The Annotation of Traditional Irish Dance Music using MATT2 and TANSEY

Bryan Duggan [1], Brendan O'Shea [2], Mikel Gainza [3], Padraig Cunningham [4]
[1,2,3] Dublin Institute of Technology, Kevin St, Dublin 8
{bryan.duggan, brendan.oshea, mikel.gainza}@dit.ie
[4] UCD School of Computer Science and Informatics
padraig.cunningham@ucd.ie

Abstract

Currents estimates put the canon of traditional Irish dance tunes at least 7,000 compositions. Given this diversity, a common problem faced by musicians and ethnomusicologists is identifying tunes from recordings. This is evident even in the number of commercial recordings whose title is gan ainm (without name). This work attempts to solve this problem by developing a Content Based Music Information Retrieval (CBMIR) System adapted to the characteristics of traditional Irish music. A system is presented called MATT2 (Machine Annotation of Traditional Tunes) whose primary goal is to annotate recordings of traditional Irish dance music with useful meta-data including tune names. MATT2 incorporates a number of novel algorithms for transcription of traditional music and for adapting melodic similarity measures to the creativity and style present in the playing of traditional music. It incorporates a new algorithm for filtering ornamentation notes and accommodating "the long note" in traditional music called Ornamentation Filtering using Adaptive Histograms (OFAH). A new algorithm is presented called TANSEY (Turn ANnotation from SEts using SimilaritY profiles) that annotates sets of tunes played segue as is the custom in traditional Irish dance music. The work presented is validated in experiments using 130 real-world field recordings of traditional music from sessions, classes, concerts and commercial recordings. Test audio includes solo and ensemble playing on a variety of instruments recorded in real-world settings such as noisy public sessions. Results are reported using standard measure from the field of information retrieval (IR) including accuracy, error, precision and recall and the system is compared to alternative approaches for CBMIR common in the literature.

Keywords: Music Information Retrieval, Traditional Irish Dance Music

1 Introduction

In common with the folk music of many countries, repertoire in Irish traditional music is primarily acquired aurally. Musicians playing Irish music learn tunes by hearing the tune played by fellow musicians in sessions, classes, workshops and from commercial recordings [1]. Similarly, organisations such as Na Píobairí Uilleann, Comhaltas Ceoltóirí Éireann and the Irish Traditional Music Archive have been acquiring field recordings of traditional music for over 60 years and these organisations now possess many thousands of hours of recordings in a variety of formats and on a variety of different media.

In order for these archives to be useful, they must be annotated with appropriate metadata, such as tune names, time signatures, key signatures and instruments. The main goal of this PhD thesis is to develop algorithms for *automatically* annotating field recordings of monophonic traditional dance music. Several recent papers address the necessity of developing MIR (Music Information Retrieval) systems that are adapted to the specific requirements of ethnic music and also to the needs of musicologists studying ethnic music [4-6]. This work presents the first attempt to develop a Content Based Music Information Retrieval (CBMIR) system adapted to the specific characteristics of Irish traditional dance music.

The algorithms and systems proposed in this work take account of characteristics as slow onset times in woodwind instruments such as the concert flute and the tin whistle, the playing of ornamentation, variation, phrasing, reversing and the playing of tunes *segue* in sets. The work also takes advantage of the ABC music notation language, which has been developed especially for the transcription of Western traditional music. There exist over 7,000 traditional Irish, Scots and Breton tunes freely available in ABC format from public databases [7-9]. Although this work focuses on traditional Irish music, it is hoped that the techniques proposed can be generalised to other genres and instruments. This work is at an advanced stage, with an expected submission at the end of 2008.

2 Research Question

This project is concerned with traditional dance music, as can be played on the concert flute and tin whistle. The most common forms of dance music are *reels, double jigs* and *hornpipes*. Other tune types include *marches, set dances, polkas, mazurkas, slip jigs, single jigs and reels, flings, highlands, scottisches, barn dances, strathspeys* and *waltzes* [12]. These forms differ in time signature, tempo and structure. Tunes are typically arranged into *sets*. A set consists of a number of tunes (commonly 2, 3 or 4) played sequentially. Each tune in a set is usually repeated 2 or 3 times [10]. The origin of many sets of tunes is unknown and musicians often compile new sets "on the fly" in traditional music sessions. Instruments used to play traditional dance music include the tin whistle, fiddle (violin), uilleann (elbow) pipes, accordion, concertina, harp and the banjo [1].

This work focuses on developing algorithms for content based music information retrieval that specifically address individual creativity and style in queries to an MIR system. In traditional music, an experienced musician rarely plays the same tune twice. Ornamentation plays a key role in the individual interpretation of traditional Irish music. The usage of ornamentation is highly personal and large variations exist in the employment of ornamentation from region to region, instrument to instrument and from musician to musician. From this brief introduction to the domain of traditional Irish dance music P1-P10 in Table 1 can be derived. These give the main challenges in MIR in the domain of traditional dance music.

P1	Support for traditional instruments
P2	Commonly used keys & modes
P3	Reversing
P4	C, C# similarity
P5	Phrasing
P6	Transposition in tinwhistles
P7	Ornamentation
P8	The long note
P9	Tempo deviation
P10	The playing of tunes in sets

Table 1: Summary of the main challenges in performing CBMIR on traditional music sources

The system should support the input of queries played in traditional instruments such as the flute, tinwhistle, fiddle and uilleann pipes or alternatively lilted queries (P1, P6). Irish traditional music is usually played legato and so any transcription system needs to support legato note onsets. Stylistic variation (P3, P5, P8, P9) is very common even within the same performance of a tune and therefore any system developed needs to be robust to melodic variations. Therefore it can be concluded that tunes played with various interpretations and ornamentation should be considered the same. Where a musician makes use of reversing (transposing by 1 octave) (P3), these should be considered the same as melodies played without reversing. This will affect melodic similarity measures that depend on exact matches.

The collection of tunes into sets played *segue* creates segmentation problems (P10). An input query to a CBMIR system for traditional music may consist of a phrase from any part of a melody, an entire melody, a entire melody played multiple times or multiple melodies played multiple times without an interval, in the same time signature and often in the same key. The challenge therefore is in segmenting a query appropriately so that each individual tune in a set can be annotated correctly.

Given the dominance of concert pitch instruments used to play traditional dance music, transposition invariance is only required for the keys and modes playable on the concert pitch instruments (P2). Polyphony also does not need to be considered as, when Irish traditional music is played in unison the same melody is simultaneously played by all the performers.

3 Related Work

Interestingly, annotation systems such as that proposed in this work do not seem to form part of the literature. The work proposed seems to fall between 2 types of MIR system. It is similar to audio fingerprinting MIR systems such as those described in [15-17] in the sense that the aim of the work is to annotate a digital recording. However these systems work entirely in the signals domain. Their aim is to identify a digital recording as being an instance of another digital recording. These systems create hashes of recordings known as *audio fingerprints* in order to decrease computational complexity and minimise memory usage. In these systems 2 versions of the same piece of music will be annotated differently. In this work, the aim is to make different interpretations of the same piece of music be annotated identically.

Several papers [18-20] report on the difficulty of extracting performance data from digital signals and hence used either MIDI data, data captured from custom instruments or on screen representations of instruments which a user must "play" using the mouse in the query by-example paradigm. It can therefore be concluded that there are additional challenges in developing MIR systems that work on audio from real instruments.

It seems reasonable to understand the aim of a typical Query by Humming (QBH) system to be to try and find a melody from a corpus that is similar to a hummed query. These approaches require adaptation to address P1-P10 from Table 1. The proposed approach maximizes similarity between pieces of music played on different instruments, in different tempos and most importantly in different regional and individual styles and *ground truth* transcriptions of the musical pieces. Gross contour representations of melodies used in many MIR systems result in too many mismatches [21-23] and will present particular problems when queries contain ornamentation.

4 Proposed Solution

MATT2 (Machine Annotation of Traditional Tunes) represents a system that addresses P1-P9, the main problems in performing MIR on traditional Irish dance music identified in Table 1. P10 is addressed using the proposed Turn ANnotation in SEts using SimilaritY profiles (TANSEY) algorithm.

First, using Gainza's onset detection algorithm [24], a frequency domain pitch detection algorithm, and a pitch spelling algorithm based on [25]'s observations about the transcription of traditional Irish dance music, tunes are transcribed to simplified strings of the reduced alphabet of the ABC music notation language. The transcriptions are *normalised* to take account of various transformations that can occur in the interpretation of traditional music, such as ornamentation, reversing and phrasing. An algorithm known as Ornamentation Filtering using Adaptive Histograms (OFAH) is proposed to normalise queries played with ornamentation. Once a transcription is made, the system compares it against a corpus of human made ground truth tune transcriptions of tunes. The corpus used is in the ABC language which has the advantage of being based on ASCII text and so tunes in ABC can be easily processed and analysed using algorithms for textual information retrieval. The transcriptions used are normalised in a 5 stage process, before comparison, to remove ornamentation, to compensate for reversing and to expand the tunes as they would be typically played by a musician interpreting the tune. [26]'s edit distance is used to calculate melodic dissimilarity, with [27]'s variation which allows for searching for substrings. Using the approach proposed, a high success rate is reported for test audio consisting of long and short phrases of music, incipits and extracts from the middle of tunes, solo and ensemble playing, (with up to 10 musicians on various traditional instruments), field recordings from concerts, informal pub sessions and badly degraded archive recordings.

MATT2 solves the problem of annotating single tunes, however in traditional music tunes are rarely played singly. More commonly tunes are played in groups of at least 2 tunes known as a *set* of tunes [10,28]. Typically each tune in the set is played twice or 3 times before musicians advance to the subsequent tune in the set. A *turn* in a set represents the point the time when a repetition of a tune begins or a second or subsequent tune is introduced. As tunes in sets are always in the same time signature and often in the same key, the challenge therefore is in segmenting sets into tunes and repetitions. The TANSEY algorithm addresses this problem by making use of melodic similarity profiles calculated using [27]'s variant of the edit distance string matching algorithm which searches for strings in substrings of a target string. TANSEY can identify the start and end of each repetition of a tune, can count the repetitions and can identify the title and associated metadata associated with each tune in a set. Experimental results are presented using precision and recall scores for the algorithm which establish its effectiveness.

5 Validation

To establish the effectiveness of ornamentation filtering and the related style compensation approaches 3 approaches are compared:

T1: A baseline, edit distance matching algorithm based on melodic contours. This approach is common in the literature and is similar to the approaches employed by [29,31,30,32,33,23,34,18].

T2: A transposition invariant edit distance matching between the corpus and transcribed queries was tested in the second scenario. For this experiment, the style compensation algorithms (OFAH, phrasing compensation, reversing and lengthening) are not employed. This was carried out to evaluate the specific effect of these algorithms. This experiment might be considered similar to the SEMEX system described in [19] (although technically SEMEX works entirely on symbols and does not have a transcription system).

T3: The complete MATT2 system. For each experiment, the system annotated the test audio. In this experiment, the style accomodation algorithms of OFAH and ABC normalisation are employed.

2 sets of audio test data are used, 50 whole tunes (WT) and 50 short excerpts (E) from tunes. Deliberately challenging audio is included, including degraded archive recordings, flute duets, flute and fiddle duets, fiddle solos, sessions with ensembles of up to 10 musicians and ensemble playing in unusual keys with background noise.

McNemar's test is used to determine statistical significance [35]. Contingency tables and X^2 (chi squared) scores are computer for (T1, T3), and (T2, T3)

To evaluate the effectiveness of the TANSEY algorithm precision and recall scores are calculated for 30 pieces of test audio containing 1 hour 27 minutes of audio with 141 turns. In carrying out this experiment, the aim is to establish if MATT2 can correctly identify the names of the tunes and if TANSEY can figure out the timings of turns. To establish a *ground truth* for the experiment, a human domain expert manually annotates the turns in the sets of tunes. A true positive *TP* is a turn annotated by the system that agrees with a human annotated turn within a threshold timeframe *tf*. The threshold used in this experiment is +/- 2 seconds. Both MATT2 and TANSEY are developed in Java.

References

[1] G. Wallis and S. Wilson, *The Rough Guide to Irish Music*, London: Rough Guides, 2001.

[2] T. Kearns and B. Taylor, *A Touchstone for the Tradition - The Willie Clancy Summer School*, 2003.

[3] Willie Clancy Summer School , "2008 Willie Clancy Summer School," 2008; http://www.setdancingnews.net/wcss/wcsst.htm.

[4] S. Doraisamy, H. Adnan, and N. Norowi, "Towards a MIR System for Malaysian Music," *7th International Conference on Music Information Retrieval, Victoria, Canada, 8 - 12 October 2006*, 2006.

[5] K. Jensen, J. Xu, and M. Zachariasen, "Rhythm-based segmentation of popular chinese music," *Proceedings of 6th International Conference on Music Information Retrieval (ISMIR 05)*, 2005, pp. 374-380.

[6] A. Nesbit, L. Hollenberg, and A. Senyard, "Towards Automatic Transcription of Australian Aboriginal Music," *5th International Conference on Musical Information Retrieval, Barcelona, Spain October 10-14, 2004*, 2004.

[7] thesession.org, "The session.org Forums," 2007; http://www.thesession.org.

[8] H. Norbeck, "ABC Tunes," 2007; http://www.norbeck.nu/abc/index.html,

[9] J. Chambers, "O' Neills Books," 2007; http://trillian.mit.edu/~jc/music/book/oneills/.

[10] F. Vallely, *The Companion to Irish Traditional Music*, New York University Press, 1999.

[11] G. Larson, *The Essential Guide to Irish Flute and Tin Whistle*, Mel Bay Publications, Inc., 2003.

[12] B. Breathnach, "Ceol Rince na hÉireann Cuid V [Dance Music of Ireland] Vol V," 1999.

[13] N. Keegan, *The Words of Traditional Flute Style*, MPhil Thesis, University College Cork, Music Department, 1992.

[14] M. Gainza, *Music Transcription within Irish Traditional Music*, PhD Thesis, Dublin Institute of Technology, Faculty of Engineering, 2006.

[15] P. Cano et al., "A Review of Audio Fingerprinting," *The Journal of VLSI Signal Processing*, vol. 41, 2005, pp. 271-284.

[16] MusicBrainz, "Welcome to MusicBrainz! - MusicBrainz," 2008; http://musicbrainz.org/.

[17] Shazam, "Shazam - The amazing music discovery engine. Join our Community," 2008; http://www.shazam.com/music/portal.

[18] L. Prechelt and R. Typke, “An interface for melody input,” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 8, 2001, pp. 133-149.
[19] K. Lemstrom and S. Perttu, “SEMEX-An Efficient Music Retrieval Prototype,” *First International Symposium on Music Information Retrieval (ISMIR)*, 2000.
[20] H. Hoos, “GUIDO/MIR an Experimental Musical Information Retrieval System based on GUIDO Music Notation,” *Symposium on Music Information Retrieval: ISMIR*, 2001, pp. 41-50.
[21] J. Schlichte, “Der automatische Vergleich von 83 243 Musikincipits aus der RISM-Datenbank: Ergebnisse-Nutzen-Perspektiven,” *Fontes artis musicae*, vol. 37, 1990, pp. 35-46.
[22] N. Adams, M. Bartsch, and G. Wakefield, “Coding of sung queries for music information retrieval,” *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.*, 2003, pp. 139-142.
[23] L. Lu, H. You, and H. Zhang, “A new approach to query by humming in music retrieval,” *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2001.
[24] M. Gainza, E. Coyle, and B. Lawler, “Onset Detection Using Comb Filters,” *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY: 2005.
[25] B. Breathnach, “Ceol Rince na hÉireann Cuid III [Dance Music of Ireland] Vol III,” 1985.
[26] V. Levenshtein, “Binary Codes Capable of Correcting Deletions, Insertions and Reversals,” *Soviet Physics Doklady*, vol. 10, 1966, p. 707.
[27] G. Navarro and M. Raffinot, *Flexible Pattern Matching in Strings: Practical On-Line Search Algorithms for Texts and Biological Sequences*, Cambridge University Press, 2002.
[28] B. Duggan, B. O'Shea, and P. Cunningham, “A System for Automatically Annotating Traditional Irish Music Field Recordings,” *Sixth International Workshop on Content-Based Multimedia Indexing, Queen Mary University of London, UK*, Jun. 2008.
[29] J. Downie, “EVALUATING A SIMPLE APPROACH TO MUSIC INFORMATION RETRIEVAL: CONCEIVING MELODIC N-GRAMS AS TEXT,” 1999.
[30] R. McNab et al., “The New Zealand Digital Library MELody inDEX,” *D-Lib Magazine*, vol. 3, 1997, pp. 4-15.
[31] A. Ghias et al., “Query by humming: musical information retrieval in an audio database,” *Proceedings of the third ACM international conference on Multimedia*, 1995, pp. 231-236.
[32] R. McNab et al., “Towards the digital music library: tune retrieval from acoustic input,” *Proceedings of the first ACM international conference on Digital libraries*, 1996, pp. 11-18.
[33] J. McPherson and D. Bainbridge, “Usage of the MELDEX Digital Music Library,” *Proceedings of the Second Annual International Symposium on Music Information Retrieval (Bloomington, IN, USA*, 2001, pp. 15-17.
[34] S. Rho and E. Hwang, “FMF (Fast Melody Finder): AWeb-Based Music Retrieval System,” *Computer Music Modeling and Retrieval: International Symposium, CMMR 2003, Montpellier, France, May 26-27, 2003: Revised Papers*, 2004.
[35] T. Dietterich, “Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms,” *Neural Computation*, vol. 10, 1998, pp. 1895-1923.