Towards a cognitive model of melodic similarity

Ludger Hofmann-Engl
Keele University UK
+44 (0) 20 87710639
ludger.hofmann-engl@virgin.net

ABSTRACT

In recent years the interest in melodic similarity has mushroomed mainly due to the increased importance of music information retrieval (MIR). A great number of similarity models and algorithms have been developed, but little or no attention has been paid to cognitive or perceptual aspects to the issue at hand. Questions, about the relevant parameters and the appropriate implementation are under-researched as are experimental data. This paper focuses on the pitch aspect of melodic similarity, scrutinising the term pitch replacing it by a less ambivalent and overused term, which we will call meloton. Based on the term meloton the paper suggests to approach the issue of 'melotonic' similarity from a transformational angle, where transformations are executed as reflections and translations. 'Melotonic' similarity then is seen as related to the transformation process in form of a transpositional and interval vector. Finally, melotonic similarity as portrait in a psychological context emerges as a multi-facet phenomenon requiring the development of flexible models.

1. INTRODUCTION

Unarguably, melodic similarity has been of great interest to composers (e.g., Schoenberg, 1967), ethnomusicologists (e.g., Adams, 1976) and music analysts (e.g., Reti, 1951). However, the issue has received new interest due to the development of the internet and the need to administrate and retrieve musical information. Early works on MIR date back to the 60's with Kassler (1966) as one of the pioneers. Not much research was done on the topic for some time, but by now the growing interest is reflected for instance in the fact that in 2000 the first international symposium on MIR was organised and attended by researchers from a great variety of fields. The interest of MIR in the issue of similarity does not necessarily add importance to the issue but certainly urgency to develop reliable and relevant similarity models. In fact, by now several models and algorithms have been proposed (Anagnostopoulou, Hörnel & Höthker, 1999; Cambouropoulos, 2000; Crawford, Iliopoulos & Raman, 1998; Dovey &Crawford, 1999; Downie, 1999; Kluge, 1996; Maidin & Fernström, 2000; Smith & McNab & Witten, 1998), but none of these models take cognitive or perceptual issues sufficiently into account, nor do they pay a closer look at as to what parameters to select and how to implement them.

Admittedly, the psychological and more specific the music-psychological research in this field leaves possibly more questions unanswered than answered (compare Goldstone, 1998), but this seems not to justify the dismissal of existing research. Notably, none of the researchers takes dynamic aspects into account and rhythmic aspects play no or little role by regarding melodic similarity exclusively as a pitch phenomenon, without considering the limitations of their models. Interestingly, the question of what the term pitch, a term which from a psychological angle is more than problematic, is not being asked, although there exists some awareness to the related issue of musical representation; that is whether we are dealing with the representation of music in form of a score, with recorded music or digital sources such as MIDI files (e.g. Wiggins & Harris & Smaill, 1989), but the central issue of pitch perception is hardly ever touched.

Most strikingly, the works of Egmond & Povel & Maris (1996) have not been referred to in a single instance to the knowledge of the author, although their findings are in agreement with previous research (Francès, 1988) and also in agreement with more recent research by Hofmann-Engl & Parncutt (1998). Their experiments indicate that transposition is a significant factor for melodic similarity judgements whereby melodic similarity decreases with increasing size of the transpositional interval. However, all the models known to the author are transpositional invariant. This demonstrates the strong tendency of researchers to borrow their tools from music theoretical teachings where transpositions of a motive are regarded as being equivalent. This is not to say that search tools within MIR should under no circumstances consider transpositions as equivalent (for instance when analysing the structure of a specific composition), but where perceptual issues are of importance such attempt will have to be seen as a shortcoming.

While some models are seemingly unaware of the evidence as produced by researchers (e.g. Dowling & Harwood, 1986; Dowling 1994; Edworthy, 1982, 1985; White, 1960) that contour is somewhat a factor determining cognitive similarity (for instance models based on dynamic programming) there are models which take contour into account (for instance Maidin's model), but still reference to psychological research is not given.

Experiments by Hofmann-Engl and Parncutt (1998) indicate that contour is in fact an imbedded factor of what they called interval difference. This is, a reference-melody raised by an interval *I* between two consecutive tones (let us say tone 1 and tone 2) produces the interval difference $D = I - I'$, with I' being the interval between the two corresponding tones (tone 1 and tone 2) of the comparison-melody. Melodies which show contour differences also produce interval differences and hence contour appears to be a factor. However, multiple regression shows that interval difference is the sole factor. Up to this date no model has accounted for these findings.

Finally, none of the models takes emotional aspects into account. True, that at this point it seems an almost unattainable task, but research by Tekman (1998) shows, that emotional aspects can be at least partly sensibly measured. Clearly, there exists a level of unawareness amongst melodic similarity researchers of psychological issues which seems hardly acceptable.

It is the intention of the present paper to contribute to the bridging of exactly this gap. Although as mentioned before dynamic, rhythmic and emotional aspects will have to be seen as factors alongside pitch, we will focus on pitch exclusively. This is, pitch is the most discussed aspect of melodic similarity, and treating all parameters would exceed the framework of this paper. However, the author is in process of developing a similarity model which takes dynamics and rhythmic features into account alongside with pitch. In the first instance we will scrutinise the term pitch arguing for its substitution by the new term, which will be called meloton, we then will consider the transformation of melodies and finally we will develop a similarity model based on the composition of two specific transformations.

2. PITCH VERSUS MELOTON

The term pitch is intriguing and perplexing at the same time, intriguing, because it is probably the most discussed musical and music psychological term, and perplexing, because it has been employed in so many different contexts that it frequently requires specification to what actually is meant by it. A situation which led Rohwer (1970) to question the usage of the term pitch altogether. However, pitch is commonly understood to be the correlate to the fundamental frequency as Rasch & Plomp (1982) explain: "Pitch is the most characteristic property of tones ... . Pitch is related to the frequency of a simple tone and to the fundamental frequency of a complex tone." Pitch is seen here exclusively as being related to a physical quantity. This is as widespread an approach as is insufficient because subtleties of pitch perception are not captured by referring to physical dimensions only. It seems this is the issue Sundberg (1991) is addressing when he suggests to see pitch as locating musical sound in terms of musical intervals and to call other aspects of pitch perception tone height (the high or lowness of sound). However, the categorical differentiation between pitch and tone height seems frail as even sounds with low pitch salience (e.g., musical chimes) can produce a distinctive pitch sensation (Askill 1997). Confronted with the phenomenon that sounds without fundamental frequency can produce pitch sensations, Schouten (1938) introduced the term residual pitch. However, from a phenomenological angle there is no difference between pitch and residual pitch — both appear to a listener in the same way. Concepts of virtual pitch (Houtsma & Goldstein 1972, Terhardt 1982, Hofmann-Engl 1999) added further complexity to the issue by demonstrating that sounds do not have one single pitch but a multiplicity of pitches with varying degrees of probabilities. Employing this concept the term pitch has to be replaced by the term most probable pitch. It seems the introduction of a new term which will have some specific psychological meaning is more than appropriate. We argue to use the term meloton as suggested by Hofmann-Engl (1989).

Without going into detail, we will consider some features of what this term is to deliver. We wish for this term to be purely of psychological meaning. This is we endeavour to understand melodic similarity from a cognitive angle. Thus, concepts of fundamental frequency and other physicalistic approaches are inadequate. Hereby, the term meloton will represent the psychological concept whereby a listener listens to a sound directing her/his attention to the sound with the intention to decide whether the sound is high or low. True, this does not deliver a quantity we could input into a similarity model, and hence we will have to define the value of a meloton somehow without using a physicalistic concept. In this context it seems most appropriate to consider an experimental setting as employed by Schouten (1938). A selected group of listeners is asked to tune in a (sinusoidal) comparison tone with variable frequency to match according to the listener's perception a test tone (for which we want to obtain a melotonic value). The logarithm of the comparison tone then will be called m-response of this listener. Assuming that the group of listeners consist of n listeners, we will obtain *n* m-responses. We will call the mean of this distribution the m-center. The mode of the distribution will be called m-peak. The relative density of the m-peak will be called melograde, and can be defined as:

$$M_g = \frac{D(p_m)}{\sum_{i=1}^{n} D(l_m)_i}$$

where M_g is the melograde, $D(p_m)$ the density of the peak of the m-distribution, $D(l_m)_i$ the density of the location lm at the place *i* and *n* the number of locations. The range of M_g is]0, 1]. Note, that models of the pitch salience (Terhardt, Stoll & Seewann 1982) are predictors for the peak of the response distribution.

We finally define the value *M* of a meloton as given by the value of the m-centre. However, if the melograde of a m-distribution is larger than 0.75 and the peak and the centre coincide with maximum deviation of 25 cents, the value of the meloton is given by the peak of the m-distribution. In this case we speak of *strong meloton* M_s. In all other cases we speak of *weak meloton* M_w.

There are several advantages to this approach as there are limitations. Firstly, the classification of melota into weak and strong melota guaranties that tones such as produced by a drum instrument will also fetch a melotonic value. This allows for the inclusion of 'drum-melodies' into a melodic similarity model. Secondly, we replaced the dogmatic attitude towards pitch perception by an understanding which is sensitive towards individual, cultural, educational and social differences; what might appear to one group as a sound with a certain melotonic value might appear to another group as a sound with a different melotonic value. This is certainly of importance when considering melodic similarity. Right at the basis melodic similarity judgement might differ due to lower level perceptual differences. Objections might be raised that this approach is impracticable in many ways, as the measurement of melotonic values is time consuming and expensive. Still, we might expect that data bases containing melotonic measurements might be established and made available in future. Another objection might be that even if measurements of single tones are available, there is no guaranty that meloton will retain their values when put into a melodic context. Although this might be true, this will have to be an issue to be investigated. However, considering the success of aural training (where listeners are required to identify tones in any context), we expect that this approach is more than promising.

Before we will base a transformation theory on melota, we will abandon the term melody due its many ambivalent connotations and replace it by the term chain. We will as mentioned only consider the melotonic component of a chain (excluding properties such as timbre, duration and loudness). A melotonic chain will be written as m-chain and as *M*(*ch*). Now, we will consider transformations of melotonic chains.

2. MELOTONIC TRANSFORMATIONS

The motivation for developing a transformation theory is driven by the proposal as put forward by Palmer (1983), where similarity is understood to be related to the transformation process involved in mapping two objects onto each other. There have been several attempts within the camp of pitch class theorists to introduce similarity measures and transformations of pitch class sets (e.g., Isaacson (1996), Lewin (1977), Morris (1979)). However, none of these approaches are of interest in this context, as a pitch class set is fundamentally a different entity than is a melotonic chain. One attempt to describe melodic transformations has been put forward by Mazzola (1987) and has been further developed by Hammel (1999) in form of matrices. Without going into detail, the main deficiency of their transformation matrices is the combination of time and pitch in one matrix leading to time and pitch appearing as mixed terms. Maybe even more important, the matrices as they stand, do not allow for general transformations. Thus, a theory of similarity based on their concept would not allow for the comparison of any melody with any melody.

Instead we will take inversions and transpositions as a starting point and generalise these two transformations. As we are dealing with melotonic transformations, we will require that two chains will have the same rhythm and the same dynamic structure. This is a restriction to the model, but this deficiency can only be overcome at a later stage including rhythm and dynamics and possibly emotional aspects in a final model. However, as mentioned, this would exceed the framework of this paper.

2.1. Inversion

It is a well known geometrical fact, that inversion can be illustrated as a reflection along a straight line. Taking a melotonic chain (melody) *M*(*ch*) to consist of an initial tone with meloton m_1 and then the subsequent intervals 100, 100, -100, -100 (where 100 might be taken to mean 100 cents), we write: $M(ch) = (m_1)$[100, 100, -100, -100] to be reflected onto the chain $M(ch') = (m_1')$[-100, -100, 100, 100] will require a straight line through the point $p = 50$ (see Figure 1)

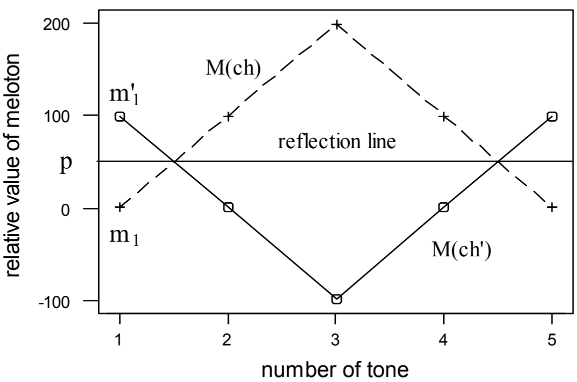


Figure 1: The m-chain M(ch) = (m_1)[100, 100, -100, -100] is mapped onto the chain $M(ch') = (m_1')$[-100,-100, 100, 100] via the reflection line through p, with $m_1 = 0$ and $p = 50$

2.2. Transposition

Executing two reflections along two different reflection lines results in transposition. Taking the example from above, where $M(ch) = (m_1)$[100, 100, -100, -100], we obtain the inversion $M(ch') = (m_1')$[-100, -100, 100, 100], when reflecting *M*(*ch*) through $p_1 = 50$ and the transposition $M(ch'') = (m_1'')$[100, 100, -100, -100] when reflecting $M(ch')$ through $p_2 = 150$. The difference between p_2 and p_1 is $p_2 - p_1 = 100$, and the transposition interval between *M*(*ch*) and $M(ch'')$ is $2(p_2 - p_1) = 200$ (Figure 2). The transposition interval is generally $2(p_2 - p_1)$ regardless where p_2 and p_1 are located. Allowing for the reflecting of single melota rather than the reflection of an entire m-chain, we can illustrate transposition as a reflection along a reflection chain, which we will call $M(\chi)$, we obtain figure 3.

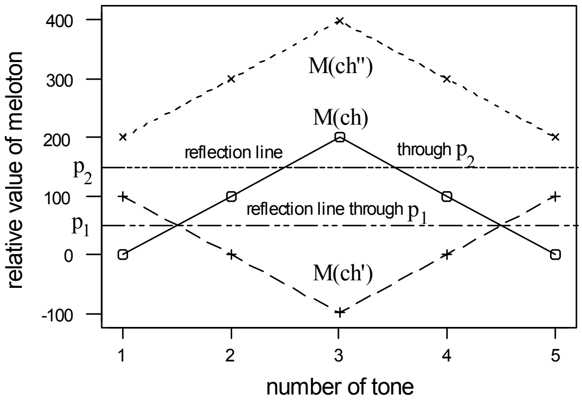


Figure 2: The m-chain $M(ch) = (m_1)[100, 100, -100, -100]$ is mapped onto the chain $M(ch'') = (m_1')[-100,-100, 100, 100]$ via the reflection line through $p_1 = 50$ and onto $M(ch'') = (m_1'')[100, 100, -100, -100]$ via the reflection line through $p_2 = 150$. The graph illustrates, that reflection along two lines results in transposition.

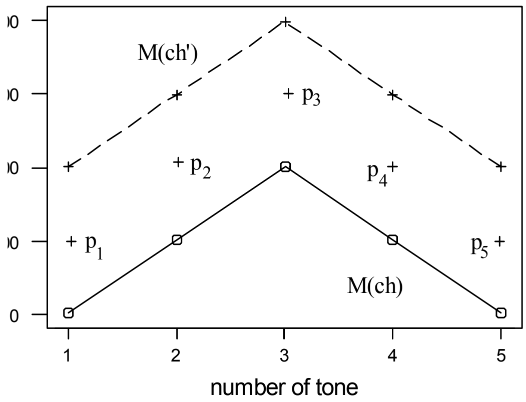


Figure 3: The m-chain $M(ch) = (m_1)[100, 100, -100, -100]$ is mapped onto the chain $M(ch') = (m_1')[-100,-100, 100, 100]$ via the sequence of reflection points p_1, p_2, ... p_5, thus effecting the transposition of $M(ch)$.

2.3. General melotonic transformations

Given a m-chain $M(ch)$ of the length n and a reflection chain $M(\chi)$ of the length n, we find that $M(ch)$ will be mapped onto $M(ch')$, where for all $m_i \in M(ch)$ and $_{all\ pi} \in M(\chi)$, that $m'_i = 2p_i - m_i$, for all $m'_i \in M(ch')$. This is important, when defining reflections and translations (we will use the mathematical term instead for transposition) within the vector-space R^{n+1}. As we will see later, the composition of two specific reflections will enable us to produce similarity measures. However, we define reflections and translations on a more formal level first. For this purpose we define the m-vector $\vec{m}$:

<u>Definition:</u>

Given a m-chain of length n, with $M(ch) = [m_1, m_2, ..., m_n]$, we define the m-vector $\vec{m}$ of length n+1 as:

$$\vec{m} = \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ m_n \\ 1 \end{pmatrix}$$

This enables us to define the reflection matrix R as:

<u>Definition</u>

$$R = \begin{pmatrix} -1 & 0 & . & . & 0 & 2p_1 \\ 0 & -1 & & & 0 & 2p_2 \\ . & & . & & & . \\ . & & & . & & . \\ 0 & 0 & & & -1 & 2p_n \\ 0 & 0 & & & 0 & 1 \end{pmatrix}$$

Multiplying the reflection matrix R by a m-vector $\vec{m}$, we obtain:

$$R \cdot \vec{m} = \begin{pmatrix} -1 & 0 & . & . & 0 & 2p_1 \\ 0 & -1 & & & 0 & 2p_2 \\ . & & . & & & . \\ . & & & . & & . \\ 0 & 0 & & & -1 & 2p_n \\ 0 & 0 & & & 0 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ m_n \\ 1 \end{pmatrix} = \begin{pmatrix} m_1 - 2p_1 \\ m_2 - 2p_2 \\ . \\ . \\ m_n - 2p_n \\ 1 \end{pmatrix}$$

Clearly, multiplying the reflection matrix by a m-vector, results in the reflection of this m-vector. As two reflections result in translation, we will define the translation matrix, where each component $m_i \in \vec{m}$ will be translated by the translation interval

$I_i = 2(p_{2i} - p_{1i})$. We define:

Definition:

$$T = \begin{pmatrix} 1 & 0 & & & 0 & I_1 \\ 0 & 1 & & & 0 & I_2 \\ . & & . & & & . \\ . & & & . & & . \\ 0 & 0 & & & 1 & I_n \\ 0 & 0 & . & . & 0 & 1 \end{pmatrix}$$

Multiplying the translation matrix T with a m-vector $\vec{m}$, we get:

$$T \cdot \vec{m} = \begin{pmatrix} 1 & 0 & . & . & 0 & I_1 \\ 0 & 1 & & & 0 & I_2 \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & & & 1 & I_n \\ 0 & 0 & . & . & 0 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ m_n \\ 1 \end{pmatrix} = \begin{pmatrix} m_1 + I_1 \\ m_2 + I_2 \\ . \\ . \\ m_n + I_n \\ 1 \end{pmatrix}$$

There exists a complex algebraic structure between reflections and translations. However, the framework of the paper exceeds a discussion of this issue. Still, it might be worth mentioning that the composition of two reflection matrices results in a translation matrix. This is the reason, why we had to introduce translations, although translations are of no significance in context of melotonic similarity. We are now equipped to consider melotonic similarity.

3. MELOTONIC SIMILARITY

It seems the best way of approaching melotonic similarity is, when we consider two m-vectors $\vec{m}_1$ and $\vec{m}_2$, such as:

$$\vec{m}_1 = \begin{pmatrix} m_{11} \\ m_{12} \\ . \\ . \\ m_{1n} \\ 1 \end{pmatrix} \text{ and } \vec{m}_2 = \begin{pmatrix} m_{21} \\ m_{22} \\ . \\ . \\ m_{2n} \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} + a \\ m_{12} + a \\ . \\ . \\ m_{1n} + a \\ 1 \end{pmatrix}$$

with *a* as a constant

We will now reflect the m-vector $\vec{m}_1$ through the 0-point R^{n+1} via the reflection matrix R_0. We obtain the m-vector $\vec{m}'_1$, with:

$$\vec{m}'_1 = \begin{pmatrix} -1 & 0 & . & . & 0 & 0 \\ 0 & -1 & & & 0 & 0 \\ . & & . & & & . \\ . & & & . & & . \\ 0 & 0 & & & -1 & 0 \\ 0 & 0 & . & . & 0 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ m_n \\ 1 \end{pmatrix} = \begin{pmatrix} -m_1 \\ -m_2 \\ . \\ . \\ -m_n \\ 1 \end{pmatrix}$$

Reflecting the m-vector $\vec{m}'_1$ onto the m-vector $\vec{m}_2$ requires the reflection matrix R^s given as:

$$R^S = \begin{pmatrix} -1 & 0 & . & . & 0 & 2\left(\frac{-m_{11} + m_{21}}{2}\right) \\ 0 & -1 & & & 0 & 2\left(\frac{-m_{12} + m_{22}}{2}\right) \\ . & & . & & & . \\ . & & & . & & . \\ 0 & 0 & & & -1 & 2\left(\frac{-m_{1n} + m_{2n}}{2}\right) \\ 0 & 0 & . & . & 0 & 1 \end{pmatrix}$$

As we find:

$$R^S \cdot \vec{m}'_1 = \vec{m}_2$$

Isolating the last column in the subspace R^n of R^{n+1}, we can define the similarity vector $\vec{V}_s$:

Definition:

$$\vec{V}_s = 2\begin{pmatrix} \frac{-m_{11} + m_{21}}{2} \\ \frac{-m_{12} + m_{22}}{2} \\ . \\ . \\ \frac{-m_{1n} + m_{2n}}{2} \end{pmatrix} = \begin{pmatrix} -m_{11} + m_{22} \\ -m_{12} + m_{22} \\ . \\ . \\ -m_{1n} + m_{2n} \end{pmatrix}$$

With $m_{2i} = m_{1i} + a$, we obtain the similarity vector:

$$\vec{V}_s = \begin{pmatrix} a \\ a \\ . \\ . \\ a \end{pmatrix}$$

Geometrically, this means that the similarity vector comes to coincide with the diagonal of the space R^n. We further find for the length of this vector:

$$\left\| \vec{V}_s \right\| = a\sqrt{n}$$

Clearly, the larger the transposition interval a is, the larger will be the length of the similarity vector. According the van Egmond, Povel & Maris (1996) melodic similarity decreases with increasing transposition interval. Thus, we will expect that the length of the similarity vector will be correlated to the transpositional component of melotonic similarity.

The intervalic component of melotonic similarity, according to Hofmann-Engl & Parncutt (1998) is a significant predictor. This is, when two m-vectors $\vec{m}_1$ and $\vec{m}_2$ are not simply transpositions of each other but deviate in shape. The similarity vector $\vec{V}_s$ will then deviate from the diagonal in R^n. Without going into a lengthily discussion, the angle between the similarity vector and the diagonal of R^n is not a suitable measure of the intervalic similarity component, as small intervalic changes can lead to a sudden increase of the angle. However, the differences between the components of the similarity vector are. Given the similarity vector $\vec{V}_s$ as:

$$\vec{V}_s = \begin{pmatrix} s_1 \\ s_2 \\ . \\ . \\ s_n \end{pmatrix}$$

We define the interval vector $\vec{V}_i$ with $I_i = s_{i+1} - s_i$ as a vector of the dimension R^{n-1}

<u>Definition:</u>

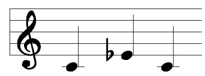
$$\vec{V}_i = \begin{pmatrix} I_1 \\ I_2 \\ . \\ . \\ I_{n-1} \end{pmatrix}$$

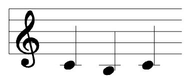
We will give an example referring to the four m-chains $M(ch_1) = (m_1)[1, -1]$, $M(ch_2) = (m_1)[3, -3]$, $M(ch_3) = (m_1)[-1, 1]$ and $M(ch_4) = (m_1)[1, 1]$ (where 1 unit might be one semitone). Setting m_1 to be $m_1 = 0$, we obtain the four m-vectors:

$$\vec{m}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{m}_2 = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{m}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \vec{m}_4 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}$$

In musical notation we obtain (setting the first tone to be c):

$M(ch_1) =$ 

$M(ch_2) =$ 

$M(ch_3) =$ 

$M(ch_4) =$ 

We find:

$$\vec{V}_s(\vec{m}_1,\vec{m}_2) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \text{ and } \vec{V}_i(\vec{m}_1,\vec{m}_2) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$\vec{V}_s(\vec{m}_1,\vec{m}_3) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \text{ and } \vec{V}_i(\vec{m}_1,\vec{m}_3) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$\vec{V}_s(\vec{m}_1,\vec{m}_4) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \text{ and } \vec{V}_i(\vec{m}_1,\vec{m}_4) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

The lengths of the similarity vectors $\vec{V}_s(\vec{m}_1,\vec{m}_2)$, $\vec{V}_s(\vec{m}_1,\vec{m}_3)$ and $\vec{V}_s(\vec{m}_1,\vec{m}_4)$ are identical (=2). Moreover, the lengths of the interval vectors $\vec{V}_i(\vec{m}_1,\vec{m}_2)$ and $\vec{V}_i(\vec{m}_1,\vec{m}_3)$ are identical (= $\sqrt{8}$), although $M(ch_1)$ and $M(ch_2)$ have the same contour, while $M(ch_1)$ and $M(ch_3)$ have different contour. However, both chains show the same interval difference. As mentioned above contour differences are imbedded within interval differences. Thus, similarity and interval vector are in agreement with experimental findings. Further, the comparison of $M(ch_1)$ and $M(ch_4)$ demonstrates, that the length of the similarity vector (=2) does not necessarily produce a length of $\sqrt{8}$. Thus, similarity and interval vector are independent similarity predictors. Because of the smaller length of the interval vector, we expect $M(ch_1)$ to be more similar when compared with $M(ch_4)$ than when compared with $M(ch_2)$.

We will expect that melotonic similarity will be correlated to the length or a derivative of the similarity vector (the longer the vector the smaller the similarity) and the deviation of the similarity vector from the diagonal measured by the length or a derivative of the interval vector (the longer the interval vector the smaller the similarity).

While the similarity vector takes the differences of two corresponding melota m_i and m'_i into account without considering any higher order (it does not matter where a pair of melota is placed within a chain), the interval vector considers higher order relationships in as much as pair-wise groupings are covered (the difference between $m_i - m_{i+1}$ and $m'_i - m'_{i+1}$). We could take even higher order relationships into account by forming the differences of the components of the interval vector in the fashion we formed the differences of the similarity vector obtaining the interval vector. We then could from the differences of these differences and so on. We then would obtain a series of vectors with decreasing dimensions starting with the similarity vector of dimension n, followed by the interval vector with the dimension n-1, followed by the differences of the interval vector producing a vector of dimension n-2 and so on till we obtain a vector of dimension 1. Thus, higher order relationships would b e covered and a predictor of melotonic similarity could be modelled around something comparable to a Taylor series. However, we expect that the similarity vector and interval vector will be sufficient to produce useful approximations.

Basing a similarity model exclusively on the lengths of the similarity and interval vector will still produce several complications. Without going into much detail, we will consider some experimental findings. Hofmann-Engl & Parncutt (1998) showed that keeping the transposition interval constant and varying the length of two melodic fragments (one to five tones), that similarity judgements increase with increasing length of the fragments. This seems to call for enveloping the components of the similarity vector by a exponential function giving more weight to earlier tones than later tones. Further, comparing two given m-chains of the length n, which are identical except one interval, Hofmann-Engl & Parncutt (1998) found that by varying the length n, that similarity judgements increase with increasing length n. Thus, a model will also have to be length sensitive. According to these researchers, tempo is not a factor in melotonic similarity, but appears as a rhythmic factor. We also might expect that aspects concerning the shape of two m-chains as covered by the interval vector will affected by the primacy/recency effect, where earlier and later tones are weighted more than are tones in the middle. This might call for enveloping the components of the interval vector by a Gauss distribution. Finally, a suitable model will require some empirical constants which will have to be determined through experimentation. However, at this point we might suggest a simple melotonic model of the form:

:

$$\left\|\vec{F}_1\right\| = \left\| \begin{pmatrix} \frac{e^{-\frac{k_1}{n^{c_1}} s_1^2}}{\sqrt{n}} \\ \frac{e^{-\frac{k_1}{n^{c_1}} s_2^2}}{\sqrt{n}} \\ \vdots \\ \frac{e^{-\frac{k_1}{n^{c_1}} s_n^2}}{\sqrt{n}} \end{pmatrix} \right\| = \sqrt{\frac{\sum_{i=1}^{n} \left(e^{-\frac{k_1}{n^{c_1}} s_i^2} \right)^2}{n}} \quad \text{and} \quad \left\|\vec{F}_2\right\| = \left\| \begin{pmatrix} \frac{e^{-\frac{k_2}{(n-1)^{c_2}} I_1^2}}{\sqrt{n-1}} \\ \frac{e^{-\frac{k_2}{(n-1)^{c_2}} I_2^2}}{\sqrt{n-1}} \\ \vdots \\ \frac{e^{-\frac{k_2}{(n-1)^{c_2}} I_n^2}}{\sqrt{n-1}} \end{pmatrix} \right\| = \sqrt{\frac{\sum_{i=1}^{n-1} \left(e^{-\frac{k_2}{(n-1)^{c_2}} I_i^2} \right)^2}{n-1}}$$

where $\left\|\vec{F}_1\right\|$ is the transpositional similarity predictor and $\left\|\vec{F}_2\right\|$ is the interval similarity predictor, k_1 and k_2 are empirical constants determining the strength of each interval component, c_1 and c_2 are empirical constants determining ho much the length of a chain affects similarity, s_1, s_2, ... s_n are the components of the similarity vector, I_1, I_2, ..., I_{n-1} are the components of the interval vector and n is the length of the compared m-chains.

An overall similarity could then be defined as:

$$S = \left\|\vec{F}_1\right\| \cdot \left\|\vec{F}_2\right\|$$

In fact, setting $c_1 = 1$ and $c_2 = -2$, we obtain a correlation of 87% with the data as produced by the two experiments as conducted by Hofmann-Engl & Parncutt (1998).

An overall similarity model taking rhythmic, dynamic and pitch features into account, might be of the form:

$$S = \alpha S_m + \beta S_d + \gamma S_r.$$

where α, β, γ are empirical constants, S_m as the pitch similarity, S_d as the dynamic similarity and S_r as the rhythmic similarity

This is not to say that this the most adequate model, but it is fashioned based on some available data, some theoretical concepts and is similar to Shepard's (1987) model. However, the model as it stands does not take into account any rhythmic or dynamic aspects nor does it pay tribute to harmonic features or emotional aspects. It also allows for the comparison of m-chains only which have equal length. Further, we might find that tones which are longer will bear more weight than shorter tones. Thus, as the model stands it might be only suitable as a predictor for short isochronous m-chains.

4. CONCLUSION

This paper set out to investigate an aspect of melodic similarity from a cognitive angle. We found that the term pitch is little satisfying and we argued for replacing it by the term meloton which was defined as a cognitive quality of sound. We further proposed that melotonic similarity is best approached by defining a set of transformations (reflections and translations). Based on the composition of two specific reflections we were able to define a similarity and interval vector which we propose to be somewhat sufficient to form the basis for a melotonic predictor. Specifically, we presented a simple similarity model which admittedly shows limitations but might demonstrate that more complex and comprehensive models can be developed. However, before a more comprehensive model will become available, many more experiments on melodic similarity will have to be conducted. Considering that we covered melotonic similarity only, we can by now conclude that the construction of sufficient models is a far more complex task than generally acknowledged, but at the same time it appears to be an achievable task.

5. REFERENCES

Adams, C. Melodic Contour Typology. Ethnomusicology, vol 20.2 (1976)

Anagnostopoulou, C., Hörnel, D. & Höthker, K. Investigating the Influence of Representations and Algorithms in Music Classification. Proceedings of the AISB '99 Symposium on Musical Creativity. (1999, Edinburgh)

Askill, J. The Subjective Pitch of Musical Chimes. (1997, Online publication).

http://faculty.millikin.edu/~jaskill.nsm.faculty.mu/chimes.html

Cambouropoulos, E. Melodic cue abstraction, similiarty and category formation:A computational appraoch. In Proceedings of ICMPC 2000. (2000, Keele University)

Crawford, T., Iliopoulos, C. & Raman, R. String-Matching Techniques for Musical Similarity and Melodic Recognition. In: Melodic Similarity - Concepts, Procedures, and Applications (ed. Hewlett, W. & Selfridge-Field, E.), Computing in Musicology 11. MIT Press, CA, 1998

Edworthy, J. Pitch and contour in music processing. Psychomusicology, 2, (1982), 44-46

Edworthy, J. Interval and contour in melody processing. Music Perception, 2, (1985), 375-388

Dovey, M. & Crawford, T. Heuristic Models of Relevance Ranking in Searching Polyphonic Music. In: Proceedings Diderot Forum on Mathematics and Music, (1999, Vienna, Austria), pp 111-123.

Dowling, W. J. Recognition of inversion of melodies and melodic contour. Perception & psychophysics, 12.5, (1971), 417-421

Dowling, W. J. & Harwood, D., (1986). Music Cognition, New York Academic Press, 1986

Downie, S., (1999). Evaluating a Simple Approach to Musical Information Retrieval: Conceiving Melodic N-Grams as Text. Ph.D. dissertation, University of Western Ontario

Egmond van, R. & Povel, D-J.. & Maris, E. The influence of height and key on the perceptual similarity of transposed melodies. Perception & Psychophysics, vol. 58, 1996, 1252-1259

Francés, R. The perception of music. (trans. Dowling), Hillsdale, New York, 1988

Goldstone, R. L. The Role of Similarity in Categorization: Providing a Groundwork. (1998, Online publication) http://cognitrn.psych.indiana.edu/rgoldsto/sim&cat.html, Indiana University

Hammel, B. Motivic Transformations & Lie Algebras. (1999, Online publication) http://graham.main.nc.us/~bhammel/Music/Liemotiv.html

Hofmann-Engl, L. J. Beiträge zur theoretischen Musikwissenschaft. M 65+1, vol. 2, 1989, Technical University Berlin

Hofmann-Engl, L. J. Virtual Pitch and pitch salience in contemporary composing.. In Preceedings the VI Brazilian Symposium on Computer Music at (1999, PUC Rio de Janeiro)

Hofmann-Engl, L. & Parncutt, R.. Computational modeling of melodic similarity judgments - two experimetns on isochronous melodic fragments. (1998, Online publication) http://www.chameleongroup.org.uk/research/sim.html

Hofmann-Engl, L. Review : Review of W.B. Hewlett & E. Selfridge-Field, eds., Melodic Similarity: Concepts, Procedures, and Applications.(Cambridge, Massachusetts: MIT Press, 1999). MTO 5.4, 1999, MIT CA

Houtsma, A J. M. & Goldstein, J. L. The central origin of the pitch of complex tones: Evidence from musical interval recognition. Journal of the Acoustical Society of America, vol. 51, 1972, 520-529

Isaacson, E. Issues in the Study of Similarity in Atonal Music. MTO 2.7, 1996, MIT CA

Kassler, M.. Toward Musical Information Retrieval. Perspectives of Music 4.2, 1966, 59-67

Kluge, R. Ähnlichkeitskriterien für Melodieanalyse. Systematische Musikwissenschaft, 4.1-2, 1996, 91-99

Lewin, D. Forte's Interval Vector, My Interval Function, and Regner's Common-note Function. Journal of Music Theory 21, 1977, 194-237

Maidin O, D. & Fernström, M. The Best of two Worlds: Retrieving and Browsing. Proceedings of COST G-6 on Digital Audio Effects, (2000, Verona)

Mazzola, G. Geometrie der Töne. Birkhäuser, Basel, 1990

Morris, R. A Similarity Index of Pitch-class Sets. Prespectives of New Music 18, 1979

Palmer, S. The psychology of perceptual organization. A transformational approach. Human and Machine Vision, New York Academic Press, 1983, 269-339.

Rasch, R. A. & Plomp, R.The Perception of Musical Tones. In: The Psychology of Music (e.g.. Deutsch), Academic Press, New York, 1982

Reti, R. The Thematic Process in Music. New York, Macmillan Company, 1951

Rohwer, J. Die harmonischen Grundlagen der Musik. Bärenreiter, Kassel, 1970

Schönberg, A. Fundamentals of Musical Composition. St. Martin's Press, New York, 1967

Schouten, J. F. The perception of subjective tones. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, 41, 1938, 1418-1424

Smith, L., McNab, J. & Witten, I. Sequence-Based Melodic Comparison: A Dynamic-Programming Approach. In: Melodic Similarity - Concepts, Procedures, and Applications (ed. Hewlett, W. & Selfridge-Field, E.), Computing in Musicology 11, 1998, MIT Press, CA

Sundberg, J.The Science of Musical Sound. Academic Press, New York, 1991

Tekman, H. G. A Multidimensional Study of Preference Judgments for Pieces of Music. Psychological Report, 82, 1998, 851-860

Terhardt, E Die Psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In: Tiefenstruktur der Musik, Technische Universität Berlin, Berlin, 1979

Terhardt, E. & Stoll, G. & Seewann, M. Algorithm for extraction of pitch and pitch salience from complex tonal signals. Journal of the Acoustical Society of America, vol. 71, 1982, 679-688

White, B. Recognition of distorted melodies. American Journal of Psychology, 73, 1960, 100-107