JOSQUINTAB: A DATASET FOR CONTENT-BASED COMPUTATIONAL ANALYSIS OF MUSIC IN LUTE TABLATURE

Reinier de Valk[1] **Ryaan Ahmed**[2] **Tim Crawford**[1]
[1] Department of Computing, Goldsmiths, University of London, [2] Digital Humanities, MIT
`r.f.de.valk@gold.ac.uk, rahmed@mit.edu, t.crawford@gold.ac.uk`

ABSTRACT

An enormous corpus of music for the lute, spanning some two and half centuries, survives today. Unlike other musical corpora from the same period, this corpus has undergone only limited musicological study. The main reason for this is that it is written down exclusively in lute tablature, a prescriptive form of notation that is difficult to understand for non-specialists as it reveals little structural information. In this paper we present JOSQUINTAB, a dataset of automatically created enriched diplomatic transcriptions in MIDI and MEI format of 64 sixteenth-century lute intabulations, instrumental arrangements of vocal compositions. Such a dataset enables large-scale content-based computational analysis of music in lute tablature hitherto impossible. We describe the dataset, the mapping algorithm used to create it, as well as a method to quantitatively evaluate the degree of arrangement (goodness of fit) of an intabulation. Furthermore, we present two use cases, demonstrating the usefulness of the dataset for both music information retrieval and musicological research. We make the dataset, the source code, and an implementation of the mapping algorithm, runnable as a command line tool, publicly available.

1. INTRODUCTION

An enormous corpus of music for the lute, roughly spanning the sixteenth, seventeenth, and first half of the eighteenth century, and containing an estimated 60,000 individual pieces in circa 860 sources, printed and manuscript, survives today [27]. Although the lute was one of the most widely used instruments during these two and a half centuries, and this corpus is thus extremely rich in information about daily musical practice, up until the present day it has not undergone the same critical musicological study as other musical corpora from the same period. The main reason for this is that the notation used to write down lute music, lute tablature, is notoriously difficult to understand for non-specialists [18, 28]. This is because lute tablature is a purely *prescriptive* form of notation: like all forms of instrumental tablature [12], it merely provides the actions a player must take—in this case, where to place the fingers on the fretboard and which strings to pluck—rather than the sounds and musical edifice these actions produce, which a *descriptive* form of notation does [20,29]. In practice, this means that lute tablature, as opposed to mensural forms of notation, conveys virtually no information about the structure, polyphonic or other, of the music it encodes, making it hard to comprehend *a prima vista*.

In this paper we propose to address the problem of the marginal position of music in lute tablature within musicological research by providing JOSQUINTAB, a dataset of 64 automatically created *enriched diplomatic transcriptions*—literal transcriptions annotated with voice, key, and mensuration information—of a selection of sixteenth-century lute and vihuela (the lute's Spanish counterpart) *intabulations*, instrumental arrangements of vocal compositions, as well as a mapping algorithm for creating such a dataset, and an implementation thereof, runnable as a command line tool. The proposed dataset enables large-scale content-based computational analysis of music in lute tablature, which, in the absence of some notion of the music's polyphonic structure, has hitherto not been possible. The command-line tool allows researchers to extend the dataset by creating their own transcriptions.

In summary, the main contributions of this paper are:

- a dataset of 64 automatically created enriched diplomatic transcriptions of lute intabulations, each in the original and in an unornamented version;
- a mapping algorithm for creating such a dataset;
- an implementation of the mapping algorithm, runnable as a command line tool.

Furthermore, we provide:

- a method to quantitatively evaluate the degree of arrangement (goodness of fit) of an intabulation, which is incorporated in the mapping algorithm;
- two use cases, demonstrating the usefulness of the dataset for both music information retrieval (MIR) and musicological research.

In the spirit of open science and reproducible research, we aim to make the resources we contribute FAIR (findable, accessible, interoperable, and reusable) [32]. We do so by using *recommended formats*—MIDI and MEI—for the dataset,[1] and by making the dataset, the source code for the mapping algorithm, and the command line tool publicly available.

Attribution: Reinier de Valk, Ryaan Ahmed, Tim Crawford. "JOSQUINTAB: A dataset for content-based computational analysis of music in lute tablature", 20th International Society for Music Information Retrieval Conference, Delft, The Netherlands, 2019.

---

[1] `https://www.loc.gov/preservation/resources/rfs/`

The remainder of the paper is structured as follows. In Section 2, we sketch the scholarly and scientific background against which this research is carried out. In Section 3 we describe the contents of JOSQUINTAB and the preprocessing that the data we build on requires, and in Section 4 the mapping algorithm by means of which we create JOSQUINTAB, the method to evaluate the dataset quantitatively, and the command line tool. Section 5 is dedicated to the use cases, and in Section 6, conclusions and directions for future work are presented.

2. BACKGROUND IN MUSICOLOGY AND MIR

2.1 Lute tablature: A brief introduction

When in the late fifteenth century lutenists began playing polyphonic music on their instrument, this entailed a need for an accommodating, score-like, notation. Several approaches were developed simultaneously, crystallising into three principal systems, whose modern names refer to their regions of origin: French, Italian, and German lute tablature [1].[2] From these three, the French system would outlive the other two, from the seventeenth century on becoming the primary system. The former two systems use letters (French) or numbers (Italian), indicating where on the fretboard the fingers must be placed, on a six-line 'staff', each line of which represents a *course* (string pair) to be plucked. The German system is more abstract—it does not use a staff—and consists of a collection of glyphs (letters, numbers, and special characters, often varying per author) representing unique course-fret combinations.

2.2 Transcription practices

Figure 1, which shows an example of the Italian system, perfectly demonstrates the problem with lute tablature. Below a deceivingly simply notational surface, structurally rather complex music may lie hidden—and it requires a significant amount of expertise to infer the information not notated. The traditional musicological solution to 'unlock' this corpus for research is the preparation of scholarly editions, commonly in paper format, in modern music notation. Some of these are 'literal' or *diplomatic* transcriptions, recording the information given in the source exactly as it appears [17] (i.e., providing only pitch and minimal duration information), thus involving a minimum of interpretation but in turn revealing little about the polyphonic structure of the music, while others are full-fledged polyphonic transcriptions, involving a maximum of interpretation, and therefore often accompanied by a critical apparatus motivating the choices made.

Preparing scholarly editions is time-consuming and specialist work—all the more so in the latter case. This is one of the reasons why recent times have seen attempts at automatic transcription of lute tablature. Computers are ideally suited for fast (and large-scale) literal transcription—in fact, this can be achieved fairly easily with many of the music notation software packages, proprietary or open-source, available today. Automatic *polyphonic* transcription of lute tablature, however, has proven to be quite challenging [8, 13, 14], and in MIR research is still considered an open problem. Thus, despite ongoing efforts in musicology and MIR—both concerning content-based [22, 23] and bibliographical research [11]—only a fraction of the lute repertory has been explored up until the present day, and much work remains to be done.

2.3 Sixteenth-century lute music and intabulations

Sixteenth-century lute music is generally considered to fall into three categories [3]: (i) original compositions, i.e., idiomatic, abstract works in varying degrees of counterpoint, such as fantasias, ricercares, or preludes; (ii) settings of dance tunes; and (iii) intabulations, i.e., instrumental arrangements of vocal compositions. It is the latter category—by far the largest of the three, by itself taking up at least approximately half of the music [3]—we focus on in this paper. Up until today, intabulations have largely escaped in-depth musicological research—mostly on the grounds that as non-original works, they would be inferior compositions, which, especially when heavily ornamented, served no other goal than to demonstrate the player's virtuosity. While it is doubtlessly true that certain arrangements are less exciting than others, a fact that should not be overlooked is that intabulations are highly informative about contemporary compositional and performance practices [6,18,19,26], giving insights into, e.g., ornamentation techniques [4] or the application of *musica ficta*, i.e., chromatic alterations not notated in vocal music [5].

Intabulations generally remain close to their vocal models: contemporary treatises on the subject of intabulating prescribe that the intabulator take over as much of the vocal model as is technically possible on the instrument [24]; several existing in-depth studies of selected intabulations [2, 4, 5, 25, 30] support this. It is exactly the resulting close relation between model and intabulation that facilitates the mapping approach presented in this paper.

3. JOSQUINTAB

As its name reveals, the dataset exclusively contains intabulations of vocal compositions by Josquin des Prez (c. 1450-1521), an earlier-generation composer highly popular among sixteenth-century intabulators [3, 21]. According to the *New Josquin Edition*, no less than 18 chansons, 17 motets (most of them in multiple parts), various parts of 17 mass sections, and even eight complete masses have been intabulated [16]—adding up to at least 60 works, depending on how one counts.

In creating the dataset, which we do by mapping machine-readable tablature encodings (in TabCode format [10, 28]) onto MIDI renderings of editions of the vocal models (as explained in detail in Section 4),[3] we build on two existing, high-quality data repositories. The tablature encodings were retrieved for a previous study [22]

---

[2] The vihuela players generally adopted the Italian system; vihuela tablature, especially the upside-down variant that equals modern guitar tablature, is sometimes referred to as *Spanish tablature*.

[3] See also `http://www.ecolm.org/`.
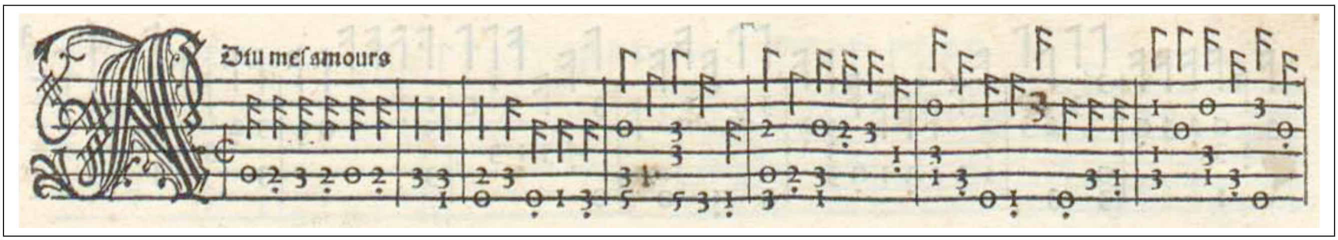
Figure 1. Francesco Spinacino, *Intabulatura de lauto, Libro primo* (Venice, 1507), *Adiu mes amours*, first system.

from Sarge Gerbode's Lute Page,[4] a curated repository of encodings made by amateur enthusiasts; the MIDI renderings of editions of the vocal models (as well as scores in PDF format of the same editions, used for reference) are retrieved from the GitHub repository for the Josquin Research Project (Center for Computer Assisted Research in the Humanities, Stanford University),[5] which aims to provide ways to search and analyse scholarly edited scores of polyphonic early music.[6]

Our eventual goal is for the dataset to contain transcriptions of *all* intabulations of Josquin compositions, including all dubious attributions (but excluding the spurious ones). This inclusive approach is taken as scholarly opinion on the matter of the authenticity of a significant number of pieces tends to change constantly [31].[7] However, while the Josquin Research Project repository provides editions of all vocal models, the Lute Page repository only contains intabulations of nine of the chansons, 10 of the motets (in 21 parts in total), and 12 of the mass sections parts, where approximately half of the chansons and motet parts have been intabulated more than once. (For practical reasons, for now we ignore the eight complete masses, all of which appear in a single print and are therefore somewhat of an exception; moreover, we only look at printed sources, facsimiles of which, needed for error checking, are generally easier to come by.) As Table 1 shows, this yields a dataset of 64 pieces (60658 data points). Each piece is represented by (i) a machine-readable encoding of the tablature in TabCode format; (ii) a MIDI rendering of the vocal model; (iii) a tuple (see below) of MIDI renderings of the created transcription; (iv) a tuple of MEI renderings (containing score and tablature) of the created transcription; and (v) a CSV file with the mapping details.[8]

In many of the intabulations, a 'net' of ornamentation has been added to the original notes, like a superstructure imposed on the music [5]. We provide each transcription both in the ornamented, original version and in a version stripped of the ornamental net (the procedure by which this is done is explained in more detail in Section 4). The latter is expected to be more useful for analysis tasks that are complicated by large amounts of ornamentation, such as the one in Use Case 2 (see Section 5.2). Furthermore, it must be noted that tablature only provides a 'minimum' duration for each note, applying to all notes in a chord (see Figure 1). We have made no attempt at reconstructing the notes' actual duration, i.e., at determining their offset—which is not a trivial problem as it depends on multiple contextual factors (e.g., counterpoint, phrasing, etc.). At this point, it is therefore important to stress that the transcriptions created are not full-fledged editions in the musicological sense of the word, and are not intended as such—rather, they serve to facilitate large-scale content-based computational analysis.

3.1 Preprocessing: Alignment issues

The mapping algorithm takes as input an encoding of the tablature in TabCode format and a MIDI rendering of the vocal model. In order for it to function optimally, these files must be aligned both in terms of pitch (i.e., the appropriate lute tuning must be used) and time (i.e., they must span the same number of bars).[9] In the ideal case, only the tuning needs to be provided—but in practice, most of the tablature encodings require a number of small corrections, and in some cases even additions, all of them necessary to ensure temporal alignment. Currently, these need to be carried out in a manual preprocessing step.

There are four types of *corrections*, three of which apply to the encodings. First, almost all encodings require the mensuration signs used to be adapted in order for the nominal beat level to correspond to that in the MIDI file (in most cases this results in augmentation—e.g., a change of a $\frac{2}{2}$ mensuration to a $\frac{2}{1}$ mensuration—or, conversely, a reduction; but in some cases they are actually incorrect); moreover, a small number of encodings require lacking mensuration signs to be added. Second, a small number of encodings require unspecified triplets to be made specific. Third, a small number of encodings require the correction of syntactic errors that render them non-parsable (technically, this type of correction does not affect any alignment issues).[10] The fourth and last type of correction applies to the MIDI renderings of the vocal models, which all require the key signature meta message to be set.

Furthermore, there is one type of *addition*, which again applies to the encodings only, and which results in lengthening them. A substantial number of encodings—29 in

---

[4] See `http://www.gerbode.net/`. The files were converted from the proprietary Fronimo format into TabCode.

[5] `https://github.com/josquin-research-project/`

[6] `http://josquin.stanford.edu/`

[7] See `https://www.cmme.org/database/projects/14/` for a project born because of this exact reason.

[8] `https://github.com/reinierdevalk/data/`

[9] The standard tuning used for the six-course sixteenth-century lute is one where the courses are tuned in perfect fourths with a major third in the middle. The lowest-sounding course was generally tuned to nominal G or A, but several other tunings occured as well.

[10] We make no attempt at correcting any incorrectly encoded notes.

Vocal model	Voices	I	Vocal model	Voices	I
Absalon fili mi	4	1	*Missa Hercules Dux Ferrarie*, Sanctus–Pleni sunt–Osanna	4, 2, 4	1, 1, 1
Benedicta es, pt. 1–3	6, 2, 6	4, 3, 3	*Missa Pange lingua*, Benedictus	2	1
Fecit potentiam	2	1	*Missa Sine nomine*, Cum sancto	4	1
In exitu Israel de Egypto, pt. 1–3	4	1, 1, 1	*Qui belles amours*	4	2
Inviolata, pt. 1–3	5	1, 1, 1	*Adieu mes amours*	4	3
Memor esto verbi tui, pt. 1–2	4	1, 1	*Comment peult avoir joye*	4	1
Pater noster, pt. 1–2	6	2, 2	*Faulte d'argent*	5	1
Preter rerum seriem, pt. 1–2	6	3, 2	*Je ne me puis tenir*	5	3
Qui habitat in adjutorio, pt. 1–2	4	2, 2	*La plus des plus*	3	1
Stabat mater, pt. 1–2	5	2, 1	*Mille regretz*	4	2
Missa De beata virgine, Cum sancto–Credo–Crucifixus	4, 5, 5	2, 1, 1	*Plus nulz regretz*	4	1
Missa Faysant regretz, Qui tollis–Sanctus–Pleni sunt–Osanna	4, 4, 3, 4	1, 1, 1, 1	*Si j'ay perdu*	4	1

Table 1. JOSQUINTAB: 36 motet part intabulations (ranging in size from 234–2238 notes; median 1257.5 notes; total 43557 notes); 13 mass section part intabulations (260–1141 notes; median 343 notes; total 5942 notes); 15 chanson intabulations (483–1250 notes; median 686 notes; total 11159 notes). I = number of intabulations.

total—contain fewer bars than their vocal models. In most cases, this is obviously due to an error on behalf of the intabulator, i.e., accidental omission. Given a sequence of bars $\{b_1, ..., b_{n-1}, b_n\}$, where bar b_1 has the same ending as bar b_{n-1}, all bars between b_1 and b_n may easily be skipped unwittingly (a scribal error known as *homoeoteleuton* in manuscript studies [9]). In exceptional cases, however, omission is intentional. Omissions are countered by inserting rests into the intabulation at the appropriate position.

Apart from an encoding and a MIDI rendering, the mapping algorithm must also be given the appropriate tuning, as well as the appropriate *reduction*: the factor by which the durations in the tablature must be multiplied (or divided) in order for its nominal beat level to correspond to that of the vocal model. The reduction can be a positive integer greater than 1, in which case it amounts to an augmentation; a negative integer, in which case it amounts to a reduction; or 1, in which case the durations do not change. Although determining the tuning and reduction is strictly speaking not a preprocessing step, this information must be given as input to the mapping algorithm, and must thus be established—currently, manually—prior to running it.

4. MAPPING ALGORITHM

The mapping algorithm itself is fairly straightforward and traverses the music chord by chord (where a chord is any event that has one or more notes in it), from left to right. Given a tablature encoding and a MIDI representation of the vocal model, first, an ordered list T is made, containing the union of the chord onset times in both tablature and MIDI. Then, two two-dimensional matrices with $|T|$ rows are constructed: the *grid* G, containing, for each onset time T_i, onset, pitch, voice, and duration information for the chord at T_i in the MIDI (or null values when there is no chord); and the *mask* M, containing, for each onset time T_i, onset, pitch, duration and note index information for the chord at T_i in the tablature (or null values when there is no chord). The grid and mask are traversed jointly, row by row, and in the process a list of *voice labels* V is created. A voice label is a binary vector encoding all voices onto which a note is mapped. Let mask row $M_{i,*}$ represent a tablature chord and grid row $G_{i,*}$ a MIDI chord at onset T_i. If $M_{i,*}$ is non-null, i.e., there is a chord in the tablature at T_i, the following steps are taken.

1. If $M_{i,*}$ contains a single note whose duration is less or equal than the *ornamentation threshold* (a pre-set parameter) and $G_{i,*}$ is null, [11] i.e., there is no chord in the MIDI at T_i (see Figure 2, horizontal brackets), the note is flagged as ornamental. Null is added to V, the note is added to a running list O, and Steps 2-4 below are skipped.
2. Each note in $M_{i,*}$ is mapped to a voice by finding the voice that goes with its counterpart in $G_{i,*}$. A note can be mapped to multiple voices.
    (a) If a match is found, a voice label is added to V.
    (b) If no match is found (see Figure 2, single brackets and non-bracketed sharps), null is added to V, and the note is added to U, the list of unmapped notes in $M_{i,*}$.
3. If U is not empty, the notes in it are mapped—in one go—to the available voices (any voices not yet occupied by a note in $M_{i,*}$). To this end, for each available voice in the MIDI, the last note with an onset that is less than T_i in that voice is added as a pitch-voice tuple to a list A. Then, the cheapest mapping of U onto A is calculated, where the cost is based on pitch distance. All corresponding null voice labels in V are replaced, and all mapped notes removed from U. (If, after this process, there are still unmapped notes—which can happen if $M_{i,*}$ contains more notes than there are voices in the model—this step is repeated, now considering *all* voices as avail-

---

[11] As ornamentation threshold we use the value two levels below beat level (e.g., in $\frac{2}{1}$ or $\frac{3}{1}$ mensuration, a quarter note).

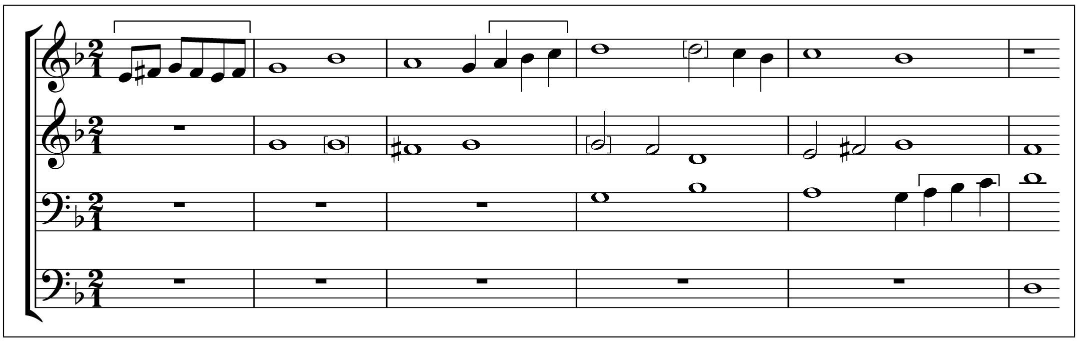

Figure 2. Francesco Spinacino, *Adiu mes amours*, opening bars (see Figure 1). Mapping algorithm output. The tablature contains all notes from the vocal model; bracketed notes and sharps are additions. Full note durations are inferred.

able. As a result, voices can have more than one note mapped onto them.)

4. (Optional.) If $M_{i-1,*}$ is ornamental, the cheapest connection of the last note in O to $M_{i,*}$ is calculated (where the cost is again based on pitch distance), and all notes in O are mapped onto the voice onto which the note in $M_{i,*}$ that yields the cheapest connection has been mapped. In case of a tie, the cheapest connection of the first note in O to $M_{i-1-|O|,*}$ (i.e., the chord immediately before that first note) is decisive. All corresponding null voice labels in V are replaced, and O is cleared.

When the grid and mask have been traversed, the algorithm combines the created list of voice labels V with an internal representation of the tablature created from the encoding, and, retaining the key and mensuration information from the MIDI, returns this as a track-separated MIDI file and an MEI rendering thereof. If Step 4 is included, these contain all ornamentation that is in the tablature; if it is skipped, they remain unornamented.

4.1 Evaluation method

There are four categories of mismatches. If in Step 1 a note is flagged, it is considered an *ornamentation*. Else, if in Step 2(b) no match is found for a note, and in Step 3 it is mapped at cost 0), it is considered a *repetition* (see Figure 2, single brackets); else, if there is a semitone discrepancy between it and a note in A, but they are the same pitch class, it is considered *musica ficta* (see Figure 2, non-bracketed sharps); else, it is considered an *alteration*. These categories are hierarchical and mutually exclusive.

This classification enables us to quantitatively evaluate the degree of arrangement (goodness of fit) of an intabulation in a granular manner. We measure the degree of arrangement by means of the *mapping ratio* $m = \frac{|N|-|M|}{|N|}$, where N is the set of all notes, and M the set of all mismatches. At the highest level of granularity, which results in the lowest m, M is equal to the total of all ornamentations, repetitions, instances of musica ficta, and alterations—i.e., to the total of mismatches across all four categories. It can be argued, however, that not all categories carry the same weight. Repetitions, for example, are idiomatic for lute music because of the instrument's short sustain, and have only little harmonic and melodic impact; and instances of musica ficta, which have a stronger harmonic and melodic impact (but no rhythmic impact), are nothing but inflections, different 'flavours' of the same pitch class. Thus, it makes sense not to include mismatches falling into these categories in M, or at least to weight them differently. We therefore redefine m as

$$m = \frac{|N| - (o|M_o| + r|M_r| + f|M_f| + |M_a|)}{|N|}, \quad (1)$$

where M_o, M_r, and M_f are the sets of all ornamentations, repetitions, and instances of musica ficta; o, r, and f their respective weighting parameters; and M_a is the set of all alterations (N, as above, is the set of all notes).

Genre	Ornamented			Unornamented	
	o, r, f			r, f	
	$1,1,1$	$1,0,0$	$0,0,0$	$1,1$	$0,0$
Motets	0.52	0.66	0.92	0.70	0.89
Mass sections	0.70	0.83	0.96	0.80	0.95
Chansons	0.57	0.69	0.94	0.75	0.92
All	0.54	0.68	0.93	0.72	0.90

Table 2. Mapping ratio m at different levels of granularity. Values are weighted averages, where each per-piece ratio is weighted by the number of notes in the piece.

Table 2 shows the mapping ratio, averaged per genre, at the highest level of granularity (i.e., the most strict case, with all parameters set to 1), at an intermediate level (with only o set to 1), and at the lowest level (i.e., the most lenient case, with all parameters set to 0) for the ornamented transcriptions, as well as at the highest and lowest level for the unornamented transcriptions, where the intermediate level does not apply. The table shows that, on average, at least half of the notes ($m = 0.54$) in the ornamented transcriptions, and approximately three quarters ($m = 0.72$) in the unornamented transcriptions, have an exact match in the vocal models. The ratios increase considerably—to 0.68 and 0.90, respectively— if we consider repetitions and musica ficta to be matches as well. The table also shows that, generally, the intabulations are fairly heavily ornamented, as the ratio at the lowest level of granularity ($m = 0.93$) is much higher than that at the other levels.

4.2 Command line tool

The Java implementation of the mapping algorithm, `TabMapper`,[12] can be run as a command line tool. A single transcription is created with the command

```
$ tabmapper tab.tc model.mid <t> <r> -n
```

where `tab.tc` is the TabCode encoding of the intabulation, `model.mid` the MIDI rendering of the vocal model, `<t>` and `<r>` are the tuning and reduction (see Section 3.1), and `-n` is an optional argument indicating that the transcription created should be unornamented. A *set* of transcriptions is created with the command

```
$ tabmapper list.csv -n
```

where `list.csv` is a CSV file specifying, for each piece, the name of the TabCode file, that of the MIDI file, the tuning, and the reduction, and `-n` is again optional.

5. USE CASES

Two use cases demonstrate the usefulness of the dataset for both MIR and musicological research.

5.1 Use Case 1: Voice separation

Use Case 1 shows how the dataset can serve to train machine learning models for voice separation in lute tablature [13, 14], which can be used for automatic polyphonic transcription of music in lute tablature. *Voice separation* can be defined as “the task of separating a musical work consisting of multi-note sonorities into independent constituent voices” [7]. From an MIR perspective, this task is considered an open problem. As with many supervised machine learning tasks in MIR, a lack of labelled data to train and evaluate models on hinders progress. This is a particularly pressing issue in the case of tasks related to music in lute tablature. Apart from our own manually created dataset containing 15 pieces (11641 data points) [13, 14], to our knowledge there currently exist no other datasets of voice-annotated tablature.

Automatic data creation can be a solution. However, data created by means of a rigid algorithm is less consistent in terms of transcription quality than data created manually, where each individual situation can be addressed on its own terms—and this may result in a model having more difficulty learning the task, and consequently, generalising worse. An experiment confirms this. In [13], we train and evaluate neural networks on nine four-voice pieces (8892 data points); our best-performing model achieves a generalisation accuracy of 79.63%. A model with the same architecture[13] trained and evaluated on the 28 four-voice pieces (26215 data points) in JOSQUINTAB yields a generalisation accuracy of only 64.47%.[14] We hypothesise this to be partly due to inconsistencies in the data; however, an analysis of the results reveals that an error propagation issue particular to imitative music, as discussed in [15], also plays a substantial role. As the dataset grows and inconsistencies are ironed out statistically, performance improvement is expected. Despite its current subpar performance, this model can already be reliably used for automatic polyphonic transcription.

5.2 Use Case 2: Cross-corpus melodic matching

Use Case 2 shows how the dataset can be applied for melodic matching across heterogeneous—in this case, instrumental versus vocal—corpora. A major goal of automatic transcription of tablature is the ability to search freely for *musical quotations* occurring between corpora in tablature and in standard notation. Our dataset provides an important stepping stone in this direction. Using the unornamented transcriptions created, we are able to do a further processing step in order to create search strings for Early Music Online Search,[15] a tool that uses OMR to enable full-text searches into sixteenth-century printed music.

Our initial results are promising. Eight pieces exist in both corpora, of which we are able to successfully search and match four. Not only are we able to match the individual pieces, but also the individual voice parts within each piece are correctly identified. In three additional cases, our search matches not with the correct piece, but rather with another piece by the same composer—showing that our search method enables the identification of common stylistic features. This shows the potential of transcribing a larger tablature corpus in order to identify hitherto unknown vocal models.

6. CONCLUSIONS AND FUTURE WORK

In this paper we present JOSQUINTAB, a dataset of automatically created enriched diplomatic transcriptions of 64 lute intabulations. We describe the contents of the dataset, the mapping algorithm used to create it, and we show how we can quantitatively evaluate it. Two use cases illustrate how the dataset can be used for content-based computational analysis within and across corpora, demonstrating its usefulness for MIR and musicological research. We make the dataset, the source code, and an implementation of the mapping algorithm publicly available.

There are many ways in which this work can be extended. One is to further automate the preprocessing (alignment, determination of tuning and reduction) that the data currently requires. Another is to parameterise certain functionality of the mapping algorithm (e.g., the value of the ornamentation threshold) in order to allow more flexibility in the transcription creation. A third is to extend the dataset so that it includes all intabulations of Josquin compositions, and those of other composers as well. It is clear that the work presented opens many new research avenues, which we plan to explore in future work.

---

[12] `https://github.com/reinierdevalk/tabmapper/`

[13] We now use the backward processing mode [13], and train using a validation set, doubling the number of training iterations and decreasing the amount of regularisation applied by a factor of 10.

[14] Values are cross-validation averages, where the dataset is partitioned along its individual pieces.

[15] `http://www.doc.gold.ac.uk/usr/265/`

7. REFERENCES

[1] W. Apel. *The notation of polyphonic music: 900-1600*. The Mediaeval Academy of America, Cambridge, MA, 5th edition, 1953.

[2] C. M. S. Bocchinfuso. Intabulations of music by Josquin des Prez in lute books published by Pierre Phalèse, 1547–1574, volumes 1–2. Master's thesis, University of Canterbury, 2009.

[3] H. M. Brown. *Instrumental music printed before 1600: A bibliography*. Harvard University Press, Cambridge, MA, 1965.

[4] H. M. Brown. Embellishment in early sixteenth-century Italian intabulations. *Proceedings of the Royal Musical Association*, 100:49–83, 1973–1974.

[5] H. M. Brown. Accidentals and ornamentation in sixteenth-century intabulations of Josquin's motets. In E. E. Lowinsky and B. J. Blackburn, editors, *Josquin des Prez: Proceedings of the International Josquin Festival-Conference*, pages 475–522. Oxford University Press, London, 1976.

[6] H. M Brown. The importance of sixteenth-century intabulations. In L. P. Grijp and W. Mook, editors, *Proceedings of the International Lute Symposium Utrecht 1986*, pages 1–29. STIMU Foundation for Historical Performance Practice, Utrecht, 1988.

[7] E. Cambouropoulos. Voice and stream: Perceptual and computational modeling of voice separation. *Music Perception*, 26(1):75–94, 2008.

[8] H. Charnassé and B. Stepien. Automatic transcription of german lute tablatures: An artificial intelligence application. In A. Marsden and A. Pople, editors, *Computer representations and models in music*, pages 143–170. Academic Press, London, 1992.

[9] R. Clemens and T. Graham. *Introduction to manuscript studies*. Cornell University Press, New York, 2007.

[10] T. Crawford. TabCode for lute repertories. *Computing in Musicology*, 7:57–59, 1991.

[11] T. Crawford, B. Fields, D. Lewis, and K. Page. Explorations in Linked Data practice for early music corpora. In *Proc. of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries, London, UK*, pages 309–312, 2014.

[12] T. Dart, J. Morehen, and R. Rastall. Tablature. In S. Sadie, editor, *The new Grove dictionary of music and musicians*, volume 24, pages 905–914. Macmillan, London, 2nd edition, 2001.

[13] R. de Valk. *Structuring lute tablature and MIDI data: Machine learning models for voice separation in symbolic music representations.* PhD thesis, City University, London, 2015.

[14] R. de Valk and T. Weyde. Bringing 'Musicque into the tableture': Machine-learning models for polyphonic transcription of 16th-century lute tablature. *Early Music*, 43(4):563–576, 2015.

[15] R. de Valk and T. Weyde. Deep neural networks with voice entry estimation heuristics for voice separation in symbolic music representations. In *Proc. of the 19th International Society for Music Information Retrieval Conference, Paris, France*, pages 281–288, 2018.

[16] W. Elders and E. Jas, editors. *Josquin des Prez: The sources*. The Collected Works of Josquin des Prez, 1. Koninklijke Vereniging voor Nederlandse Muziekgeschiedenis, Utrecht, 2013.

[17] J. Grier. *The critical editing of music: History, method, and practice*. Cambridge University Press, Cambridge, 1996.

[18] J. Griffiths. The lute and the polyphonist. *Studi Musicali*, 31(1):89–108, 2002.

[19] J. Griffiths. Songs without words: The motet as solo instrumental music after Trent. In E. Rodríguez-García and D. V. Filippi, editors, *Mapping the motet in the post-Tridentine era*, pages 206–227. Routledge, Abingdon, 2019.

[20] M. Kanno. Prescriptive notation: Limits and challenges. *Contemporary Music Review*, 26(2):231–254, 2007.

[21] H. Y. Kwee. Sixteenth-century printed instrumental arrangements of works by Josquin des Prez: An inventory. *Tijdschrift van de Vereniging voor Nederlandse Muziekgeschiedenis*, 22(1):43–66, 1971.

[22] D. Lewis, T. Crawford, and D. Müllensiefen. Instrumental idiom in the 16th century: Embellishment patterns in arrangements of vocal music. In *Proc. of the 17th International Society for Music Information Retrieval Conference, New York, NY, USA*, pages 524–530, 2016.

[23] R. J. Lewis, T. Crawford, and D. Lewis. Exploring information retrieval, semantic technologies and workflows for music scholarship: The Transforming Musicology project. *Early Music*, 43(4):635–647, 2015.

[24] H. Minamino. *Sixteenth-century lute treatises with emphasis on process and techniques of intabulation*. PhD thesis, University of Chicago, 1988.

[25] W. H. Mook. De bewerkingen voor luit en vihuela da mano van het motet *Benedicta es caelorum regina* van Josquin des Prez, volumes 1–2. Master's thesis, Rijksuniversiteit Utrecht, 1993.

[26] A. Moths. Mit oder ohne Text? Was Diminutionen und Intabulierungen zur Analyse vokaler Musik des 16. und 17. Jahrhunderts beitragen können. *Dutch Journal of Music Theory*, 16(1):39–46, 2011.

[27] A. J. Ness and C. A. Kolczynski. Sources of lute music. In S. Sadie, editor, *The new Grove dictionary of music and musicians*, volume 24, pages 39–63. Macmillan, London, 2nd edition, 2001.

[28] C. Rhodes and D. Lewis. An editor for lute tablature. In R. Kronland-Martinet, T. Voinier, and S. Ystad, editors, *Computer Music Modeling and Retrieval: Third international symposium, CMMR 2005*, pages 259–264. Springer, Berlin, 2006.

[29] C. Seeger. Prescriptive and descriptive music-writing. *The Musical Quarterly*, 44(2):184–195, 1958.

[30] G. Thibault. Instrumental transcriptions of Josquin's French chansons. In E. E. Lowinsky and B. J. Blackburn, editors, *Josquin des Prez: Proceedings of the International Josquin Festival-Conference*, pages 455–474. Oxford University Press, London, 1976.

[31] R. C. Wegman. The other Josquin. *Tijdschrift van de Koninklijke Vereniging voor Nederlandse Muziekgeschiedenis*, 58(1):33–68, 2008.

[32] M. D. Wilkinson et al. A design framework and exemplar metrics for FAIRness. *Scientific Data*, 5(180118), 2018.