Henrik Ekeus, Samer A. Abdallah, Mark D. Plumbley, Peter W. McOwan Queen Mary University of London
Centre for Digital Music
School of Electronic Engineering and Computer Science
black{hekeus,samer.abdallah}@eecs.qmul.ac.uk

The Melody Triangle - Pattern and Predictability in Music

Abstract

The Melody Triangle is an interface for the discovery of melodic materials, where the input – positions within a triangle – directly map to information theoretic properties of the output. The measures are the entropy rate, redundancy and *predictive information rate* [?] of the random process used to generate the sequence of notes. These are all related to the *predictability* of the sequence and as such address the notions of expectation and surprise in the perception of music. We describe some of the relevant ideas from information dynamics, how the Melody Triangle is defined in terms of these, and describe two physical incarnations of the Melody Triangle. 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 cognitively-informed compositional aid for the generation of musical textures; the user's control at the abstract level of randomness and predictability. Finally we outline a pilot 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.

Information Dynamics

The relationship between Shannon's [?] information theory and music and art in general has been the subject of some interest since the 1950s [?, ?, ?, ?, ?]. The general thesis is that perceptible qualities and subjective states like uncertainty, surprise, complexity, tension, and interestingness are closely related to information-theoretic quantities like entropy, relative entropy, and mutual information.

Music is an inherently dynamic process. The idea that the musical experience is strongly shaped by the generation and playing out of strong and weak expectations was put forward by, amongst others, music theorists L. B. Meyer [?] and Narmour [?]. Central to this is the idea that music is not a static object presented as a whole, but as a phenomenon that 'unfolds' and is experienced *in time*; as listeners we continually build and re-evaluate expectations of what is to come next.

Information dynamics [?] considers several different kinds of predictability in musical patterns, how these might be quantified using the tools of information theory, and how they shape or affect the listening experience. Our working hypothesis is 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 as well as using previous musical experience, such as a familiarity with musical styles and conventions. As the music unfolds, listeners continually revise their model; in other words, they revise their own, subjective probabilistic belief state. These changes in probabilistic beliefs can be associated with quantities of information; these are the focus of information dynamics.

The Melody Triangle

The use of stochastic processes in music composition has been widespread for decades—for instance Iannis Xenakis applied probabilistic mathematical models to the creation of musical materials [?]. While such processes can drive the *generative* phase of the creative process, information dynamics can serve as a novel framework for a *selective* phase, by providing a set of criteria to be used in judging which of the generated materials are of value. This alternation of generative and selective phases as been noted before [?]. Information-dynamic criteria can also be used as *constraints* on the generative processes, for example, by specifying a certain temporal profile of suprisingness and uncertainty the composer wishes to induce in the listener as the piece unfolds.

The Melody Triangle enables the discovery of melodic content matching a set of information theoretic criteria. Positions within the triangle correspond with pairs of values of entropy rate and redundancy. The physical interface to the Triangle has so far been realised in two forms: as an interactive installation and as a screen based interface.

Given coordinates corresponding to a point in the triangle, we select from a pre-built library of random processes, choosing one whose entropy rate and redundancy match the desired values. The implementations discussed in this paper use first order Markov chains as the content generator, since it is easy to compute the theoretically exact values of entropy rate, redundancy and predictive information rate given the transition matrix of the Markov chain. However, in principle, any generative system could be used to create the library of sequences, given an appropriate probabilistic listener model supporting the estimation of entropy rate and redundancy.

The Markov chain based implementation generates streams of symbols in the abstract; the alphabet of symbols is then mapped to a set of distinct sounds, such as pitched notes in a scale or a set of percussive sounds. Further by layering these streams intricate musical textures can be created. The selection of notes or sounds is arbitrary, as long as they are all distinguishable. Indeed, the symbols could be mapped to even non sonic outputs such as visible shapes, colours, or movements.

Any sequence of symbols can be analysed and information theoretic measures estimated from it. The novelty of the Melody Triangle lies in that we reverse this mapping: given desired values for these measures, as determined from the user interface, we return a stream of symbols with the desired properties. In the next section we describe the three information theoretic measures that we use.

Sequential Information Measures

The *entropy rate* of a random process is a basic measure of its randomness or unpredictablity. Consider the viewpoint of an observer at a certain time, and split the sequence into an infinite *past*, as single symbol in the *present*, and the infinite *future*. The entropy rate is a conditional entropy; informally:

$$\mathrm{EntropyRate} = H(\mathrm{Present}|\mathrm{Past}), \tag{1}$$

that is, it represents our average uncertainty about the present symbol *given* that we have observed everything before it. Processes with zero entropy rate can be predicted perfectly given enough of the preceeding context.

The *redundancy* of the a process, in the sense we are using the term here, is a measure of how much the predictability of the process depends on knowing the preceeding context. It is the difference between the entropy of a single element of the sequence in isolation (imagine chosing a note from a musical score at random with your eyes closed and then trying to guess the note) and its entropy after taking into account the preceeding context:

$$\mathrm{Redundancy} = H(\mathrm{Present}) - H(\mathrm{Present}|\mathrm{Past}). \tag{2}$$

If the previous symbols reduce our uncertainty about present symbol a great deal, then the redundancy is high. For example, if we know that a sequence consists of a repeating cycle such as $\ldots b, c, d, a, b, c, d, a \ldots$, but we don't know which was the first symbol, then the redundancy is high, as $H(\mathrm{Present})$ is high (because we have no idea about the present symbol in isolation, but $H(\mathrm{Present}|\mathrm{Past})$ is zero, because knowing the previous symbol immediately tells us what the present symbol is.

The *predictive information rate* (PIR) brings in our uncertainty about the future. It is a measure of how much each symbol reduces our uncertainty about the future as it is observed, *given* that we have observed the past:

$$\mathrm{PIR} = H(\mathrm{Future}|\mathrm{Past}) - H(\mathrm{Future}|\mathrm{Present}, \mathrm{Past}). \tag{3}$$

It is a measure of the *new* information in each symbol. Notice that if the past completely determines both the present and the future (as in the cyclic pattern above) the PIR is zero,

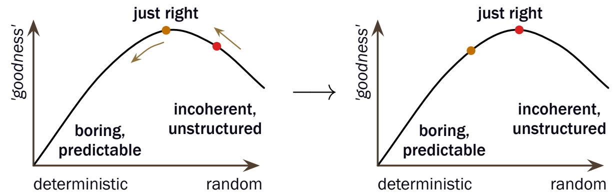

Figure 1: The Wundt curve relating randomness/complexity with perceived value. Repeated exposure sometimes results in a move to the left along the curve [**?**].

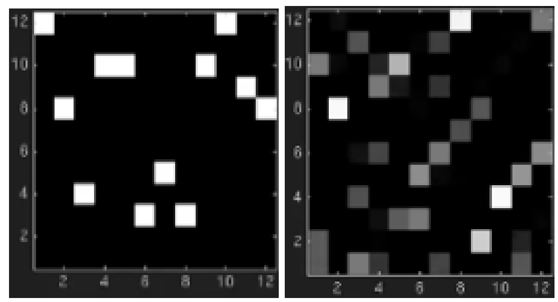

Figure 2: Two transition matrixes. The shade of white represents the probabilities of transition from one symbol to the next (black=0, white=1). The current symbol is along the bottom, and in this case there are twelve possibilities (mapped to a chromatic scale). The left hand matrix has no uncertainty; 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.

since the present symbol brings no new information. However, if the symbols in a sequence are generated completely independently, e.g. by rolling a die for each one, then again, the present symbol provides no information about the future and the PIR is zero.

Processes with high PIR maintain a certain kind of balance between predictability and unpredictability in such a way that the observer must continually pay attention to each new observation as it occurs in order to make the best possible predictions about the evolution of the sequence. This balance between predictability and unpredictability is reminiscent of the inverted 'U' shape of the Wundt curve (see fig. **??**), which summarises the observations of Wundt [**?**] that stimuli are most pleasing at intermediate levels of novelty or disorder, where there is a balance between 'order' and 'chaos'.

Populating the triangle

Before the Melody Triangle can used, it has to be 'populated' with possible parameter values for the melody generators. These are then plotted in a 3d statistical space of redundancy, entropy rate and predictive information rate. In our case we generated thousands of transition matrixes, representing first-order Markov chains, by a random sampling method. In figure we see a representation of how these matrixes are distributed in the 3d statistical space; each one of

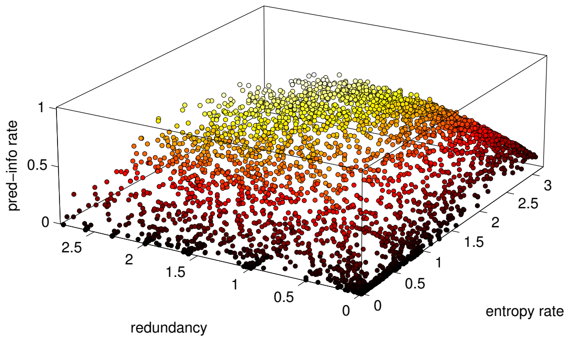

Figure 3: The population of transition matrices in the 3D space of entropy rate, redundancy and PIR, all in bits. The concentrations of points along the redundancy axis correspond to Markov chains which are roughly periodic with periods of 2 (redundancy 1 bit), 3, 4, etc. all the way to period 7 (redundancy 2.8 bits). The colour of each point represents its PIR—note that the highest values are found at intermediate entropy and redundancy, and that the distribution as a whole makes a curved triangle. Although not visible in this plot, it is largely hollow in the middle.

these points corresponds to a transition matrix.

When we look at the distribution of transition matrixes 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.

Though the interface is 2D, the third dimension (PIR) is implicitly present, as transition matrices retrieved from along the centre line of the triangle will tend to have higher PIR. We hypothesise that, under the appropriate conditions, these will be perceived as more 'interesting' or 'melodic.'

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. When the user, through the interface, selects a position within the triangle, the corresponding transition matrix is returned. Figure 2 shows how the triangle maps to different measures of redundancy, entropy rate and predictive information rate.

Each corner corresponds to three different extremes of predictability and unpredictability, which could be loosely characterised as 'periodicity', 'noise' and 'repetition'. In our experiments with visualising and sonifying sequences sampled from first order Markov chains [**?**], we found that the measures of redundancy rate, entropy rate and predictive information rate correspond to perceptible characteristics, and that the transition matrices maximising or minimising each of these quantities are quite distinct. High entropy rates are associated with completely uncorrelated sequences with no recognisable temporal structure. High values of redundancy rate are associated with long periodic cycles (and low PIR and entropy rate). High values of predictive information rate are associated with intermediate values of redundancy rate and entropy rate, and recognisable, but not completely predictable, temporal structures.

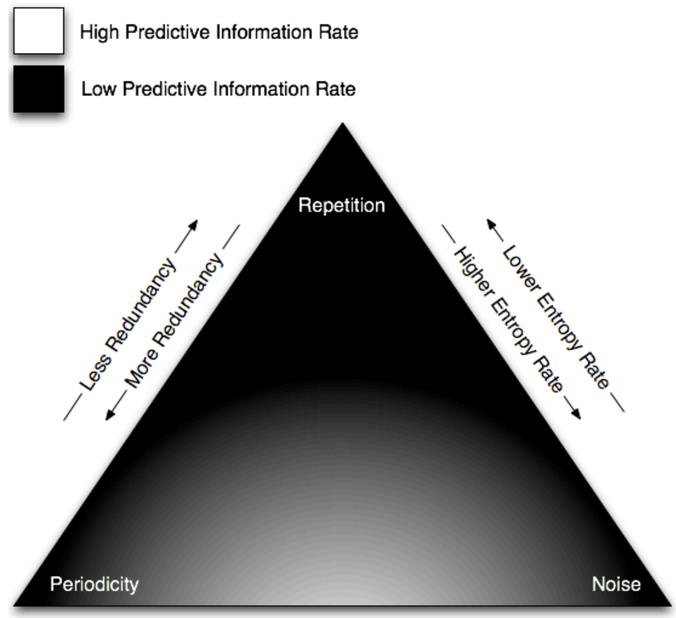

Figure 4: The Melody Triangle

User Interfaces

Any number of interfaces could be developed for the Melody Triangle[1]. We have developed two; 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 Kinect[2] camera tracks individuals in a space and maps their positions in the space to the triangle.

The Multi-User Installation

As a Kinect camera overlooks a space, its 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 changes the predictability of the generated melodic content. When multiple people are in the space they can cooperate to create interweaving melodies, forming intricate 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 changes and volume (see Table 1, Figure 4).

Tracking and Control Tracking and control was done using the OpenNI libraries' API[3] and high level middle-ware

[1]The Melody Triangle was developed in Prolog and MatLab. It can be controlled with OpenSoundControl messages, and thus is independent of any specific interface implementation.

[2]http://www.xbox.com/en-GB/Kinect

[3]http://OpenNi.org/

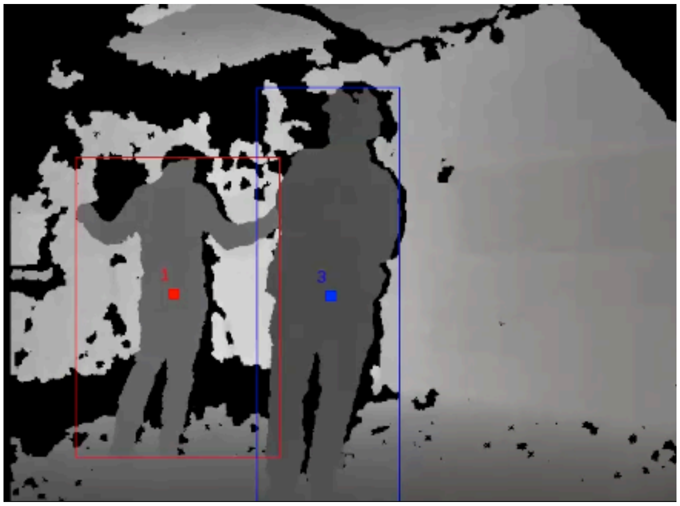

Figure 5: The depth map as seen by the Kinect, and the bounding box outlines the blobs detected by OpenNI.

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

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.

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 the instrument being 'played'.

Observations Although visitors would need an initial bit of training they would then quickly be able 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 an area that merits attention. By not having one user be able to control the whole narrative, the participants would communi-

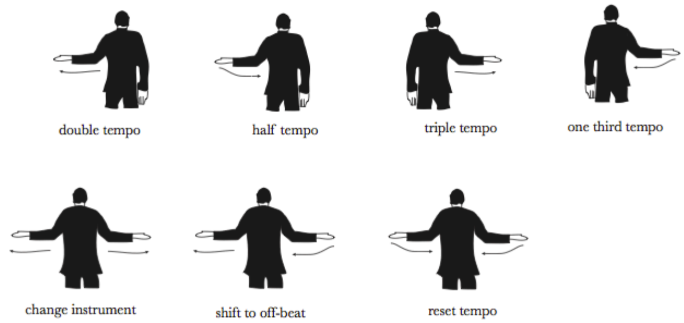

Figure 6: Gestures and their resulting effect

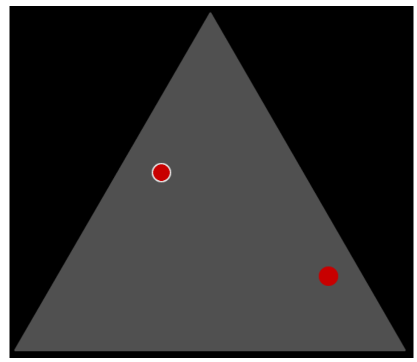

Figure 7: Screen shot of the screen based interface for the Melody Triangle

cate verbally and direct each other in the goals of learning to use the system and finding interesting musical textures. This collaboration added an element of playfulness and enjoyment that was clearly 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, art installation and technical demonstration. It is clear however, that as a vehicle for communicating ideas related to the expectation, pattern and predictability in music to the public, it is very effective.

The Screen Based Interface

The screen based interface can serve as a compositional tool. A number of tokens, each representing a sonification stream or 'voice', can be dragged in and around the triangle. For each token, a sequence of symbols is sampled using the corresponding transition matrix, which are then mapped to notes of a scale or percussive sounds[4]. Keyboard commands give control over other musical parameters such as pitch register, inter-onset interval, tempo and dynamics. The system is capable of generating intricate musical textures when multiple tokens are in the triangle.

[4]The sampled sequence could easily be mapped to other musical processes, possibly over different time scales, such as chords, dynamics and timbres. It would also be possible to map the symbols to visual or other outputs.

In this mode the Melody Triangle is a cognitively-informed compositional aid; unlike other computer aided composition tools or programming environments, here the composer exercises control at the abstract level of information-dynamic properties. The use of Markov Chains for the generation of musical content is not anything new, rather the novelty lies in the ability to define criteria in the selection of generated materials that relate to how a listener might perceive the output.

Information Dynamics and Musical Preference Study

We are currently in the process of using the screen-based Melody Triangle user interface to investigate the relationship between the information-dynamic characteristics of sonified Markov chains and subjective musical preference. We carried out a pilot study with six participants, who were asked to use a simplified form of the user interface (a single controllable token, and no rhythmic, registral or timbral controls) under two conditions: one where a single sequence was sonified under user control, and another where an additional sequence was sonified in a different register, as if generated by a fixed invisible token in one of four regions of the triangle. In addition, subjects were asked to press a key if they 'liked' what they were hearing.

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

We recorded subjects' behaviour as well as points which they marked with a key press. Some results for three of the subjects are shown in fig. **??**. Though we have not been able to detect any systematic across-subjects preference for any particular region of the triangle, subjects do seem to exhibit distinct kinds of exploratory behaviour. Our initial hypothesis, that subjects would linger longer in regions of the triangle that produced aesthetically preferable sequences, and that this would tend to be towards the centre line of the triangle for all subjects, was not confirmed. However, it is possible that the design of the experiment encouraged an initial exploration of the space (sometimes very systematic, as for subject c) aimed at *understanding* how the system works, rather than finding musical patterns. It is also possible that the system encourages users to create musically interesting output by *moving the token*, rather than finding a particular spot in the triangle which produces a musically interesting sequence by itself.

Comments collected from the subjects suggest that the information-dynamic characteristics of the patterns were readily apparent to most: several noticed the main organisation of the triangle, with repetitive notes at the top, cyclic patterns along one edge, and unpredictable notes towards the opposite corner. Some described their systematic exploration of the space. Two felt that the right side was 'more controllable' than the left (a consequence of their ability to return to a particular distinctive pattern and recognise it as one heard previously). Two reported that they became bored towards the end, but another felt there wasn't enough time to 'hear out' the patterns properly. One subject did not 'enjoy'

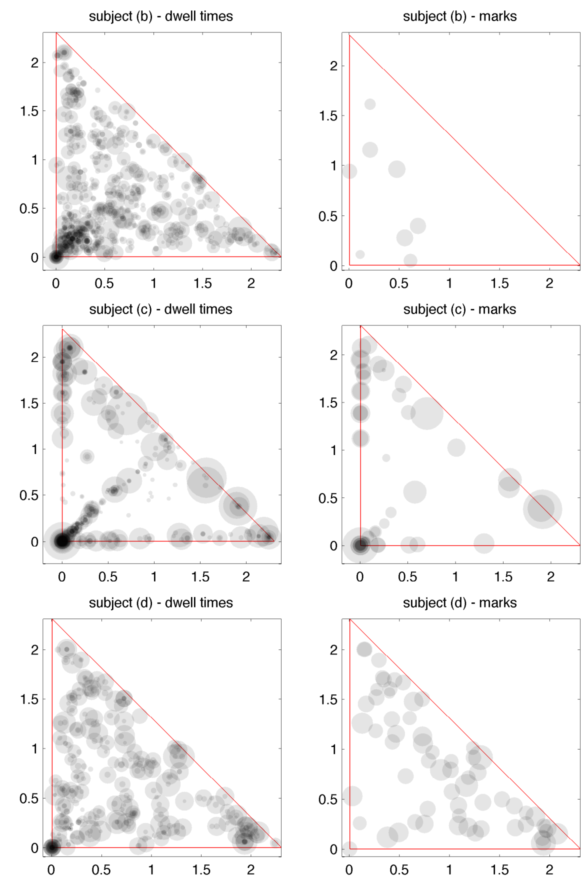

Figure 8: Dwell times and mark positions from user trials with the on-screen Melody Triangle interface, for three subjects. The left-hand column shows the positions in a 2D information space (entropy rate vs multi-information rate in bits) where each spent their time; the area of each circle is proportional to the time spent there. The right-hand column shows point which subjects 'liked'; the area of the circles here is proportional to the duration spent at that point before the point was marked.

the patterns in the lower region, but another said the lower central regions were more ‘melodic’ and ‘interesting’.

We plan to continue the trials with a slightly less restricted user interface in order make the experience more enjoyable and thereby give subjects longer to use the interface; this may allow them to get beyond the initial exploratory phase and give a clearer picture of their aesthetic preferences. In addition, we plan to conduct a study under more restrictive conditions, where subjects will have no control over the patterns other than to signal (a) which of two alternatives they prefer in a forced choice paradigm, and (b) when they are bored of listening to a given sequence.

Further Work

We are currently investigating how higher-order Markov models can be mapped to information theoretic measures and adapting the Melody Triangle to those models. This would generate higher level patterns and provide more long-term structures. Further more sophisticated listener models [**?**] [**?**] could be used for computing information measures for more conventional or ecologically valid music.

As it stands, the streams of symbols generated are only mapped to note values. However they could just as well be applied to any other musical property, such as intervals, chords, dynamics, timbres, structures and key changes. The possibilities for the Melody Triangle to be compositional guide in these other domains remains to be investigated.

We are investigating the possibility of turning the Melody Triangle into a mobile phone based music making application. It is hoped that by collecting usage statistics we may have a rich source of data that can help determine any relationship between the information dynamics measures and aesthetic preference. Although our initial data on aesthetic preference are inconclusive, there is still plenty of work to be done in this area: where-ever there are probabilistic models, information dynamics can shed light on their behaviour.

acknowledgments

This work is supported by an 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 McCallum and Davie Smith from QMUL EECS for Kinect programming support.