A Musical Expression Model Using Tonal Tension
Tetsuya MIZUTANI* and Yamato IWAMI
Department of Computer Science, University of Tsukuba, Ibaraki 3058573, Japan
*Corresponding author

Keywords: Tonal pitch space, Tonal tension.

Abstract. Musical expression is made by repetition of relaxation and tension. The tension of a score is made from its chord progression, melodic motion, beat structure, etc. In this paper, a simple model of musical expression is investigated that is obtained from expression symbols in the score of music with the harmonic tension values introduced by Tonal Pitch Space.

Introduction
Musical expression is made by repetition of relaxation and tension. The tension in a score is made from its chord progression, melodic motion and progression, beat structure, rhythmic structure, etc., which depends on not only the structure of the score of music but also the historical interpretation of the piece.

Many conventional musical theories insist that the tension associated with music scores tends to accelerate its performance speed (agogics) and to be up its sound (dynamics), and vise versa [1]. However, in the music informatics or the artificial intelligence field, there is a few researches to investigate the relationship between such tension with scores and musical expression of actual performances, e.g. [2].

A generative theory of tonal music (GTTM) [3] is a cognitive theory of tonal music that analyzes music by some strong hypotheses and structural rules of music which introduces hierarchical musical structures. Tonal pitch space (TPS) [4] is succession of it introducing structures of pitch, tone, chord, etc. Each theory models hypothetical experienced listeners who are supposed to listen music structurally and hierarchically. These theories introduce the harmonic tension structures of music scores, by the notions harmonic distances, prolongational reduction structures, etc., the latter of which express hierarchical relationship, called the prolongational tree, between each harmonic and melodic event from a musical score, and they represent the repetition of tension, relaxation, continuity and progression. From them, the tonal tension structure of chord progression is proposed, by which the tension value is defined.

It must be noted that these theories are not for expression of music performance but for the structures of scores, i.e. they do not mention any actual expression model.

In this paper, a simple model of musical expression obtained from a relationship between an actual performance of a music and the structures of the score of the piece of music, which is calculated by the multiple regression analysis using expression symbols in the score with the harmonic tension values as the independent variables of the analysis.

Tonal Tension Values
Chord Distance
The chord distance between two chords expresses the distance of them in the musical theory.

Definition 1
The chord distance $\delta(x \rightarrow y)$ between two chords is defined as

$$\delta(x \rightarrow y) = i + j + k$$ (1)
where \( i \) is the distance in the cycle of fifth of the diatonic scale of \( x \) and that of \( y \); \( j \) the distance of two chords in the cycle of fifth; \( k \) is the weighted sum of \( y \) minus that of \( x \). The weighted sum of a chord is calculated as the following weight for each note in the chord; weight 4 if it is the root note, 3 if the fifth note in the chord, 1 if the other note in the chord, 0 otherwise, i.e. it does not consist of the chord.

It must be noted that the distance is not in the metric space of mathematics, i.e. it does not hold the symmetric relation nor triangle inequality.

**Harmonic Tension Values**

There are 3 kinds of harmonic tension values for a chord in the music score, the surface tension value, the inheritance tension value and the hierarchical global tension value, defined by the prolongational tree.

**Definition 2**

The surface tension value \( T_{\text{diss}}(y) \) for a target chord \( y \) is defined as:

\[
T_{\text{diss}}(y) = sd + inv + nt
\]

where \( sd \) (scale degree) is 1 if the melodic note is the third or the fifth of \( y \), and otherwise 0; \( inv \) (inversion) is 2 if it is inversion and 0 otherwise; \( nt \) (nonharmonic tone) 1 for seventh, 3 for diatonic nonharmonic and 4 for chromatic nonharmonic tone, 0 otherwise.

**Definition 3**

The hierarchical global tension value \( T_{\text{glob}}(y) \) for a chord \( y \) is defined as:

\[
T_{\text{glob}}(y) = T_{\text{diss}}(y) + \delta(x_{\text{dom}} \rightarrow y) + T_{\text{inh}}(x_{\text{dom}})
\]

where \( y \) is the target chord, \( x_{\text{dom}} \) the chord that directly dominates \( y \) in the prolongational tree, \( T_{\text{diss}}(y) \) the surface tension associated with \( y \), \( \delta(x_{\text{dom}} \rightarrow y) \) the chord distance from \( x_{\text{dom}} \) to \( y \) and \( T_{\text{inh}}(x_{\text{dom}}) \), called the inheritance tension value of \( y \), the sum of the chord distance values inherited by \( y \) from chords that dominate \( x_{\text{dom}} \).

**Multiple Regression Analysis**

The multiple regression analysis [5] is a statistical technique for estimating the relationship among a dependent variable and some independent variables.

Let \( y \) be a dependent variable, \( x_1, \ldots, x_n \) be independent variables, \( a_1, \ldots, a_n \) be partial regression coefficients, and \( a_0 \) be an intercept. Suppose that the size of the set of samples is greater than \( n \). Then, the optimum solution for the coefficients is obtained by the least squares method from the multiple regression functions

\[
y = a_0 + a_1 x_1 + \ldots + a_n x_n
\]

The \( p \)-value indicates the probability that the partial regression coefficient is 0. In general, it is the strong evidence against null hypothesis, i.e. the coefficient is meaningful, if the value is less than 0.05, when the significant level is 5%.

The samples used in the analysis are events, each of which is defined as the set of notes sounding simultaneously. Notice that event must be distinguished from chord. The value of the dependent variable in each sample is the BPM (Beats Per Minute) of the corresponding event, those of independent variables are surface and global tension value of the event and the existence of each music symbols, e.g. accent, crescendo, etc.

**Analysis**

In order to investigate expression models of music, the multiple regression analysis has been done for Nocturne in E-flat Major, Op. 9, No. 2 by Frédéric François Chopin, performed by V.D. Ashkenazy, whose score and performance data are obtained from CrestMuseDB [6], one of the largest music
A database for researches of musical informatics. Figure 1 shows the score of the first 8 bars of the piece and Table 1 shows the list of independent variables used in the analysis.

![Figure 1. The score of nocturne in e-flat major, op. 9, No. 2, by chopin.](image)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
<th>Type of Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_surface</td>
<td>The surface tension value</td>
<td>nonnegative integer</td>
</tr>
<tr>
<td>T_hierarchical</td>
<td>The hierarchical global tension value</td>
<td></td>
</tr>
<tr>
<td>T_total</td>
<td>The sum of the above two values</td>
<td></td>
</tr>
<tr>
<td>crescendo</td>
<td>The event with crescendo</td>
<td></td>
</tr>
<tr>
<td>diminuendo</td>
<td>The event with diminuendo</td>
<td></td>
</tr>
<tr>
<td>Met_slurST1</td>
<td>The starting event of a slur for a downbeat</td>
<td></td>
</tr>
<tr>
<td>Met_slurST2</td>
<td>The starting event of a slur for an upbeat</td>
<td></td>
</tr>
<tr>
<td>pre_grace_note</td>
<td>The event just before a grace note</td>
<td></td>
</tr>
<tr>
<td>suc_grace_note</td>
<td>The event just after a grace note</td>
<td></td>
</tr>
<tr>
<td>piano</td>
<td>The event with piano (as an expression symbol)</td>
<td></td>
</tr>
<tr>
<td>staccato</td>
<td>The event with a staccato</td>
<td></td>
</tr>
<tr>
<td>turn</td>
<td>The event with a turn</td>
<td></td>
</tr>
<tr>
<td>pralltriller</td>
<td>The event with a pralltriller</td>
<td></td>
</tr>
<tr>
<td>accent</td>
<td>The event with an accent</td>
<td></td>
</tr>
<tr>
<td>trill</td>
<td>The event with a trill</td>
<td></td>
</tr>
</tbody>
</table>

The multiple regression analysis has been done with the following 3 cases for agogics and dynamics, respectively:

- **Case 1.** Using all independent variables but those for tension values;
- **Case 2.** Case 1 with the two tension values as two independent variables (T_surface and T_hierarchical);
- **Case 3.** Case 1 with the sum of the surface and global tension values as one independent variable (T_total).

Figure 2 shows the experimental results of the coefficients by the multiple regression analysis, which are computed by a program written in R [7]. Table 2 shows the influence of the tension values expressing for the values of the coefficient and the p-values, as a summary of Figure 2.

Table 3 is the correlation coefficients between the original and the estimated performances. It is shown that the estimated performances have weak positive correlations for agogics and rather strong for dynamics.
Table 2. The Coefficients of tension values and their p-values

<table>
<thead>
<tr>
<th>Case</th>
<th>Agogics</th>
<th>Dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>p-value</td>
</tr>
<tr>
<td>2. T_surface</td>
<td>-0.515</td>
<td>0.266</td>
</tr>
<tr>
<td>2. T_hierarchical</td>
<td>0.286</td>
<td>0.051</td>
</tr>
<tr>
<td>3. T_total</td>
<td>0.183</td>
<td>0.158</td>
</tr>
</tbody>
</table>

Table 3. Correlational coefficients

<table>
<thead>
<tr>
<th></th>
<th>Agogics</th>
<th>Dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>0.437</td>
<td>0.523</td>
</tr>
<tr>
<td>Case 2</td>
<td>0.476</td>
<td>0.710</td>
</tr>
<tr>
<td>Case 3</td>
<td>0.455</td>
<td>0.680</td>
</tr>
</tbody>
</table>

From the results of the analysis, the model introduced in this article succeeds to capture some tendency of expression of the piece, however, it is rather weak and coarse in order to synthesis actual musical performances. The possible reasons are:

1. The relationship among independent variables is expressed by linear combinations, although the actual change of agogics and dynamics may be nonlinear.
2. There may exist some hidden parameters which effect expression.
3. The sample data are very few. If numerous data of music pieces and performance data are used in the analysis, stronger and finer model may be obtained.
4. The music expression strongly depends on the music pieces and the personality of each performer, especially his own interpretation of them.

**Related Works**

There are many researches in this area. Especially, a system for the multiple regression analysis in order to apply the musical expression has been developed [8].

There are also many researches for music performance systems. For example, a performance system to modify parameters of the performance rules is introduced in [9].

The relationship between score information and the performance expression has been reported in [10], whose result is applied to the automatic ensemble system to create expressive secondo performances [11].

**Conclusion**

A simple musical expression model using harmonic tension values with expression symbols is introduced, which captures expression of music piece, however, which is coarse in order to synthesis actual musical performances. More analysis with huge datasets of music pieces and performance data must be done in our future work, and moreover, additional parameters for music expression, like melodic attraction mentioned in [4], must be investigated.

**References**