EVIDENCE FOR PIANIST-SPECIFIC RUBATO STYLE IN CHOPIN NOCTURNES

Miguel Molina-Solana
Dpt. Computer Science and AI
University of Granada, Spain
miguelmolina at ugr.es

Maarten Grachten
IPEM - Dept. of Musicology
Ghent University, Belgium

Gerhard Widmer
Dpt. of Computational Perception
Johannes Kepler Univ., Austria

ABSTRACT

The performance of music usually involves a great deal of interpretation by the musician. In classical music, the final ritardando is a good example of the expressive aspect of music performance. Even though expressive timing data is expected to have a strong component that is determined by the piece itself, in this paper we investigate to what degree individual performance style has an effect on the timing of final ritardandi. The particular approach taken here uses Friberg and Sundberg's kinematic rubato model in order to characterize performed ritardandi. Using a machine-learning classifier, we carry out a pianist identification task to assess the suitability of the data for characterizing the individual playing style of pianists. The results indicate that in spite of an extremely reduced data representation, when cancelling the piece-specific aspects, pianists can often be identified with accuracy above baseline. This fact suggests the existence of a performer-specific style of playing ritardandi.

1. INTRODUCTION

Performance of music involves a great deal of interpretation by the musician. This is particularly true of piano music from the Romantic period, where performances are characterized by large fluctuations of tempo and dynamics. In music performance research it is generally acknowledged that, although widely used, the mechanical performance (with a constant tempo throughout the piece) is not an adequate norm when studying expressive timing, since it is not the way a performance should naturally sound.

As an alternative, models of expressive timing could be used, as argued in [18]. However, only few models exist that deal with expressive timing in general [2, 16]. Due to the complexity and heterogeneity of expressive timing, most models only describe specific phenomena, such as the timing of grace notes [15] or the final ritardando.

Precisely, the final ritardando —the slowing down toward the end of a musical performance to conclude the piece gracefully— is one of the clearest manifestations of expressive timing in music. Several models have been proposed [3, 14] in the related literature to account for its specific shape. Those models generally come in the form of a mathematical function that describes how the tempo of the performance changes with score position.

In a previous empirical study by Grachten *et al.* [4] on the performance of final ritardandi, a kinematic model [3] was fitted to a set of performances. Even though some systematic differences were found between pianists, in general the model parameters tend to reflect primarily aspects of the piece rather than the individual style of the pianist (i.e. expressive timing data is expected to have a strong component that is determined by piece-specific aspects).

This fact is relevant in a recurrent discussion in the field of musicology, about which factor (the piece or the performer) mostly influences a performance [9]. Some experts argue that the performance should be preceded of a thorough study of the piece; while others indicate that the personal feeling of music is the first and main point to be considered. Works supporting both views can be found in [12]. A study by Lindström *et al.* [7] involving a questionnaire, showed that music students consider both the structure of the piece and the feelings of the performer as relevant in a performance.

The current paper extends that previous work by Grachten *et al.*, by investigating whether or not canceling piece-specific aspects leads to a better performer characterization. Musicologically speaking, the validation of this hypothesis implies that performers' signatures do exist in music interpretation regardless of the particular piece. We present a study of how final ritardandi in piano works can be used for identifying the pianist performing the piece. Our proposal consists in applying a model to timing data, normalizing the fitted model parameters per piece and searching for performer-specific patterns.

Performer characterization and identification [8, 13] is a challenging task since not only the performances of the same piece by several performers are compared, but also the performance of different pieces by the same performer. Opposed to performer identification (where performers are supposed to have distinctive ways of performing) is piece identification —which requires the structure of the piece to imply a particular expressive behavior, regardless of the performer.

A further implication of this work would be that, when an estimation can be made of the prototypical performance based on the musical score, this estimation could be a useful reference for judging the characteristics of performances. This knowledge can also allow the artificial interpretation of musical works by a computer in expressive and realistic ways [17].

This paper is organized as follows: Section 2 describes the dataset used for this study, including the original timing data and the model we fit them to. Section 3 deals with the data processing procedure. Results of the pianist classification task are presented and discussed in Section 4, while Section 5 states conclusions and future work.

2. DATA

The data used in this paper come from measurements of timing data of musical performances taken from commercial CD recordings of Chopin's Nocturnes. This collection has been chosen since these pieces exemplify classical piano music from the Romantic period, a genre that is characterized by the prominent role of expressive interpretation in terms of tempo and dynamics. Furthermore, Chopin's Nocturnes is a well-known repertoire, performed by many pianists, and thus facilitating large scale studies.

As explained before, models of expressive timing are generally focused in a certain phenomenon. In our study, we will focus on the final ritardando of the pieces. Hence, we select those Nocturnes whose final passages have a relatively high note density and are more or less homogeneous in terms of rhythm. With these constraints we avoid the need to estimate a tempo curve from only few interonset intervals, and reduce the impact of rhythmic particularities on the tempo curve.

In particular, we used ritardandi from the following pieces: Op. 9 nr. 3, Op. 15 nr. 1, Op. 15 nr. 2, Op. 27 nr. 1, Op. 27 nr. 2 and Op. 48 nr. 1. In two cases (Op. 9 nr. 3 and Op. 48 nr. 1), the final passage consists of two clearly separated parts, being both of them performed individually with a ritardando. These ritardandi were treated separately — namely *rit1* and *rit2*. So that, we have 8 different ritardandi for our study.

The data were obtained in a semi-automated manner, using a software tool [10] for automatic transcription of the audio recordings. From these transcriptions, the segments corresponding to the final ritardandi were then extracted and corrected manually by means of *Sonic Visualiser*, a software tool for audio annotation and analysis [1].

The dataset in this paper is a subset of that used in previous work [4], as we are only considering those pianists from whom all eight recordings are available. Table 1 shows the names of these pianists and the year of their recordings. Hence, the dataset for the current study contains a total amount of 136 ritardandi from 17 different pianists.

2.1 Friberg & Sundberg's kinematic model

As mentioned in Section 1, we wish to establish to what degree the specific form of the final ritardando in a musical

Arrau (1978)	Falvai (1997)	Pires (1996)
Ashkenazy (1985)	Harasiewicz (1961)	Pollini (2005)
Barenboim (1981)	Hewitt (2003)	Rubinstein (1965)
Biret (1991)	Leonskaja (1992)	Tsong (1978)
d'Ascoli (2005)	Mertanen (2001)	Woodward (2006)
Engerer (1993)	Ohlsson (1979)	

Table 1. Performer and year of the recordings analyzed in the experiments

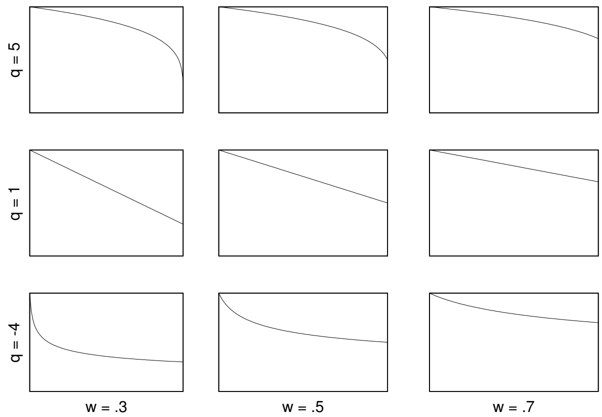

Figure 1. Examples of tempo curves generated by the model using different values of parameters w and q. In each plot, the x and y axis represent score position and tempo respectively, both in arbitrary units.

performance is dependent on the identity of the performing pianist. We address this question by fitting a model to the data, and investigating the relation between the piece/pianist identity and the parameter values of the fitted model. To such a task, we employ the kinematic model by Friberg & Sundberg [3].

This model is based on the hypothesized analogy of musical tempo and physical motion, and is derived from a study of the motion of runners when slowing down. From a variety of decelerations by various runners, the decelerations judged by a jury to be most aesthetically pleasing turned out to be those where the deceleration force is held roughly constant. This observation was implying that velocity was proportional to square root function of time, and to a cubic root function of position. Equating physical position to score position, Friberg and Sundberg used this velocity function as a model for tempo in musical ritardandi. Thus, the model describes the tempo $v(x)$ of a ritardando as a function of score position x:

$$v(x) = (1 + (w^q - 1)x)^{1/q} \tag{1}$$

The parameter q is added to account for variation in curvature, as the function is not necessarily a cubic root of position. The parameter w represents the final tempo, and was added since the tempo in music cannot reach zero. The model can be fitted to ritardandi performed by particular pianists by means of its parameters.

Parameters w and q generate different plots of tempo curves (see Figure 1). Values of $q > 1$ lead to convex tempo curves, whereas values of $q < 1$ lead to concave

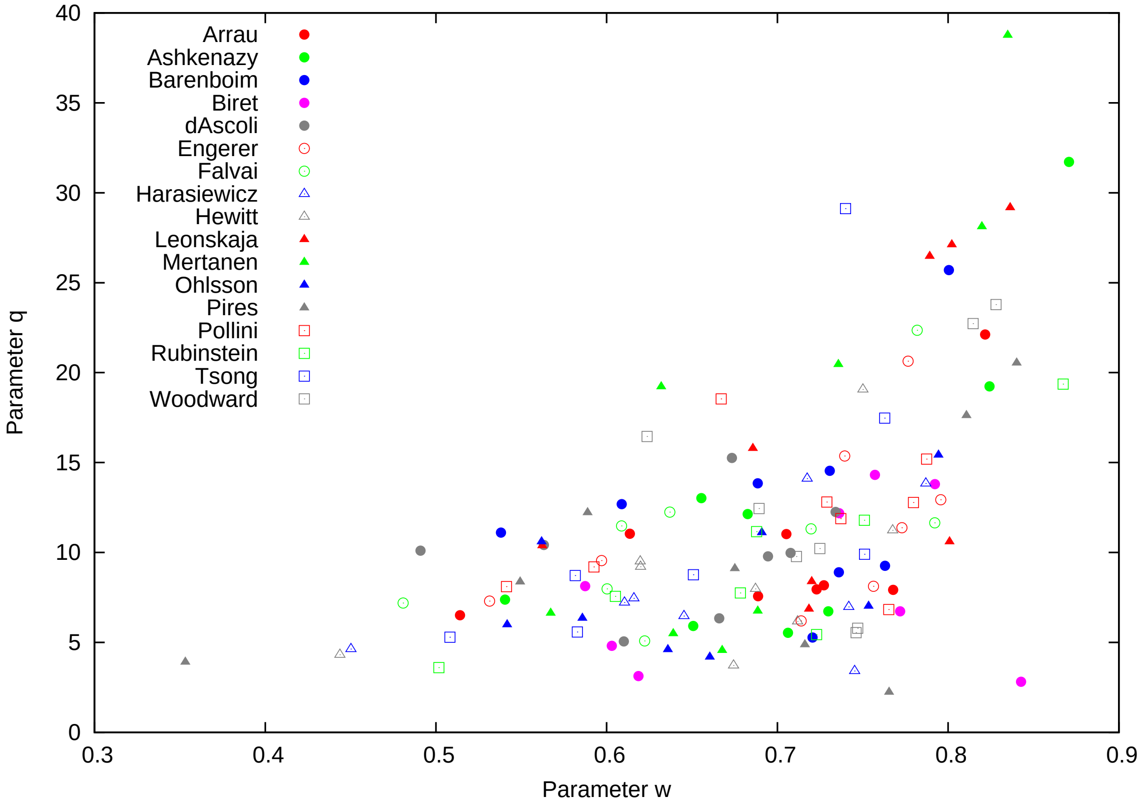

Figure 2. Original data representation in the w-q plane

curves. The parameter w determines the vertical end position of the curve.

Even though this kind of models are incomplete as they ignore several musical characteristics [6], the kinematic model described above was reported to predict the evolution of tempo during the final ritardando quite accurately, when matched to empirical data [3]. An additional advantage of this model is its simplicity, both conceptually (it contains few parameters) and computationally (it is easy to implement).

The model is designed to work with normalized score position and tempo. More specifically, the ritardando is assumed to span the score positions in the range [0,1], and the initial tempo is defined to be 1. Although in most cases there is a ritardando instruction written in the score, the ritardando may start slightly before or after this instruction. When normalizing, we must assure that normalized position 0 coincide with the actual start of the ritardando. A manual inspection of the data showed that the starting position of the ritardandi strongly tended to coincide among pianists. For each piece, the predominant starting position was determined and the normalization of score positions was done accordingly.

The model is fitted to the data by non-linear least-squares fitting through the Levenberg-Marquardt algorithm [1], using the implementation from *gnuplot*. The model fitting is applied to each performance individually, so for each combination of pianist and piece, three values are obtained: w, q and the root mean square of the error after fitting (serving this value as a *goodness-of-fit* measure).

At this point, we can represent each particular ritardando in the corpus as a combination of those two attributes: w and q. In Figure 2 [2], the values obtained from fitting are displayed as a scatter plot on the two-dimensional attribute space q versus w. The whole dataset —136 instances— is shown in this plot. Each point location correspond to a certain curve with parameters w and q. We refer the reader to Figure 1 to visualize the shape of different combination of parameters.

As can be seen from Figure 2, there are no clusters that can be easily identified from this representation. Hence, the performer identification task using these original data is expected to have a low success rate.

3. METHOD

In Section 1, we already mentioned that the expressive timing data is expected (as stated in [4]) to have a strong component that is determined by piece-specific aspects such as rhythmical structure and harmony. In order to focus on pianist-specific aspects of timing, it would be helpful to remove this piece-specific component.

Let X be the set of all instances (i.e. ritardando performances) in our dataset. Each instance $x \in X$ is a duple (w, q). Given a ritardando i, X_i is the subset of X that

[1] The fitting must be done by numerical approximation since the model is non-linear in the parameters w and q

[2] this figure is best viewed in color

contains those instances $x \in X$ corresponding to that particular ritardando.

In order to remove the piece-specific components, we propose to apply a linear transformation to the 2-attribute representation of ritardandi. This transformation consists in calculating the performance norm for a given piece and subtracting it from the actual examples of that piece. To do so, we first group the instances according to the piece they belong. We then calculate the centroid of each group (e.g. mean value between all these instances) and move it to the origin, moving consequently all the instances within that group.

We are aware that modelling the performance norm of a given ritardando as the mean of the performances of that ritardando is not the only option and probably not the best one. In fact, which performance is the best and which one is the most representative is still an open problem with no clear results. Moreover, several performance norms can be equally valid for the same score. In spite of these difficulties, we chose to use the mean to represent the performance norm, for its simplicity and for the lack of an obvious alternative.

Two approaches were then devised in order to calculate that performance norm. In the first one, the mean performance curve is calculate as a unweighted mean of the attributes w and q (see Equation 2); whereas in the second one, *fit* serves to weight the mean (see Equation 3).

In the first approach, the performance norm for a given ritardando i can be calculated as:

$$norm_i = \frac{\sum\limits_{x_i \in X_i} x_i}{|X_i|} \tag{2}$$

In the second approach, it is calculated as a weighted mean, where fit_i stands for the *fit* value of instance x_i:

$$norm_i = \frac{\sum\limits_{x_i \in X_i} x_i \mathit{fit}_i}{\sum \mathit{fit}_i} \tag{3}$$

In either case, all instances x_i are then transformed into x'_i by subtracting the corresponding performance norm:

$$x'_i = x_i - norm_i \tag{4}$$

X' would be then the dataset that contains all x'. After this transformation, all x' contain mainly information about the performer of the ritardando, as we have removed the common component of the performances per piece.

4. EXPERIMENTATION

In order to verify whether pianists have a personal way of playing ritardandi, independent of the piece they play, we have designed a classification experiment with different conditions, in which performers are identified by their ritardandi. The ritardandi are represented by the fitted model parameters. In one condition, the data instances are the set X, i.e. the fitted model parameters are used as such, without modification. In the second and third conditions,

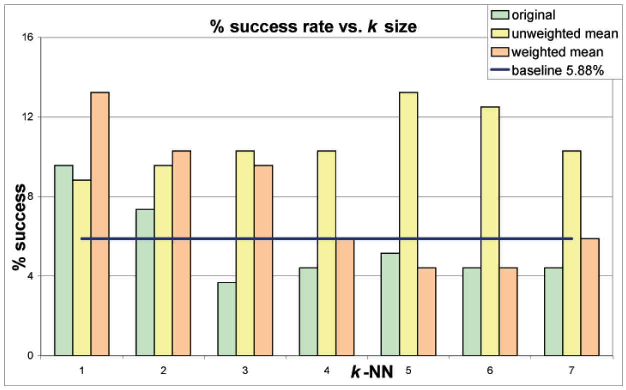

Figure 3. % success rate in the performer identification task using the whole dataset, with different k-NN classifiers. Baseline value (5.88%) from random classification is also shown

the piece-specific component in every performance is subtracted (data set X'). The second condition uses the unweighted average as the performance norm, the third condition uses the weighted average.

Note that accurate performer identification in this setup is unlikely. Firstly the current setting, in which the number of classes (17) is much higher than the number of instances per class (8), is rather austere as a classification problem. Secondly, the representation of the performer's rubato by a model with two parameters is very constrained, and is unlikely to capture all (if any) of the performer's individual rubato style. Nevertheless, by comparing results between the different conditions, we hope to determine the presence of individual performer style independent of piece.

As previously explained, the training instances (ritardandi of a particular piece performed by a particular pianist) consist of two attributes (w and q) that describe the shape of the ritardando in terms of timing. Those attributes come from matching the original timing data with the kinematic model previously cited.

The pianist classification task is executed as follows. We employ k-NN (Nearest Neighbor) classification, with $k \in \{1, \ldots, 7\}$. The target concept is the pianist in all the cases, and two attributes (w and q) are used. For validation, we employ leave-one-out cross-validation over a dataset of 136 instances (see Section 2). The experiments are carried out by using the Weka framework [5].

Figure 3 shows the results for the previously described setups, employing a range of k-NN classifiers with different values of $k \in \{1, \ldots, 7\}$. We also carry out the classification task using the original data (without the transformation) that were shown in Figure 2, in order to compare the effect of the transformation.

The first conclusion we can extract from the results is that the success rate is practically always better when transforming the data than when not. In other words, by removing the (predominant) piece-specific component, it gets easier to recognize performers. This is particularly interesting as it provides evidence for the existence of a performer-specific style of playing ritardandi, which was our initial hypothesis.

Note however, that the success rate is not so good to allow this representation for being a suitable estimation of the performer of a piece, even in the best case. A model with only two parameters cannot comprise the complexity of a performer expressive fingerprint. Although improving performer identification is an interesting problem, that is not the point of this work.

As can be seen, employing a weighted mean of w and q for calculating the performance norm of a piece —being *fit* the weight— leads to better results when k is small (i.e. $k < 3$). However, this approach, which is methodologically the most valid, does not make a remarkable difference with respect to the original data for larger values of k.

An interesting and unexpected result is that the transformation with the unweighted mean (see equation 2), gives better results for medium-large k values. The lower results for smaller k could be explained by the fact that instances with a low fit (which are actually noisy data), interfere with the nearest-neighbor classification process. The better results for higher k suggest that in the wider neighborhood of the instance to be classified, the instances of the correct target dominate —and thus that the noise due to low fit is only limited.

Note also that this approach is more stable with respect to the size of k than the original or the weighted ones. It also outperforms the random classification baseline —that is 5.88% with 17 classes— for all the different values of k.

Further experiments show that those are the trends for those two different transformation of the data. Employing the weighted mean leads to the highest accuracy using a 1-NN classifier, but it quickly degrades as k is increased. On the other hand, an unweighted mean leads to more stable results, with the maximum reached with an intermediate number of neighbors.

Although (as expected with many classes, few instances and a simplistic model) the classification results are not satisfactory from the perspective of performer identification, the improvement that transforming the data (by removing piece-specific aspects) gives in classification results, suggests that there is a performer-specific aspect of rubato timing. Even more, it can be located specifically in the curvature and depth of the rubato (w and q parameters).

5. CONCLUSIONS AND FUTURE WORK

Ritardandi in musical performances are good examples of the expressive interpretation of the score by the pianist. However, in addition to personal style, ritardando performances tend to be substantially determined by the musical context they appeared in. Because of this fact, we propose in this paper a procedure for canceling these piece-specific aspects and focus on the personal style of pianists.

To do so, we make use of collected timing variations during ritardando in the performances of Chopin Nocturnes by famous pianists. We obtain a two-attributes (w,q) representation of each ritardando, by fitting Friberg and Sundberg's kinematic model to the data.

A performer identification task was carried out using k-Nearest Neighbor classification on, comparing the (w,q) representation to another condition in which average w and q values per piece are subtracted from each (w,q) pair.

The results indicate that in even in this reduced representation of ritardandi, pianists can often be identified by the tempo curve of the ritardandi above baseline accuracy. More importantly, removing the piece-specific component in the w and q values leads to better performer identification.

This suggests that even very global features of ritardandi, such as its depth (w) and curvature (q), carry some performer-specific information. We expect that a more detailed representation of the timing variation of ritardandi performances will reveal more of the individual style of pianists.

A more detailed analysis of the results is necessary to answer further questions. For instance, do all pianists have a quantifiable individual style or only some? Also, there is a need for alternative models of rubato (such as the model proposed by Repp [11]), to represent and study ritardandi in more detail.

Finally, we intend to relate our empirical findings with the musicological issue of the factors affecting music performances. Experiments supporting whether or not the structure of the piece and the feelings of the performer are present in renditions could be of interest to musicologists.

6. ACKNOWLEDGMENTS

This research is supported by the Austrian Research Fund FWF under grants P19349 and Z159 ('Wittgenstein Award'). M. Molina-Solana is supported by the Spanish Ministry of Education (FPU grant AP2007-02119).

7. REFERENCES

[1] Chris Cannam, Christian Landone, Mark Sandler, and Juan Pablo Bello. The sonic visualiser: A visualisation platform for semantic descriptors from musical signals. In *Proc. Seventh International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, October 8-12 2006.

[2] Anders Friberg. Generative rules for music performance: A formal description of a rule system. *Computer Music Journal*, 15(2):56–71, 1991.

[3] Anders Friberg and Johan Sundberg. Does musical performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *Journal of the Acoustical Society of America*, 105(3):1469–1484, 1999.

[4] Maarten Grachten and Gerhard Widmer. The kinematic rubato model as a means of studying final ritards across pieces and pianists. In *Proc. Sixth Sound and Music Computing Conference (SMC 2009)*, pages 173–178, Porto, Portugal, July 23-25 2009.

[5] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA Data Mining Software: An update. *SIGKDD Explorations*, 11(1):10–18, 2009.

[6] Henkjan Honing. When a good fit is not good enough: a case study on the final ritard. In *Proc. Eighth International Conference on Music Perception & Cognition (ICMPC8)*, pages 510–513, Evanston, IL, USA, August 3-7 2004.

[7] Erik Lindström, Patrik N. Juslin, Roberto Bresin, and Aaron Williamon. "expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, 20:23–47, 2003.

[8] Miguel Molina-Solana, Josep Lluis Arcos, and Emilia Gomez. Using expressive trends for identifying violin performers. In *Proc. Ninth Int. Conf. on Music Information Retrieval (ISMIR2008)*, pages 495–500, 2008.

[9] Miguel Molina-Solana and Maarten Grachten. Nature versus culture in ritardando performances. In *Proc. Sixth Conference on Interdisciplinary Musicology (CIM10)*, Sheffield, United Kingdom, July 23-24 2010.

[10] Bernhard Niedermayer. Non-negative matrix division for the automatic transcription of polyphonic music. In *Proc. Ninth International Conference on Music Information Retrieval (ISMIR 2008)*, Philadelphia, USA, September 14-18 2008.

[11] Bruno H. Repp. Diversity and commonality in music performance - An analysis of timing microstructure in Schumann's "Träumerei". *Journal of the Acoustical Society of America*, 92(5):2546–2568, 1992.

[12] John Rink, editor. *The Practice of Performance: Studies in Musical Interpretation*. Cambridge University Press, 1996.

[13] Efstathios Stamatatos and Gerhard Widmer. Automatic identification of music performers with learning ensembles. *Artificial Intelligence*, 165(1):37–56, 2005.

[14] Johan Sundberg and Violet Verrillo. On the anatomy of the retard: A study of timing in music. *Journal of the Acoustical Society of America*, 68(3):772–779, 1980.

[15] Renee Timmers, Richard Ashley, Peter Desain, Henkjan Honing, and W. Luke Windsor. Timing of ornaments in the theme of Beethoven's Paisiello Variations: Empirical data and a model. *Music Perception*, 20(1):3–33, 2002.

[16] Neil P. Todd. A computational model of rubato. *Contemporary Music Review*, 3(1):69–88, 1989.

[17] Gerhard Widmer, Sebastian Flossmann, and Maarten Grachten. YQX plays Chopin. *AI Magazine*, 30(3):35–48, 2009.

[18] W. Luke Windsor and E.F. Clarke. Expressive timing and dynamics in real and artificial musical performances: Using and algorithm as an analytical tool. *Music Perception*, 15(2):127–152, 1997.