a similar manner to adults. The mechanism for such processing is musically specific, with certain neurons directly responsible for pitch perception, rhythm and melodic contour (Johnsrude, Penhune, & Zatorre, 2000; Weinberger & McKenna, 1988) being found only in the right hemisphere of the brain (Trehub et al., 1984).

Perception is a subjective manner of assessment, as by definition differences in perception account for subjective opinion and hence do not easily conform to standardisation. The pitch, loudness or location of sounds can help define their similarity- as can their individual timbres. Also the temporal variations of sounds (such as modulations over time or even their initial onset), can lead to sounds being perceived as grouped or separate- relative to their occurrence and subsequent change (Bregman, 1993). Physically, the fundamental frequency of a sound (and its associated harmonic series) is important in distinguishing between separate sources, as sounds of different fundamental frequency can be detected as separate rather than fused. The rhythmic components of a source also play a major role in its detection (Deutsch, 1980) and recognition, and different rhythmic patterns allow sounds of often similar timbre and pitch to be perceived as separate rather than fused (Bregman, 1993).

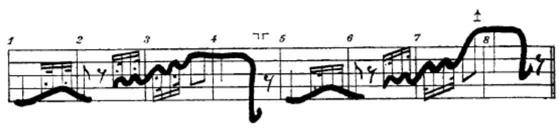
Some studies of the mechanics of human audio perception suggest that the requirements for detection and recognition of melodic patterns are different (Hébert & Peretz, 1997), where long-term memory pattern recognition is biased more towards melodic factors than the rhythmic elements required by pattern detection. Although not an arrhythmic condition by any means, a preference is exhibited for melodic criteria when testing the ability of participants to recognise previously introduced patterns. For this reason, the work presented in this paper distinguishes between recognition using contour (shape) and detection using rhythm, aiming to illustrate the crucial role of both criteria in human perception of sound and music.

2. CONTOUR PATTERN RECOGNITION

Melodic contour has been considered by many musicologists as a means of defining relative changes in pitch (Toch, 1948) (with respect to time), rather than the definition of absolute values. In this manner, the shape, direction and range of a melody can all be summarised by its overall contour. Graphical contour representations were considered by composers such as Schoenberg (Schoenberg & Strang, 1967) as a means of supplementing a musical score (Figure 1):



Menuetto, String Quartet in D, K. 575, mvt. III, mm. 1-16



Andante, Symphony 39, K. 543, mvt. II, mm 1-8

Figure 1. Contour Graphs of Selected Mozart Compositions, taken from Schoenberg (Schoenberg & Strang, 1967)

Contour can be considered an important part of musical memory. Dowling (Dowling, 1978) suggests that contour information functions separately and independently from scalar information in memory. Experiments by Edworthy (Edworthy, 1983) showed that single pitch alterations in a melody could be detected by participants as changes in contour- even when they were unable to define what pitch had been actually altered in the pattern. This capability is believed to be present in infancy (Chang & Trehub, 1977) (around 5 months), at a stage of development where changes in pitch cannot be recognised. It has also been shown that different brain cells are used in the processing of melodic contour (Weinberger & McKenna, 1988) than are used in the detection of temporal or harmonic (Sutter & Schreiner, 1991) components of music. This aspect of neural activity would again suggest that different parts of the brain are used (Zatorre, 1999) in the detection and recognition of musical events: rhythmic factors being paramount in detection, while melodic contour and range (Dowling, 1991; Massaro, Kallman, & Kelly, 1980) and being more important in the recognition of familiar and recently learned melodies.

In previous research into the use of contour (Cullen & Coyle, 2005, 2006), multimodal patterns defined as contour icons were developed to exploit gestalt concepts of good continuation and belongingness (Bregman, 1993) (Figure 2):

Figure 2. Example Up and Down Contour Icons, with associated musical score representations

Testing was then performed to assess whether contour icons were more memorable than low-level earcon pattern designs (Hankinson, John & Edwards, 2000) within a data sonification, to determine the effect of shape on pattern recognition (Figure 3):

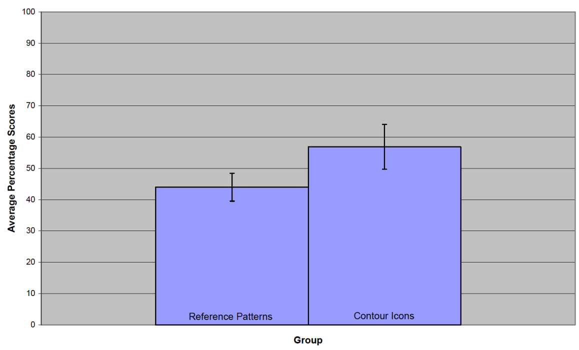


Figure 3. Graph showing overall average percentage scores for recognition of low-level patterns and contour icons in a data sonification, showing standard deviations

Results showed that performance had improved from 44% in the low-level (earcon) reference pattern condition to 56.87% in the contour icon condition, a significant improvement ($T(20) = -3.68$, $p=0.0007$) that suggests contour icons are more memorable than low level reference patterns that do not employ shape as a melodic feature. Post-test Task Load Index testing (Hart, Sandra, 2006) that examines participant workload during a task showed a significant reduction ($T(20) = 4.53$, $p<0.0001$) in overall workload from 50.33 to 36.25 for the contour icon condition.

Though no reduction was significant in any individual category, the scores were lower for the contour icon condition in each case. Having said this, higher data to pattern combinations had proven less effective, and it was observed on several occasions that whilst participants could recognise a particular contour icon they were subsequently unable to remember its data mapping. This suggests that the abstract nature of the mapping between value and contour icon was difficult to remember for some participants, though this may not necessarily interfere with the use of shape as an aid to recognition.

Although significant for data sonification, the role of contour in musical pattern recognition requires further investigation in relation to its potential role in computational music analysis. Some consideration has been given to the concept of stream analysis of musical segments (Rafailidis et al., 2008), whilst Karydis et al (Karydis, Nanopoulos, A., Papadopoulos, & Cambouropoulos, 2007) define a computational model of the musical score that includes the concept of a perceptual 'voice' within the overall auditory stream. It is argued that contour may play a significant role within such models, given it's demonstrable effect on human musical pattern recognition.

3. RHYTHMIC PATTERN DETECTION

Rhythm is a fundamental building block of musical composition (Taylor, 1989) that serves to group various sonic events within a piece for aesthetic purposes. In sonification research, rhythm can be employed to group patterns used to represent data for analysis so that they may be more efficiently processed by the listener. It is argued that rhythm is a fundamental component of all human interactions (Jones, 1976), and so is similarly fundamental to the communication of effective musical patterns to a listener.

In the case of infants, the role of rhythm is the most fundamentally important aspect of early cognitive development (Zentner & Eerola, 2010), and is believed to begin in the womb (where the child is often observed to move in response to rhythms in speech or music). Infants display several common rhythms (Fridman, 1991), which are used to seek attention from their parents or other adults. This use of rhythmic patterns is both frequent and essential (Kempton, 1980) in the communication between infant and adult, communication that is dictated by a pulse common to all parties. Indeed, the variation or absence of such rhythmic components is observed to engender disinterest and negative responses from the child involved (Drake, Jones, & Baruch, 2000).

Rhythm dictates the structure of a piece of music, from the individual sequence of notes to the hierarchical groupings of different musical phrases or passages. The ability of musicians to detect and convey complex structures (Jongsma, Desain, & Honing, 2004) within a piece is a direct result of training and experience, the lack of which effectively reduces rhythmic patterns to sequential processes. This means of structuring music relies heavily on the metrical organization (Essens, 1995) of such rhythmic patterns into regular frameworks, utilising the time signature of the piece to define different sections. Thus rhythm allows a piece of music to be organised into sections- sections of differing levels of complexity. By defining the bar (or measure) in terms of the beat, the basic organisational structure of a piece of music is decided. When this bar structure is then further organised into sections (such as the simple verse and chorus of popular music) it allows differing pieces of related musical information to be conveyed in a structured manner.

In previous research (Cullen & Coyle, 2003, 2006), rhythm was investigated as part of a strategy to sonify data, and the specific role of rhythmic parsing was subsequently investigated in the sonification of (fictitious) exam results (Cullen, Coyle, & Russell, 2005). Test participants were informed they would be asked questions on a sonification of 20 exam results, which contained 4 distinct course groups (with 5 members each) in sequential order. The test used rest notes between course groups in the parsing condition, compared to a single grouping of musical events in the control condition, as a means of using rhythm to delineate groupings (or structures) within the data. Participants were asked questions that compared the data of each group (e.g.

which group had a higher pass rate) to determine the effect of adding rhythmic gaps to the processing of information in the sonification (Figure 4):

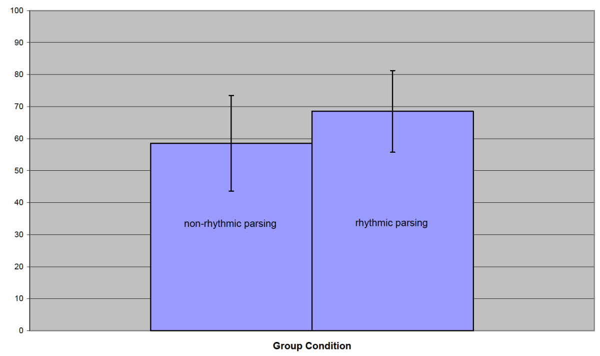


Figure 4. Graph showing overall average percentage scores (by test condition) for rhythmic parsing of a data sonification, showing standard deviations

Overall results showed performance improved to 75.3% in the rhythmic parsing condition from 67.6% in the non-rhythmic parsing condition. This improvement was significant ($T(20) = -2.79$, $p=0.008$), suggesting that rhythmic parsing had a positive effect on performance in multiple stream sonification. In addition, post-test TLX questions relating to the workload involved in analyzing a data sonification showed a significant reduction in overall workload from 60.75 to 41.33 in the rhythmic parsing condition ($T(20) =7.45$, $p<0.001$), with significant reductions in temporal demand (16.33 to 7.65, $T(20) =6.236$, $p<0.001$), effort (9.95 to 4.583, $T(20) =4.435$, $p<0.001$), and frustration (7.983 to 4.05, $T(20) =2.966$, $p=0.005$).

These results suggested that participants had found the rhythmic parsing condition a more effective method of representing sub-groups in a data sonification, though the use of a rest note to parse the data arguably serves only to indicate a change in the current context within the sonification. A more effective method of rhythmic parsing could employ features such as markers and labels (Smith & Walker, 2002), in combination with rest notes to better mimic the compositional use of rhythm as a means of grouping motifs and patterns into distinct structures within a larger piece (Barry, Gainza, & Coyle, 2007).

4. DISCUSSION & FUTURE WORK

This section is still to be completed, but will consider the following 3 areas:

- Hierarchical models for short-term/long-term structures- Contour?
- Measuring relevance of different musical properties and structure principles- Dan & Mikel (Barry et al., 2007)
- Developing taxonomies/ontologies for structure annotation- Rhythm & Contour.

5. REFERENCES

Barry, D., Gainza, M., & Coyle, E. (2007). Music Structure Segmentation using the Azimugram in conjunction with Principal Component Analysis. In *Audio Engineering Society, 123rd Convention* (pp. 1–8).

Bregman, A. S. (1993). Auditory scene analysis: hearing in complex environments. *Thinking in Sound: The Cognitive Psychology of Human Audition*.

Chang, H. W., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, *24*(2), 324–331. doi:10.1016/0022-0965(77)90010-8

Cullen, C., Coyle, D. E., & Russell, D. N. (2005). *The Sonic Representation of Mathematical Data. Faculty of Engineering and Faculty of Applied Arts*. Dublin Institute of Technology.

Cullen, C., & Coyle, E. (2003). Rhythmic Parsing of Sonified DNA and RNA Sequences. *Irish Signals and Systems Conference, ISSC 2003*.

Cullen, C., & Coyle, E. (2005). Musical Pattern Design Using Contour Icons. *Eleventh Meeting of the International Conference on Auditory Display (ICAD 05)*.

Cullen, C., & Coyle, E. (2006). Harmonically Combined Contour Icons for Concurrent Auditory Display. In *Proc. IET Irish Signals and Systems Conf* (pp. 501–506).

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception {&} Psychophysics*, *28*(5), 381–389. doi:10.3758/BF03204881

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*(4), 341–354. doi:10.1037/0033-295X.85.4.341

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception &: Psychophysics*, *50*(4), 305–313. doi:10.3758/BF03212222

Drake, C., Jones, M. R., & Baruch, C. (2000). *The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending*. *Cognition* (Vol. 77). doi:10.1016/S0010-0277(00)00106-2

Edworthy, J. (1983). The Acquisition of Symbolic Skills. In D. Rogers & J. A. Sloboda (Eds.), (pp. 263–271). Boston, MA: Springer US. doi:10.1007/978-1-4613-3724-9_30

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics*, *57*(4), 519–532. doi:10.3758/bf03213077

Fagen, J., Prigot, J., Carroll, M., Pioli, L., Stein, A., & Franco, A. (1997). Auditory context and memory retrieval in young infants. *Child Development*, *68*(6), 1057–1066. doi:10.1111/j.1467-8624.1997.tb01984.x

Fridman, R. (1991). Proto-rhythms: Basis for the birth of musical intelligence and language expression. *Journal of Prenatal & Perinatal Psychology & Health*.

Hankinson, John, C. ., & Edwards, A. D. N. (2000). Musical Phrase-Structured Audio Communication. *Proceedings of the 6th International Conference on Auditory Display, Atlanta, GA, USA, 2000.*

Hart, Sandra, G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Human Factors and Ergonomics Society Annual Meting*, 904–908. doi:10.1037/e577632012-009

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: are melodic and temporal patterns equal partners? *Memory & Cognition*, *25*(4), 518–533. doi:10.3758/BF03201127

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000).

Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain : A Journal of Neurology*, *123 (Pt 1*, 155–163. doi:10.1093/brain/123.1.155

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323–355. doi:10.1037/0033-295X.83.5.323

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological Psychology*, *66*(2), 129–152. doi:10.1016/j.biopsycho.2003.10.002

Kaas, J. H., Hackett, T. A., & Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Current Opinion in Neurobiology*, *9*(2), 164–170. doi:10.1016/S0959-4388(99)80022-1

Karydis, I., Nanopoulos, A., Papadopoulos, A., & Cambouropoulos, E. (2007). Visa : the Voice Integration / Segregation Algorithm, (April).

Kempton, W. (1980). The rhythmic basis of interactional micro-synchrony. *The Relationship of Verbal and Nonverbal Communication*, 67–75.

Massaro, D. W., Kallman, H. J., & Kelly, J. L. (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(1), 77–90. doi:10.1037/0278-7393.6.1.77

Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *The Journal of the Acoustical Society of America*, *82*(2), 454–464. doi:10.1121/1.395446

Rafailidis, D., Nanopoulos, A., Cambouropoulos, E., Manolopoulos, Y., Science, C., & Studies, M. (2008). Detection of stream segments in symbolic musical data. In *The International Society of Music Information Retrieval (ISMIR 2008)* (pp. 83–88).

Schoenberg, A., & Strang, G. (1967). *Fundamentals of music composition*. St. Martin's Press.

Smith, D. R., & Walker, B. N. (2002). Tick-marks, Axes, and Labels: The Effects of Adding Context to Auditory Graphs. *International Conference on Auditory Display*, 1–6.

Sutter, M. L., & Schreiner, C. E. (1991). Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J Neurophysiology*, *65*(5), 1207–1226. Retrieved from http://jn.physiology.org/content/65/5/1207

Taylor, E. (1989). *The AB Guide to Music Theory: Part 1*. Associated Board of the Royal Schools of Music.

Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, *25*(1), 122–127. doi:10.1037/0012-1649.25.1.122

Toch, E. (1948). *The shaping forces in music: An inquiry into harmony, melody, counterpoint, form*. Criterion Music Corporation.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, *55*(3), 821–830. doi:10.1016/S0163-6383(84)80430-0

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *43*(2), 217–229. doi:10.1037/h0084223

Weinberger, N. M., & McKenna, T. M. (1988). Sensitivity of Single Neurons in Auditory Cortex to Contour: Toward a Neurophysiology of Music Perception. *Music Perception: An Interdisciplinary Journal*, *5*(February), 355–389. doi:10.2307/40285407

Zatorre, R. J. (1999). Brain imaging studies of musical perception and musical imagery. *Journal of New Music Research*, *28*(3), 229–236. doi:10.1076/jnmr.28.3.229.3112

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5768–5773. doi:10.1073/pnas.1000121107