

INDIRECT ACQUISITION OF FINGERINGS OF HARMONIC NOTES ON THE FLUTE

Corey Kereliuk[#], Bertrand Scherrer[#], Vincent Verfaillie^{#, b}, Philippe Depalle[#]
and Marcelo M. Wanderley^{#, b}

[#]Sound Processing and Control Laboratory

^bInput Devices and Music Interaction Laboratory
CIRMMT – McGill University – Montréal, QC, Canada

ABSTRACT

In this paper we present an approach for the indirect acquisition of specific fingerings that produce harmonic notes on the flute. We analyse both temporal and spectral characteristics of the attack of harmonic notes produced by specific control gestures involving fingering and potentially overblowing. We then show that it is possible to acquire this effect through signal analysis using a principal component analysis on spectral data. An 8-fold cross-validation showed this approach to be successful for a single performer playing isolated notes with *mf* dynamics.

1. INTRODUCTION

In order to create interactive music systems, it is necessary to acquire control information from a performer's actions (for example, fingering on the flute). One possibility is the use of *augmented instruments* constructed by attaching sensors to traditional instruments. For the flute, four examples can be found in the literature: the Hyper-Flute [1], the McGill air jet sensor [2], the LMA Flute [3] and the MIDI Flute [4]. Tab. 1, adapted from [5], compares these systems in terms of the variables they extract and, as is the concern of this paper, their ability to detect specific harmonic notes. This particular technique allows a flutist to play notes with the same pitch using different fingerings by changing the properties of the air jet [6].

For instance, a D6¹ can be obtained using D6 fingering as well as D5 and D4 fingerings by overblowing. Musicians usually refer to these cases as D6, 1st harmonic note of D5 and 3rd harmonic note of D4 respectively (*cf.* Tab. 2), while physicists would identify them as harmonic series of the 1st harmonic of D6, harmonic series of the 2nd harmonic of D5 (plus some sub-harmonics and partials) and harmonic series of the 4th harmonic of D4 (plus some sub-harmonics and partials). The score notation for these three configurations is given in Fig. 1.

None of the systems in Tab. 1 can detect this performance parameter even though it is of common use for flutists playing contemporary music, jazz and improvised music. One could combine the McGill air jet pressure sensor [2] with the LMA Flute [3] to detect both fingering and

overblowing. The intrusive nature of the pressure sensor, however, motivates the use of alternative methods. Herein lies the interest of an approach relying mainly on an analysis of the sound. This type of approach, known as *indirect acquisition* [7], has many benefits, the main one being that no alterations to the instrument are required (apart from the need of a microphone). On the other hand, this method requires complex algorithms which can be computationally intensive.

Device	Variables	Fing.	Air jet	Over.
MIDI flute [4]	all key pos. (on/ off)	all	—	—
LMA flute [3]	all key pos. (cont.), sound amplitude	all	—	—
Hyper-flute [1]	2 key pos. (cont.), inclination, flute rotation, distance to computer	2	—	—
McGill Air-Jet Sensor [2]	total air pressure around mouthpiece, flute weight around thumb	—	pressure	—

Table 1. Comparison of various augmented flutes according to the variables they extract, the possibility to detect fingerings, air jet and overblowing.

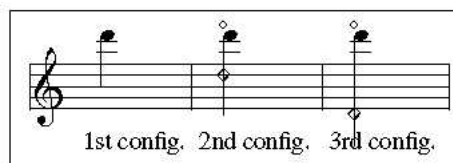


Figure 1. Score for three fingering configurations for D6. A diamond denotes the required fingering and a note with a circle above denotes the required pitch.

Viewpoint	Config. 1	Config. 2	Config. 3
Fingering	D6	D5	D4
Air Jet Pressure	normal	overblow	overblow
Score	D6	1st harmonic	3rd harmonic
F0	f_{D6}	$2 f_{D5} \approx f_{D6}$	$4 f_{D4} \approx f_{D6}$

Table 2. Naming convention for configurations in Fig. 1.

¹ We choose the convention where A4 corresponds to 440 Hz.

2. METHODOLOGY

We first present the data set we collected for indirect acquisition of fingerings of harmonic notes on the flute, and then discuss the choice of appropriate sound descriptors for future realtime analysis.

2.1. Data Collection

We recorded 20 isolated samples of each fingering (normal plus one or two harmonic notes) listed in Tab. 3, all with a normal attack and a *mf* dynamic. Thus a total of $18 \times 20 = 360$ samples were recorded. Depending on the point of view, the three configurations illustrated in Fig. 1 can be expressed with a different lexicon (*cf.* Tab. 2). In this paper, we refer to configurations 1, 2 and 3 as D6 with normal fingering, D6 with D5 fingering and D6 with D4 fingering. We focused on fingerings corresponding to pitches one or two octaves below the performed pitch (see [8] for diagrams of the different fingerings). All of our recordings came from a single performer and were made using an EarthWorks SR 77 microphone (positioned approximately 10 cm above the flute mouthpiece) and an Apogee Rosetta 800 sound card (16-bit, 44.1kHz) on a Mac G5.

Grp	Note	Fingerings	Grp	Note	Fing.
1	D#5	D#5, D#4	5	F6	F6, F5
2	D6	D6, D5, D4	6	F#6	F#6, F#5
3	D#6	D#6, D#5, D#4	7	G6	G6, G5
4	E6	E6, E5	8	G#6	G#6, G#5

Table 3. Experimental data set: note and fingerings used to play this note for each of the 8 groups.

2.2. Strategy for Analysing Fingerings of Harmonic Notes

When a flutist plays a harmonic note using a given fingering and overblowing, several changes appear in the sound. For example, the fundamental frequency is not exactly the same as with the normal fingering (*cf.* Tab. 2). Also, differences arise in the spectral envelope of both the harmonic and residual components of the sound, as well as in the temporal and spectral structures of the attack. We did not use the slight difference in fundamental frequency since experienced flutists can correct it, for instance, by adjusting the air flow and tilting the flute². Additionally, detecting changes in the spectral envelope of the harmonics and residual noise requires non-realtime analysis³. We decided to focus on the temporal and spectral structures of the attack which seem to allow for realtime detection of fingerings of harmonic notes; the temporal and spectral analyses are explained in the next two sections.

² “[T]he player must be sensitive to the subtleties of each fingering and must compensate appropriately for any inherent defects in intonation, dynamics, or tone quality”, [9] p. 143.

³ Audioscupt 2.8 β has some capabilities for realtime envelope and noise estimation, meaning this approach might be feasible.

3. TEMPORAL ANALYSIS OF THE ATTACK

3.1. Observations

We examined the evolution of the short-time energy of the signal during the attack via the RMS profile. The RMS was computed using Hann windows 4 times the length of the period of the lowest possible pitch (C4), with a hop size of 20 samples. In Fig. 2, we display the average and standard deviation of the RMS profiles computed for the fingering configurations of D6 presented in Fig. 1.

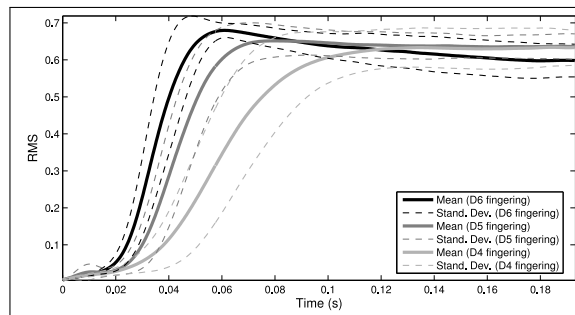


Figure 2. Mean and standard deviation of the RMS profile for three different fingerings of D6.

We noticed that the RMS profile increased faster for normal fingerings than for alternate fingerings. Also, the RMS rises faster for configuration 2 (D6 with D5 fingering) than for configuration 3 (D6 with D4 fingering).

3.2. Results

In order to quantify the previous observations, we collected the inflexion point of the attack profiles for all the sounds in our data set. Fig. 3 represents the inflexion points corresponding to three different fingerings for D6, together with the mean and standard deviation of the inflexion point for each fingering. The x-axis represents the time at which the inflexion occurs with respect to the onset. The y-axis represents the value of the slope at the inflexion point, the maximum slope of the RMS profile. It appears that, on average, the RMS profiles of the different fingerings tend to cluster in different regions of this 2-dimensional representation. Nevertheless, we observe a lot of overlap between the different fingerings. Therefore it appears that, while the RMS profile may not be perfectly suited for the identification of harmonic note fingerings, it still provides information about the attack (which could be useful in combination with spectral analyses for other applications such as attack type detection).

4. SPECTRAL ANALYSIS OF THE ATTACK

4.1. Observations

We carried out a spectral analysis on the attack portion of our recordings using Audioscupt [10]. Fig. 4 shows three spectrograms of the note D6 played with alternate fingerings. As the fingering changes from D6 to D5 and then to

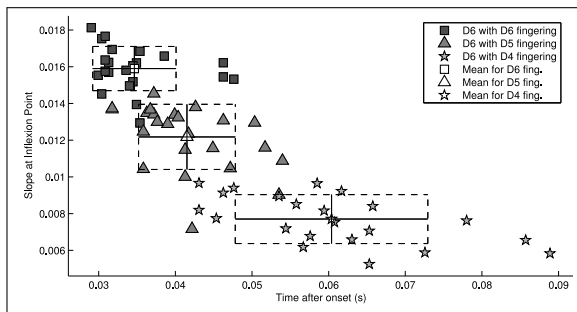


Figure 3. Slope vs. time of the inflexion point for D6.

D4, we notice the presence of energy at frequencies that are not in the harmonic series of D6. Moreover we see that these additional peaks emerge at different frequencies for each fingering. Quite logically, these secondary peaks correspond to the minima of the acoustic impedance [8]. In the case of D6 with D4 fingering these peaks seem to correspond to sub-harmonics of the fundamental. In the other two configurations we can see a combination of sub-harmonics and partials.

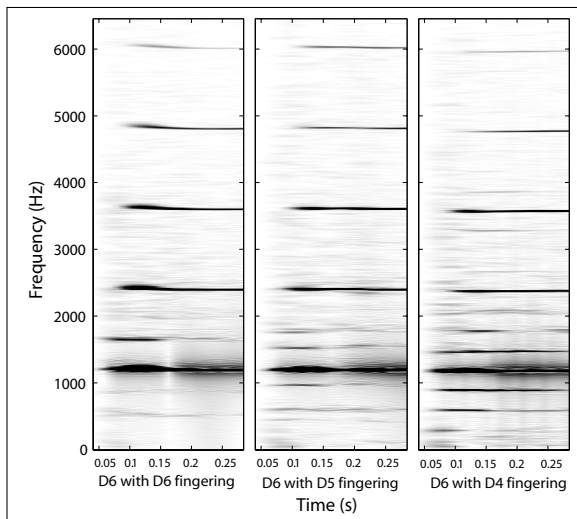


Figure 4. Spectrograms for the three configurations.

4.2. Feature Extraction

Based on these observations we decided to develop a classification scheme for fingerings of harmonic notes by using the power in frequency bands centered on sub-harmonic intervals. To simplify the feature extraction we assumed *a priori* knowledge of the pitch and onset time of each note. For a monophonic instrument like the flute, these parameters would be relatively easy to extract in real-time. We performed an FFT on each flute recording using a 2048-point Hamming window and 1024-point hop size. Using *a priori* pitch information, we extracted the maximum magnitude peaks in frequency bands of width $f_0/8$, centered on $k \cdot f_0/4$ where k is the rank of the sub-harmonic in the series ($k = [1, 2, \dots, 6]$). The power of these 6 peaks was then averaged in time over a small num-

ber of windows (typically 15, representing around 370 ms at 44.1 kHz) starting from the onset of the note.

4.3. Principal Component Analysis

We applied a principal component analysis to the 6-dimensional feature vector just described in order to isolate the elements responsible for the greatest variance.

The principal component analysis (PCA) technique decomposes a data set onto the eigenvectors⁴ of its covariance matrix [11]. A reduction in dimensionality can often be achieved using PCA since the first few eigenvectors (*principal components*) usually account for a high percentage of variance in the analysed data (in which case non-principal components may be discarded with minimal information loss). PCA can aid in the interpretation of data because it concentrates information previously spread across several interrelated variables (see [12] for its use in the analysis of guitar timbres). Since separating the dimensions of a data set according to variance will often cause clustering in the eigenspace, PCA can also be used as a classifier. Here we used it to classify harmonic notes fingerings similarly to how the embouchure pressure and attack types were identified on the clarinet in [13].

4.4. PCA Results

We performed a separate PCA on each configuration in our data set and, to verify our results, we performed an 8-fold cross-validation. Cross-validation is the most widely used method for obtaining unbiased estimates of model performance in machine learning applications [14].

Thus, for each group listed in Tab. 3 we generated 8 subsets, each with 14 training samples, and 2 test samples (generated using a random permutation). The 8-fold cross-validation lends confidence to our results since each test is performed 8 times using different training/test sets (instead of just a single time).

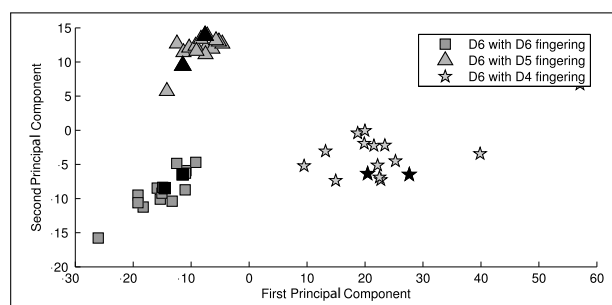


Figure 5. First two principal components for 3 configurations of D6 (gray: training data; black: test data).

Fig. 5 and 6 show typical results of the PCA. In each figure the gray shading represents training data and the black shading represents testing data. The first two principal components were found to account for over 90% of the

⁴ These eigenvectors define a linear transformation between the original feature space and an *eigenspace*.

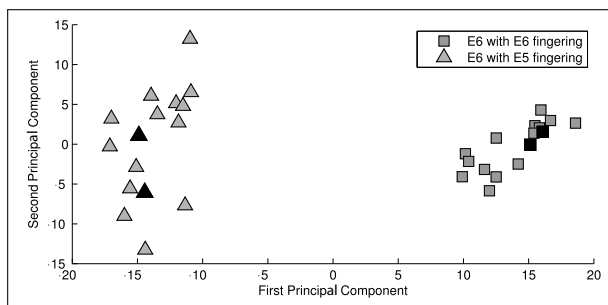


Figure 6. First two principal components for 2 configurations of E6 (gray: training data; black: test data).

variance in each group. Referring back to the figures, notice that each configuration forms a distinct cluster in the eigenspace. We used a Euclidean distance measure in order to classify the harmonic note fingerings used on each test sample. In other words, the squared distance between each test sample and the center of gravity of each training cluster was measured, and used to classify the test samples. We found that all of the test samples from our data set were correctly classified using this metric. After training the system off-line, the system was used in realtime (with a priori onset detection and pitch estimation).

5. CONCLUSION AND FUTURE WORK

This work has examined techniques for indirect acquisition of fingerings of harmonic notes on the flute⁵. Although the RMS profile is very useful for other analyses such as attack type classification, it was not sufficient for classification purposes. On the other hand, we have demonstrated that it is possible to identify which fingering was used to produce a given harmonic note by applying a PCA to the energy in frequency bands centered on sub-harmonic intervals. The results for a single performer were very robust, giving 100% correct classification on eight different notes, using an 8-fold cross-validation.

A natural next step would be to extend this study to multiple performers. Indeed, it remains to be seen whether a performer invariant system would be possible, or whether a PCA calibration would be required on a performer by performer basis. Another logical step would be to allow for truly realtime indirect acquisition, by combining realtime attack detection and pitch estimation, as well as by reducing the analysis window length. We would also like to test the performance of this technique in more realistic musical conditions, for example, on a series of articulated notes, as well as notes with variations in dynamics.

6. ACKNOWLEDGMENTS

This research is supported by grants from the FQRNT (*Fond Québécois pour la Recherche en Nature et Technologies*), the FQRSC (*Fond Québécois pour la Recherche*

⁵ see http://www.music.mcgill.ca/musictech/spcl/ia_flute_harmonics for additional material.

sur la Société et la Culture), the NSERC (Natural Sciences and Engineering Research Council of Canada), Québec's Ministry of Economic Development (PSIIRI grant), and CIRMMT (Centre for Interdisciplinary Research on Music, Media and Technology).

7. REFERENCES

- [1] C. Palacio-Quintin, "The hyper flute," in *Proc. Int. Conf. on New Interfaces for Musical Expression*, Montreal, Qc, Canada, May 2003, pp. 206–7.
- [2] A. D. Silva, M. M. Wanderley, and G. Scavone, "On the use of flute air jet as a musical control variable," in *Proc. Int. Conf. on New Interfaces for Musical Expression*, 2005.
- [3] S. Ystad and T. Voinier, "A virtually real flute," *Computer Music Journal*, vol. 25, no. 2, pp. 13–24, 2001.
- [4] D. Pousset, "La flute-midi, l'historique & quelques applications," Master's thesis, Université de Paris-Sorbonne, 1992.
- [5] M. M. Wanderley and E. Miranda, *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. A-R Editions, 2006, pp. 45–52.
- [6] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2nd ed. Springer, 2005, pp. 503–551.
- [7] M. M. Wanderley and P. Depalle, "Gestural control of sound synthesis," *Proc. IEEE*, vol. 92, no. 4, 2004.
- [8] J. Wolfe, "Flute acoustics." [Online]. Available: <http://www.phys.unsw.edu.au/music/flute/modernB/D6.html>
- [9] N. Toff, *The Flute Book: a complete guide for students and performers*, 2nd ed. New York: Oxford University Press, 1996.
- [10] N. Bogaards, A. Roebel, and X. Rodet, "Sound analysis and processing with Audiosculpt 2," in *Proc. Int. Computer Music Conf.*, 2004.
- [11] A. Daffertshofer, C. J. Lamoth, O. G. Meijer, and P. J. Beek, "PCA in studying coordination and variability: a tutorial." *Clin Biomech (Bristol, Avon)*, vol. 19, no. 4, pp. 415–28, 2004.
- [12] N. Orio, "The timbre space of the classical guitar and its relationship with the plucking techniques," in *Proc. Int. Computer Music Conf.*, 1999, pp. 391–4.
- [13] E. B. Egozy, "Deriving musical control features from a real-time timbre analysis of the clarinet," Master's thesis, MIT, 1995.
- [14] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*. Springer, 2001.