Automatic musical transcription of the Tatar folk song: comparative analysis of AI-powered programs

Liliya Zelimkhanovna Borodovskaya, Ziliya Mukhtarovna Yavgildina*, Elena Aleksandrovna Dyganova, Larisa Stanislavovna Maykovskaya, Irina Aleksandrovna Medvedeva

Assoc. Prof., Ethno-artistic Creativity and Music Education, Higher School of Arts Kazan State Institute of Culture, Kazan, Russia. ORCID: 0000-0002-5680-1187 Email: lilianotka@yandex.ru
* Corresponding Author: Prof., Vice-rector on Scientific work Department of Ethno-artistic Creativity and Music Education, Kazan State Institute of Culture, Kazan, Russia. ORCID: 0000-0002-4193-6126 Email: zilia.javgi@gmail.com
Assoc.Prof., Department of Tatar Studies and Cultural Studies, Kazan Federal University, Kazan, Russia. ORCID: 0000-0003-2875-5109 Email: dirigerdea@mail.ru
Prof., Department of Music Education, Faculty of Arts Moscow State Institute of Culture, Khimki, Moscow region, Russia. ORCID: 0000-0002-7144-0260 Email: maykovskaya@gmail.com
Prof., Dean of the Faculty of Art and Music Education, Yakovlev Chuvash State Pedagogical University, Cheboksary, Russia. ORCID: 0000-0003-3132-4078 Email: medvedevaia@gmail.com

DOI 10.12975/rastmd.20221018 Submitted January 28, 2022 Accepted March 9, 2022

Abstract

This article is relevant due to the loss of the carriers of folk music that needs to be recorded in digital audio formats and requires music transcription for the subsequent creation of collections for the purposes of scientific research by ethnomusicologists. The study aims at determining the need to use software for the automatic music transcription of audio recordings of folk music. The main research method is the comparative analysis of the music transcription of the Tatar Kryashen songs performed by people and three AI-powered programs (Celemony Melodyne, AudioScore Ultimate and Cubase). Then we compared the scores we prepared and the visual data of three programs: wave, spectral, "piano roll" and traditional music scores. According to five evaluation parameters (the accuracy of displaying a melody, rhythm, key, time signature and subjective assessment), the Cubase program was recognized as the most user-friendly. It is still controversial whether to use artificial intelligence for the music transcription of folk songs since music researchers decide for themselves. The undoubted benefit of the automatic music transcription of folk music is the rapid analysis of audio recordings, the ability to create more music notations in a shorter time, assist in the analysis of fragments that are difficult to hear by ear and restore damaged audio recordings.

Keywords

automatic music transcription, computational ethnomusicology, Tatar folk song

Introduction

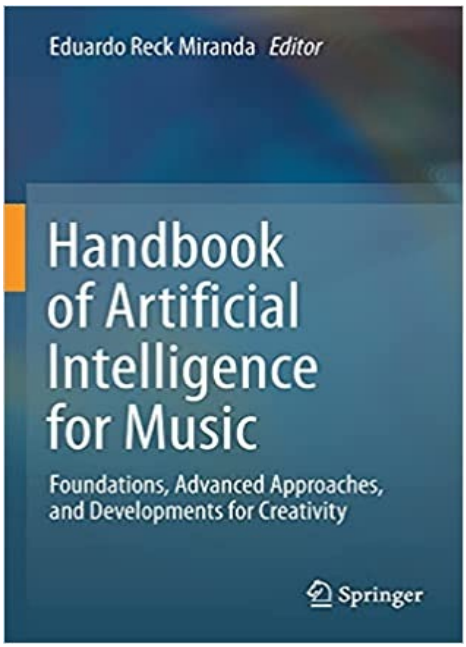
The article is relevant since numerous folk music archives located in various state scientific institutions and personal collections of folklorists need to be digitized. Musical transcriptions and catalogs are necessary for further scientific, educational and creative use by a wide range of people. The high value of folk music within cultural heritage is evidenced by the fact that folk songs serve as sources of different styles and trends in modern music (Briot, 2021).

Nowadays mass musical culture is in a state of degradation or "melodic stagnation" (Zhang, 2020). Successful creative projects are often based on a modern interpretation of folk songs, which distinguishes them from one-day hits. This is associated with the lack of music collections and teaching aids on the music heritage of different peoples that could be used by soloists and ensembles, as well as students of creative specialties (Pavlov, 2021). Published notes have been mastered by singers and have already entered their repertoire, thereforearchival audio recordings of folk songs should be deciphered and brought to the consumer in a musical format (Grebosz-Haring & Weichbold, 2020).

Within the framework of this article, we will consider only one of the above-mentioned topics, namely the automatic music transcription of audio recordings of folk songs using artificial intelligence (AI) technologies (Sviyazova, 2019). The thing is that there is a large number of audio recordings (folk music "big data") that need automatic music transcription to speed up processing and to assist folklorists and ethnomusicologists working with songs in audio format.

Literature Overview

Both Russian and foreign scientific studies are concerned with the topic of automatic music transcription but there are not many foreign surveys on working with folklore material. For example, there are only articles by Rao et al. (2012), Kroher and Gomez (2016). A consistent approach in the field of automatic music transcription is demonstrated by the International Society for Music Information Retrieval (ISMIR), which has been holding annual conferences for more than 20 years and publishes scientific collections of articles with free access for review (ISMIR, 2022). Among the main directions of the ISMIR's work is the creation of software for music analysis based on the European 12 equal temperament system. However, these programs do not provide reliable results for vocal folk music since they do not capture all the subtleties of intonation less than a semitone and do not recognize complex rhythm-and-meter patterns of improvisation-based musical genres (Heil, 2017). Therefore, the ISMIR community will continue its work in the direction of automatic music transcription.



Publisher: Springer; 1st ed. 2021 edition (July 3, 2021)

Editor: Prof. Eduardo Reck Miranda

Title: Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity

Language: English

Hardcover:1024 pages

ISBN-13: 978-3030721152

This book presents comprehensive coverage of the latest advances in research into enabling machines to listen to and compose new music. It includes chapters introducing what we know about human musical intelligence and on how this knowledge can be simulated with AI. The development of interactive musical robots and emerging new approaches to AI-based musical creativity are also introduced, including brain-computer music interfaces, bio-processors and quantum computing.

Figure 1. "Handbook of Artificial Intelligence for Music" Book (Amazon.com, 2022)

Summarizing the scientific issues of automatic music transcription and the use of artificial intelligence technology in music, "Handbook of Artificial Intelligence for Music" was published in 2021 (Miranda, 2021). This book contains all the most relevant and recent world research in the field of artificial intelligence for musicians: a wide range of scientific topics affecting sociological, philosophical and musicological issues; music cognition and perception; improvisation and composition; orchestration and studio production; sound synthesis and signal processing; music transcription. This guide shows how artificial intelligence technologies have entered the music field by modern scholars, programmers, biologists and engineers, which reveals interdisciplinary perspectives in this digital sphere (Turchet et al., 2018).

In relation to AI-based automatic music transcription in this "Handbook", we would like to highlight a long article by Lele Liu (2021) "From Audio to Music Notation". It examines in detail the history of automatic music transcription, music information retrieval (MIR) and the current state of automatic music transcription (Liu & Benetos, 2021). While exploring different AI-based methods of automatic music transcription, Lele Liu called convolutional neural networks (CNN) the most promising. The author claimed that there were unresolved problems of AI-based automatic music transcription, which is in line with the main topic of this article: "Although AMT is still very active as a topic within MIR, the performance of current AMT systems is still far from satisfactory, especially when it comes to polyphonic music, multiple instruments, *non-Western music and 'complete' transcription*" (Liu & Benetos, 2021: 707-708).

Speaking about the music transcription of "non-Western music", many scholars mean vocal folklore and conclude that modern AI-driven software cannot deal with complex rhythmic and metric structures, melismatic vocals, deciphering nuances of expressive performance or making full music notations (Kroher & Gómez, 2016: 901; Liu & Benetos, 2021: 709).

Among some studies and achievements in the field of automatic music transcription, it is necessary to highlight the works of Arakelyan (2011), Gorbunova (2016), Gubaidullin (2016), Danshina and Fillipovich (2014), Kharuto (2019, 2020), Yunusova (2020), Yakimuk (2016, 2020), etc. In their works, these scientists and developers present the results of research in the field of automatic music transcription, both based on AI-powered software (Speech Analyzer, Praat) and their own inventions (SPAX, Yunusova & Kharuto, 2020) in this area.

unusova and Kharuto (2020) represented a relatively new scientific field – Computational Ethnomusicology (CE) – that covers several areas, including automatic music transcription. In general, the article by Yunusova and Kharuto is a fundamental analysis of the historical development and the current state of automatic music transcription in the field of traditional folk music. The authors called *computer-assisted music oriental studies* a separate direction in the Computational Ethnomusicology in Russia, including this study on the automatic music transcription of the Tatar folk songs (Yunusova & Kharuto, 2020).

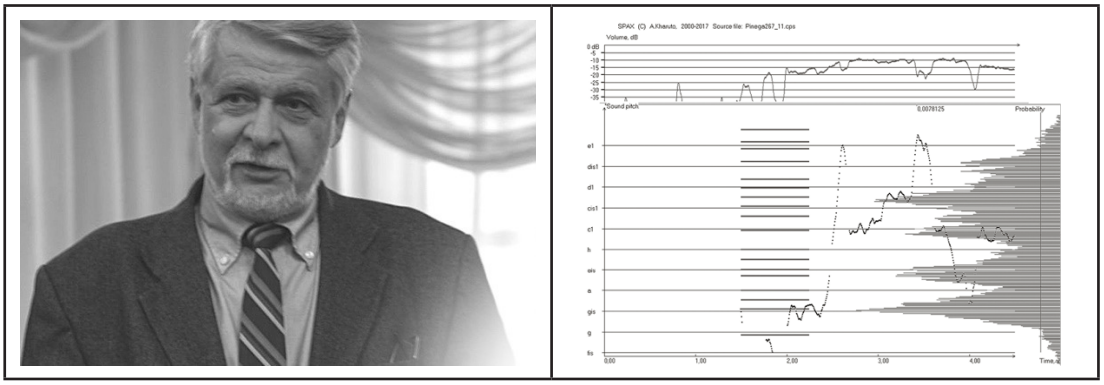


Figure 2. A. V. Kharuto and SPAX program developed by him/her

A certain scientific and technological breakthrough was made by Kharuto (2019). He developed the SPAX program to decipher a monophonic example of a folk song belonging to a "non-European" 12 equal temperament system with high accuracy, when the exact type of scale, temperament and interval structures are not set. We find it important that this program does not perform complete automatic music transcription but serves only as an intermediate result for the study and notation of a folk song as the authors mentioned (Kharuto, 2019).

Each of the above-mentioned authors analyzed one of the existing problems in the field of automatic music transcription. Thus, a group of scientists (A.A. Konev, A.Yu. Yakimuk, A.O. Osipov) "developed a software package capable of working with audio recordings of vocal performances in wav format and recordings made through the interface" (Konev et al., 2015; Yakimuk, 2020). However, the practical development of A.Yu. Yakimuk is not associated with automatic music transcription for the purposes indicated in this article, therefore it can only be an additional tool. The main objective of the software developed by A.Yu. Yakimuk is to improve the training of vocalists in order to practice off-key singing.

It is worth mentioning that there is no specialized AI-driven software for the complete music transcription of audio recordings of folk songs and ethnomusicologists have to use various programs in their work. It is noteworthy that many scientific works used the Melodyne program (Celemony Melodyne) (Celemony, n.d.) within comparative research methods. This software is not intended for automatic music transcription but is used in studio work as a tool for sound processing and correcting performance errors. Although there are many specialized programs for automatic music transcription and the effectiveness of their work is described in scientific papers (Sviyazova, 2019), the Melodyne program remains one of the most popular among ethnomusicologists. We consider the results of song analysis as exemplified by this and other programs, which will give the opportunity to identify the pros and cons of using artificial intelligence in the process of automatic music transcription.

Objective of Study

The article aims at conducting a comparative analysis of several AI-driven programs used for the automatic music transcription of the one-voiced folk song of the Tatar Kryashens "Och Bagana" ("Three Pillars"). The results should reveal pros and cons of each program, and help musicians choose the best digital assistant for the musical transcription of folk songs powered by AI technology. The main criterion for choosing such a program is to facilitate musical transcription in the process of creating folk music scores and singing tutorials. Thus, it is necessary to make not complex ethnomusicological and analytical notations but simpler forms for a wide

range of musicians. The collection, storage, processing and use in education and musical creativity of numerous folk songs written by various peoples of the Volga region is one of the tasks set for students and teachers of the Department of Ethno-Artistic Creativity and Music Education of the Kazan State Institute of Arts and Culture.

Importance of the Study

Based on the above-mentioned objective, we set the following tasks:

- To process a digitized audio recording in wave format using such programs as AudioScore Ultimate, Celemony Melodyne and Steinberg Cubase with the output in a MIDI file for the final editing of the score in the Sibelius program;

- To compare the results of "human" music transcription with automatic music transcription in three programs according to the criteria of maximum accuracy of melody and rhythm; the ability to determine scale and time signature;

- To determine pros and cons of each program in the process of automatic music transcription and their possible use in the mass processing of folk songs at the stage of finalizing music collections for publication.

The theoretical results will be useful for ethnomusicologists working with large arrays of audio recordings belonging to folk music in the process of their analysis and notation, as well as software developers in the sphere of AI technology to understand the related problems and needs.

The practical results can be used in music colleges and universities within the framework of such disciplines as "Collecting and transcribing folk music" and "Music informatics" as an example of using artificial intelligence technologies in the educational process. As a modern pedagogical approach, one can use the competitive method we proposed (Dyganova et al., 2017). For example, we divided students into groups to analyze songs in two or three different programs, compared the results on the speed and quality of music transcription. Thus, we would create the effect of competitive training in studies and research activities.

Thus, the tasks of this study lie in several related scientific fields, including computational ethnomusicology (computer-assisted music oriental studies), automatic music transcription, as well as artificial intelligence technologies for musicians and teachers.

Methods

Research Model

In this research, phenomenology and case study, which are qualitative research methods, were used together. According to this paradigm, the experience of the experts in the field of musicology with regard to existing software and these experiences are compared according to certain criteria. In the research, it was made to compare the performances of the software suitable for western music against Tatar folk songs and to describe the current situation from different perspectives. The authors did the analysis of visual music display in three popular programs having free versions: AudioScore, Melodyne and Cubase. The following visual forms of display have been chosen: wave, spectral, "piano scroll", traditional music score and notation. These formats are most suited for a comparative analysis of programs with the results of "human" music transcription since the main objective of this study is to select the most convenient digital tool to assist an ethnomusicologist.

Procedure

The practical scientific basis for the study was the digital music laboratory of the Kazan State University of Culture and Arts and folk songs recorded from their carriers by a Bachelor student Lyubov Mikhailovna Timofeeva, who graduated in the direction of "Artistic Folk Culture" under the profile "Ethnocultural Center Administration" in 2021. For example, we selected the "Och

Bagana" song ("Three Pillars") recorded from a resident of the village of Staroe Tyaberdino of the Republic of Tatarstan, Galina Efremovna Medova (born in 1961). This is a song of the Tatar Kryashens in a drawling manner sung by unmarried girls during the summer games of the Pitrau holiday (Peter's Day).

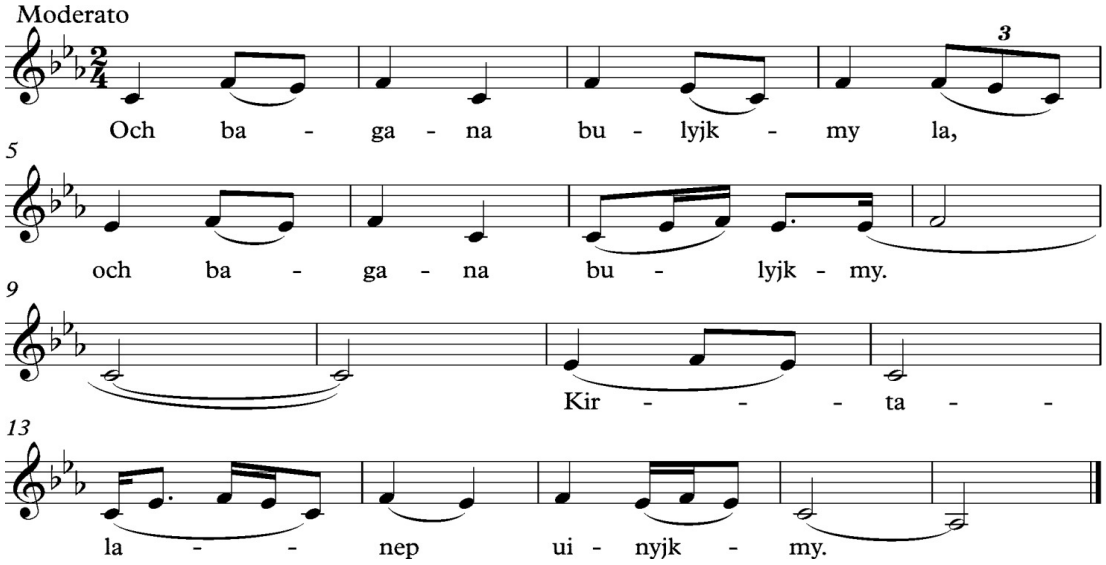
Considering the rhythm-and-meter complexity of such tunes, the presence of melisma and improvisation, we face a rather difficult task to reveal the capabilities of three programs. The automatic detection of a digitized audio recording comprises the following file formats: the original audio recording of the song in wave format, the final output in MIDI file and the upload of MIDI file to the Sibelius program (or any other notation editor, for example, MuseScore).

At the preliminary stage, we have not edited out any overtones and noises from the audio recording. This is also an important point in evaluating the functioning of programs since the singer's breathing or extraneous noise during recording manifest themselves in spectrograms that cannot remove useless sound signals. This folk song is presented as a "field" recording, i.e. recorded on a smartphone in a lively atmosphere.

Results

Variety of AI-powered Programs

The automatic music transcription of the Tatar Kryashen lingering folk song "Ech Bagana" ("Three Pillars") conducted in each AI-powered program should be compared with the notation we performed. It is worth mentioning that our notation was not made in an analytical form but rather a simpler format suitable for a wide range of musicians and students of music schools.



Notation 3. The notation of the "Och Bagana" song we compiled

The song lyrics can be literally translated in the following way:

"We'll have three pillars, we'll have three pillars
Let us play with this obstacle;
Let all the girls get together, let all the girls get together
Shall we play until dawn?"

Cubase 5 Program

Firstly, we will analyze this song in the world's most popular Cubase 5 program (Steinberg, n.d.). Although there is already version 11, we want to show that this software has been able to create scores from audio using artificial intelligence for many years. It is a multi-functional program for studio workers, composers and phonogram producers used to improve the purity of vocal performance. We will use this function called Vari Audio (hereinafter referred to as Pitch & Warp) and is located on the Sample Editor tab. Figure 4 presents the song simultaneously in wave, spectral and "piano roll" formats, which clearly shows each part of the singing. In Figure 4, we see the vocal vibration of the performer. This is a good indicator of automatic music transcription. Although not reflected in the music notation, this can be done in comments.

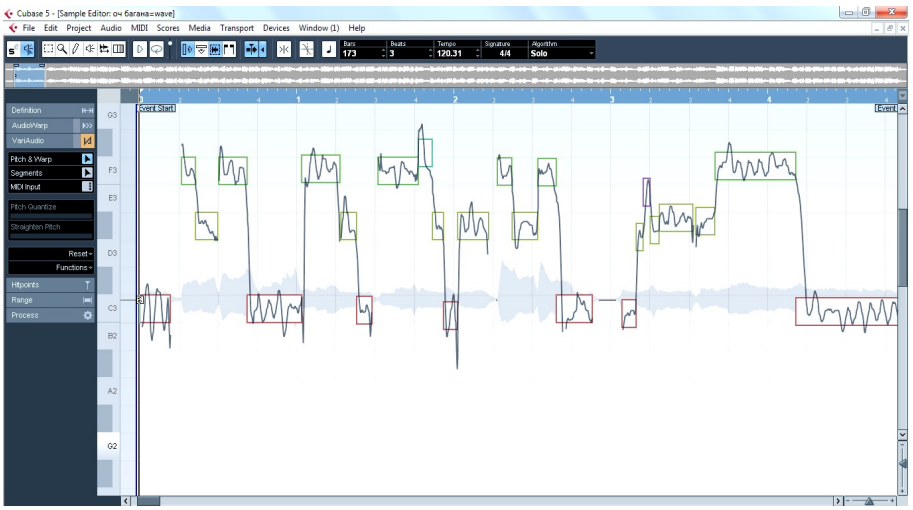


Figure 4. The visual display of the 1st to 10th measures of the "Och Bagana" song in Cubase 5 (wave, spectral and "piano roll" formats)

The next screenshot demonstrates the process of converting the same fragment of the song to MIDI file. For that, the user should click the Functions > Extract MIDI on the Sample Editor tab (Figure 4 on the left).

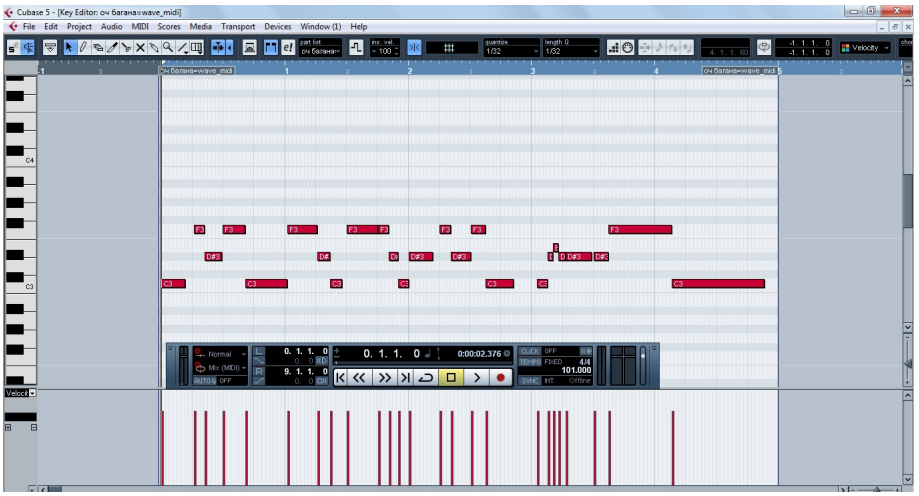


Figure 5. The visual display of the 1st to 10th measures of the "Och Bagana" song in Cubase 5 (MIDI file and "piano roll")

This format is as close as possible to the notation. When the user switches to the Score tab in Cubase, the traditional music notation of the song appears on the screen (Figure 6).

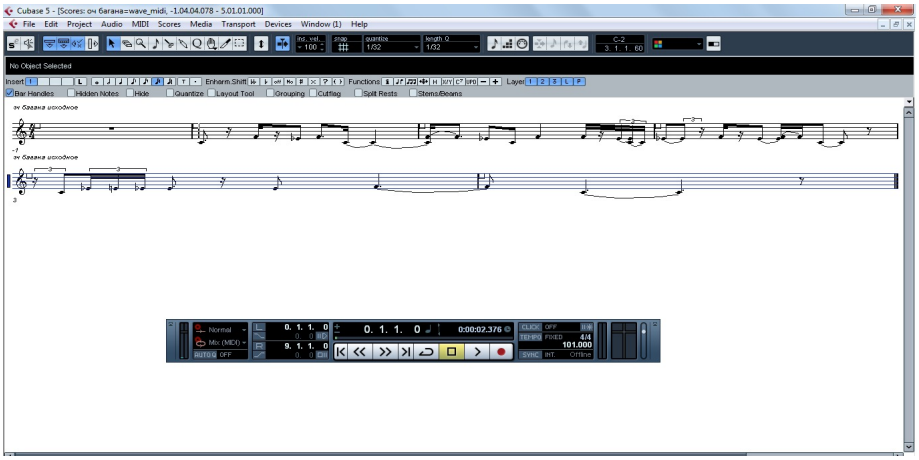


Figure 6. The visual display of the 1st to 10th measures of the "Och Bagana" song in Cubase 5 (MIDI file and "piano roll")

Figure 6 already shows some problems of the notation: the excessive fusion of notes and the difficulty of displaying the rhythm for the subsequent reading of notes. This is due to the special accuracy of measuring the duration of notes in the Cubase program (up to 1/128 of the duration in this version). Automated quantization (aligning rhythm patterns of the song) was not performed. From the viewpoint of music theory, the program is "illiterate" since the key signatures of the song (in C minor) were not set. This function is available in the other programs but this software requires pre-settings of the tone system on the Midi tab.

AudioScore Program

Let us consider the possibilities of AudioScore (Neuratron AudioScore Ultimate) (Neuratron, 2020) in transcribing the same song. In Figure 7, we see that the AudioScore program automatically detects that the given sample is performed by a human voice (see the Voice line at the bottom right). Moreover, this program can detect musical instrument timbres and polyphonic music with up to 16 simultaneous sounds. Rhythmic figures and note values are available up to 1/32 note. The spectrogram in Figure 7 is rather conditional and less detailed than in the Cubase program. It does not reflect voice vibration, which is also significant for automatic music transcription.

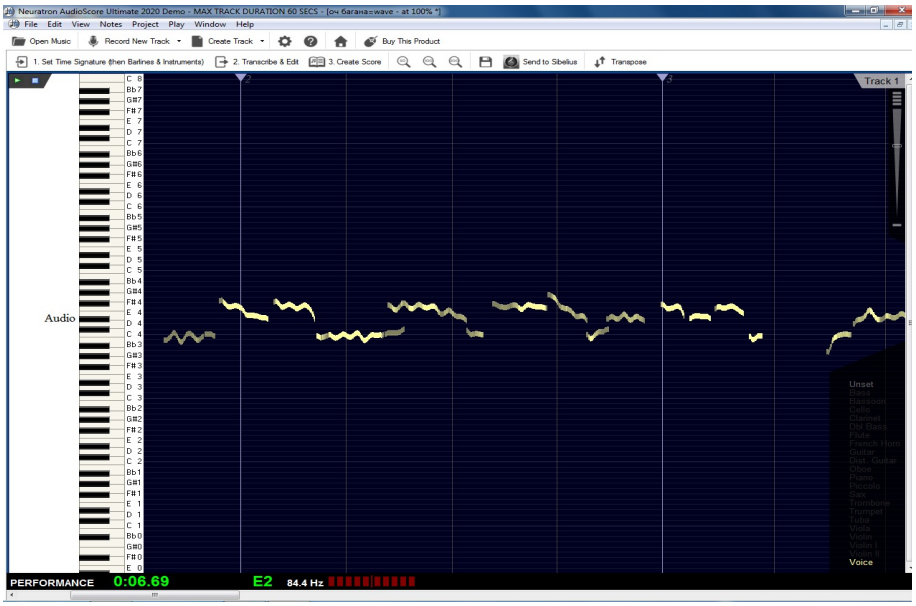


Figure 7. The visual display of the 1st to 10th measures of the "Och Bagana" song in AudioScore (spectral format and "piano roll")

In Figure 8, we can see the "flaws" of the program in the music notation of a monophonic song, where it is recorded as a two-voice song. This happened because AudioScore combined notes that are less than 1/32. They are layered on top of each other like in a two-voice song.

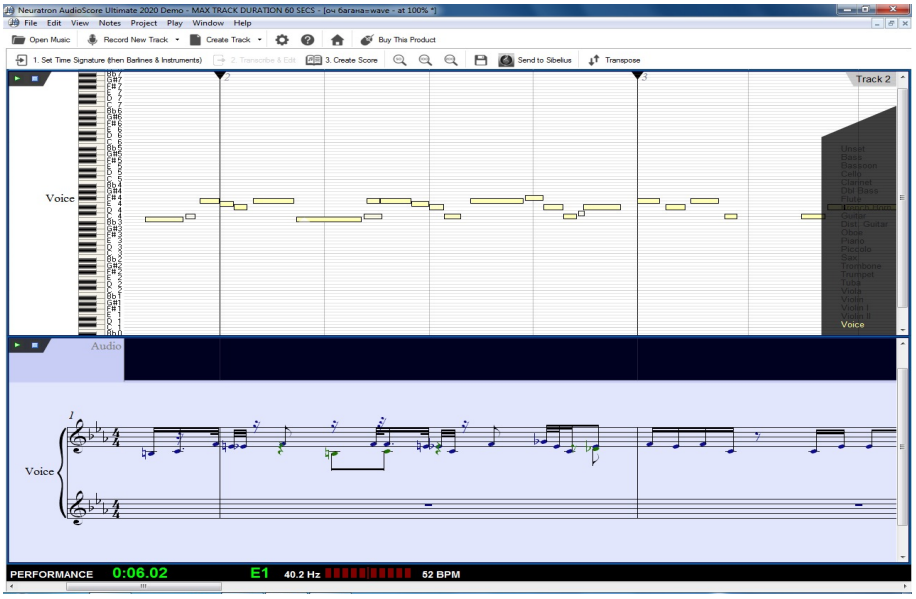


Figure 8. The visual display of the 1st to 10th measures of the "Och Bagana" song in AudioScore (keyboard music sheets and music scores).

Melodyne 4 Program

The results of the Melodyne 4 program (Celemony Melodyne) can be seen in Figure 9. For comparison, the results of Melodyne 3 are shown in Figure 10.

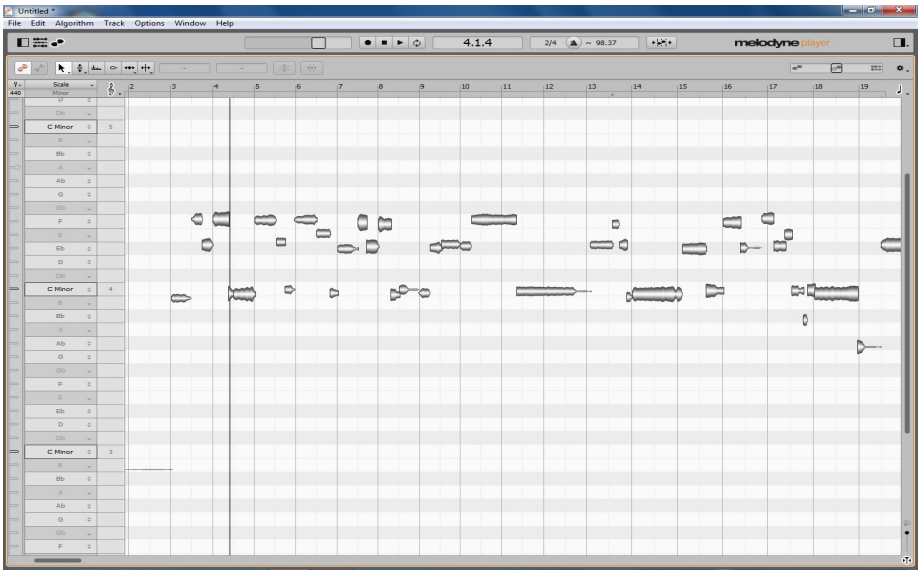


Figure 9. The visual display of the 1st to 17th measures of the "Och Bagana" song in Melodyne 4 ("piano roll" and spectral format)

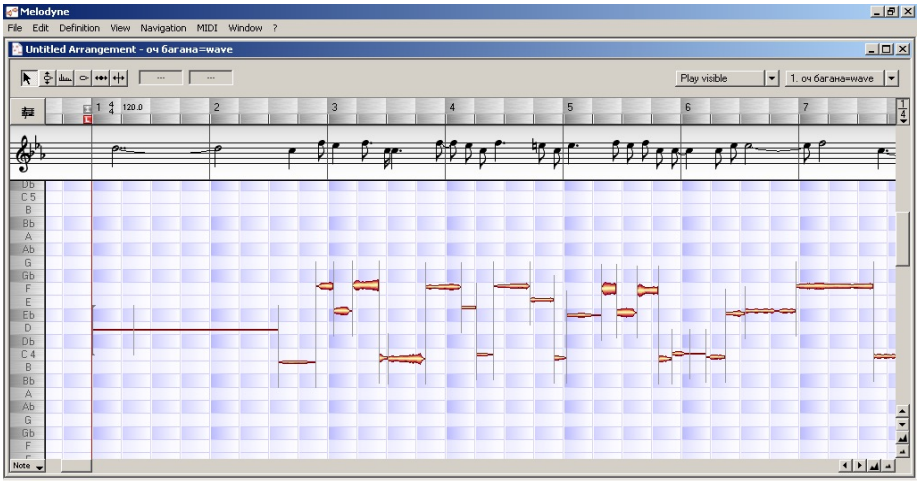


Figure 10. The visual display of the 1st to 10th measures of the "Och Bagana" song in Melodyne 3 ("piano roll" and spectral format)

Figure 9 and Figure 10 demonstrate screenshots from the Melodyne program that is mainly intended for the studio correction of vocal performance. It also copes with automatic music transcription and saves the results in a MIDI file, which is convenient for the further song editing. However, the analysis of notes in this program is limited to 1/32, which significantly reduces the results of automatic music transcription. A distinctive feature of its latest versions is automatic noise reduction, improved analysis of rhythm-and-meter features (bar division), key definition, the ability to analyze chords and polyphony. All this puts the program on a par with the most effective tools used by musicians for automatic music transcription.

Comparison of Software

After using three programs, we compared "human" music transcription and automatic music transcription according to the following criteria: the maximum melodic and rhythmic accuracy; the ability to determine the key and time signature. Our subjective assessment on a 10-point rating scale is presented in Table 1.

Table 1. The comparison of the automatic music transcription of the "Och Bagana" song performed in three programs

Evaluation criterion	Cubase	Melodyne	AudioScore
Melodic accuracy	9	8	7
Rhythmic accuracy	9	8	7
Key definition	0	9	9
Time signature and bar division	9	8	8
Final score	27	33	31
Subjective assessment	Best	Good	Poor

Table 1 clearly shows the results of all the programs. In our opinion, the evaluation criterion "Key definition" is not the most important since a competent musician can easily determine the key using accidentals. In relation to folk songs, this is not a crucial element for automation as there are often modal variability, non-European modal systems and other modal factors typical of folk music that are still unknown to artificial intelligence. In this case, the fact that Cubase did not set chromatic signs is not regarded as a disadvantage. This program has many settings and functions for each editor, including the choice of mode and key.

The above-mentioned screenshots and personal experience indicate that Tatar folk songs are characterized by numerous melismas, short notes, complex rhythms and vocal improvisation. Under these conditions, the Cubase program is deemed the most suitable. The sensitivity of recognition and quantization of note values is 1/128 – this fact provides this software with a competitive advantage over the other two programs. Being a full-fledged virtual studio, this program has the ability to process and edit audio recordings (for example, noise reduction) with a high degree of accuracy before automatic music transcription, which increases the quality of music transcription. The pros of Cubase in the automatic music transcription of a folk song. The user can add audio files of various formats, when the other two programs have limitations (support only mp3 and wave).

The Melodyne program provides the high-quality and detailed transcription of vocal music, displays the key and divides into measures. The only drawback is the inaccurate display of rhythmic features in music transcription (the developers did not make it for ethnomusicological automatic music transcription). If we evaluate programs in accordance with their ease of use, Melodyne takes the top spot. Cubase has a more complex interface, requires a powerful computer and the skills of a musical sound engineer, which is an obstacle for ethnomusicologists.

Discussion

Many scholars assessing software for automatic music transcription prefer using Melodyne for the purposes of comparative analysis because it allowed automatic vocal recognition earlier than Cubase and provided good results (Gubaidullin, 2016; Yakimuk, 2020). The developers of Melodyne are constantly improving their product, and its current version recognizes polyphonic music. In this regard, Melodyne and Cubase

are on the same level, struggling for market leadership.

The AudioScore program is the best tool for transcribing the European 12 equal temperament system, with rhythmic figures of at least 1/32 notes. It recognizes polyphony and divides the score into voices. In relation to the Tatar Kryashen folk song, the program failed to provide automatic music transcription in a suitable form.

One of the practical issues that ethnomusicologists might face when selecting one of the above-mentioned programs is the high price for a license, varying from $100 to $2,000 for different versions. Their free demo versions are limited to one month.

According to the main criterion for choosing a digital assistant to speed up and facilitate music transcription for a person creating collections of folk music and teaching aids for singers, all three programs can be used with due regard to the above-mentioned disadvantages. We have been conducting similar studies for several years, and other works on this topic are presented in the article (Borodovskaja, 2021).

Conclusion

Summing up the study results, we have determined pros and cons of using AI-powered programs for the music transcription of the Tatar folk music. The traditional "human" notation of folk songs is a rather complicated and time-consuming process that requires a good ear for music and literacy in the field of music theory, harmony and solfeggio. However, there are such complex audio examples of folk songs in which small notes of melismas are barely perceptible to the ear, performance-specific rhythmic and other musical nuances are very difficult to catch or there are damaged fragments in an audio recording. In this case, AI-powered programs serve as digital assistants in the field of computational ethnomusicology. The visual display of a song in the form of a spectrogram and a MIDI keyboard allows to "see" difficult parts of the song and helps to edit notes. The pros of such programs include the high speed of audio processing in the process of automatic music transcription, which is required for a large number of folklore archives.

Artificial Intelligence is playing huge role in music creation and music processing. Computational ethnomusicology has great prospects for development when digital technologies are rapidly changing the world and all areas of human life. AI-based technologies for processing big music data (folk song recordings) need joint efforts of both programmers and musicians.

The cons of AI-driven software for music transcription are the need for basic "music computer" skills; the high cost of licensed programs and equipment for computational ethnomusicology; the inaccuracy of automatic music transcription. It is much easier and faster for many musicians to notate an audio recording of a folk song using their hearing, mind and knowledge.

The research objective was fully achieved. Using AI-driven programs, we have determined some melismas (hard to hear by ear) in the 7th and 13th measures of the Tatar Kryashen song "Och Bagana" ("Three Pillars") and added them to the final score. The comparative analysis demonstrated that Cubase or Melodyne should be selected as digital assistants for the purposes of automatic music transcription since they fully reflect all the vocal nuances in a spectrogram. The Cubase program showed the best results in the visualization of notes in the process of music transcription. The final decision on the use of programs for automatic music transcription rests with musicians.

Prospects

This research will be useful for scholars studying the folk music of the Eastern peoples, whose melodies are often complex in terms of rhythm and melody. The

experience of using three programs can be beneficial for Music Informatics classes where students of music universities process the results of folklore expeditions.

Limitations

Computational ethnomusicology has great prospects for development when digital technologies are rapidly changing the world and all areas of human life. AI-based technologies for processing big music data (folk song recordings) need joint efforts of both programmers and musicians. A significant limitation is a fact that most programs for automatic music transcription are made by European developers and aim at samples of European and popular music. Furthermore, they are expensive and inaccessible to Russian and Eastern ethnomusicologists. At the present stage, AI-driven software for automatic music transcription cannot accurately score a folk song in comparison with a person, which is the main factor limiting the use of such digital assistants.

Acknowledgments

This paper is performed as part of the implementation of the Development Program of the Kazan State Institute of Culture and as part of the implementation of the Kazan Federal University Strategic Academic Leadership Program.

Contributions of the Authors

Liliya Zelimkhanovna Borodovskaya and Ziliya Mukhtarovna Yavgildina - conceived of the presented idea, developed the theory and performed the computations. Elena Aleksandrovna Dyganova and Larisa Stanislavovna Maykovskaya - verified the analytical methods, encouraged to investigate and supervised the findings of this work. Irina Aleksandrovna Medvedeva - contributed to the design and implementation of the research, to the analysis of the results. All authors provided critical feedback and helped shape the research, analysis and manuscript. All authors discussed the results and contributed to the final manuscript.

References

Arakelyan, A.E. (2011). Est li u nekrasovskikh kazakov mnogogolosie? (opyt issledovaniya tekhnicheskimi sredstvami) [Is there the chorus of the Nekrasov Cossacks? (the experience of research through technical means)]. Yuzhno-Rossiiskii Muzykalnyi Almanakh, 2(9), 3-7.

Borodovskaja, L.Z. (2021). Istorija razvitija informatizacii discipliny “Sbor i rasshifrovka muzykalnogo folklore” [The historical development of discipline “The collection and transcription of music folklore”]. In: L.Z. Borodovskaja (Compl.), Mnogogrannyj mir tradicionnoj kultury i narodnogo hudozhestvennogo tvorchestva: Proceedings of the All-Russian scientific conference within the All-Russian competition AR/VR “Hackathon in the sphere of culture”, October 12, 2020, Kazan, Russia (pp. 251-254). Kazan: Kazan State Institute of Culture.

Briot, J. P. (2021). From artificial neural networks to deep learning for music generation: history, concepts and trends. Neural Computing and Applications, 33(1), 39–65. https://doi.org/10.1007/s00521-020-05399-0

Celemony. (n.d.). Melodyne 5. https://www.celemony.com/en/melodyne/new-in-melodyne-5

Danshina, M.V., Filippovich, A.Yu. (2014). Metodika avtomatizirovannoi rasshifrovki znamennykh pesnopenii [The methods of automatic transcription of Znamenny Chants]. Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. N.E. Baumana. Seriya “Priborostroenie”, 4(97), 55-69.

Dyganova, E.A., Yavgildina, Z.M., Shirieva, N.V. (2017). The competition training method in the formation of professional competence of the future music teacher. Man in India, 97(20), 403-414.

Gorbunova, I.B. (2016). Metodicheskie aspekty tolkovaniya funktsionalno-logicheskikh zakonomernostei muzyki i muzykalno-kompyuternye tekhnologii: sistemy muzykalnoi notatsii [The methodological aspects of explaining functional-logical regularities of music and music computational technologies: music notation systems]. Obshchestvo: Sotsiologiya, psikhologiya, pedagogika, 10, 69-77.

Grebosz-Haring, K., & Weichbold, M. (2020). Contemporary art music and its audiences: Age, gender, and social class profile. Musicae Scientiae, 24(1), 60–77. https://doi.org/10.1177/1029864918774082

Gubaidullin, F.F. (2016). Tsifrovaya konvertatsiya analogovykh zapisei selkupskikh instrumentalnykh naigryshei [The digitalization of analog recordings of instrumental folk tunes]. Vestnik Tomskogo gosudarstvennogo universiteta. Kulturologiya i iskusstvovedenie, 1(21), 99-105.

Heil, L. (2017). Teaching Improvisation through Melody and Blues-Based Harmony: A Comprehensive and Sequential Approach. Music Educators Journal, 104(1), 40–46. https://doi.org/10.1177/0027432117711484

ISMIR (2022). International Society for Music Information. Retrieval web: https://ismir.net/

Kharuto, A. (2019). Zvukoryad v traditsionnykh muzykalnykh kulturakh: Vozmozhnosti kompyuternogo analiza [Tone series in traditional musical cultures: Possibilities of computer-assisted analysis]. In: Muzykalyқetnografiyadaғy tұlғa, dəstүr, mədeniet: Aleksandr Zataevich 150 zhyldyғyna arnalғan Khalyқaralyқ ғylymi-praktikalyқ konferentsiyanyң materialdary [Personality, tradition, culture in musical ethnography: Proceedings of the International scientific-practical conference dedicated to the 150th anniversary of Alexander Zataevich] (pp. 77-82). Almaty: Kazakh National Conservatory named after Kurmangazy.

Konev, A.A., Onishchenko, A.A., Kostyuchenko, E.Yu., Yakimuk, A.Yu. (2015). Avtomaticheskoe raspoznavanie muzykalnykh not [The automatic detection of music scores]. Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta, 3, 32-47.

Kroher, N., Gómez, E. (2016). Automatic transcription of flamenco singing from polyphonic music recordings. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24(5), 901-913.

Liu, L., & Benetos, E. (2021). From audio to music notation. In: E.R. Miranda (Ed.), Handbook of artificial intelligence for music (pp. 693-714). Cham: Springer.

Miranda, E.R. (Ed.). (2021). Handbook of artificial intelligence for music: Foundations, advanced approaches, and developments for creativity. Cham: Springer Nature.

Neuratron. (2020). AudioScore 2020. https://www.neuratron.com/audioscore.htm

Pavlov, D. N. (2021). Methods for development of creative expression of college students in music composition. Bulletin of Nizhnevartovsk State University, (1 (53)), 66–73. https://doi.org/10.36906/2311-4444/21-1/09

Rao, B.T., Chinnam, S., Kanth, P.L., Gargi, M. (2012). Automatic Melakarta Raaga Identification System: Carnatic Music. International Journal of Advanced Research in Artificial Intelligence, 1(4), 43-48.

Steinberg. (n.d.). Cubase. https://www.steinberg.net/cubase/

Sviyazova, E.R. (2019). Problema avtomaticheskoi muzykalnoi transkriptsii [The issue of automatic music transcription]. Molodezhnyi vestnik Ufimskogo gosudarstvennogo aviatsionnogo tekhnicheskogo universiteta, 1, 159-162.

Turchet, L., Fischione, C., Essl, G., Keller, D., & Barthet, M. (2018). Internet of Musical Things: Vision and Challenges. IEEE Access, 6, 61994–62017. https://doi.org/10.1109/ACCESS.2018.2872625

Yakimuk, A.Yu. (2016). Algoritmy analiza chastoty osnovnogo tona vokalnogo ispolneniya [The algorithms of analyzing the basic frequency of vocal performance]. In: Nauchnaya sessiya TUSUR-2016: Proceedings of the International scientific conference of students, postgraduate students and young scientists (pp. 245-248). V-Spektr.

Yakimuk, A.Yu. (2020). Otsenka raboty programmnogo kompleksa po raspoznavaniyu not [Assessing the functioning of programs deciphering music scores]. In: I.A. Kurzina, G.A. Voronova (Eds.), Perspektivy razvitiya fundamentalnykh nauk (pp. 141-143). Tomsk: Tomsk State University of Control Systems and Radioelectronics.

Yunusova, V.N., Kharuto, A.V. (2020). Kompyuternaya etnomuzykologiya: Zadachi, metody, rezultaty [Computational ethnomusicology: Tasks, methods, results]. Muzykalnaya akademiya, 3, 162-177.

Zhang, L. (2020). The genre of folk song settings: the ways of interaction of traditional and professional musical art. Aspects of Historical Musicology, 21(21), 113–123. https://doi.org/10.34064/khnum2-21.07