

ANALYSIS OF SAXOPHONE PERFORMANCE FOR COMPUTER-ASSISTED TUTORING

Matthias Robine
SCRIME - LaBRI, Université Bordeaux 1
351, cours de la Libération,
F-33405 Talence cedex, France
robine@labri.fr

Graham Percival, and Mathieu Lagrange
University of Victoria
Finnerty Road
BC, Canada
[gperciva, lagrange]@uvic.ca

Abstract

We present a software tool which analyses the technical ability of saxophonists. In an academic practicing context, control of the air pressure is an important aspect which is strongly correlated to the academic level of the performer. We propose several methods for evaluating the performer by considering the evolution of the fundamental frequency during the performance of specially designed exercises. We show how those metrics can be computed in real time and discuss the pedagogical feedback ability of the proposed tool.

1. INTRODUCTION

We propose a new tool for music tutoring which provides easy feedback of the technical ability of a musician. Following the work presented in [1], we examine several low-level audio features in a manner which closely matches the evaluation of a professional teacher. We focus at first on the saxophone, but these methods may be applied to other wind instruments and even other instrument families such as bowed strings. The *Piano Tutor* [2] and *IMUTUS* [3, 4] / *VEMUS* [5] projects present some interactive music tuition multimedia systems for training users on traditional instruments. These systems include performance analysis, detection of the errors, and visual feedback designed for novices. We aim to provide such an application which can be used by all performers from beginners to experts, and which can classify them as a function of their technical performance.

We begin in Section 2 by explaining the motivation of our work. Section 3 presents our method of analysis and perceptual scaling. In Section 4 we explain the exercises which students are asked to perform, and Section 5 presents the results of those exercises. The interface of a new tool for technical evaluation of musicians is presented in Section 6. In Section 7 we discuss potential uses of this tool, and in Section 8 we discuss the perspective of a complete tuition tool.

2. EVALUATION OF THE MUSIC PERFORMANCE

In this study we focus on the technical part of the music performance and build a special protocol with adapted exercises and metrics to analyse them. This allows us to propose an interface which provides feedback of technical ability.

The technical evaluation of a performer is a vital portion of his education. To focus on improving their technical control, musicians often practice “technical exercises”: material which has little expressive value. These exercises are evaluated principally on the accuracy of intonation (pitch), rhythms (durations), and tempo (speed). According to the internationally famous saxophone teacher Jean-Marie Londeix [6], there is one type of exercise that is strongly correlated to the technical level of a saxophonist in an academic context: long tones. These exercises are often practiced by saxophonists, and consist principally in controlling the air pressure. The performance of these exercises reflects very well the capacity to control the instrument.

Such exercises also avoid problem faced by other systems such as *IMUTUS*: distinguishing between technical errors and deliberate expressive decisions. When performing a piece of music, a good performer may slightly alter the intonation or rhythm in order to achieve a certain expressive effect, while a bad performer will alter the intonation and rhythm unintentionally. Analysing the performance of long tones avoids this problem, since there is no room for deliberate expression. Moreover, performance of long tones on every note of the saxophone can accurately differentiate a wide range of technical levels, from novice to expert.

3. ANALYSIS OF THE PERFORMANCE

Most sounds produced by wind and bowed string instruments can be defined as pitched sounds. During the production of long tones, the sound can be conveniently decomposed into a sum of sinusoids whose frequencies are harmonically related. The sound can be expressed as

$$s(t) = \sum_{k=0}^{K-1} a_k \cos(2\pi f_0 k t + \Phi_k) \quad (1)$$

where f_0 is called the fundamental frequency. The harmonic relation expressed by this formula is tightly related to the perceptive notion of pitch. Another important notion is the perceived loudness which is related the acoustical intensity. Following the ANSI definition of timbre [7], these two notions are not related to the physical properties of instrument / instrumentalist combination, which ensures that proposed approach applies generally.

3.1. Fundamental Frequency Estimation

The estimation of the fundamental frequency from a monophonic signal is a widely studied area; see [8] for a review. As proposed by Rabiner [9], we take advantage of the autocorrelation function to efficiently estimate the f_0 .

3.2. Amplitude estimation

Since we are interested in the evolution of the amplitude over time, we integrate the signal power over small intervals:

$$a(i) = \sqrt{\sum_{n=i\Delta}^{i\Delta+I} s(n)^2} \quad (2)$$

where Δ is the hop size in samples. This parameter is estimated at the same frame rate as $f_0(i)$.

3.3. Perceptive Scaling

The evolution of the sound parameters reflects the performance of a musician while following constraints expressed perceptively. The Fechner law applies to every sensory organ and states that the sensation is proportional to the logarithm of the excitation. For example, we can consider a *crescendo / decrescendo* tone (one of the exercises; see Section 4). In Figure 3, we can see that the curve of the amplitude of an expert performance follows a linear evolution in the dB scale. It is therefore convenient to express the parameters using perceptive scales, such as in Figure 2. The amplitude is then expressed in deciBels. Similarly, the fundamental frequency is expressed using the Equivalent Regular Bandwidth (ERB).

4. EXERCISES

We consider the metrics proposed in [1] to extract some evaluation criteria from the performance of the saxophonist. Specifically, we evaluate the ability of the saxophonist to control his air pressure during the performance by considering the evolution of the pitch and the amplitude while playing simple notes such as in Figure 1.



Figure 1. Sample saxophone exercise.

4.1. Straight Tones

When performing a straight tone, the instrumentalist is requested to produce a sound with constant frequency and amplitude. To evaluate the quality of its performance, it is natural to consider the standard deviation of the observations. However, if the amplitude is very high, a slight deviation of the fundamental frequency will be perceptively important. On contrary, if the amplitude is very low, a major deviation

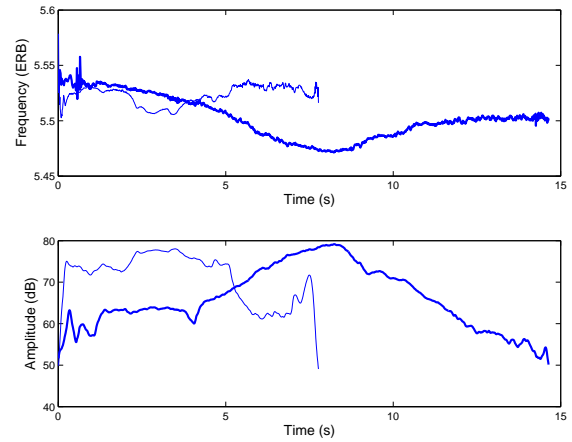


Figure 2. Pitch and amplitude vectors of a long tone *crescendo / decrescendo* played by two performers. In double solid line, the performer is an expert and in solid line, the performer is a mid-level student.

of the frequency parameter will not be perceptible. To cope with this issue, we consider a standard deviation of the observation vector weighted by the amplitude. This computation is performed in a sliding fashion, using fixed-size blocks to be able to compare several performances, see Figure 2.

4.2. Long Tones *crescendo / decrescendo*

When the instrumentalist performs a long tone *crescendo / decrescendo*, the amplitude should start from an amplitude close to 0, linearly increases to reach a maximum value M at index m , and linearly decreases to reach an amplitude close to 0. From the evolution of the amplitude of a partial A , we compute the piecewise linear evolution L and compare the analysed evolution against this ideal evolution. Two examples of the difference between A and its piecewise linear version L are shown in Figure 3.

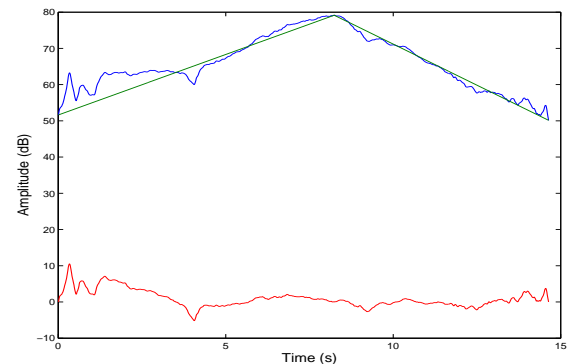


Figure 3. Amplitude vector A and piecewise linear vector L of a partial for two long tones *crescendo / decrescendo*. The difference between the two vectors is plotted on the bottom.

Amplitude results					
	α	p	mf	f	<>
experts (3)	21	124 (12)	108 (30)	197 (100)	111 (13)
confirmed (8)	19	100 (33)	100 (28)	100 (25)	100 (23)
mid (6)	20	55 (24)	61 (14)	67 (17)	57 (11)
elementary (5)	10	54 (13)	46 (5)	53 (12)	43 (12)
beginners (8)	9	50 (27)	39 (20)	47 (19)	35 (9)

Frequency results					
	α	p	mf	f	<>
experts (3)	21	146 (44)	100 (26)	136 (47)	127 (33)
confirmed (8)	19	100 (36)	100 (34)	100 (51)	100 (37)
mid (6)	20	48 (18)	57 (19)	63 (19)	62 (20)
elementary (5)	10	33 (12)	39 (14)	37 (8)	32 (3)
beginners (8)	9	35 (19)	32 (15)	34 (17)	40 (19)

Table 1. Results for note G. 5 classes (levels) of performers are represented with the number of performers in each class within parentheses. The results are scaled such that the confirmed class has 100 marks, and the standard deviation is within parentheses. The technical marks correspond to the supposed technical level as illustrated by the amplitude results for the straight tone mezzo forte.

5. RESULTS

The technique of an instrumentalist is principally evaluated according to the best performers in his class or music school. This is why we have chosen to give technical marks with respect to the best performances. Saxophonists were clustered in five classes according to their academic level validated by school teachers. We have chosen the *confirmed* class as mark reference (mark 100). It groups high level students and teachers, and contains 8 elements. Although the *expert* class could be a better reference due to the better marks obtained by its elements, it does not contain enough elements (only 3). We distinguish amplitude and frequency results. Since these values computed using the metrics introduced in [1] are errors, we consider as marks the inverse of the values multiplied by 100. These marks are then divided by the mean of the marks obtained by instrumentalists of the *confirmed* class.

Table 1 shows the results of the performance of long tones on the note G by saxophonists, where α is the multiplier coefficient of amplitude from *piano* straight tone to *forte* straight tone. *p*, *mf*, and *f* correspond to the straight tones played re-

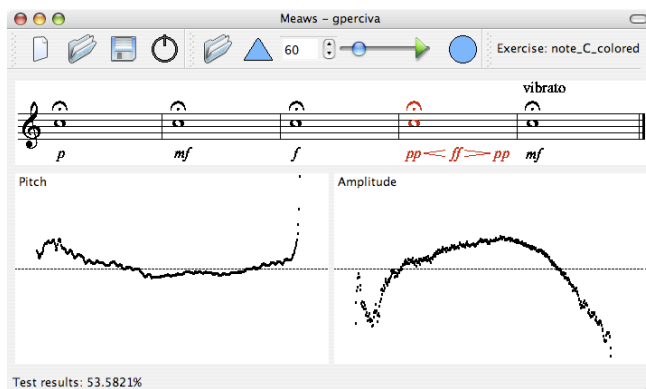


Figure 4. Meaws with pitch and amplitude displayed.

spectively with low, medium and high amplitude. The tone <> correspond to the long tone *crescendo* / *decrescendo*. This study comes after results presented recently in [1], computed using sinusoidal modeling technique instead of fundamental frequency estimation. We note that the use of perceptual scales leads to lower standard deviation for the results in the most cases than before. The marks reflect fairly the ranking of the performers since level classes are homogeneous with reasonable standard deviations.

6. INTERFACE

We provide a user-friendly tool for administering the tests: Meaws (Musician Evaluation and Audition for Winds and Strings). This software was written with Marsyas [10] and Trolltech Qt4 [®] and runs on Windows, MacOS X, and Linux.

After selecting a username and exercise, musicians perform the exercises. When each exercise has been recorded, Meaws analyzes the audio. Pitches and amplitude are extracted and expressed using perceptive scaling (Meaws represents pitches in MIDI note values instead of ERB). These pitch and amplitude values are displayed in the bottom portion of the screen as is shown in Figure 4.

In the next version of this software, this information will be stored for future viewing and – with the musician’s permission – uploaded to a central location. Personal data such as the musician’s name would be removed; the server would only retain user-supplied information such as the number of years that instrument was studied, whether the musician has studied any other instrument, and how the seriously the musician studies (casual amateur, university music student, etc.). This would let musicians from geographically remote locations compare their technical abilities.

7. CASE STUDIES

We now present two case studies where the tool we propose could be useful in a pedagogic manner. The first case concerns the regular evaluation of a single performer: a saxophonist could regularly use our tool to evaluate his technical performance. He can identify his points of weakness and can therefore choose to focus on particular technical exercises in

function of them. In some cases, he may choose to perform certain exercises because they explicitly stress his ability to perform that particular technical skill. But in many cases, simply concentrating on a particular skill will result in good improvements – no special exercises are necessary. The saxophonist can verify the results of his practice by measuring his progress with the same tool.

A second case is the use of our tool for a class of saxophonists. It can be used by the teacher every so often to make a evaluation of the whole class. He may use this objective tool to classify the pupils according to their technical ability and therefore to verify the correspondence with their academic level. The results give also some pedagogical information of the technical points he can choose for the class to work particularly before next evaluation, or perhaps even influence the choice of a particular technical book for the whole class.

Our new tool is particularly useful because the metrics may be clearly visualised. Musicians can easily grasp the relationship between the audio and the graphical display, along with their particular strengths or weaknesses.

8. DISCUSSION

We proposed in this paper a tool for visualizing and evaluating the performance of saxophonists and which produces results which are close to the judgment of a saxophone teacher. It takes advantages of several improvements made on top of the previous work proposed in [1]. Namely, it considers the evolution of the fundamental frequency and the global amplitude over time allows us to obtain more robust observations. Expressing those evolutions in some perceptual scales is also useful to be closer to the requirements expressed by the teachers. As far as the analysis is concerned, we plan to work more on the evaluation of vibrato tones. Since the pitch changes in a regular manner, we cannot evaluate such tones in the same way we evaluated the straight long tones.

In a pedagogical perspective, a lot of work may still be done to exploit the data extracted from the performance for meaningful feedback. For example, another way of representing results shown on Table 1 in a graphical manner is plotted on Figure 5. For a specific performer, we show his actual level compared to the mean performance achieved to the class in which he belongs, and the mean performance of the next (higher) class. This way, he can easily see which note(s) he must improve before he can move into the next class. Figure 5 is a possible visualisation for the results of Lilian with respect to the exercises he realised. p_f stands for the frequency result of a straight tone played *piano* and $\langle\langle_a$ for the amplitude result of a long tone *crescendo / decrescendo*. The grey zone is delimited by two lines: the bottom line is the mid technical level of the Lilian’s class and the top line represents the level for his upper class. The results of Lilian are the red points that can be evaluated with respect to the grey zone. We can note here that Lilian is better than the mid technical level of his class regarding the pitch control of the straight tone *mezzo forte* – in fact, Lilian’s result is better than the mid level of the upper class too. However, he must improve his control of the amplitude in *forte* long tones before he can join the next class.

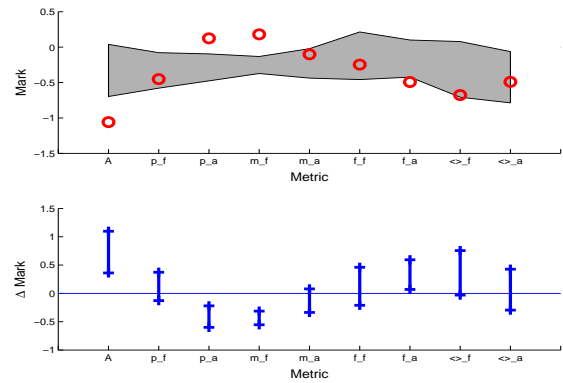


Figure 5. Visualisation of the results of Lilian with respect to the exercise practiced.

9. REFERENCES

- [1] Matthias Robine and Mathieu Lagrange, “Evaluation of the technical level of saxophone performers by considering the evolution of spectral parameters of the sound,” in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, Victoria, Canada, 2006, pp. 79–84.
- [2] Roger B. Dannenberg, Marta Sanchez, Annabelle Joseph, Robert Joseph, Ronald Saul, and Peter Capell, “Results from the piano tutor project,” in *Proceedings of the Fourth Biennial Arts and Technology Symposium*, 1993, pp. 143–150.
- [3] Dominique Fober, Stéphane Letz, Yann Orlarey, Anders Askenfeld, Kjetil Hansen, and Erwin Schoonderwaldt, “IMUTUS – an interactive music tuition system,” in *Proceedings of the Sound and Music Computing conference (SMC)*, Paris, 2004, pp. 97–103.
- [4] Erwin Schoonderwaldt, Anders Askenfeld, and Kjetil Hansen, “Design and implementation of automatic evaluation of recorder performance in IMUTUS,” in *Proceedings of the International Computer Music Conference (ICMC)*, Barcelona, Spain, 2005, pp. 97–103.
- [5] Dominique Fober, Stéphane Letz, and Yann Orlarey, “VEMUS - une école de musique européenne virtuelle,” in *Proceedings of the Journées d’Informatique Musicale*, Lyon, France, 2007, pp. 57–63, in french.
- [6] James C. Umble, *Jean-Marie Londeix: Master of the Modern Saxophone*, RONCORP Publications, Cherry Hills, USA, 2000.
- [7] American National Standard Institute, “USA Standard Acoustical Terminology,” 1960.
- [8] Alain de Chevigné, *Computational Auditory Scene Analysis. Principles, Algorithms, amid Applications*, chapter Multiple F0 Estimation, pp. 45–79, John Wiley and sons, Hoboken, New Jersey, 1991.
- [9] L. R. Rabiner, “On the use of autocorrelation analysis for pitch detection,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, pp. 24–33, 1977.
- [10] George Tzanetakis and Perry Cook, “Marsyas: a framework for audio analysis,” *Org. Sound*, vol. 4, no. 3, pp. 169–175, 2000.