

The Melody Triangle - Pattern and Predictability in Music

Henrik Ekeus
Queen Mary University of
London
Media and Arts Technology
School of Electronic
Engineering and Computer
Science
hekeus@eecs.qmul.ac.uk

Mark D. Plumbley
Queen Mary University of
London
Center for Digital Music
School of Electronic
Engineering and Computer
Science
mark.plumbley@
eecs.qmul.ac.uk

Samer Abdallah
Queen Mary University of
London
Center for Digital Music
School of Electronic
Engineering and Computer
Science
samer.abdallah@
eecs.qmul.ac.uk

Peter W. McOwan
Queen Mary University of
London
Computer Vision Group
School of Electronic
Engineering and Computer
Science
Peter.McOwan@
eecs.qmul.ac.uk

ABSTRACT

The Melody Triangle is a Markov-chain based melody generator where the input - positions within a triangle - directly map to information theoretic measures of its output. The measures are the entropy rate, redundancy and *predictive information rate*[1] in the melody. Predictive information rate is an information measure developed as part of the Information Dynamics of Music project (IDyOM)¹. It characterises temporal structure and is a way of modelling expectation and surprise in the perception of music.

We describe the Information Dynamics model and how it forms the basis of the Melody Triangle. We outline two interfaces and uses of the system. The first is a multi-user installation where collaboration in a performative setting provides a playful yet informative way to explore expectation and surprise in music. The second is a screen based interface where the Melody Triangle becomes a compositional tool for the generation of intricate musical textures using an abstract, high-level description of predictability. Finally we outline a study where the screen-based interface was used under experimental conditions to determine how the three measures of predictive information rate, entropy and redundancy might relate to musical preference. We found that...

Keywords

Information dynamics, Markov chains, Collaborative performance, Aleatoric composition

1. INFORMATION DYNAMICS

¹<http://www.idyom.org/>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.
Copyright remains with the author(s).

Music involves patterns in time. When listening to music we continually build and re-evaluate expectations of what is to come next. Composers commonly, consciously or not, play with this expectation by setting up expectations which may, or may not be fulfilled. This manipulation of expectation and surprise in the listener has been articulated by music theorist Meyer[3] and Narmour[5]. Core to this is the idea that music is not a static object presented as a whole, as a Lerdahl and Jackendo analysis[2] would imply, but as a phenomenon that ‘unfolds’ and is experienced *in time*.

The information dynamic approach[1] considers several different kinds of predictability in musical patterns, how human listeners might perceive these, and how they shape or affect the listening experience. Central to this is the idea that listeners maintain a dynamically evolving statistical model that enables them to make predictions about how a piece of music will continue. They do this using both the immediate context of the piece, and on previous musical experience, and as the music unfolds they continually revise their model.

2. THE MELODY TRIANGLE

The Melody Triangle is based on first order Markov chains, represented as transition matrixes, that generate streams of symbols. By mapping the symbols to individual notes, melodies are generated. Further by layering these streams of notes can result in intricate musical textures. The choice of notes or scale is not a part of the Melody Triangle’s core functionality, in fact the symbols could be mapped to anything, even non sonic outputs.

Any sequence of symbols can be analysed and information theoretic measures taken from it. The novelty of the Melody Triangle lies in that we go ‘backwards’ - given desired values for these measures, as determined from the user interface, we return a stream of symbols that match those measures. The information measures used are redundancy, entropy rate and predictive information rate.

2.1 Information measures

2.1.1 Redundancy

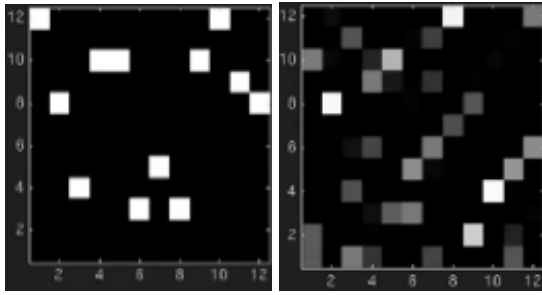


Figure 1: Two transition matrixes. The shade of white represents the probabilities of transition from one symbol to the next (black=0, white=1). The current symbol is along the bottom, and in this case there are twelve possibilities (mapped to a chromatic scale). The left hand matrix has no uncertainty; it represents a periodic pattern. The right hand matrix contains unpredictability but nonetheless is not completely without perceivable structure, it is of a higher entropy rate.

[todo - a more formal description] Redundancy tells us the difference in uncertainty before we look at the context (the fixed point distribution) and the uncertainty after we look at context. For instance a transition matrix with high redundancy, such as one that represents a long periodic sequence, would have high uncertainty before we look at the context but as soon as we look at the previous symbol, the uncertainty drops to zero because we now know what is coming next.

2.1.2 Entropy rate

[todo - a more formal description] Entropy rate is the average uncertainty for the next symbol as we go through the sequence. A looping sequence has 0 entropy, a sequence that is difficult to predict has high entropy rate. Entropy rate is an average of ‘surprisingness’ over time.

2.1.3 Predictive Information Rate

[todo - a more formal description] Predictive information rate tell us the average reduction in uncertainty upon perceiving a symbol; a system with high predictive information rate means that each symbol tells you more about the next one. If we imagine a purely periodic sequence, each symbol tells you nothing about the next one that we didn’t already know as we already know how the pattern is going. Similarly with a seemingly uncorrelated sequence, seeing the next symbol does not tell us anymore because they are completely independent anyway; there is no pattern. There is a subset of transition matrixes that have high predictive information rate, and it is neither the periodic ones, nor the completely un-corellated ones. Rather they tend to yield output that have certain characteristic patterns, however a listener can’t necessarily know when they occur. However a certain sequence of symbols might tell us about which one of the characteristics patterns will show up next. Each symbols tell a us little bit about the future but nothing about the infinite future, we only learn about that as time goes on; there is continual building of prediction.

2.2 Making the triangle

We generate hundreds of transition matrixes, representing first-order Markov chains, by a random sampling method. These are then plotted in a 3d statistical space of redundancy, entropy rate and predictive information rate. In figure 2 we see a representation of how these matrixes are

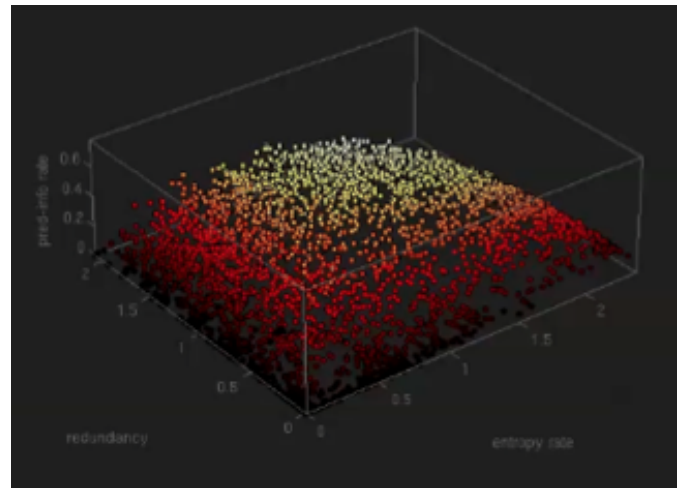


Figure 2: The population of transition matrixes distributed along three axes of redundancy, entropy rate and predictive information rate. Note how the distribution makes a curved triangle-like plane floating in 3d space.

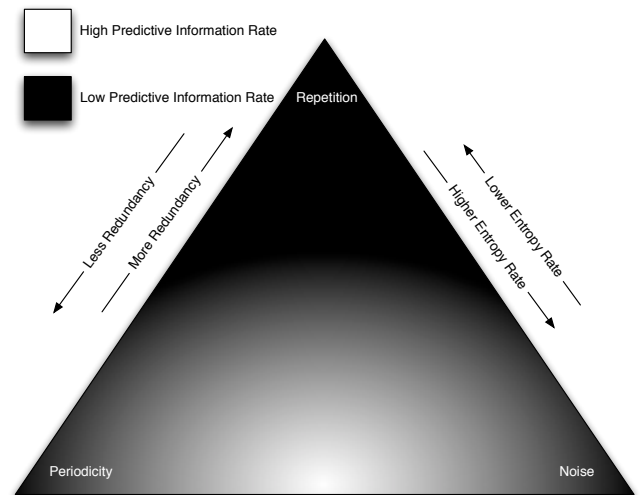


Figure 3: The Melody Triangle [todo fix shading to be more triangular]

distributed; each one of these points corresponds to a transition matrix.

When we look at the distribution of randomly generated transition matrixes and plotted in this space, we see that it forms an arch shape that is fairly thin. It thus becomes a reasonable approximation to pretend that it is just a sheet in two dimensions; and so we stretch out this curved arc into a flat triangle. It is this triangular sheet that is our ‘Melody Triangle’ and forms the interface by which the system is controlled.

When the Melody Triangle is used, regardless of whether it is as a screen based system, or as an interactive installation, it involves a mapping to this statistical space. Then a transition matrix corresponding to this position in statistical space is returned. As can be seen in figure 3, a position within the triangle maps to different measures of redundancy, entropy rate and predictive information rate.

[todo improve, include example melodies?] Each corner corresponds to three different extremes of predictability and

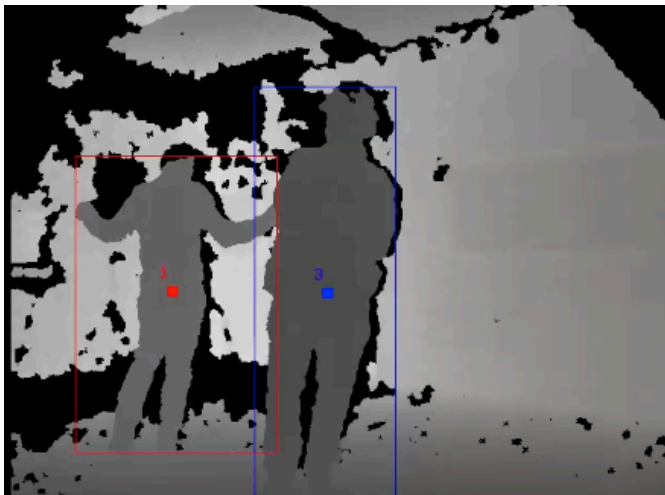


Figure 4: The depth map as seen by the Kinect, and the bounding box outlines the blobs detected by OpenNI.

unpredictability, which could be loosely characterised as periodicity, noise and repetition. Melodies from the ‘noise’ corner have no discernible pattern; they have high entropy rate, low predictive information rate and low redundancy. These melodies are essentially totally random. A melody along the ‘periodicity’ to ‘repetition’ edge are all deterministic loops that get shorter as we approach the ‘repetition’ corner, until it becomes just one repeating note. It is the areas in between that provide the more interesting melodies, those that have some level of unpredictability, but are not completely random and conversely that are predictable, but not entirely so. This triangular space allows for an intuitive exploration of expectation and surprise in temporal sequences based on a simple model of how one might guess the next event given the previous one.

3. USER INTERFACES

The Melody Triangle engine² is controlled over OSC messages and thus any number of interfaces could be developed to for it. Currently two different interfaces exist; a standard screen based interface where a user moves tokens with a mouse in and around a triangle on screen, and a multi-user interactive installation where a Kinect³ camera tracks individuals in a space and maps their positions in the space to the triangle.

3.1 The Multi-User Installation

As a Kinect camera overlooks a space, its the range naturally forms a triangle. As visitors/users comes into the range of the camera, they start generating a melody, the statistical properties of this melody determined by the mapping of physical space to statistical space as discussed above. Thus by exploring the physical space the participant explores the predictability of the generated melodic content. When multiple people are in the space, they can cooperate to create musical polyphonic textures.

The streams of symbols are mapped to MIDI and then played with software instruments in Logic. The tracking system was capable of detecting gestures, and these were mapped to different musical effects such as tempo changes, periodicity changes (going to the off-beat), instrument/register

Table 1: Gestures and their resulting effect

left arm	right arm	meaning
out	static	double tempo
in	static	halve tempo
static	out	triple tempo
static	in	one-third tempo
out	in	shift to off-beat
out	out	change instrument
in	in	reset tempo

changes and volume (see Figure 1).

3.1.1 Tracking and Control

Tracking and control was done using the OpenNI libraries’ API⁴ and high level middle-ware for tracking with Kinect. This provided reliable blob tracking of humanoid forms in 2d space. By triangulating this to the Kinect’s depth map it became possible to get reliable coordinate of visitors positions in the space.

This system was extended to detect gestures. By detecting the bounding box of the 2d blobs of individuals in the space, and then normalising these based on the distance of the depth map it became possible to work out if an individual had an arm stretched out or if they were crouching.

With this it was possible to define a series of gestures for controlling the system without the use of any controllers(see table 1). Thus for instance by sticking out one’s left arm quickly, the melody doubles in tempo. By pulling one’s left arm in at the same time as sticking the right arm out the melody would shift onto the offbeat. Sending out both arms would change instrument.

3.1.2 Observations

Although visitors would need an initial bit of training they could then quickly be made to collaboratively design musical textures. For example, one person could lay down a predictable repeating bass line by keeping themselves to the periodicity/repetition side of the room, while a companion can generate a freer melodic line by being nearer the ‘noise’ part of the space.

The collaborative nature of this installation is one area that merits attention. By not having one user be able to control the whole narrative, the participants would communicate verbally and direct each other in the goals of learning to use the system, and eventually towards finding interesting musical textures. The collaborative nature added an element of playfulness and enjoyment that was obviously apparent.

As an artefact this installation is an exploratory prototype, and occupies an ambiguous role in terms of purpose; it is in a nebulous middle ground between instrument and art installation. One thing is clear is that as a vehicle for communicating ideas related to the expectation, pattern and predictability in music it is very effective.

3.2 The Screen Based Interface

The Melody Triangle can also be explored with a standard keyboard and mouse interface. A triangle is drawn on the screen, screen space thus mapped to the statistical space of the Melody Triangle. A number of round tokens, each representing a melody can be dragged in and around the triangle. When a token is dragged into the triangle, the system will start generating the sequence of notes with statistical properties that correspond to its position in the triangle.

²developed in Prolog and MatLab

³<http://www.xbox.com/en-GB/Kinect>

⁴<http://OpenNi.org/>

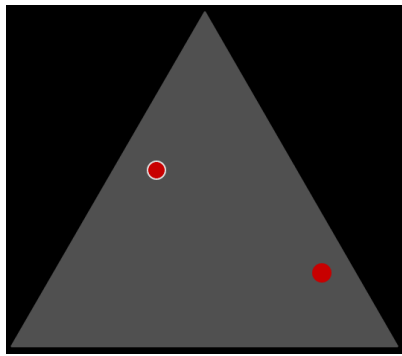


Figure 5: Screen shot of the screen based interface for the Melody Triangle

Additionally there are a number of keyboard controls. These include controls for changing the overall tempo, for enabling and disabling individual voices, changing registers, going to off-beats and changing the speed of individual voices. The system gives some feedback by way of colour changes to indicate when a token has locked on to a new melody, and contains a buffer zone for allowing tokens to be pushed right to the edges of the triangle without falling out.

[TODO: discussion on its use as a composition assistant.. some comments on the aesthetics of the output (why it all sounds like minimalism.) why interesting]

4. MUSICAL PREFERENCE AND INFORMATION DYNAMICS STUDY

We carried out a preliminary study that sought to determine any correlation between aesthetic preference and the information theoretical measures of the Melody Triangle. In this study participants were asked to use the screen based interface of the Melody Triangle. It was simplified so that all they could do was move tokens around. The axes of the triangle would be randomly rearranged for each participant.

The study was divided in to two parts, the first investigated musical preference with respect to single melodies at different tempos. In the second part of the study, a background melody is playing and the participants are asked to find a second melody that 'works well' with the background melody. For each participant this was done four times, each with a different background melody from four different areas of the Melody Triangle.

After the study the participants were surveyed with the Goldsmiths Musical Sophistication Index[4] to elicit their prior musical experience.

[TODO!!]

At the end of the study the participants were asked to fill in a questionnaire to elicit their prior musical experience.

4.1 Results

X participants took part in the study (mean median age). (Prior musical experience?)

4.2 Observation/Discussion

5. FURTHER WORK

6. ACKNOWLEDGMENTS

This work is supported by EPSRC Doctoral Training Centre EP/G03723X/1 (HE), GR/S82213/01 and EP/E045235/1(SA), an EPSRC Leadership Fellowship, EP/G007144/1 (MDP) and EPSRC IDyOM2 EP/H013059/1. Thanks to Louie Mc-

Callum and Davie Smith from QMUL for Kinect programming support.

7. REFERENCES

- [1] S. Abdallah and M. Plumbley. Information dynamics: patterns of expectation and surprise in the perception of music. *Connection Science*, 21(2):89–117, 2009.
- [2] F. Lerdahl and R. Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA, Jan. 1983.
- [3] L. Meyer. Music, the arts, and ideas: Patterns and Predictions in Twentieth-Century Culture. University of Chicago Press, Chicago, 1967.
- [4] D. Müllensiefen, B. Gingras, L. Stewart, and J. Musil. Goldsmiths Musical Sophistication Index (Gold-MSI): Technical Report and Documentation v0. 9. 2011.
- [5] E. Narmour. *Beyond Schenkerism*. the need for alternatives in music analysis. Univ of Chicago Press, 1977.