KDC: AN OPEN CORPUS FOR COMPUTATIONAL RESEARCH OF *DASTGĀHI* MUSIC

Babak Nikzat
Kunstuniversität Graz
b.nikzat@kug.ac.at

Rafael Caro Repetto
Kunstuniversität Graz
rafael.caro-repetto@kug.ac.at

ABSTRACT

Iranian *dastgāhi* music is considered as the classical repertory of contemporary Iran. In the 19th century, the melodic modes that developed during its long history were grouped in categories, each of them known as *dastgāh*. The *dastgāhi* system presents unique features, that have been object of musicological study since its inception. However, computational methods for its research are still scarce, due in good part to the lack of open, well curated corpora. The aim of the KUG *Dastgāhi* Corpus (KDC) is to contribute to the development of computational corpus driven research for this tradition. KDC is created following the FAIR principles, and in close collaboration with performers and scholars, who contribute to it with annotations and qualitative evaluations. Besides presenting the first version of KDC, in this paper we explore the possibilities that Iranian *dastgāhi* music offers to computational research. In order to test the performance of state-of-the-art technologies applied to this music tradition, we present preliminary results for several analytical tasks, and discuss their opportunities and limitations learnt in the process.

1. INTRODUCTION

The potential of computational methods for supporting musicological research tasks is increasingly being recognized and adopted by musicologists. Due to the high degree of specialization required both from the technical and musicological sides of this kind of research, collaboration between specialists from both disciplines is one of the most fruitful approaches for computational musicology. Developed at the Institute for Ethnomusicology of the Kunstuniversität Graz (KUG), the KUG *Dastgāhi* Corpus (KDC) aims to offer a growing collection of well curated and annotated data of Iranian *dastgāhi* music around which engineers, musicologists and researchers from other disciplines interested in this music tradition can collaborate. In this first section we describe the main features of the *dastgāhi* musical system and offer an overview of the state-of-the-art of its computational research. Then, we introduce and describe in detail the first version of KDC. In section 3 we discuss the potential of Iranian *dastgāhi* music for computational research. In the following sections, preliminary analyses carried out to test such potential are presented and discussed. The paper closes with our plan for expanding KDC and future work.

1.1 Iranian *dastgāhi* music

Iranian *dastgāhi* music is modal. The basic modal entities in this tradition are known as *gušе*s and they are defined by a model melody built on a specific intervallic structure, a series of characteristic melodic patterns, and a modal centre known as *šāhed* [1]. Related *gušе*s are grouped in specific collections, and in a specific order. These collections are known either as *dastgāh* or *āvāz*, each of them with its own specific name. A *dastgāh* is larger than an *āvāz* in terms of number of *gušе*s, and an *āvāz* is considered to be a subcategory of a specific *dastgāh*. Nowadays, the canonical repertoire consists of 7 *dastgāh*s and 5 *āvāz*es (see Table 1). For simplicity, in this paper we are using *dastgāh* to refer to both of them collectively. The *gušе*s in a particular *dastgāh* are organized in a fixed order, and even though the performance of a *dastgāh* does not required all its *gušе*s to be performed, the order has to be maintained. The first *guše* in all *dastgāh*s is called *darāmad*, and it has a very significant meaning since it introduces the *dastgāh*. Most performances start and end with *darāmad*.

Along the history of Iranian *dastgāhi* music, great masters (*ostād*s) of this tradition defined their own canonical repertory of the 7 *dastgāh*s and 5 *āvāz*es, fixing the number of *gušе*s in each of them, their particular order, their specific melodic contour and their rhythmic rendition. These canonical repertoires are known as *radif*s, always associated with their corresponding *ostād*. These fixed *radif*s are nowadays only used for learning the tradition. Contemporary musicians perform their own interpretation of the repertory, but through the rules they experientially extracted from the *radif* they learnt. As a more encompassing concept, different performance schools are recognized, each of them with idiosyncratic performance styles. The most extended ones today are Tabriz, Tehran and Isfahan.

Iranian *dastgāhi* music uses a microtonal tuning system, resulting in specific intervallic structures defined for each *guše*. Three basic interval categories are considered, namely tone, semitone and an intermediate one between the previous two, whose specific width is not fixed but flexible [1], depending on *guše* or even performer, and that is the result of altering one natural tone less than a semitone

up (*sori*) or down (*koron*). These intervallic structures form specific scales, in which particular tones receive predominant roles, such as *šāhed* and *ist*, which are fundamental for defining the modal features of a *guše*.

Specific techniques for vocal and instrumental performance, including an extensive repertory of ornamentation, is another essential aspect of this music, showing a great diversity depending on school, instrument, and personal style. The most common instrumental ornaments include: *ešāreh*, *tekiyeh*, *qalt*, *dorrāb*, *riz*, and *šalāl*. *Tahrir* is the most significant singing ornament, which requires a specific technique in which the singer switches between chest and head voice (see Fig. 5).

*Guše*s can be performed in free rhythm or using rhythmic cycles. Even though this is fixed per each *guše* in the *radif*s, it is decided by each musician in contemporary performance. When a *guše* is sung in free rhythm, the prosodic rules, known as *aruz*, of Persian classical poetry, to which most of the sung lyrics in Iranian *dastgāhi* music belong, informs the rhythmic rendition of the singing. This results in a characteristic rhythmic style that is also adopted when a *guše* is perfomed only instrumentally [2, 3]. *Aruz* not only informs musical rhythm through the sequence of long and short syllables, but also through the qualitative features of the vowels (*hejā*) [4, 5]. For metred performance, musicians can draw on rhythmic cycles known as *adwār-e iqāi*, the most common ones being *kerešme*, *čāhār-pāreh*, *zanguleh* and *gereyli* [4]. Each *dowr-e iqāi* is defined by a number of pulses, known as *naqareh*, and a structure of accents.

A typical ensemble of Iranian *dastgāhi* music is formed by *santur* (hammered zither), *tār* (long-neck lute), *setār* (long-neck lute), *ney* (bamboo flute), *kamanče* (spike fiddle), *tonbak* (goblet drum) and a singer. The music, however, is also commonly performed by a solo singer or instrumentalist.

Dastgāhi music is an oral tradition and conventionally the students learn it by memorizing a specific *radif* taught in private lessons by an *ostād*. The learning process is based on observation and imitation. In the contemporary context some teachers use transcriptions in staff notation as teaching material. However the notation functions just as memory aid, since many details cannot be represented in the staff notation system.

1.2 Computational analysis of Iranian *dastgāhi* music

Considering the magnitude of Iranian *dastgāhi* music, in terms of history, wide performance realm, complex theorization and extensive musicological research, this music tradition has received comparatively little attention from computational approaches. Pioneering work has been carried out by the researcher and *santur* performer Peyman Heydarian. Since his master thesis [6], Heydarian is working on the computational analysis of this instrument, especially its acoustic features [7–9], as well as on *dastgāhi* music performed in this instrument, with a special focus on *dastgāh* classification [10–12]. In fact, this second task, automatic *dastgāh* identification and classification, has been the most common one by other researchers who approached this tradition from computational methods [13–16].

Computationally aided musicological research on *dastgāhi* music is still very rare. Shafiei used score to audio alignment for the analysis of *tahrir* [17], while Sanati analysed intervals from *ney* recordings [18]. There has also been some work on *dastgāhi* music generation using deep neural networks [19].

We argue that one of the reasons that explains that the computational analysis of *dastgāhi* music is not yet been fully established in an unified trend, instead of occasional isolated works, is the lack of a comprehensive, publicly accessible datasets upon which continuous research can be built. All the publications referenced in the previous paragraphs gathered their own research datasets, but none of them offer any indication about possibilities for accessing them. Only Heydarian has been able to sustain a decades long research on this tradition based on his own dataset [20], but it is not public, and it only includes recordings of *santur*. Hence, the goal of KDC precisely is to contribute to the establishment of a coherent line of computational research on Iranian *dastgāhi* music by providing an open, well curated corpus of high quality recordings, upon which this research can be built.

2. DESCRIPTION OF KDC

KDC is conceived as an ever growing corpus, so in this paper we present its first version. Although far from its ideal size in terms of coverage, it already offers a coherent and comprehensive representation of the *dastgāhi system*. Currently, KCD contains 213 solo recordings by four professional musicians, 65 of which consist of complete *guše* performances and hence form the core of the corpus.

In order to analyse the corpus, we draw on the five criteria defined by Serra [21]:

Purpose: As described in section 1, KDC is created for the analysis of Iranian *dastgāhi* music using corpus driven computational methods. Due to the lack of a notational system originally devised for this music tradition, and the difficulties of staff notation for capturing key elements of this music, especially regarding tuning, intonation and ornamentation, KDC primarily aims at gathering audio recordings. However, the inclusion of other data types, such as scores, lyrics, videos, images, etc., that might contribute to the general purpose of the corpus, is not excluded. A key point in the creation of KDC is the close collaboration with musicians. They do not only contribute with recordings and annotations, but also with their knowledge, evaluation, and research ideas. Musicians are thus considered as active collaborators of KDC.

Coverage: According with KDC's purpose, we focus on voice and melodic instruments. In this first version of the corpus, being one of our goals to test state-of-the-art technologies for the analysis of this music tradition, we have focused on solo performances, and consciously excluded ensemble performances, that would present a greater challenge for computational analysis. In order to

		Performance		*Šāhed*	*Ist*	Pitch Set	*Radif*	
Dastgāh	*Čāhārgāh*	7	9:02	5	3	2	5	5:21
	Homāyun	5	7:35	5	4	2	3	3:50
	Māhur	6	8:37	5	2	2	3	2:03
	Navā	6	9:19	5	4	2	3	4:01
	Rāst-Panjgāh	4	5:56	4	3	2	2	2:15
	Segāh	5	8:57	5	2	2	3	2:12
	Šur	7	9:55	5	2	2	3	2:12
Āvāz	*Abuatā*	5	6:47	5	3	1	2	2:01
	Afšāri	5	9:36	5	4	1	2	1:56
	Bayāt-e Tork	5	8:03	5	3	1	2	1:47
	Dašti	5	6:53	5	4	1	2	2:10
	Esfahān	5	9:21	5	3	2	2	2:40
	Total	65	100:01	59	38	20	31	31:40

Table 1. Number of recordings in KDC according to *dastgāh* and *āvāz*. For performance and *radif* recordings, the duration is also given.

Inst.	Musician	Rec.	Dur.
voice	F. Sahebghalam, R. Zalpour	24	37:50
ney	R. Zalpour, M. Khodadadi	37	45:54
setār	M. Shaari	35	47:57
Total		96	131:41

Table 2. Number of performance and *radif* recordings in KDC according to instrument.

build a first dataset that provides a coherent and comprehensive representation of the *dastgāh* system, this first version of KDC contains the first *guše* of all 7 *dastgāh*s and 5 *āvāz*es, known as *darāmad*, which holds great significance (see section 1.1). Four professional musicians (see Table 2) have contributed to KDC with original recordings of complete performances [1] of these 12 *darāmad*s. Hence, the corpus currently contains five full renditions of the 12 *darāmad*s, namely, two vocal ones, one female and one male (with the exception of the *darāmad* of *dastgāh rāst-panjgāh*, whose female vocal rendition could not be recorded), two by *ney*, and one by *setār*. According to their own decision, the musicians occasionally recorded more than one version of a particular *darāmad*, and all of them are included in the corpus. This results in 65 performance recordings, which is the core of KDC (see Table 1).

Completeness: KDC includes a csv file with metadata and annotations for all recordings, including the corresponding *dastgāh*, *guše*, artist, instrument, recording type, and duration. Considering the methodological purpose of the corpus, the musicians who collaborate with KDC were asked to record the *šāhed* of the performed *darāmad* as a single, stable pitch. The 59 resulting recordings are included in KDC, and they were used to compute the frequency of the *šāhed*, [2] which is added as an annotation in the metadata file. The musicians who collaborate with KDC show a great interest in the project and actively propose contributions to it, according to their understanding of which data could benefit the purpose of KDC. As a consequence, some musicians recorded the *ist* of the performed *darāmad* in a separate file as a single, stable pitch, resulting in 38 recordings, the corresponding pitch set, resulting in 20 recordings, and one or more related performances according to *radif* (see Table 1). Since pattern analysis is a key task for the study of Iranian *dastgāh*s, two of the collaborator musicians manually annotated the characteristic patterns that identify the performed *guše* using the software Praat [23]. Each of them annotated two full renditions of the 12 *darāmad*s, so that only the *setār* recordings are still not annotated in terms for patterns.

Quality: all the recordings in KDC are provided in the lossless compression format flac. The recordings by M. Shaari and F. Sahebghalam were recorded in a professional recording studio at KUG, those by M. Khodadadi were recorded by the first author using a Zoom H4 recorder, and those by R. Zalpour by himself using his cell phone. The musicians, all having previous recording experience, evaluated the quality of their own recordings both in terms of sound and performance quality and agreed to include them in KDC as good representatives of their art. To validate the pitch frequencies of *šāhed*, mp3 files were generated, and R. Zalpour assessed their correctness.

Reusability: All the collaborator musicians contributed their recordings to KDC for them to be published under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). [3] In order to ensure the reusability of KDC, we follow the FAIR principles. [4] To ensure that all recordings are findable and reusable, they are made available through the KUG's Phaidra repository. [5] Regarding accessibility, the

[1] Given its improvisational nature, there is no standard rendition of a particular *guše*. By complete performance we mean a rendition that completely conveys the performed *guše*.

[2] The pitch of the *šāhed* was computed with the pYin: Notes plugin [22] of Sonic Visualiser

[3] https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode

[4] https://www.go-fair.org/fair-principles/

[5] The following Phaidra collection is created for KDC: https://phaidra.kug.ac.at/view/o:127195

metadata of all recordings are also stored in MusicBrainz [6] in its original Farsi and with English aliases, and linked to their corresponding Phaidra objects. A specific protocol for interoperable structuring of Iranian *dastgāhi* music is still being developed, and is expected to be evolving along with the expansion of the corpus.

3. RESEARCH POTENTIAL OF KDC

The hypothesis that motivated the KDC project is the conviction that musicological study of Iranian *dastgāhi* music can deeply benefit from computational methods. Conversely, we argue that the characteristics of this music raise interesting research tasks for music information retrieval. In this section we point out some of the most relevant of these tasks. Since KDC was created with the purpose of researching the modal aspect of this music, tasks related to this domain are described in more detail.

3.1 Melody and mode

The study of melody in Iranian *dastgāhi* music is one of the most essential tasks, especially for the purpose of KDC, therefore pitch track extraction is essential. According to our preliminary analyses, state-of-the-art algorithms perform satisfactorily for monophonic recordings as those by voice or *ney*. However, further research is still needed to obtain equally satisfactory results for instruments with high resonance as the *setār* (see section 4). A relevant research task based on these pitch tracks is the corpus driven analysis of vibrato and *tahrir* (see Fig. 5). Their automatic identification and measurement would contribute to their deeper understanding, but also to better define school styles, or idiomatic performance for specific instruments.

Tasks conducing to the characterization of *guše*s are also of key importance. A particularly relevant one is automatic pattern analysis. A main challenge for this task from the computational point of view is evaluation. In order to contribute to that, KDC includes manual annotations by expert musicians. Analyses of each *guše*'s tonal hierarchy are also essential, especially those contributing to the understanding of concepts like *šahed* or *ist* (see section 1.1).

Given the classification system of *guše*s, *dastgāh*s and *āvāz*es, this music tradition is well suited for automatic recognition and classification tasks, which in fact are the ones most commonly explored to date (see section 1.2).

Finally, even though Iranian *dastgāhi* music is not easily represented on staff notation, a tradition of transcriptions using this notation system exists. This music then can be also taken as a challenging case for automatic transcription, for which existing scores (currently only in print) can be used for evaluation.

3.2 Other musical aspects

Iranian *dastgāhi* music offers interesting tasks for rhythm research. In metred performance, automatic *dowr-e iqāi* detection becomes a significant challenge, given the heterophonic nature of ensemble performance and the timbral and stylistic characteristic of solo performance. Expressive microtiming deviations from isochrony is a meaningful analysis for characterising performance style. Unmetred performance of *dastgāhi* music offers a great opportunity for multimodal analysis combining audio and lyrics analysis (this would require incorporating text data to KDC). Natural Language Processing methods can be applied to model the prosodic features of *aruz*, that can later be mapped to music performance.

Timbre is a very diverse and expressive feature of Iranian *dastgāhi* music performance. Combined with nuances in dynamics, singers use vowel shades to introduce variability in sustained notes. Equally, *ney* performers use a wide range of timbres from bright, clean ones to raspy, breathy ones. Automatic analysis of the timbral categories can inform about stylistic preferences and if a relationship exists with specific *guše*s. The use of dynamics is another important element in Iranian *dastgāhi* music expressivity, and its computational analysis can shed light about its systematic or stylistic use.

4. PRELIMINARY ANALYSES

In order to test the potential of KDC as described in the previous section, we carried out a series of preliminary analyses. The current size of KDC does not allow to obtain general conclusions about Iranian *dastgāhi* music. Our aim is to explore the possibilities for computational methods, and therefore to suggest further lines of research for future stages of KDC.

Since our main interest is the melodic dimension of *dastgāhi* music, our analysis started with pitch track extraction. To that aim, we applied the CREPE algorithm [24] to all monophonic recordings of KDC, that is, those by voice and *ney*. After visual and aural examination, we concluded that state-of-the-art algorithms, both for monophonic and polyphonic signals, [7] do not produce results usable for further analysis on the recordings of *setār*, an instrument with high, and sought after, resonance and common use of chords. Consequently, the recordings by *setār* were excluded for these analyses. In order to ensure reproducibility, the data, metadata and annotations used for these analyses are shared as a dataset. [8] The code and resulting plots are also shared in a public repository. [9]

4.1 *Šāhed*

In order to study *šāhed*, we compute pitch histograms from the previously extracted pitch tracks using the algorithm developed in [27]. The histograms are computed in cents aligned to the *šāhed* that KDC has annotated for each individual recording. We also compute pitch histograms for the aggregated pitch tracks of all recordings of the same

[6] The following MusicBrainz collection is created for KDC: https://musicbrainz.org/collection/6ee214c7-f116-4691-b46d-36033c8b17b0

[7] Pitch extraction was attempted on the *setār* recordings using the algorithms CREPE [24], pYin [22] and Melodia [25], the last two in their Sonic Visualiser [26] plug-in version.

[8] https://phaidra.kug.ac.at/o:127202

[9] https://github.com/Rafael-Caro/KDC-v1.0

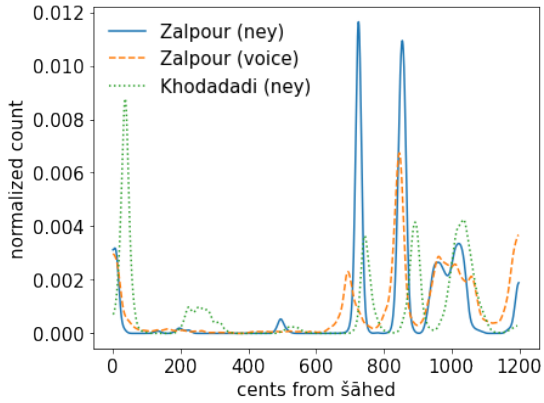

Figure 1. Folded pitch histograms of three recordings of *darāmad abuatā*

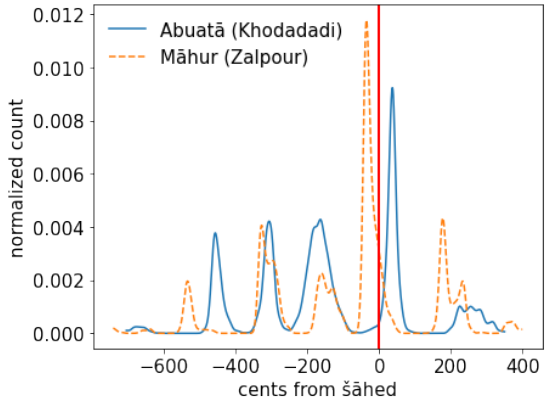

Figure 2. Pitch histograms of two *ney* recordings. The vertical line indicates the *šāhed*

dastgāh. Besides, histograms folded to one octave are also computed for individual recordings and for *dastgāh*s.

One of the most discussed issues in Iranian classical music is the concept of *šāhed*. Different definitions for this concept can be found in the existing literature. According to Farhat, "[i]n most of the Persian modes one tone assumes a conspicuously prominent role. It may or may not be the finalis" [1]. Nooshin argues that *šāhed* is "[t]he most prominent pitch, functioning as the tonal center of the mode" [28], while Talai defines it as "[t]he most frequently repeated pitch" [29]. We argue that corpus driven analyses can help to examine which of the definitions for *šāhed*, namely, "prominent" tone, "tonal center" or "most frequently repeated" tone better represents our results.

Fig. 1 shows the folded pitch histograms of three different performances of the same *guše*, *darāmad abuatā*. It can be seen how the *šāhed* is the most frequent tone in M. Khodadadi's performance, but that is not the case in the two ones by R. Zalpour, both vocal and *ney*. This preliminary result, together with other examples not mentioned here, suggests a more nuanced understanding of *šāhed*, that would escape a unique, universally applicable definition.

Mus. (inst.)	Mean	SD	Lowest	Highest
Saheb. (voice)	4.19	10.28	-10.11	28.10
Zalp. (voice)	6.92	7.26	-5.02	16.46
Zalp. (*ney*)	-3.19	16.87	-35.43	23.99
Khod. (*ney*)	13.95	13.70	-7.69	37.07

Table 3. Mean and SD of the differences in cents between the *šāhed*'s pitch in their isolated recordings and in performance per musician and instrument. The difference between the lowest and highest performance of the *šāhed* compared with the isolated recording is also given in cents.

4.2 Intonation

One interesting issue is the musicians' abstract notion of the *šāhed*'s pitch and its actual intonation in performance. Since all the collaborator musicians recorded the *šāhed* of their performances as an extra, isolated file according to their aforementioned abstract notion, we can compare its pitch in both contexts. Fig. 2 shows how M. Khodadadi performs the *šāhed* 37.07 cents higher than the isolated version he recorded himself, whilst R. Zalpour performs it 35.43 cents lower. We computed the difference between isolated *šāhed* and its performed version (obtained from the value of the closest peak in the histogram) for all recordings (excluding *setār*), and calculated the mean and SD for each musician and instrument, as it can be seen in Table 3. The results show a great variability in the divergence between conceptualized and performed *šāhed*, frequently reaching easily audible deviations.

These observations points to an interesting line of future research regarding the perception of intervallic structures by musicians. The specific width of the intervals of a particular *guše*, an essential element for the conceptualization of Iranian *dastgāhi* music, is part of the implicit knowledge of each individual musician, enacted mainly during performance. Their intellectual formulation, as the examples here analysed suggest, might no correspond with this implicit knowledge.

4.3 Pitch alteration

In some *gušes* specific degrees can be rendered by two neighbouring pitches. For example, in the case of *darāmad afšāri*, the degree above the *šāhed* can be performed at 119 cents above it, which is considered the main tone for this degree, or at 204 cents over the *šāhed*, which is considered an altered version of the same degree, as it can be seen in Fig. 3 (cent values computed according to [1]). These altered pitches are known as *not-e moteġayer* [1,30]. The use of this pitch alteration can be observed in pitch histograms as flat peaks, as it can be seen in Fig. 4 around 200 cents. These results point to another interesting line of research, for the computational detection of *not-e moteġayer* that can contribute to the better characterization of *gušes*.

4.4 Vibrato and *tahrir*

In order to analyse vibrato, we ran the Vibrato algorithm available in the Essentia library [31] on all recordings, and

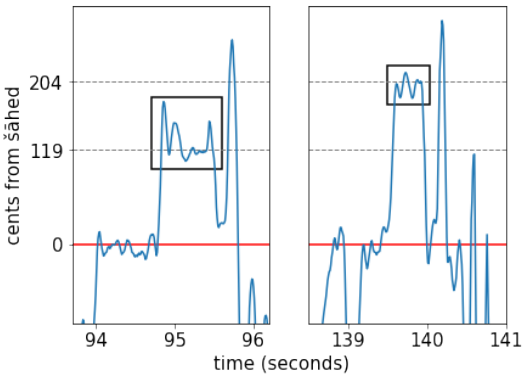

Figure 3. Two excerpts of the the pitch track of the of the vocal recording of *darāmad afšāri* by F. Sahebghalam. The horizontal red line indicates the *šāhed*, the gray line at 119 cents indicates the main pitch, whose corresponding *not-e moteġayer* is indicated by the gray line at 204 cents. The boxes mark the performance of these two notes.

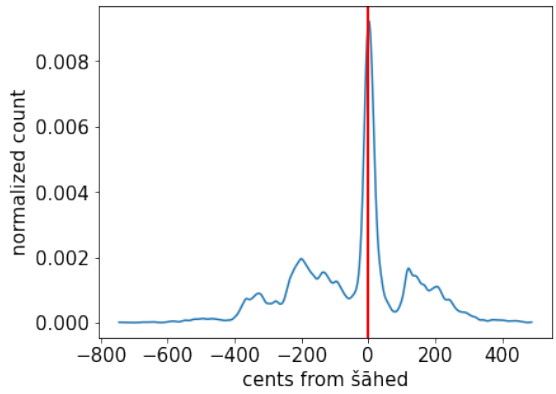

Figure 4. Pitch histogram of the vocal recording of *darāmad afšāri* by F. Sahebghalam. The vertical line indicates the *šāhed*

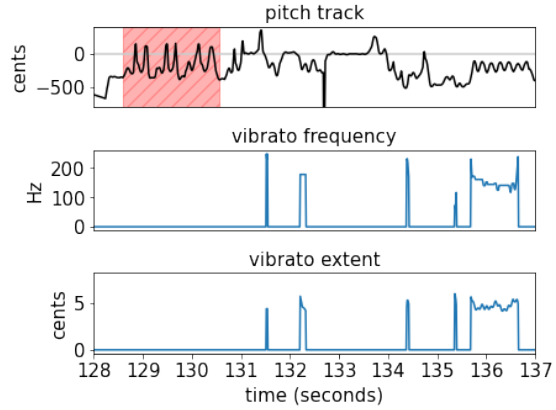

Figure 5. Excerpt of the recording of *darāmad abuatā* by F. Sahebghalam. Upper plot shows pitch track (horizontal line indicates *šāhed*), middle plot detected vibrato frequency, and lower plot detected vibrato extent. Highlighted section shows *tahrir*.

the results were plotted against the pitch track (see Fig. 5 for an example). These results, stored as csv files, together with the plots for all recordings are available in the code repository. Through visual evaluation of the results we argue that further research is required in order to improve automatic detection of vibrato for this tradition.

Dastgāhi music is an idiomatic tradition. Vocal and instrumental performances are characterized by specific techniques and sound qualities, being vibrato a relevant one. Its extent, frequency and occurrence are idiomatically specified by instrument, school or personal style. Automatic vibrato detection and measurement can contribute to characterizing idiomatic styles. Another important vocal technique is *tahrir*, which consists in wide cyclic pitch alteration of a particular tone in terms of pitch and timber. The pitch fluctuation is not realized as glissandi but jumps, and this might explain why it is not detected by the used algorithm, as shown in Fig. 5. Analysing the difference between these two techniques would help to understand the idiomatic practices in Iranian *dastgāhi* music.

5. CONCLUSIONS AND FUTURE WORK

With KDC we set a very ambitious goal: to create an open, well curated, ever growing corpus of Iranian *dastgāhi* music data, metadata and expert annotations which can contribute to the development of a coherent line of computational research for this music tradition. This is a task that requires collaborative efforts from musicologists, musicians and computer engineers. KDC is still in its infancy, but its first version presented in this paper is complete enough to test state-of-the-art methodologies, obtain preliminary results, and consequently propose new directions for future research.

The expansion of KDC is and is going to be a constant future task. As any other aspect of our lives in the past two years, the COVID-19 pandemic has hindered the development of KDC with the suspension of international travels (economic sanctions against Iran prevent working with musicians living in the country from outside). Besides the addition of recordings that cover new *gušes*, instruments and schools, increasing the number of expert annotations is one of our short term goals. In a first stage, we aim at having the 5 renditions of the whole set of 12 *daramād*s annotated at least by three different musicians, so that we can start studying this phenomenon from the perception of the performers.

Regarding computational analysis, a short term goal is to seek collaborations to perform automatic pattern analysis. This is a fundamental element for the characterization of *gušes*, and most of our collaborator musicians have highlighted the importance of this task and their interest in it. The annotations currently existing in KDC are aimed to this goal. Other tasks that are of our interest, and for which collaboration with computer engineers is also necessary, is the improvement of pitch track extraction for instruments with great resonance as *setār*. Finally, collaboration will also be sought for the development of methods for automatic *tahrir* detection and analysis.

6. ACKNOWLEDGEMENTS

The authors would like to express their sincerest gratitude and acknowledgement to their collaborator musicians Masoud Shaari, Farahnaz Sahebghalam, Reza Zalpour and Mohammad Khodadadi, not only for contributing their recordings and ideas to KDC, but also for their enthusiastic support to the project. We would also like to thank Dr. Negar Bouban and Reza Zalpour for their very valuable manual annotations of melodic patterns. Finally, we thank Kunstuniversität Graz for the financial support that makes KDC possible.

7. REFERENCES

[1] H. Farhat, *The Dastgāh Concept in Persian Music*. Cambridge and New York: Cambridge University Press, 1990.

[2] T. Tsuge, "Rhythmic aspects of the Âvâz in persian music," *Ethnomusicology*, vol. 14, no. 2, pp. 205–227, 1970.

[3] F. Amoozegar-Fassie, "The poetics of persian music: The intimate correlation between prosody and persian classical music," Ph.D. dissertation, Faculty of Graduate Studies, University of British Columbia, UK, 2008.

[4] M. Azadehfar, *Rhythmic structure in Iranian music*. Tehran: Tehran University Press, 2017.

[5] N. Bouban, "Barresi-ye Āvāi-ye ritm dar pāyehā-ye vājegan-e š'er va nasr-e fārsi-ye rasmi [the phonetic study of rhythm in units of poems and writings in standard farsi]," *Majaleh-ye pajuheš-hā-ye zabānšenasi*, vol. 1, no. 1, pp. 101–122, 2010.

[6] P. Heydarian, "Music note recognition for santoor," Master's thesis, Tarbiat Modarres University, Iran, 2000.

[7] P. Heydarian and J. D. Reiss, "The persian music and the santur instrument," in *Proc. of the 6th Int. Society for Music Information Retrieval Conf.*, London, UK, 2005, pp. 524–527.

[8] P. Heydarian and L. Jones, "Measurement and calculation of the parameters of santur," in *Proc. of the Annual Conf. of the Canadian Acoustical Association (CAA)*, Vancouver, Columbia, 2008.

[9] ——, "Tonic and scale recognition in persian audio musical signals," in *12th Int. Conf. on Signal Processing (ICSP)*, Hangzhou, China, 2014.

[10] P. Heydarian, L. Jones, and A. Seago, "Automatic mode estimation of Persian musical signals," in *133rd Audio Engineering Society Convention*, San Francisco, USA, 2012.

[11] P. Heydarian, "Automatic recognition of persian musical modes in audio musical signals," Ph.D. dissertation, Faculty of Art, Architecture and Design, London Metropolitan University, UK, 2016.

[12] P. Heydarian and D. Bainbridge, "Dastgàh recognition in Iranian music: Different features and optimized parameters," in *6th International Conference on Digital Libraries for Musicology (DLfM '19), November 9, 2019, The Hague, Netherlands.*, 2019, pp. 53–57.

[13] N. Darabi, N. H. Azimi, and H. Nojumi, "Recognition of dastgah and maqam for persian music with detecting skeletal melodic models," in *Proc. SPS-DARTS 2006-The second annual IEEE Benelux/DSP Valley Signal Processing Symposium*, Antwerp, Belgium, 2006.

[14] S. Abdoli, "Iranian traditional music dastgah classification," in *Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011*, 2011, pp. 275–280.

[15] M. A. Layegh, S. Haghipour, and Y. N. Sarem, "Classification of the radif of mirza abdollah a canonic repertoire of persian music using svm method," *Gazi University Journal of Science*, vol. 1, no. 4, pp. 57–256, 2013.

[16] L. Ciamarone, B. Bozkurt, and X. Serra, "Automatic Dastgah Recognition Using Markov Models," in *Proceedings of the 14th International Symposium on Computer Music Multidisciplinary Research*, 2019, pp. 51–58.

[17] S. Shafiei, "Analysis of vocal ornamentation in iranian classical music," in *Proc. of the 16th Sound and Music Computing Conf.*, Málaga, Spain, 2019, pp. 437–441.

[18] F. Sanati, "An investigation on the value of intervals in persian music," Master's thesis, Department of Music, University of Jyväskylä, Finland, 2020.

[19] S. Malekzadeh, M. Samami, S. R. Azar, and M. Rayegan, "Classical music generation in distinct dastgahs with alimnet ACGAN," *CoRR*, vol. abs/1901.04696, 2019.

[20] P. Heydarian and J. D. Reiss, "A database for persian music," in *Proc. of the Digital Music Research Network Summer Conference (DMRN 2005)*, Glasgow, UK, 2005.

[21] X. Serra, "Creating research corpora for the computational study of music: the case of the compmusic project," in *53rd AES International Conference on Semantic Audio*, New York, USA, 2014, pp. 1–9.

[22] M. Mauch and S. Dixon, "pYIN: A fundamental frequency estimator using probabilistic threshold distributions," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014)*, Florence, Italy, 2014, pp. 659–663.

[23] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]," 1992–2022. [Online]. Available: https://www.praat.org

[24] J. W. Kim, J. Salamon, P. Li, and J. P. Bello, "Crepe: A convolutional representation for pitch estimation," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, 2018, pp. 161–165.

[25] J. Salamon and E. Gómez, "Melody extraction from polyphonic music signals using pitch contour characteristics," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, pp. 1759–1770, 2012.

[26] C. Cannam, C. Landone, and M. Sandler, "Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files," in *Proceedings of the ACM Multimedia 2010 International Conference*, Firenze, Italy, October 2010, pp. 1467–1468.

[27] G. K. Koduri, V. Ishwar, J. Serrà, and X. Serra, "Intonation analysis of rāgas in carnatic music," *Journal of New Music Research*, vol. 43, pp. 73–94, 2014.

[28] L. Nooshin, "The song of the nightingale: Processes of improvisation in dastgāh Segāh (Iranian classical music)," *British Journal of Ethnomusicology*, vol. 7, no. 1, pp. 69–116, 1998.

[29] D. Talai, "A New Approach to the Theory of Persian Art Music: The Radīf and the Modal System," in *Garland Encyclopedia of World Music Volume 6: The Middle East*, V. Danielson, S. Marcus, and D. Reynolds, Eds. New York: Routledge, 2001, pp. 894–903.

[30] H. Asadi, "Theoretical foundation of persian classical music: Dastgāh as a muti-modal cycle," *Mahoor*, vol. 6, pp. 43–56, 2004.

[31] D. Bogdanov, N. Wack, E. Gómez Gutiérrez, S. Gulati, H. Boyer, O. Mayor, G. Roma Trepat, J. Salamon, J. R. Zapata González, and X. Serra, "Essentia: An audio analysis library for music information retrieval," in *Proc. of the 14th Int. Society for Music Information Retrieval Conf.*, Curitiba, Brazil, 2013, pp. 493–498.