

УДК 004:523

Автоматическое распознавание музыкальных нот*

А.А. КОНЕВ¹, А.А. ОНИЩЕНКО², Е.Ю. КОСТЮЧЕНКО³, А.Ю. ЯКИМУК⁴

¹ 634050, РФ, г. Томск, пр. Ленина, 40, Томский государственный университет систем управления и радиоэлектроники, кандидат технических наук, доцент. E-mail: kaa1@keva.tusur.ru

² 634050, РФ, г. Томск, ул. Гагарина, 7, ООО «Тензор», инженер органа криптографической защиты. E-mail: veganwin@gmail.com

³ 634050, РФ, г. Томск, пр. Ленина, 40, Томский государственный университет систем управления и радиоэлектроники, кандидат технических наук, доцент. E-mail: key@keva.tusur.ru

⁴ 634050, РФ, г. Томск, пр. Ленина, 40, Томский государственный университет систем управления и радиоэлектроники, аспирант. E-mail: yau@keva.tusur.ru

В данной работе рассматривается процесс распознавания музыкальных нот. В качестве входных данных может выступать звуковой файл, в котором находится запись мелодии. Источником звука при записи мелодии может выступать одиночный музыкальный инструмент, их набор, а также голос человека, напевающего мелодию. Входной файл изначально представлен в формате .wav или записывается в этом формате с микрофона. Обработка осуществляется путем последовательного определения частоты основного тона сигнала и преобразованием значения частоты основного тона в соответствующую ноту. Проводится сравнение программ, реализующих алгоритмы выделения частоты основного тона. Специфика исследования заключается в использовании на первом этапе алгоритмов выделения частоты основного тона с учетом особенностей слуховой системы человека, воспринимающей звук. Применение этих алгоритмов, с одной стороны, позволяет достичь более точного определения значения по сравнению с аналогами, в частности с пиковыми методами. С другой стороны, используемые в настоящий момент в модели слуховой системы человека гребенки фильтров имеют ограничение верхней частоты основного тона в 400 Гц, так как приспособлены для работы с обычной речью. Это является недостатком, поскольку при исполнении мелодий, например, оперными певцами частота основного тона может достигать 1400 Гц. Однако, это ограничение не является принципиальным и требует дополнительного исследования на предмет возможности расширения гребенки фильтров, моделирующих слуховую систему человека. Второй этап представляет собой преобразование выделенной частоты основного тона к ноте на основе разработанного алгоритма. Алгоритм базируется на таблицах соответствия частот нотам с учетом их минимальной продолжительности. Представлены результаты тестирования программы, реализующей эти алгоритмы. В заключении представлены основные результаты работы и ее последующее развитие: расширение диапазона допустимых к распознаванию частот.

* Статья получена 01 июня 2015 г.

Работа выполнена при поддержке Министерства образования и науки Российской Федерации в рамках базовой части государственного задания ТУСУР на 2015 г. (проект № 3657).

Ключевые слова: распознавание нот, частота, основной тон, спектр звука, обработка звука, фильтрация звука, вокал, нотная запись

DOI: 10.17212/1814-1196-2015-3-32-47

ВВЕДЕНИЕ

Мир человека наполнен звуками различного происхождения. Среди них голоса людей и животных, звуки технологических приборов и агрегатов, природных явлений, бытовой шум и т. д. Большую часть звуков человек производит сам при разговорах или пении. Выражая свои эмоции, он использует свой голос. Это и стало в древности предпосылками создания музыкальной грамоты – системы отображения звуков на бумаге.

Перед многими непрофессиональными музыкантами стоит проблема перевода своего вокального и музыкального творчества в партитуры (ноты), соответствующие правилам сольфеджио (нотной грамоты). Автоматическое распознавание звучащих нот при помощи специального программного обеспечения позволило бы ускорить и повысить удобство записи партитур. Существует множество решений для преобразования воспроизводимой мелодии в нотную запись [1–3], однако подавляющее большинство из них предназначено для работы с музыкальными инструментами и не способно качественно работать с мелодиями, напеваемыми только голосом. Решения, позволяющие работать только с голосом, тоже существуют [4], однако они предназначены именно для идентификации мелодии, а не для получения нотной записи.

Человеческий голос при пении занимает гораздо больший диапазон частот, нежели при разговоре. Так, голос типичной женщины при разговоре имеет частоту от 165 до 255 Гц, а при пении варьируется от 130 до 739 Гц. Профессиональные певцы имеют голосовой диапазон куда больше, чем обычный человек. В среднем голосовой диапазон составляет 73,91...1396,90 Гц [5].

На основе этого можно сделать заключение, что, с одной стороны, для автоматического преобразования можно использовать стандартные подходы для выделения частоты основного тона; с другой стороны, у таких подходов могут возникнуть проблемы в области верхних частот и может потребоваться модификация алгоритмов выделения частоты основного тона.

1. СУЩЕСТВУЮЩИЕ РЕШЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ НОТ И ОПИСАНИЕ ИСПОЛЬЗУЕМОГО БАЗОВОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Рынок программного обеспечения, предлагающего пользователю функции распознавания частот и нот, достаточно обширен, но не всегда программы соответствуют заявленному качеству или вовсе не справляются с поставленными задачами. Из множества программ были выбраны PRAAT, разработанная в Университете Амстердама [6], SR-0, разработанная в ТУСУР, и Melodine [7], специализирующаяся непосредственно на распознавании нот и разработанная в Мюнхене.

В качестве базовой программы использованы программа исследования речевых сигналов SR0, разработанная в ТУСУРе, а именно один из ее моду-

лей – модуль определения частоты основного тона и его дальнейшая модификация pitch-track.

Учитывая, что базовая программа и использованные в ней алгоритмы определения частоты основного тона [8] были рассчитаны на работу с речью, а не с вокалом, диапазон «видения» разрабатываемой программы на данный момент ограничивается возможностями базового программного обеспечения, а именно от 73 до 400 герц.

Для дальнейшей работы из базового программного обеспечения взят модуль выделения частоты основного тона на основе системы фильтров, использующих особенности слуховой системы человека. В качестве входных данных этого модуля при условии постоянства его параметров (они были выбраны на ранних этапах исследований, связанных с оценкой точности определения частоты основного тона [9, 10], и позволяют получать результаты с погрешностью не более 0,6 % [11], что существенно превышает точность, получаемую на основе пиковых методов, использующих быстрое преобразование Фурье [12] и основанных на применении мел-кепстральных коэффициентов [13] и других классических методов [14]) выступает только звуковой сигнал, подаваемый в формате .wav или приводимый к этому формату в режиме реального времени в процессе записи звука. На выходе этого модуля – частота основного тона в любой (в текущий в случае обработки в режиме реального времени) момент времени. Экранная форма данного модуля Pitch-track представлена на рис. 1.

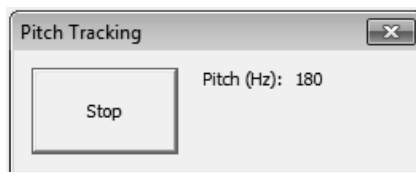


Рис. 1. Экранная форма Pitch-track

При этом модуль может быть использован и без графического интерфейса при использовании его в разрабатываемом программном обеспечении.

2. СРАВНЕНИЕ С АНАЛОГАМИ ЧАСТИ, ОТВЕЧАЮЩЕЙ ЗА ВЫДЕЛЕНИЕ ЧАСТОТЫ ОСНОВНОГО ТОНА И ОБОСНОВАНИЕ ВЫБОРА БАЗОВОЙ ПРОГРАММЫ

Для тестирования выбранных программ было определено несколько стратегий:

- тестирование коротких звуков (стаккато),
- тестирование звуков, находящихся на минимальном интервале,
- тестирование звуков, находящихся в среднем интервале,
- тестирование звуков с произношением слов,
- тестирование мелодии с произношением слов.

По заданным стратегиям был составлен план тестирования. В него входят 9 записей с разными вариациями воспроизведения:

- локация: малая октава. Диапазон: до – си. 12 нот. Интервал: полутон. Длительность: две ноты в секунду;

- локация: первая октава. Диапазон: до – си. 12 нот. Интервал: полутон. Длительность: одна нота в секунду;
- локация: малая октава. Ноты: до, ми, соль, до, соль, ми, до. Длительность: две ноты в секунду;
- локация: первая октава. Ноты: до, ми, соль, до, соль, ми, до. Длительность: одна нота в секунду;
- локация: малая октава. Ноты: до, ми, соль, до, соль, ми, до. Стаккато;
- локация: первая октава. Ноты: до, ми, соль, до, соль, ми, до. Стаккато;
- локация: малая октава. Ноты: до, ми, соль, до, соль, ми, до. Длительность: две ноты в секунду. С произношением названий нот;
- локация: первая октава. Ноты: до, ми, соль, до, соль, ми, до. Длительность: две ноты в секунду. С произношением названий нот;
- мелодия со словами.

Каждая аудиозапись записана женским голосом и имеет следующие параметры:

- расширение *wav*,
- частота дискретизации 12 кГц,
- контрольная сумма 16 бит,
- канал моно.

Примеры полученных результатов представлены на рис. 2–7.

Эталонные частоты для программ Praat и Melodine выделены пунктирной линией.

1) Локация: малая октава. Диапазон: до – си. 12 нот. Интервал: полутон. Длительность: две ноты в секунду.

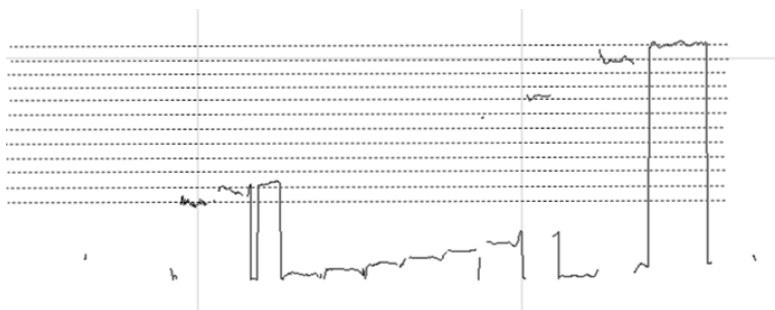


Рис. 2. Тест № 1. Программа Praat

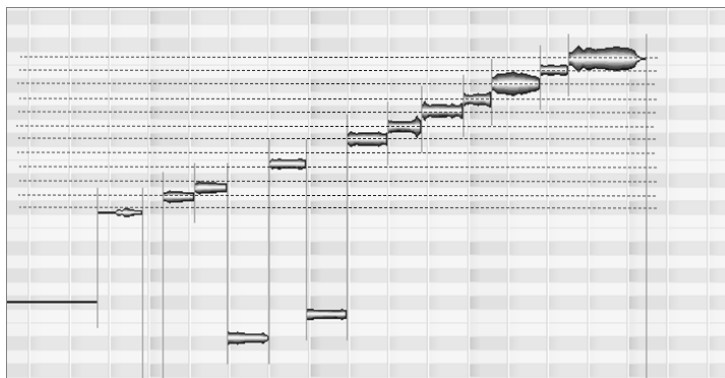


Рис. 3. Тест № 1. Программа Melodine

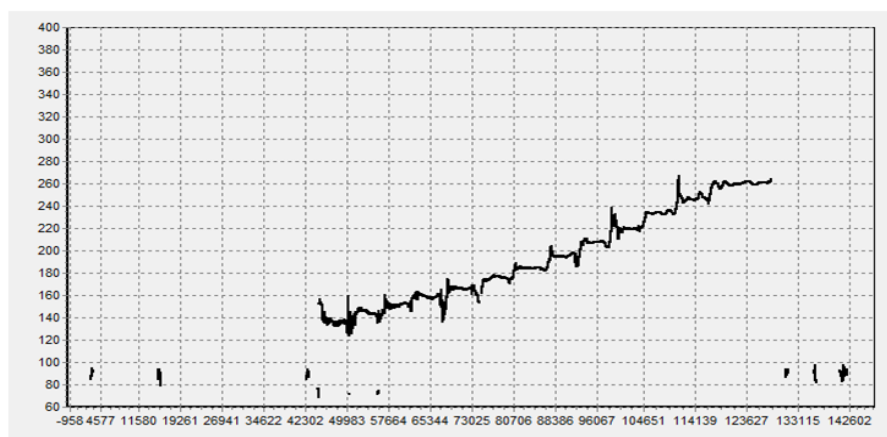


Рис. 4. Тест № 1. Программа Sr-0

2) Мелодия со словами.

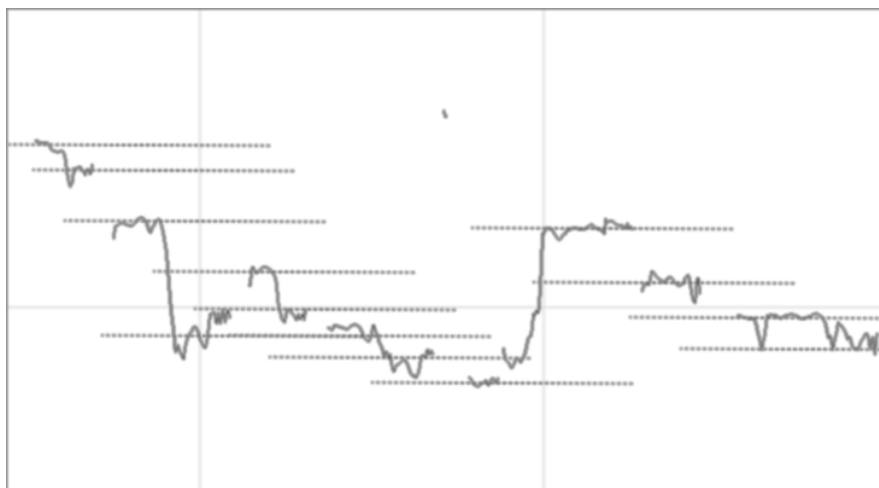


Рис. 5. Тест № 9. Программа Praat

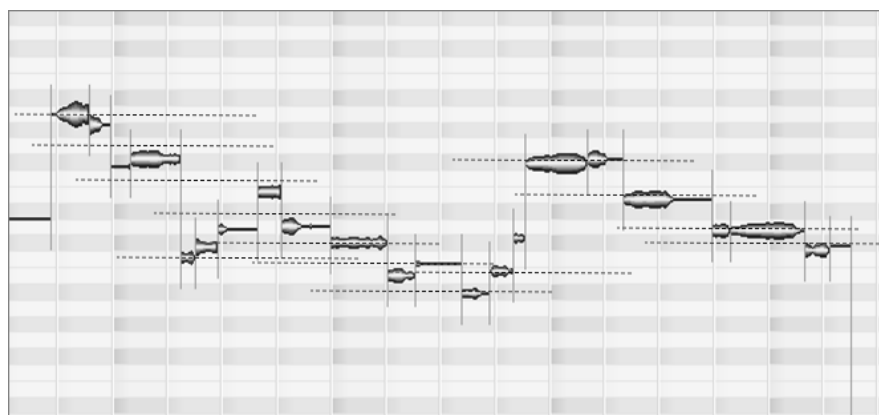


Рис. 6. Тест № 9. Программа Melodine

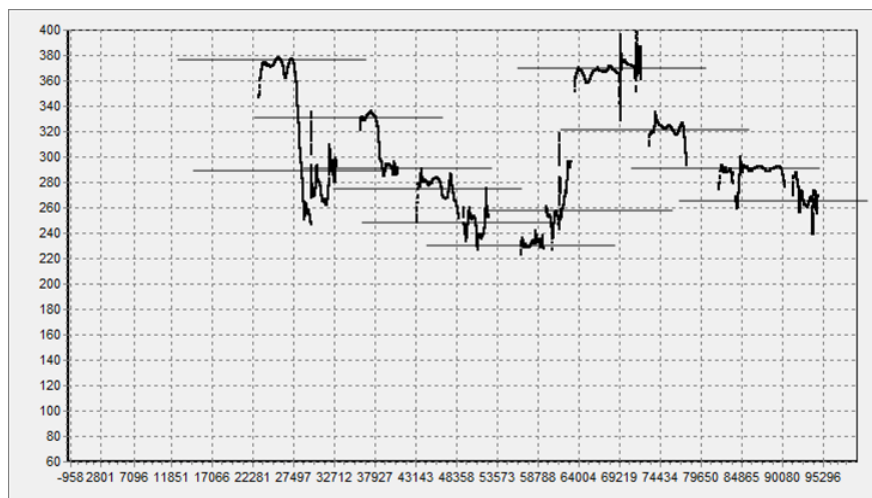


Рис. 7. Тест № 9. Программа Sr-0

В результате исследовательских работ было доказано, что для получения параметров вокального сигнала, необходимых для вычисления частоты тона, лучше всего использовать программный комплекс, реализующий модель периферической части слуховой системы человека. Было сделано 9 аудиозаписей, которые содержали 82 ноты для определения. В месте с продуктом Sr-0, который, собственно, и реализует модель периферической части слуховой системы человека, были протестированы программные комплексы Praat и Melodine. На графике можно увидеть результат (рис. 8). Sr-0 справился с определением нот лучше всех, определив 85 % от общего количества. На основе его и был построен алгоритм определения нот.

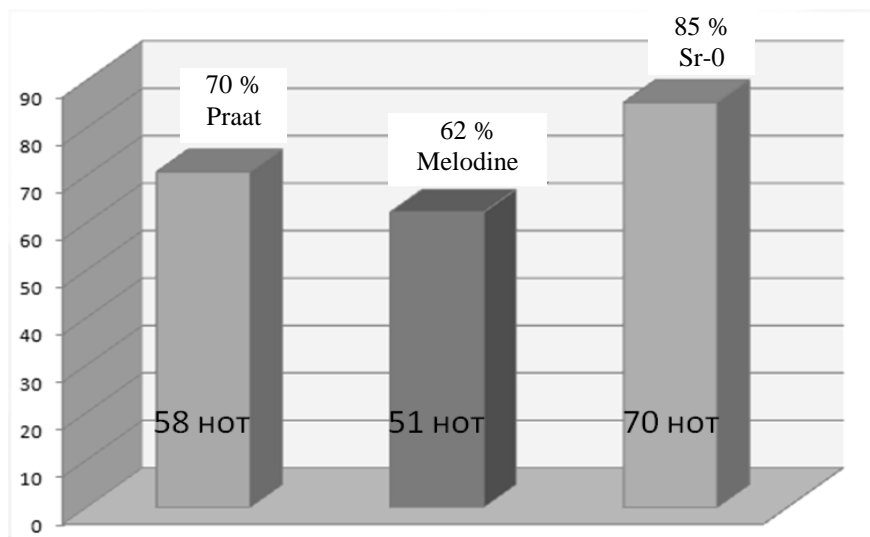


Рис. 8. Результат исследования

3. СТРУКТУРА И ОСНОВНОЙ АЛГОРИТМ РАЗРАБАТЫВАЕМОГО ПРОГРАММНОГО КОМПЛЕКСА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ НОТ

В основу разрабатываемого программного обеспечения положен подход последовательного получения частоты основного тона голоса при помощи базового программного обеспечения и преобразование этой частоты в ноту. Между нотами и частотами существует зависимость. Каждой ноте соответствует определенная частота. Это утверждение не является обратимым – существуют некоторые частоты, которым не соответствуют ноты (таблица) [15].

Соответствие нот частотам

| Частота, Гц | Обозначение ноты | Суб- контр- октава | Контр- октава | Большая октава | Малая октава | 1.00 октава | 2.00 октава | 3.00 октава | 4.00 октава | 5.00 октава |
|----------------|---------------------|--------------------------|------------------|-------------------|-----------------|----------------|----------------|----------------|----------------|----------------|
| Нота | | | | | | | | | | |
| До | C | – | 32.70 | 65.41 | 130.82 | 261.63 | 523.25 | 1046.50 | 2093.00 | 4186.00 |
| До-диез | C | – | 34.65 | 69.30 | 138.59 | 277.18 | 554.36 | 1108.70 | 2217.40 | 4434.80 |
| Ре | D | – | 36.95 | 73.91 | 147.83 | 293.66 | 587.32 | 1174.60 | 2349.20 | 4698.40 |
| Ре-диез | D | – | 38.88 | 77.78 | 155.56 | 311.13 | 622.26 | 1244.50 | 2489.00 | 4978.00 |
| Ми | E | 20.61 | 41.21 | 82.41 | 164.81 | 329.63 | 659.26 | 1318.50 | 2637.00 | 5274.00 |
| Фа | F | 21.82 | 43.65 | 87.31 | 174.62 | 349.23 | 698.46 | 1396.90 | 2793.80 | – |
| Фа-диез | F | 23.12 | 46.25 | 92.50 | 185.00 | 369.99 | 739.98 | 1480.00 | 2960.00 | – |
| Соль | G | 24.50 | 49.00 | 98.00 | 196.00 | 392.00 | 784.00 | 1568.00 | 3136.00 | – |
| Соль-диез | G | 25.95 | 51.90 | 103.80 | 207.00 | 415.30 | 830.60 | 1661.20 | 3332.40 | – |
| Ля | A | 27.50 | 55.00 | 110.00 | 220.00 | 440.00 | 880.00 | 1720.00 | 3440.00 | – |
| Си-бемоль | B | 29.13 | 58.26 | 116.54 | 233.08 | 466.16 | 932.32 | 1864.60 | 3729.20 | – |
| Си | H | 30.87 | 61.74 | 123.48 | 246.96 | 493.88 | 987.75 | 1975.50 | 3951.00 | – |

Алгоритм модуля, реализующего предложенную идею, представлен на рис. 9.

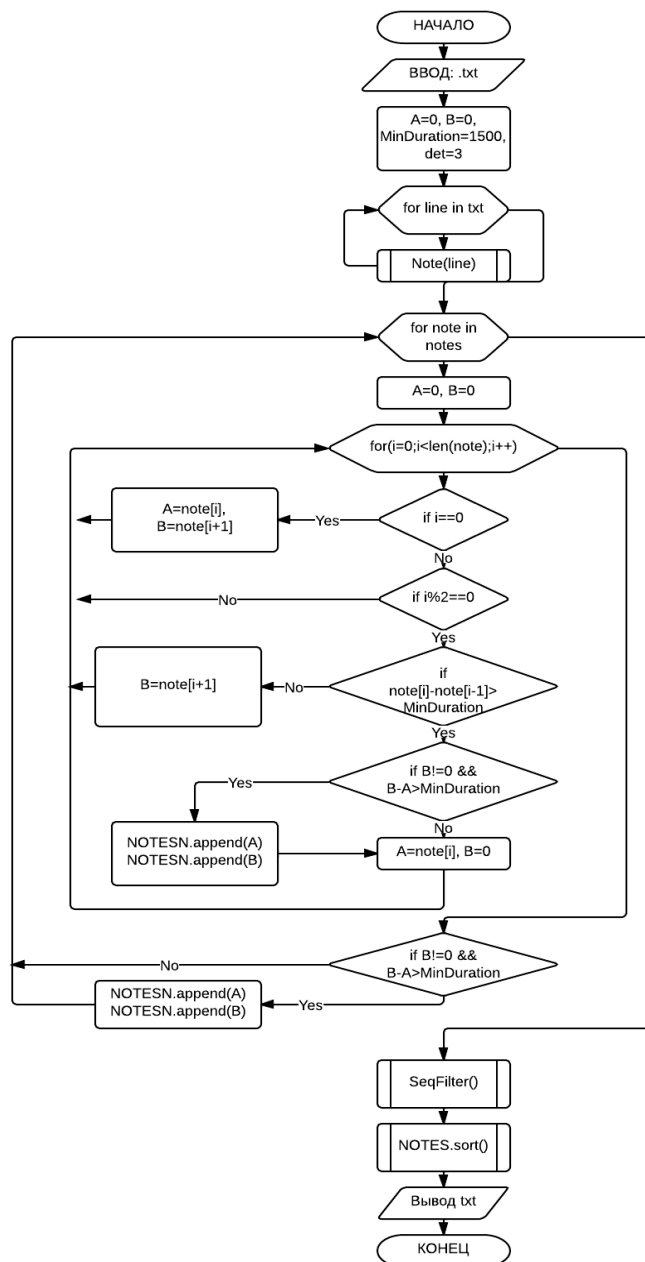


Рис. 9. Общая блок-схема алгоритма определения нот

В рамках данного алгоритма используются следующие обозначения:

A – начало ноты;

B – конец ноты;

MinDuration – минимальная длительность ноты;

Det – диапазон ноты / 2;

Line – строка из файла;

Note() – функция определения принадлежности частоты к ноте;

notes – массив найденных нот;

note – ноты из массива notes;

NOTESN – массив нот после проверки на минимальную длительность;

SeqFilter() – подпрограмма фильтрации 1-м и 2-м способами.

NOTES.sort() – подпрограмма сортировки нот по появлению в записи.

Готовая программа состоит из нескольких модулей, структуру можно увидеть на рис. 10.

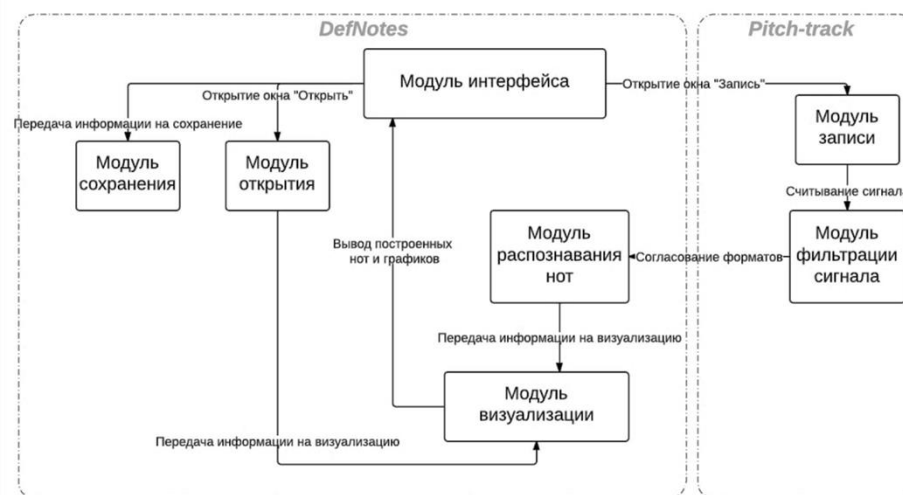


Рис. 10. Модули программы

4. ТЕСТИРОВАНИЕ РАЗРАБОТАННОГО ПРОГРАММНОГО КОМПЛЕКСА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ НОТ

Для тестирования программы были сделаны 16 записей: 8 – женским голосом, 8 – мужским.

8 записей женского голоса включают в себя следующее:

локация 1: первая октава. Ноты: до, ре, ми, фа, соль, фа, ми, ре, до. Стаккато (отрывисто), без произношения нот (только звуки);

локация 2: первая октава. Ноты: до, ре, ми, фа, соль, фа, ми, ре, до. Легато (связанно), без произношения нот (только звуки);

локация 3: первая октава. Ноты: си, ре#, фа#, ре#, си. Стаккато, без произношения нот;

локация 4: малая октава, первая октава. Ноты: си, ре#, фа#, ре#, си. Легато, без произношения нот;

локация 5: малая октава, первая октава. Ноты: ля, ля#, си, до, до#, до, си, ля#, ля. Стаккато, без произношения нот;

локация 6: малая октава, первая октава. Ноты: ля, ля#, си, до, до#, до, си, ля#, ля. Легато, без произношения нот;

локация 7: первая октава. Ноты: до, ми, ре, фа, ми, соль. Стаккато, с произношением нот;

локация 8: первая октава. Ноты: до, ми, ре, фа, ми, соль. Легато, с произношением нот.

Отличие мужского набора записей заключалось в отсутствии в записи № 5 Стакато двух последних нот ля# и ля ввиду проблем их воспроизведения отрывисто неподготовленным мужским голосом.

Результаты для ряда тестов представлены на рис. 11–13, сверху – выход частоты основного тона, середина – постоянная составляющая частоты основного тона (среднее значение), снизу – полученная нотная запись.

Локация 1: первая октава. Ноты: до, ре, ми, фа, соль, фа, ми, ре, до. Стакато (отрывисто), без произношения нот (только звуки).

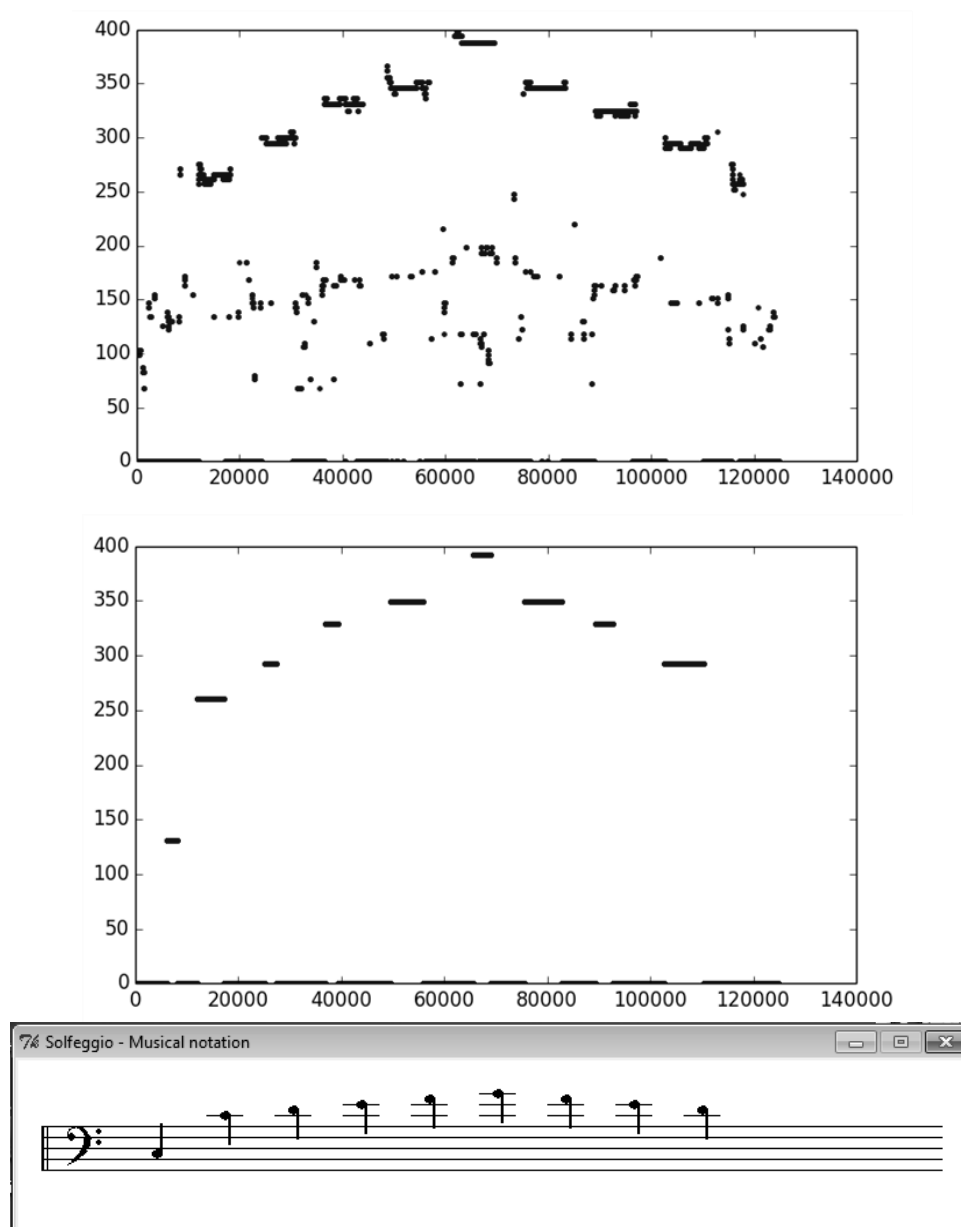


Рис. 11. Запись № 1, женский голос

Локация 4: малая октава, первая октава. Ноты: си, ре#, фа#, ре#, си. Легато, без произношения нот.

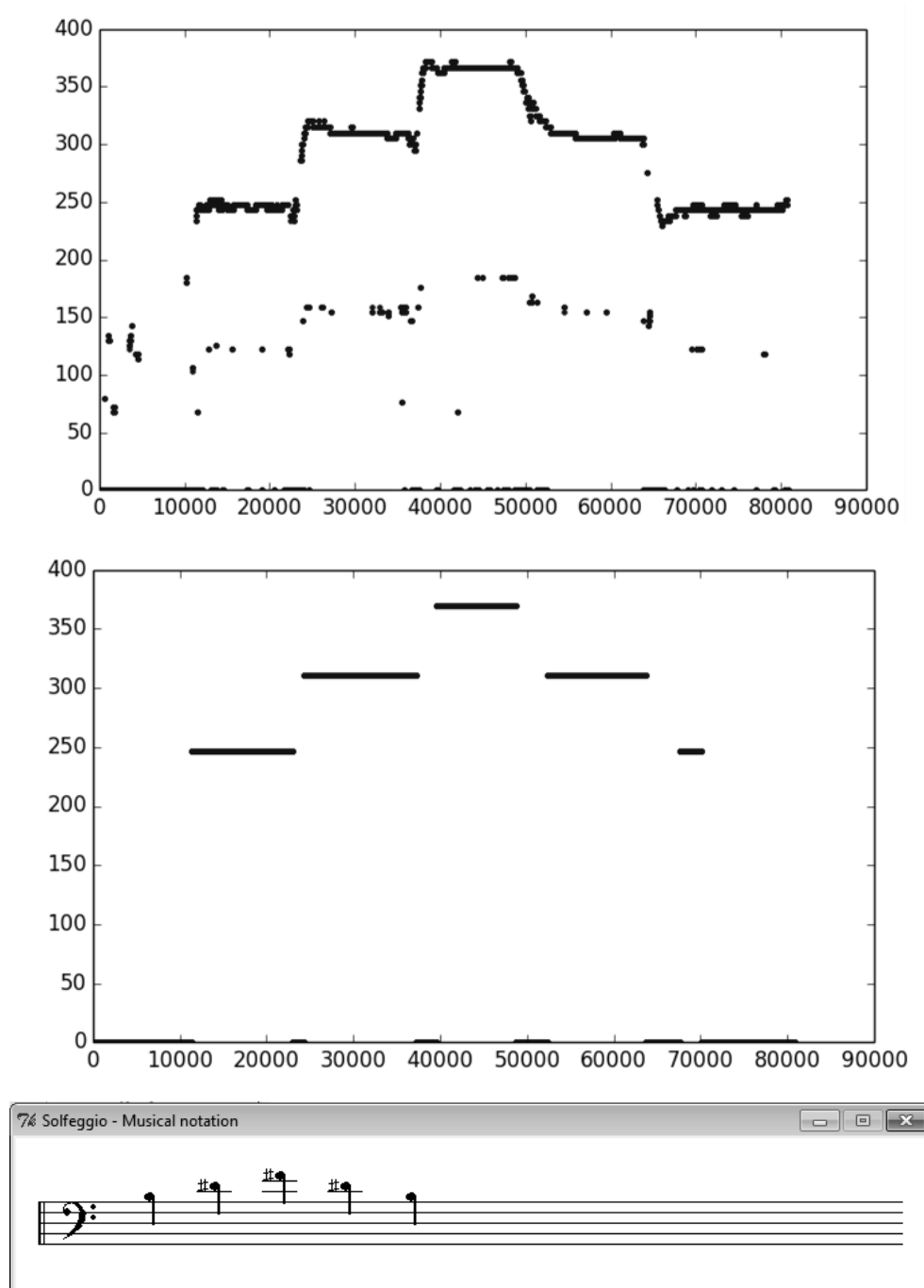


Рис. 12. Запись № 4, женский голос

Локация 1: малая октава. Ноты: до, ре, ми, фа, соль, фа, ми, ре, до. Стаккато (отрывисто), без произношения нот (только звуки).

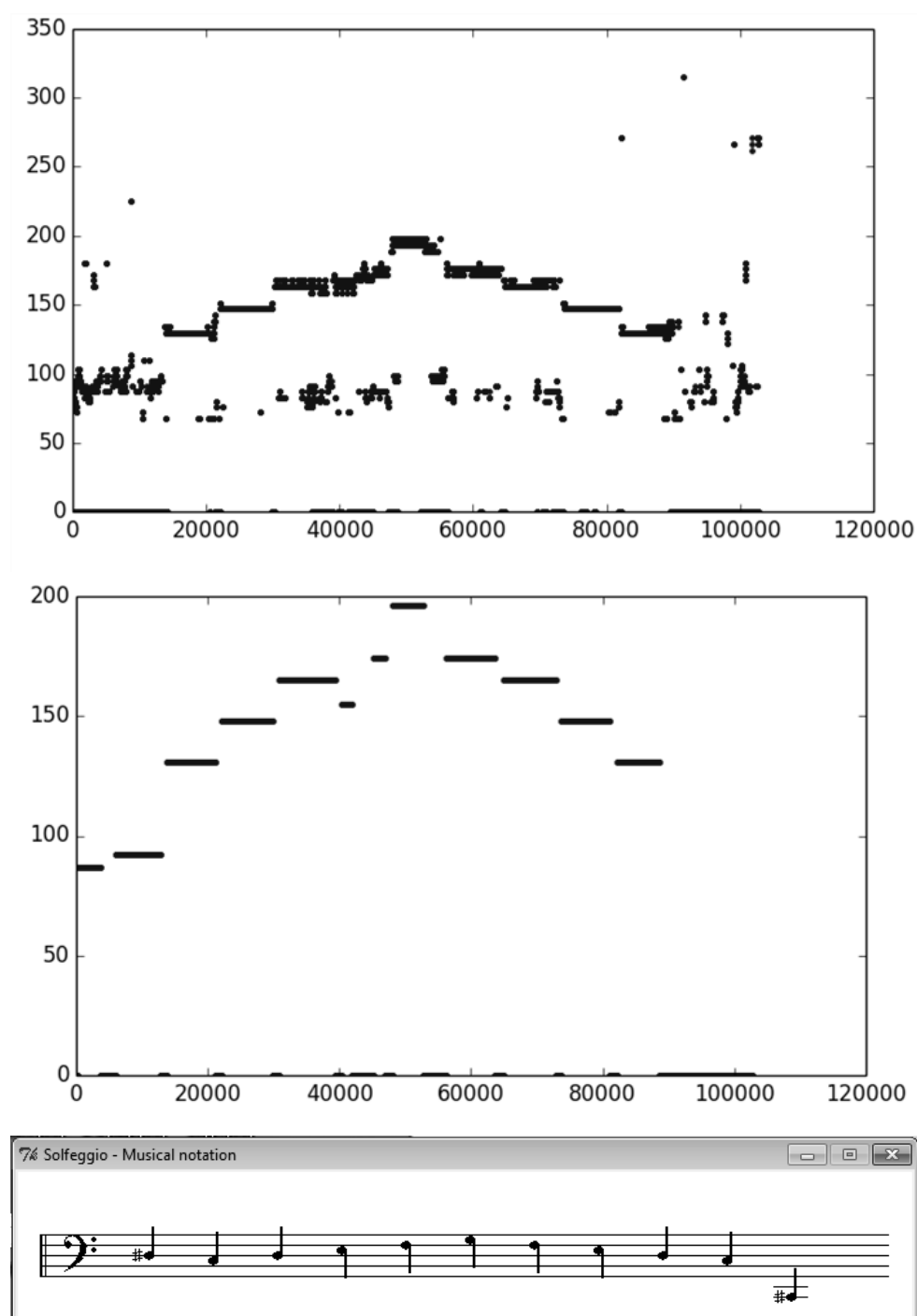


Рис. 13. Запись № 1, мужской голос

В 16 тестах были воспроизведены 114 нот, 58 из которых были спеты женским голос, а 56 – мужским. На рис. 14 представлена результирующая диаграмма тестирования.

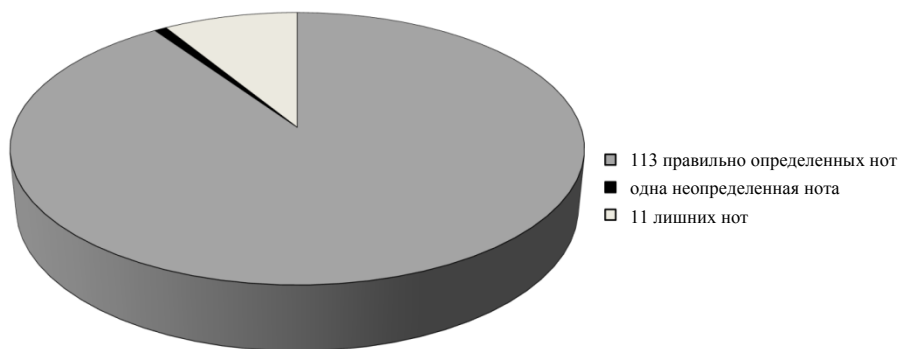


Рис. 14. Результирующая диаграмма тестирования

ЗАКЛЮЧЕНИЕ

Разработанная программа Defnotes позволяет пользователю автоматически переводить вокал в нотную запись. Нотная запись на выходе программы соответствует музыкальной грамоте, поэтому очень удобна для музыканта. Также программа выводит графики частот, что будет удобно для технически образованного пользователя.

Было проведено сравнение между программами, определяющими ЧОТ. Были проанализированы 9 аудиофайлов, содержащих 92 ноты для определения. Программа Sr-0, основанная на модели периферической части слуховой системы человека, справилась с тестированием лучше остальных, определив 85 % частот. Именно эта программа использовалась как базовая.

Для проверки программы был составлен план тестирования, включающий в себя типичные вокальные ситуации. Так же для полноты тестов исследовались как мужской, так и женский голос. Всего было сделано 16 записей, содержащих 114 нот для определения. Верно были определены 113 нот, одна нота не определилась и 11 нот были определены ошибочно.

В целом разработанное программное обеспечение справляется со своей задачей в области нижней части спектра, однако для применения в области частоты выше 400 Гц необходима дополнительная модификация базового программного обеспечения для расширения диапазона используемых каналов частот.

Работа выполнена при поддержке Министерства образования и науки Российской Федерации в рамках базовой части государственного задания ТУСУР на 2015 г. (проект № 3657).

СПИСОК ЛИТЕРАТУРЫ

1. Connect with the music you love [Electronic resource]. – Available at: <http://www.shazam.com/> (accessed: 07.10.2015).
2. The open music encyclopedia [Electronic resource]. – Available at: <http://www.musipedia.org/> (accessed: 05.10.2015).
3. Love that song but don't know what it is? Tunatic can tag it for you! [Electronic resource]. – Available at: <http://www.wildbits.com/tunatic/> (accessed: 05.10.2015).

4. The ultimate music search [Electronic resource]. – Available at: <http://www.midomi.com/> (accessed: 05.10.2015).
5. *Способин И.В.* Элементарная теория музыки. – М.: Музыка, 1968. – 203 с.
6. Praat: doing phonetics by computer [Electronic resource]. – Available at: <http://www.fon.hum.uva.nl/praat/> (accessed 07.10.2015).
7. The musical view of audio [Electronic resource]. – Available at: <http://www.celemony.com/en/melodyne/> (accessed: 05.10.2015).
8. *Черных Д.В., Конев А.А., Мещеряков Р.В.* Элементы программного комплекса для оценки биометрических параметров в защищенных системах // Электронные средства и системы управления. – 2011. – № 1. – С. 188–190.
9. *Конев А.А., Тихонова В.И.* Выделение вокализованных звуков в слитной речи // Акустика речи. Медицинская и биологическая акустика. Архитектурная и строительная акустика. Шумы и вибрации. Аэроакустика: сборник трудов XVI сессии Российского акустического общества. – М.: ГЕОС, 2005. – Т. 3. – С. 47–50.
10. Сравнение параметров для выделения вокализованных сегментов и классификации гласных фонем / С.А. Елистратов, М.А. Косенко, Е.Ю. Костюченко, А.А. Чичерин // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2012. – Т. 1, № 2. – С. 171–174.
11. *Конев А.А.* Модель и алгоритмы анализа и сегментации речевого сигнала: автореф. дис. ... канд. техн. наук. – Томск, 2007. – 20 с.
12. *Mitre A., Queiroz M., Faria R.* Accurate and efficient fundamental frequency determination from precise partial estimates // Proceedings of the 4th AES Brazil Conference, May 2006. – Sao-Paulo, 2006. – P. 113–118.
13. Estimation of pitch from speech signals [Electronic resource]. – Available at: iitg.vlab.co.in/?sub=59&brch=164&sim=1012&cnt=1 (accessed: 05.10.2015).
14. *Gerhard D.* Pitch extraction and fundamental frequency: history and current techniques: technical report TR-CS 2003-06. – Regina Saskatchewan, Canada: Department of Computer Science University of Regina, 2003. – 22 p.
15. *Тюлин Ю.Н.* Краткий теоретический курс гармонии. – М.: Музыка, 1978. – 168 с.

Конев Антон Александрович, кандидат технических наук, доцент Томского государственного университета. Основные направления научных исследований: анализ речевых сигналов, защита информации. Имеет более 60 публикаций. E-mail: kaa1@keva.tusur.ru

Онищенко Александра Александровна, инженер органа криптографической защиты ООО «Тензор». Основное направление научных исследований – анализ звуковых сигналов. Имеет одну публикацию. E-mail: veganwin@gmail.com

Костюченко Евгений Юрьевич, кандидат технических наук, доцент Томского государственного университета. Основные направления научных исследований: анализ речевых сигналов, идентификация биометрических сигналов. Имеет более 60 публикаций. E-mail: key@keva.tusur.ru

Якимук Алексей Юрьевич, аспирант Томского государственного университета. Основное направление научных исследований – анализ звуковых и речевых сигналов. Имеет 5 публикаций. E-mail: yay@keva.tusur.ru

*Automatic recognition of music**

A.A. KONEV¹, A.A. ONISCHENKO², E.YU. KOSTYUCHENKO³, A.Yu. YAKIMUK⁴

¹ 634050, Russia, Tomsk, 40, Lenin Prospekt, Tomsk State University of Control Systems and Radio Electronics, Ph.D., associate professor. E-mail: kaa1@keva.tusur.ru

² 634050, Russia, Tomsk, 7, Gagarina St., Tensor Ltd Company, engineer, centre of cryptographic protection. E-mail: veganwin@gmail.com

³ 634050, Russia, Tomsk, 40, Lenin Prospekt, Tomsk State University of Control Systems and Radioelectronics, Ph.D (Eng.), associate professor. E-mail: key@keva.tusur.ru

⁴ 634050, Russia, Tomsk, 40., Lenin Prospekt, Tomsk State University of Control Systems and Radioelectronics, a graduate student. E-mail: yay@keva.tusur.ru

In this paper the process of recognizing musical notes is described. A sound file which contains a record of a melody can be used as input data. A source of the sound in recording tunes can be a single musical instrument, a set of them as well as the voice of a man crooning a tune. Originally, an input file is presented in the .wav format or recorded in this format using a microphone. Processing is carried out by sequentially determining the fundamental frequency of the signal and converting the value of the pitch frequency into the corresponding A comparison of programs that implement algorithms for pitch frequency selection is made. Using fundamental frequency selecting algorithms which take into account the peculiarities of the human auditory system at the first stage is a specific character of the study. Using these algorithms, on the one hand, allows determining values more accurately in comparison with analogues, in particular with peak methods. On the other hand, filter banks currently used in models of the human auditory system, have an upper limit of the pitch frequency of 400 Hz as they are adapted for use with conventional speech.

This is a disadvantage because a pitch frequency can be up to 1400 Hz for opera singers. However, this limitation is not critical and requires further study for possible extension of the filter bank simulating the human auditory system. The second step is the conversion of the selected frequency pitch into a note on the basis of the algorithm. The algorithm is based on the tables of frequency correspondence to notes and note minimum duration. The results of testing the program implementing these algorithms are presented. In conclusion, the main results of the work and its subsequent development, namely, the expansion of the range acceptable for frequency identification are presented.

Keywords: music recognition, frequency, pitch, sound spectrum, sound processing, sound filtering, vocal, musical notation

DOI: 10.17212/1814-1196-2015-3-32-47

REFERENCES

1. *Connect with the music you love*. Available at: <http://www.shazam.com/> (accessed 05.10.2015)
2. *The open music encyclopedia*. Available at: <http://www.musipedia.org/> (accessed 05.10.2015)
3. *Love that song but don't know what it is? Tunatic can tag it for you!* Available at: <http://www.wildbits.com/tunatic/> (accessed 05.10.2015)
4. *The ultimate music search*. Available at: <http://www.midomi.com/> (accessed 05.10.2015)
5. Sposobin I.V. *Elementarnaya teoriya muzyki* [Elementary music theory]. Moscow, Muzyka Publ., 1968. 203 p.

* Received 1 June 2015.

This work was supported by the Ministry of education and science of the Russian Federation within the framework of the base part of the state task TUSUR (TUCSR) 2015 (project number 3657).

6. Praat: doing phonetics by computer. Available at: <http://www.fon.hum.uva.nl/praat/> (accessed 05.10.2015)
7. The musical view of audio. Available at: <http://www.celemony.com/en/melodyne/> (accessed 05.10.2015)
8. Chernykh D.V., Konev A.A., Meshcheryakov R.V. Elementy programmnoy kompleksa dlya otsenki biometricheskikh parametrov v zashchishchennykh sistemakh [Elements of software for the evaluation of biometric parameters in protected systems]. *Elektronnye sredstva i sistemy upravleniya – Electronic instrumentation and control systems*, 2011, no. 1. pp. 188–190.
9. Konev A.A., Tikhonova V.I. Vydelenie vokalizovannykh zvukov v slitnoi rechi [Isolation voiced sounds in continuous speech]. *Akustika rechi. Meditsinskaya i biologicheskaya akustika. Arkhitekturnaya i stroitel'naya akustika. Shumy i vibratsii. Aeroakustika: sbornik trudov XVI sessii Rossiiskogo akusticheskogo obshchestva* [Acoustics of speech. Medical and biological acoustics. Architectural and building acoustics. Noise and vibration. Aeroacoustics. Proceedings of the XVI Session of the Russian Acoustical Society]. Moscow, GEOS Publ., 2005, vol. 3, pp. 47–50.
10. Elistratov S.A., Kosenko M.A., Kostyuchenko E.Yu., Chicherin A.A. Sravnenie parametrov dlya vydeleniya vokalizovannykh segmentov i klassifikatsii glasnykh fonem [Comparison of parameters for selection of voiced segments and classification of vowel phonemes]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics*, 2012, vol. 1, no. 2, pp. 171–174.
11. Konev A.A. *Model' i algoritmy analiza i segmentatsii rechevogo signala*. Avtoref. diss. kand. tekhn. nauk [The model and algorithms for the analysis and segmentation of the speech signal. Author's abstract of PhD eng. sci. diss.]. Tomsk, 2007. 20 p.
12. Mitre A., Queiroz M., Faria R. Accurate and efficient fundamental frequency determination from precise partial estimates. *Proceedings of the 4th AES Brazil Conference 2006*. Sou-Paulo, Brazil, 2006, pp. 113–118.
13. *Estimation of pitch from speech signals*. Available at: iitg.vlab.co.in/?sub=59&brch=164&sim=1012&cnt=1 (accessed 05.10.2015)
14. Gerhard D. *Pitch extraction and fundamental frequency: history and current techniques: technical report TR-CS-2003-06*. Regina, Saskatchewan, Canada, Department of Computer Science University of Regina, 2003. 22 p.
15. Tyulin Yu.N. *Kratkii teoreticheskii kurs garmonii* [A brief theoretical course of harmony]. Moscow, Muzyka Publ., 1978. 168 p.