Autocorrelation and the Study of Musical Expression

Peter Denis and Siebe de Vos

Music Department  Centre for Knowledge Technology
City University  Utrecht School of the Arts
Nagumpion Square  Lange Visstraat 38
GB-London ECIV OHB  NL-3511 BK Utrecht

ABSTRACT: In performance musical structure is conveyed as variations of timing and other parameters. A method was designed to analyze these variations using autocorrelation. Peaks in the autocorrelation function are interpreted as periods of repeated components in the musical structure. Care has to be taken in using the standard autocorrelation function in this domain. Partial autocorrelation was used to remove the effects of a fundamental period.

Introduction

In musical performances the performer uses variations of timing, dynamics and articulation. An often posed hypothesis is that these expressive variations are closely linked to—the performer's interpretation of it [Clarke 1987].

In a research project on expressive timing we designed a set of tools, called POCO, to analyze, modify and generate musical expression [Honing 1990]. One of the tools analyzes expressive data using autocorrelation in order to find regularities which, according to the hypothesis above, correspond to musical structure. In this article we focus on expressive timing, but other expressive parameters can be dealt with using the same method.

Any deviation from a strict metronomical performance is regarded as expressive timing. We assume that it mainly stems from a multiplicative combination of tempo factors at several structural levels and the exact metrical note durations in a score. By dividing the durations observed in the performance (the inter-onset intervals) by the durations in the score a measure of local tempo is obtained. This function from time (onset-time in the score) to relative duration, the expressive timing signal, the logarithm of which can depends linearly on the components.

As an example we will use Bach's C major prelude (WFC 1), which was the subject of many other studies [Cook 1987, Lehrer and Jackendoff 1983, Povel 1977]. All notes in this piece are of equal duration. The main structural units are half-bar, bar (16 notes) and 2 bars, at higher levels the metrical grouping is not trivial. Performers generally exhibit an amazing consistency in expressive timing over performances. The expressive timing of one of the performances is shown in Figure 1.

Todd's approach in analyzing this kind of timing data is to look at the local maxima [Todd 1987]. They indicate a slowing down and Todd's analysis relates the relative height of these peaks directly to the structural boundary strength. Although it is not so difficult to spot obvious phrase-final lengthening, it is unlikely that a robust classification of peaks into structural levels can be made. Therefore we look for more global methods to detect structure.

![Figure 1. Expressive timing of the Bach Prelude](image-url)
Regularity in musical structure will be reflected as periodicity in the expressive timing signal. We use autocorrelation as a statistical method to find periodicity; the periods found are interpreted as the length of unequal components. We will assume here that musical structure is more or less homogeneous, at least for some time span and at some level.

**Autocorrelation**

If a signal is periodic with a period \( P \), it will resemble itself after an interval \( P \). A well-known statistical measure of resemblance is correlation. By calculating correlations between a signal and the same signal delayed by different lags we obtain a series of values, the autocorrelation curve. When the signal contains a periodic component with period \( P \), a peak in the autocorrelation curve occurs at this value. Considering our domain we have to be careful in the use of the standard autocorrelation [Bowen-Magnus 1979, Priestley 1981]. The function must depend on time to be able to show changes in periodic structure. We realize this by placing a window on the samples. Then the autocorrelation at \( t \) is the autocorrelation in the window that ends in \( t \) or \( X(t-w+1) \ldots X(t) \), where \( W \) is the window size and \( X \) the signal. The window size should depend on the lag, otherwise a change in the level of a component with a small period will go unnoticed since there are still many 'old' periods contained in the window. We choose a window proportional to the lag, in the examples we used a factor \( \mu \). A second reason to use relatively small windows is that the signal cannot be assumed to be stationary, which means that its statistical properties like mean and variance change over time. But using small time intervals the error introduced may be neglected. This leads to the following definitions of mean and autocovariance:

\[
X_{w} = \frac{1}{W} \sum_{i=0}^{W-1} X(i),
\]

\[
R_{w}(r) = \frac{1}{W} \sum_{i=0}^{W-r-1} (X(i+r)-X_{w})(X(i)-X_{w}).
\]

The factor \( \frac{1}{W} \) instead of \( \frac{1}{W-r} \) corrects the values for greater lags which are calculated with only a fraction of the samples, as in the commonly used biased autocovariance estimator [Priestley 1981].

The autocorrelation is defined in terms of the autocovariance as:

\[
\rho_{w}(r) = \frac{R_{w}(r)}{R_{w}(0)}
\]

and the time dependent autocorrelation function with proportional window size is given by:

\[
\rho(r) = \rho_{w}(r)
\]

Figure 2 shows the autocorrelation of the signal in Figure 1. Note the prominent peaks at the lags corresponding to the length of metrical units. We found these peaks only in data of expert performers, showing their ability to produce consistent timing patterns. Note also that the autocorrelation definition used is not very meaningful in the smaller lags, because it depends on a very small number of measurements.

![Figure 2. Autocorrelation of the data of Figure 1 (beginning of bar 31, \( \mu=9 \)).](image)

ICMC GLASGOW 1990 PROCEEDINGS

35
Partial autocorrelation
A problem occurs in interpreting the autocorrelation curve. When a signal repeats itself after period \( P \), it will also be the same after period \( 2P \), \( 3P \), etc. To detect if there is additional regularity at these levels over and above the regularity originating from their 'fundamental', \( P \), we use partial autocorrelation. Partial correlation determines the correlation between two variables, cancelling out the influence of other variables on both of them. In the case of autocorrelation, it removes the effect of smaller periodicities on the autocorrelation for a certain lag. The partial autocorrelation at lag \( r \), \( \rho(r|s) \), is defined as (Bowman 1979):

\[
\rho(k, k) = \frac{\sum_{j=1}^{k-1} p(k-1,j)p(k,j)}{1 \sum_{j=1}^{k-1} p(k-1,j)^2} \quad \text{if } k > 1,
\]

\[
\rho(k, j) = \rho(k-1,j) \cdot \rho(k, k) \cdot \rho(k-1,k-j).
\]

This formula depends on a statistically sound autocorrelation function. We cannot use the modified autocorrelation directly, but it is possible to retain the dependence on time and lag when we recompute the autocorrelation function for each lag of the partial autocorrelation:

\[
\rho(r|s) = \rho(r,s) \quad \text{where } \rho(s) = \rho_{yy}(s).
\]

The advantage of partial autocorrelation is seen in Figure 3: e.g., the peak at the 3.5 bar lag in Figure 2, which arose only because it is a multiple of the half bar length vanishes in the partial autocorrelation.

![Figure 3. Partial autocorrelation of the data of Figure 1 (beginning of bar 31, p = 4).](image)

Making use of the time dependency of the analysis we can show the relative stability of merical units of 2 bars or smaller, throughout the piece (Figure 4). In this picture the data is truncated to zero for correlation smaller then 0.05. Note the conflicting evidence of a 3 bar and a 4 bar unit arising because of the inhomogeneity of the piece at higher levels.

Conclusion
Although we have not yet studied many performances of the Bach Prelude, nor other pieces, autocorrelation seems a promising method to study expressive timing and detecting merical structure from expression. The method can also be used to detect structural changes in a performance and to investigate how absolute tempo determines the focus of attention of the performer to particular structural levels, changing the relative heights of the peaks in the partial autocorrelation. Although we only showed an example with notes of equal duration, other kinds of music can be treated with an interpolation scheme.
However, the method has severe intrinsic limitations. It is based on the assumption that the expressive components at each structural level are more or less stable and periodic and it assumes independent combination of the components, an assumption that clearly limits the applicability of this method. Furthermore, no phase information is retained. Statistical reliability is questionable for small windows and we can use the method only for generating hypotheses about the structure of the piece, not for testing them statistically.

Figure 4. Partial autocorrelation of the data of figure 1 through time ($p = 4$).

In further research we will work on different measures of expressive timing. We want to use the result of autocorrelation (and another kind of analysis) to separate the independent structural components of the expressive timing signal. This will open up the possibility of 'micro surgery' on expressive timing, in which modifications can be made at each structural level. Another application of the analysis method described is forecasting, in which the expressive timing curve can be extrapolated from a known segment. This might result in more robust methods of score following, tempo tracking and quantization. As this research tries to unravel the internal structure of the expressive timing signal, we hope to gain more insight in the musical and cognitive reality of tempo curves and their representation.

Acknowledgements
We would like to thank Eric Clarke, Henkjan Honing, Steve McAdams, Klaus de Rijck, and Luc van Vugt for their help and we are especially grateful to Chris Mould for his performances.

References
Bowenman, B.L. and R.T. O’Connell. 1979 "Time Series and Forecasting, an Applied Approach" Boston, MA: PWS.
Honing, H. 1990 "POCO, An Environment for Analysing, Modifying and Generating Expression in Music." ICMC 90, San Francisco: CMA.

ICMC GLASGOW 1990 PROCEEDINGS