Cross Entropy as a Measure of Musical Contrast

Robin Laney[1], Robert Samuels[1], and Emilie Capulet[2]

[1] Open University,
Milton Keynes MK7 6AA, UK
[2] London College of Music,
University of West London, W5 5RF, UK

Abstract. We present a preliminary study of using the information theoretic concept of cross entropy to measure musical contrast in a symbolic context, with a focus on melody. We measure cross entropy using the Information Dynamics Of Music (IDyOM) framework. Whilst our long term aim is to understand the use of contrast in sonata form, in this paper we take a more general perspective and look at a broad spread of Western art music of the common practice era. Our results suggest that cross entropy has a useful role as an objective measure of contrast, but that a fuller picture will require more work.

Keywords: Contrast, Similarity, Cross entropy, N gram model, Markov model

1 Introduction

In Western art music of the common practice era, similarity and contrast play key roles. Similarity to a corpus is a way of establishing style, and similarity within a work provides coherence. Contrast, on the other hand, not only adds interest, but is essential to the aesthetic illusion of dramatic resolution of conflict. In terms of empirical understanding, similarity has been widely studied in terms of music information retrieval techniques, but contrast much less so.

Conklin and Witten [4], Pearce [8], and Whorley [12], in principled approaches to generation from a corpus, use a low value for cross entropy between a generated piece and a model of the corpus as a measure of stylistic coherence. Thus low cross entropy is used as a model of similarity. In this paper we investigate whether high entropy between motifs or themes might be a good measure of contrast, on the basis that both similarity and contrast are related to listeners expectations being respectively fulfilled or denied. We build on the Information Dynamics Of Music (IDyOM) framework [8] [9].

In this paper we focus on melody. Whilst it is clear that many other features play a key role in establishing contrast, in particular dynamics, texture, harmony, timbre, and articulation, there is a sense in which melody is central, for example in analysing a theme a musicologist will typically look at the shape of the melody.

In the results section we take five pieces from composers of the common practice period and measure cross entropy between patterns representing themes, subjects or motifs that might be expected to contrast. As a comparison we also look at cross entropy between the pieces and the patterns. Our measurements are of intervals and durations, leaving the investigation of other features, including combinations, to later work. A long term motivation is to see whether musically meaningful contrast between themes and motifs, ascribed or assumed by musicological writing, can be correlated with objectively measurable properties.

2 Background

In this section we briefly introduce the use of entropy in the context of music, and consider the relationship of similarity and contrast, and the extent to which a study of the latter can draw on work on the former.

Shannon's notion of entropy in the context of his theory of information [10] has been influential in the field of computational musicology. The use of cross entropy in a probabilistic predictive/generative context was introduced by Conklin and Witten [4] who also note that Meyer [6] relates musical experience and entropy. Collins et al. [3] use entropy as part of a stylistic consistency metric. The IDyOM model of Pearce [8] [9] combines the predictions of multiple length n gram models, using Prediction by Partial Match, a technique developed by Cleary and Witten [1]. The following equation follows Pearce [8].

Let a sequence $(a_1...a_n)$ be represented as a_1^n. For a given a value of n, an n gram model predicts the probability of the next event a_n given a previous sequence a_1^{n-1}. That is $p(a_n \mid a_1^{n-1})$. If we let m be a model, that can predict, using variable lengths of initial sequences, probabilities p_m, then the cross entropy $H(p_m, a_1^j)$ is:

$$\begin{aligned} H_m(p_m, a_i^j) &= -\frac{1}{j}\log_2 p_m(a_1^j) \\ &= -\frac{1}{j}\sum_{i=1}^{j}\log_2 p_m(a_i \mid a_1^{i-1}) \end{aligned} \tag{1}$$

Cross entropy, is a measure of how likely the sequence $a_1...a_n$ is relative to the model m. Low values suggest that if the model was built from a corpus, then the sequence is similar to the corpus, and we explore in this paper whether high values correspond to a sequence that contrasts with the corpus. In fact we will compare individual phrases by using one phrase as the corpus, and similarly use the piece as a corpus to measure the contrast of a phrase with the overall piece. We are interested in whether cross entropy is a useful model of contrast.

Similarity in music is non-trivial, as musical segments can be analysed from many perspectives, using a variety of models. Wiggins et al. [13] argue that music is a subjective experience that in a sense only exists in the mind of listeners. So what models and measures correspond to human experience is a non-trivial question requiring empirical work. A similar argument can be applied to contrast, but more initial work is required. Given the multi-dimensional nature of both similarity and contrast, and the subjective nature of music, we should not assume that contrast and similarity are simply opposites. Tversky [11] in seminal work taking a feature based approach to similarity makes a similar point, as well as showing that similarity (and by implication) contrast are context specific.

In a similar vein, Marsden [5] has questioned whether musical similarity is a definitive phenomenon or a product of interpretation. He reviews a number of approaches: the edit distance between melodies, possibly under Time-Warping; structural models based on reduction; and feature extraction combined with machine learning or statistical analysis. Whilst contrast is not simply the opposite of similarity, it is worth considering how similarity measures can be applied to contrast. Reduction approaches remove decorative notes leaving structural ones, though from the point of view of contrast such a dichotomy might be misleading in that deviance from a structure might be a way of establishing a contrast. Edit distance methods appear appropriate if we are dealing with segments of the same length or where one is an elongation of the other. The advantage of a cross entropy based approach is that it allows us to compare the character of two segments regardless of length considerations. We can ask for example, if a particular short motif provides contrast after a long passage.

3 Corpus

We used the contents of the Johannes Kepler University Patterns Development Database training corpus [2]: Orlando Gibbons' The Silver Swan (1612), Johann Sebastian Bach's Fugue in A minor, BWV 889, Wolfgang Amadeus Mozart's Minuet from Piano Sonata in E flat major, K. 282, Ludwig van Beethoven's Minuet and Trio from Piano Sonata in F minor, op. 2, no. 1, and Frédéric Chopin's Mazurka in B minor, op. 24, no. 4. This gives a broad spread across music of the common practice period. The training corpus contains a number of patterns (labelled A,B...), that according to accepted conventions would contrast as follows:

- A and C in the Gibbons (ascending v descending),
- A and B in the Bach (in the sense of subject and countersubjects),
- C and D in the Mozart (first and second themes in a minuet),
- C and D in the Beethoven (themes from a minuet and trio),
- B and D in the Chopin (themes from a mazurka).

The corpus database includes scores for the pieces and patterns.

4 Methodology

Using the IDyOM software we measured the cross entropy between a number of combinations of patterns and the whole piece. IDyOM constructs models that are multi order Markov models. We label the phrases under consideration a and b and the whole piece p, and denote the cross entropies of interest as:

- CE(a,b) Cross entropy for pattern a with respect to a model of pattern b
- CE(b,a) Cross entropy for pattern b with respect to a model of pattern a
- CE(a,p) Cross entropy for pattern a with respect to a model of the piece
- CE(b,p) Cross entropy for pattern b with respect to a model of the piece

Additionally, for completeness and transparency we measured:

- CE(p,a) Cross entropy for piece with respect to a model of pattern a
- CE(p,b) Cross entropy for piece with respect to a model of pattern b

We also take averages of CE(a,b) and CE(b,a) (labelled pattern in tables in section 5) and averages of CE(a,p) and CE(b,p) (labelled pattern-piece in tables in section 5) in order to give an overall idea of the cross entropy between patterns and the cross entropy between the patterns and the whole piece respectively. A working hypothesis is that if:

- a) cross entropy is a good model of contrast (and similarity) and
- b) pieces are coherent through similarity with interest generated by contrast between the identified themes

then we expect that the average of CE(a,b) and CE(b,a) to be greater than the averages of CE(a,p) and CE(b,p). However we also need to take into account that contrast might, for example, be established through pitch whilst keeping durational content more similar.

5 Results

In the following two tables we see our results for the cross entropy of pitch predicted using a model of intervals and for the cross entropy of durations predicted using a model of durations.

Composer	CE(a,b)	CE(b,a)	CE(a,p)	CE(b,p)	CE(p,a)	CE(p,b)	pattern	pattern-piece
Gibbons	3.03	3.57	2.09	2.95	4.45	4.42	3.30	2.52
Bach	5.23	4.64	1.94	0.83	6.02	3.61	4.93	1.38
Mozart	3.10	3.00	1.24	2.79	4.48	4.26	3.05	2.02
Beethoven	3.13	3.16	2.19	0.80	4.17	3.87	3.14	1.50
Chopin	3.23	3.39	0.71	1.03	4.16	4.53	3.31	0.87

Table 1. Cross entropy of pitch predicted using a model of intervals

For both pitch and durations, the average entropies between patterns (labelled pattern) is larger than that between the patterns and the whole piece (labelled pattern-piece). So the data is consistent with our working hypothesis. However, an important confounding factor relates to differences in the sizes of the corpus used in the comparisons. Clearly, the corpus for a comparison between two patterns is smaller than when using a corpus of the whole piece and

Composer	CE(a,b)	CE(b,a)	CE(a,p)	CE(b,p)	CE(p,a)	CE(p,b)	pattern	pattern-piece
Gibbons	1.83	2.23	1.23	1.34	3.11	1.98	2.03	1.28
Bach	3.87	4.23	1.43	0.70	4.16	2.32	4.05	1.07
Mozart	2.59	2.56	0.76	1.62	2.54	2.90	2.57	1.19
Beethoven	3.40	4.40	0.27	0.28	3.15	2.26	3.90	0.28
Chopin	2.61	2.40	0.77	1.01	3.49	3.91	2.50	0.89

Table 2. Cross entropy of durations predicted using a model of durations

there will often be a fall in cross entropy as corpus size increases. We also need a clearer idea of what magnitude of ratio suggests a given strength of contrast. One approach might be to look at a range of patterns across a piece and use an average as a baseline.

6 Conclusions

There is much existing work on similarity, although that work is far from complete. However, contrast, has received little attention. Given the centrality of identifying contrasts when carrying out analysis, a better understanding of relevant objective features would enrich the analytical process, and widen the scope of digital tools. It would also support a fresh look at historical musicological writings on the role of contrast in a range of forms.

Why should we wish to provide digital tools? Currently the degree of training required for a competent analysis is very steep. Moreover, once one commits to a particular way of viewing a piece in terms of its structure, it is expensive time-wise to consider alternatives. Yet, often it is the tension between alternative interpretations of similarity and contrast that leads to the sublime nature of a work. See Marsden [5] for a highly pertinent example. Automated tools would widen the likelihood of working with alternative readings and make it feasible for a wider group to engage with analysis, such as instrumental teachers, conservatoire staff, and professional performers. Currently there is much interest in electronic editions such as the Online Chopin Variorum Edition [7] which allow a much more flexible attitude towards score editions, their alternatives, and blended versions. Easy to use tools would further open up and democratise an area of work which has sometimes been seen as the preserve of a small elite.

For music in the common practice period, coherence depends on adhering to the style, but interest is generated through contrast. From this assumption it is easy to see a potential weakness of generating a piece from a single Markov model. The problem is that contrasts in pieces within a corpus might effectively get averaged out in the model and when generating pieces this will be reflected.

The picture painted by our numerical data was consistent with our working hypothesis. However, differences in the sizes of corpus used in each comparison is a compounding issue that needs further work. The overall corpus, whilst being wide-ranging historically, was of limited size. Further work is needed to give a firmer idea of how cross entropy and contrast relate and also how the measure compares with others, or indeed could be meaningfully combined with them. Further work might also look at layers of contrast: If a motif A contrasts with a motif B, this might be perceived as a contrast at a micro-level. But if A and B recur, they are still contrasting, but the question is whether A and B together create a new level of contrast with another, more large-scale, element of the music, lets say C, and so forth.

Acknowledgements. We are grateful for discussions and advice to Byron Dueck, Tom Collins, Alan Marsden, Marcus Pearce, Raymond Whorley, Geraint Wiggins, and Alistair Willis.

References

1. Cleary, J. G. and Witten, I. H.: Data compression using adaptive coding and partial string matching. IEEE Trans. Communications 32(4), 396-402 (1984)
2. Collins T., Böck S., Krebs F., and Widmer G.: Bridging the audio-symbolic gap: the discovery of repeated note content directly from polyphonic music audio. In Proceedings of the Audio Engineering Society's 53rd Conference on Semantic Audio, 12 pages, London, (2014)
3. Collins, T., Laney, R., Willis, A. and Garthwaite, P.H.: Developing and evaluating computational models of musical style. Artificial Intelligence for Engineering Design, Analysis and Manufacturing. in press
4. Conklin, D. and Witten, I.H.: Multiple viewpoint systems for music prediction Journal of New Music Research. 24(1), 51-73 (1995)
5. Marsden, A.: Interrogating melodic similarity: a definitive phenomenon or the product of interpretation? Journal of New Music Research. 41(4), 323-335 (2012)
6. Meyer, L.B.: Meaning in music and information theory. Journal of Aesthetics and Art Criticism 15, 412-424 (1957)
7. Online Chopin Variorum Edition `http://www.ocve.org.uk/index.html`
8. The construction and evaluation of statistical models of melodic structure in music perception and composition. Pearce, M.T. PhD Thesis. City University London (2005)
9. Pearce M.T., GA Wiggins, G.A.: Auditory Expectation: The Information Dynamics of Music Perception and Cognition. Topics in Cognitive Science. 4(4), 625-652 (2012)
10. Shannon C.E: A mathematical theory of communication. Bell Syst. Tech. J., vol. 27, 379-423, (1948)
11. Tversky, A.: Features of similarity. Psychological Review 84,327-352. (1977)
12. The Construction and Evaluation of Statistical Models of Melody and Harmony. Whorley, R.P. PhD. thesis. Goldsmiths, University of London. (2013)
13. Wiggins G.A., Müllensiefen, D, Pearce, M.T.: On the non-existence of music: Why music theory is a figment of the imagination. Musicae Scientiae vol. Discussion Volume 5, 231-255 (2010)