Highland Piping Ornament Recognition Using Dynamic Time Warping

Duncan W. H. Menzies
Centre for Digital Music
Queen Mary University of London
London, UK
d.w.h.menzies@qmul.ac.uk

Andrew P. McPherson
Centre for Digital Music
Queen Mary University of London
London, UK
a.mcpherson@qmul.ac.uk

ABSTRACT

This work uses a custom-built digital bagpipe chanter interface to assist in the process of learning the Great Highland Bagpipe (GHB). In this paper, a new algorithm is presented for the automatic recognition and evaluation of the various ornamentation techniques that are a central aspect of traditional Highland bagpipe music. The algorithm is evaluated alongside a previously published approach, and is shown to provide a significant improvement in performance. The ornament detection facility forms part of a complete hardware and software system for use in both tuition and solo practice situations, allowing details of ornamentation errors made by the player to be provided as visual and textual feedback. The system also incorporates new functionality for the identification and description of GHB fingering errors.

Author Keywords

Great Highland Bagpipe, computer assisted instruction, automatic ornament detection, dynamic time warping

ACM Classification

H.5.5 [Information Interfaces and Presentation] Sound and Music Computing, J.5 [Arts and Humanities] Performing Arts, H.5.2 [Information Interfaces and Presentation] User Interfaces

1. INTRODUCTION

The traditional repertoire and playing style of the Great Highland Bagpipe (described in [8]) are in many ways quite distinct from other genres of Western folk and classical music. In particular, the limited range of available pitches and absence of timbral or dynamic control have led to the development of a wide array of ornamentation techniques, which take the form of specific combinations of one or more short gracenotes. Such embellishments are rigorously and formally defined, and are an essential element of Highland piping practice. This paper presents a novel approach for the automatic detection and evaluation of piping ornamentation performed on a digital chanter interface (Figure 1), using an iterative pattern matching process based on Dynamic Time Warping (DTW). In Section 4, the algorithm is tested alongside a previous method from NIME 2012 [8].

Extending previous work on assistive graphical user interface (GUI) design for one-to-one piping lessons [9], the ornament recognition algorithm forms part of a complete hardware and software system to support the GHB learning process in both expert tuition and solo practice situations. Based on expert feedback obtained during a pilot study with an early version of the GUI, the system described in this paper also incorporates new functionality to highlight and describe piping-specific fingering errors.

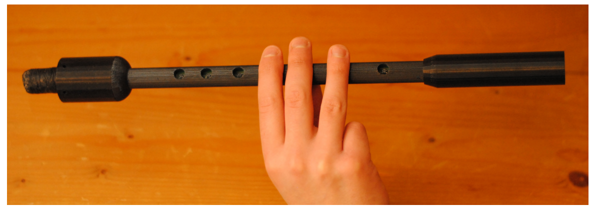


Figure 1: Digital chanter interface in 3D printed casing.

2. BACKGROUND

2.1 Previous Work

2.1.1 Digital Chanter Interface

This work uses a custom-built digital chanter interface [9], which employs infrared reflectance sensors mounted inside the holes of a cylindrical chanter shell to detect the continuous movements of the player's fingers. This provides a physical playing experience much closer to that of an acoustic chanter than the capacitive contacts used in commercially available electronic bagpipes such as the Redpipes[1]. The interface also incorporates an air pressure sensor [7, 12] in place of the chanter reed, allowing it to be connected to a traditional set of pipes. This enables the user to practice the breathing and bag pressure aspects of piping without the high volume levels associated with acoustic bagpipes.

2.1.2 Automatic GHB Ornament Detection

The formalised nature of GHB ornamentation makes it an ideal candidate area for automatic recognition and evaluation. The previous method presented in [8] employs a rule-based algorithm to detect potential ornaments and compare them to a series of templates, allowing errors in the execution of the movement to be identified. However, the accuracy of this approach is imperfect, particularly for student players whose technique can be inconsistent.

2.1.3 GUI Program for One-to-One Piping Lessons

The development of technological tools for music pedagogy is an active field of research. A significant proportion of existing work in this area concerns piano tuition. These projects use MIDI input from a digital keyboard to capture multiple aspects of the player's technique and generate illustrative visualisations, either as solo practice tools [2, 4,

[1]http://redpipes.eu/

13], or in order to assist a human tutor in describing their observations to the pupil [14]. Visualisations to compare student and expert recordings have also been investigated [6]. The prototype system presented in [9] first applied these concepts to GHB tuition, using sensor input from the digital chanter to enable the recording, playback, visualisation and comparison of instructor and pupil performances.

2.2 Musical Pattern Matching using DTW

Dynamic Time Warping (DTW) is a popular technique for musical pattern recognition. In [15], DTW is employed to recognise melodic fragments in MIDI keyboard performances of a range of music from Bach fugues to bebop. Paulus and Klapuri [10] use DTW to assess the similarity between temporal rhythmic patterns extracted directly from audio signals. [11] presents a DTW-based method for the classification of monophonic Greek traditional clarinet recordings according to 12 pre-defined reference patterns.

2.3 Ornament Detection in Other Genres

Detection of musical ornamentation in genres other than Highland piping has been addressed in several recent studies. Brown and Smaragdis [1] use independent component analysis to examine trills in piano and flute recordings, in order to compare trill rates between performances. Gomez et al. [5] present a method based on the Smith-Waterman algorithm to identify a range of pre-defined ornamentation techniques in *a cappella* flamenco pieces. [3] concerns the detection of ornamentation in Irish folk music. The system uses onset detection, audio segmentation and pitch recognition to find instances of single and multi-note ornaments.

3. METHODS

The goal of the system is to achieve robust detection of mistakes made by beginning players, and to provide meaningful feedback on the nature of these errors using the GUI. This section describes the implementation of these functions.

3.1 Ornament Recognition With DTW

This work extends the method presented in [8] to an iterative pattern matching approach using DTW. The software includes an XML file containing 64 ornament templates, detailing the pitches and approximate durations of each gracenote in the movement, and all permitted previous and subsequent notes (e.g. a *birl* must always end on low A).

The first step in the process is to identify any series of one or more short notes or *gracenotes* as a potential ornament. A gracenote is defined here as any note whose duration falls between two specified lengths L_{min} and L_{max}. However, it is often the case that the first note of certain ornaments (e.g. *throw on D*) is elongated for emphasis. For this reason, the algorithm begins detecting a possible ornament when any note shorter than a higher limit $L_{poss} > L_{max}$ is reached. If the next note is within the bounds of a normal gracenote, the longer first note is included in the sequence. Once a complete gracenote sequence is detected, it is marked as a potential ornament and compared to each of the templates.

Figure 2 depicts a flow chart of the ornament recognition process. Firstly, the complete gracenote sequence is considered. If both the melody notes immediately before and after the potential ornament are valid, then the DTW algorithm attempts to match the gracenote pitches to those in the template. DTW allows the duration of the performed gracenotes to differ greatly from the template and still be deemed correct. However, pitches that are either surplus to or missing from the template incur a penalty of one point for each millisecond sample that cannot be matched.

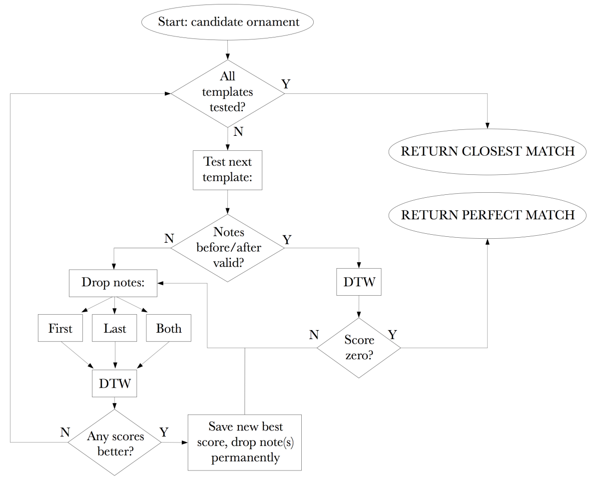
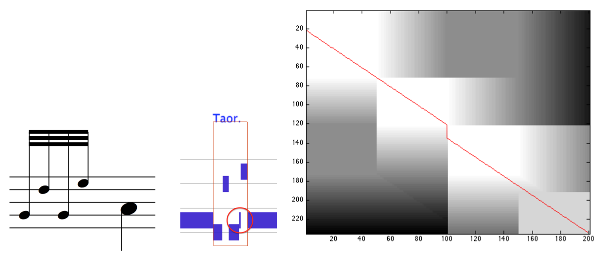


Figure 2: Ornament recognition flow chart.



(a) (b) (c)

Figure 3: (a) *Taorluath* ornament notation, (b) detected *taorluath* with erroneous note, (c) DTW plot.

Should the DTW comparison return a penalty score of zero, this ornament is designated a perfect match and the detection is complete. If the score is non-zero, or the previous and/or subsequent notes are invalid, the algorithm follows an iterative process, in which three alternative solutions are tested by dropping a gracenote from (a) the beginning of the sequence, (b) the end, and (c) both. If any of these alternatives provides a better score then this gracenote is dropped permanently (adding a fixed penalty to the score) and the iteration continues until no improvement is found, before repeating for the next template.

This approach allows embellishments performed with significant deviations in both pitch and timing to be identified accurately, even in fast tunes where the durations of melody notes can be comparable to ornament notes. Figure 3 shows a correctly detected *taorluath* movement containing an erroneous note (circled in red), and the resulting DTW plot.

3.2 GUI for Student Feedback

Figure 4 shows the GUI displaying a performance recorded with the digital chanter. The solid green bars represent notes on a stave, with durations indicated using a proportional notation similar to the familiar piano roll format. Detected embellishments are enclosed in rectangular boxes, below which the ornament name is written. The red box in the top right of the display indicates that the *tachum* movement contains an error (an extra note between the G and B gracenotes) which has been highlighted by the system with a red circle. Detailed feedback on the execution of a particular embellishment is available in the form of a pop-up text window by clicking on the ornament in the display.

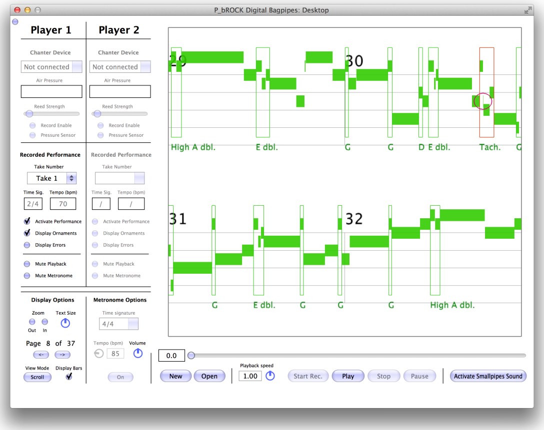


Figure 4: GUI showing ornaments detected in performance. Red box (top right) indicates an error.

3.3 False Fingering Detection

False fingering refers to the act of playing a note on the top hand of the chanter without executing the proper bottom hand fingering, and is seen as a serious technical flaw in traditional piping circles. Nonetheless, to the inexperienced player, the comparatively subtle differences in sound between correct and incorrect fingerings can be difficult to discern. During the pilot study described in [9], the instructor suggested that the ability to highlight instances of false fingering would be a valuable addition to the system.

Since there is only one correct fingering for each of the nine notes in the traditional piping scale, this facility can be implemented conveniently using a simple lookup table approach, in which the correct state of the eight chanter holes is stored for each possible pitch. The GUI includes the option to highlight any note (or section thereof) which is fingered incorrectly in red. Details of the false fingering can be displayed in a text window (Figure 5), allowing the user to identify, recreate and rectify the error.

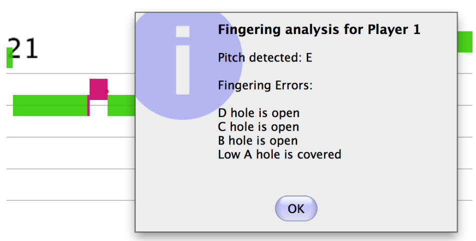


Figure 5: False fingering feedback.

It should be noted that the concept of false fingering applies only to melody notes; gracenotes in an embellishment are usually performed with one finger at a time, and hence do not correspond to the correct fingering for a given pitch. It is therefore a prerequisite to meaningful false fingering detection that the ornament recognition algorithm performs effectively, to avoid labelling gracenotes as false fingerings.

4. EVALUATION

This section describes a quantitative evaluation of the performance of the DTW ornament detection algorithm (hereafter referred to as OR_{dtw}), alongside the original approach [8] presented at NIME 2012 (termed OR_{2012}). The algorithms were tested on a dataset of 30 performances recorded using the digital chanter interface: a first set comprised of 5 performances each by 3 professional bagpipers, and a second group of 15 recordings made by 11 piping students (1 or 2 pieces by each player). The students were aged 11-17 years, and had been learning the bagpipes for 1-4 years.

The two algorithms were tested using identical settings for gracenote sequence detection: L_{min} = 15ms, L_{max} = 100ms and L_{poss} = 170ms. These values were determined empirically during the development of the system, and were not altered at any point during the evaluation.

4.1 Annotation of Ground Truth Ornaments

Prior to the evaluation, the recordings were manually annotated to provide a ground truth reference for the type and location of each of the embellishments attempted by the player. In some cases, the incorrect execution of one ornament can manifest itself as a slightly distorted instance of a different technique. The aim of the algorithm is to identify ornaments, however poorly executed, without any prior knowledge of the performer's intention. For this reason, the criterion for annotation of the ground truth ornaments was whether or not an experienced human listener would be able to determine from the context which ornament was attempted, without necessarily knowing the correct ornamentation of the tune. Over all 30 performances, a total of 3629 ground truth ornament annotations were made.

4.2 Results

For each algorithm, the detected embellishments were compared to the ground truth annotations, giving a number N_C of correct matches in each case. The algorithms were evaluated for *precision* P and *recall* R, which are given by:

$$P = \frac{N_C}{N_D} \quad \text{and} \quad R = \frac{N_C}{N_A}$$

where N_D is the total number of ornaments detected, and N_A is the number of ground truth annotations. The P and R values can then be combined into a single *F-measure* statistic by which to compare the algorithms:

$$F = 2\left(\frac{PR}{P+R}\right)$$

The P, R and F values obtained by the two algorithms are presented in Table 1. Across all 30 test recordings, the OR_{dtw} algorithm achieved an improvement of 6.8% over OR_{2012}. To assess the statistical significance of the results obtained, paired-sample t-tests were computed using the F-measures for each recording (Table 2). In all categories, the improvement in performance was found to reject the null hypothesis at a significance level of 99.9% ($p < 0.001$).

4.3 Discussion

For the OR_{dtw} algorithm to be valuable to students, it must provide an accurate account of which ornaments were performed, and which contained mistakes. Of the 1240 ornaments detected in the student recordings, 249 (20%) were found to contain errors. 209 (84%) of these 249 ornaments were correctly matched to the ground truths. This is an encouraging result; however, there are still instances in which the player's technique leads to incorrect recognition.

Ornament recognition errors generally occur for one of two reasons. The first takes place in the gracenote sequence detection step, when the duration of one or more notes in a performed embellishment falls outwith the predefined bounds. In this case, single note ornaments are ignored entirely, and multi-note ornaments are often identified as some combination of their constituent gracenotes.

The second cause of mis-identification is that poor execution can result in the detected sequence more closely re-

Group	Algorithm	No. Annotations N_A	No. Detected N_D	No. Correct N_C	P	R	F-Measure
Experts	OR_{dtw}	2340	2274	2226	0.979	0.951	0.9649
	OR_{2012}		2276	2047	0.899	0.8748	0.887
Students	OR_{dtw}	1289	1240	1094	0.882	0.849	0.8652
	OR_{2012}		1296	1055	0.814	0.818	0.816
All	OR_{dtw}	3629	3514	3320	0.9448	0.9149	0.930
	OR_{2012}		3572	3102	0.868	0.8548	0.862

Table 1: Comparison of ornament detection algorithms across all pieces in expert and student groups.

Group	Num Pieces	t-value	p-value
Experts	15	4.5614*	4.4394^{-4}
Students	15	4.3431*	6.7492^{-4}
All	30	6.3840*	5.5854^{-7}

Table 2: Paired-sample t-tests for performance of OR_{dtw} and OR_{2012} ornament detection algorithms ($*p < 0.001$).

sembling a different ornament. For this reason, the process of annotating ground truth ornaments involved some ambiguity, particularly for the student performances. The annotations were made based on the contextual knowledge an expert piper would use to discern the player's intention. This high-level understanding of the wider context of the piece is not implemented in the detection algorithm itself.

5. OBSERVATIONS FROM USER STUDY

The complete system was used in an extensive user study with an experienced piping instructor and 17 students, a full discussion of which is beyond the scope of this paper. This section highlights some observations that are of particular relevance to the developments described above.

Firstly, it was observed that the instructor generally chose not to consult the textual description facility when discussing the students' execution of ornaments, opting instead to provide his own feedback based on the sound of the performed embellishment, and its appearance on the GUI. This is unsurprising, as the system can not provide the level of detail of an expert tutor; indeed, this function was developed specifically for solo practice to avoid the introduction of bad habits in the absence of an instructor's supervision.

In contrast, the tutor made frequent use of the text window to recreate and describe instances of false fingering, many of which had not been identified during the original performance. Moreover, the instructor quickly became adept at distinguishing genuine false fingerings (i.e. incorrectly performed melody notes) from gracenotes which were too long to be detected by the algorithm, and were hence labelled as false fingerings. In such cases, the instructor was able to provide feedback tailored to the student's level of experience. Beginning players were advised that it is better to exaggerate than to rush movements, and hence they should not consider the mis-detection as a mistake in their playing at this stage, while more advanced students were simply instructed that the gracenote was "too long".

6. CONCLUSION

The system described in this paper is an example of a digital interface designed to connect to a long established and highly formalised musical tradition. The success of such systems is dependent not only on practical considerations such as appropriate sensing and mapping strategies, but also, critically, on ensuring that the particular constraints and implications of the cultural context are inherent in the design. By integrating support for the ornamentation and fingering techniques that are an integral aspect of traditional Highland piping practice, this work demonstrates how digital technologies can provide a meaningful contribution to even the most conservative musical genres.

Acknowledgements

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) as part of the Centre for Doctoral Training in Media and Arts Technology at QMUL.

7. REFERENCES

[1] J. C. Brown and P. Smaragdis. Independent Component Analysis for Automatic Note Extraction from Musical Trills. *JASA*, 115(5):2295–2306, 2004.

[2] R. B. Dannenberg, M. Sanchez, A. Joseph, R. Joseph, R. Saul, and P. Capell. Results from the Piano Tutor Project. In *Proc. Fourth Biennial Arts and Technology Symposium*, Mar 1993.

[3] M. Gainza and E. Coyle. Automating Ornamentation Transcription. In *Proc. ICASSP*, April 2007.

[4] W. Goebl and G. Widmer. Unobtrusive Practice Tools for Pianists. In *Proc. Music Perception and Cognition*, August 2006.

[5] F. Gómez, A. Pikrakis, J. Mora, J. Díaz-Báñez, E. Gómez, and F. Escobar. Automatic Detection of Ornamentation in Flamenco. In *Int. Workshop on Machine Learning and Music*, 2011.

[6] T. Knight, N. Boulliot, and J. R. Cooperstock. Visualization feedback for musical ensemble practice: A case study on phrase articulation and dynamics. In *IS&T/SPIE Electronic Imaging*, 2012.

[7] D. Menzies and D. Howard. The CyberWhistle, An Instrument For Live Performance. In *Colloquium on Musical Informatics XII*, Gorizia, Italy, Sept 1998.

[8] D. W. H. Menzies and A. P. McPherson. An Electronic Bagpipe Chanter for Automatic Recognition of Highland Piping Ornamentation. In *Proc. NIME*, Ann Arbor, MI, USA, May 2012.

[9] D. W. H. Menzies and A. P. McPherson. A Digital Bagpipe Chanter System to Assist in One-to-One Piping Tuition. In *Proc. SMC*, July 2013.

[10] J. Paulus and A. Klapuri. Measuring the Similarity of Rhythmic Patterns. In *Proc. ISMIR*, October 2002.

[11] A. Pikrakis, S. Theodoridis, and D. Kamarotos. Recognition of Isolated Musical Patterns using Context Dependent DTW. *IEEE Trans. Speech and Audio Processing*, 11(3):175–183, 2003.

[12] G. P. Scavone. THE PIPE: Explorations with Breath Control. In *Proc. NIME*, May 2003.

[13] S. Shirmohammadi, A. Khanafar, and G. Comeau. MIDIATOR: A Tool for Analysing Students' Piano Performance. *Revue de Recherche en Éducation Musicale*, 24:35–48, 2006.

[14] S. W. Smoliar, J. A. Waterworth, and P. R. Kellock. pianoFORTE: A System for Piano Education Beyond Notation Literacy. In *Proc. ACM MM*, Nov 1995.

[15] D. R. Stammen and B. Pennycook. Real-time Recognition of Melodic Fragments Using the Dynamic Timewarp Agorithm. In *Proc. ICMC*, 1993.