

Министерство образования Республики Беларусь  
Учреждение образования  
«Полоцкий государственный университет»

**ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ:  
ДОСТИЖЕНИЯ, ПРОБЛЕМЫ, ИННОВАЦИИ  
(ИКТ-2018)**

Электронный сборник статей

I Международной научно-практической конференции,  
посвященной 50-летию Полоцкого государственного университета

(Новополоцк, 14–15 июня 2018 г.)

Новополоцк  
Полоцкий государственный университет  
2018

**Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ-2018)** [Электронный ресурс] : электронный сборник статей I международной научно-практической конференции, посвященной 50-летию Полоцкого государственного университета, Новополоцк, 14–15 июня 2018 г. / Полоцкий государственный университет. – Новополоцк, 2018. – 1 электрон. опт. диск (CD-ROM).

Представлены результаты новейших научных исследований, в области информационно-коммуникационных и интернет-технологий, а именно: методы и технологии математического и имитационного моделирования систем; автоматизация и управление производственными процессами; программная инженерия; тестирование и верификация программ; обработка сигналов, изображений и видео; защита информации и технологии информационной безопасности; электронный маркетинг; проблемы и инновационные технологии подготовки специалистов в данной области.

*Сборник включен в Государственный регистр информационного ресурса. Регистрационное свидетельство № 3201815009 от 28.03.2018.*

Компьютерный дизайн М. Э. Дистанова.

Технические редакторы: Т. А. Дарьянова, О. П. Михайлова.

Компьютерная верстка Д. М. Севастьяновой.

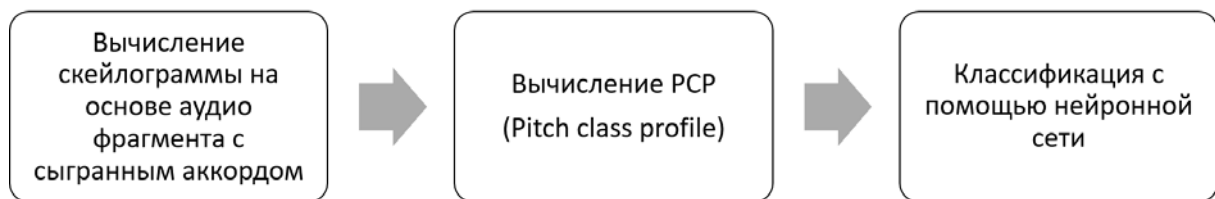
211440, ул. Блохина, 29, г. Новополоцк, Беларусь  
тел. 8 (0214) 53-21-23, e-mail: irina.psu@gmail.com

**РАСПОЗНАВАНИЕ МУЗЫКАЛЬНЫХ АККОРДОВ  
НА ОСНОВЕ ВЕЙВЛЕТ ПРЕОБРАЗОВАНИЯ**

**И.А. БРАГИНЕЦ**

*(Белорусский государственный университет, Минск)*

Аккорд – это сочетание трех и более музыкальных звуков разной высоты сыгранных одновременно. Задача их распознавания может возникнуть при разработке систем для обучения игре на музыкальных инструментах, систем идентификации музыкальных произведений. В данной работе предлагается базовый алгоритм для решений данной задачи, структура алгоритма изображена на рисунке 1.



**Рисунок 1. - Структура алгоритма**

Важно понимать, какие музыкальные звуки существуют, их спектральный состав и как они обозначаются.

Для обозначения музыкальных звуков в работе использована научная нотация (англ. Scientific Pitch Notation), где музыкальный звук обозначается:

- Буквами латинского алфавита – C, D, E, F, G, A, B;
  - Знак альтерации, если требуется;
  - Арабскими цифрами, обозначающими октаву, к которой звук принадлежит.
- Пример: C4, C#4, D4, D#4, E4, F4, F#4, G4, G#4, A4, A#4, B4 – все звуки 4 октавы.

Как известно, человеческий слух характеризуется логарифмическим восприятием – два звука, для которых соотношение частот является степенью двойки, воспринимаются похоже и называются одинаково. Октава (интервал между такими звуками) делится на 12 интервалов, называемых полутонами, где соотношение частот между звуками на расстоянии одного полутона равно  $2^{1/12}$ , в качестве опорной частоты обычно берут частоту ноты A4 – 440 Гц, если обозначить n количество полутонов относительно этой ноты, то частота:

$$f(n) = 440 * 2^{n/12}$$

**Вычисление скейлограммы.** Для вычисления скейлограммы используется непрерывное вейвлет-преобразование [1].

$$g \text{ материнская вейвлет – функция, } \int_{-\infty}^{+\infty} g(u)du = 0, g(\pm\infty) \rightarrow 0$$

$$WT(f(a,t)) = \int_{-\infty}^{+\infty} f(u) \overline{g_{(a,t)}(u)} du, g_{(a,t)} = \frac{1}{\sqrt{a}} g\left(\frac{u-t}{a}\right) \text{ – дочерние вейвлеты}$$

*a – масштаб, t – время*

Анализируемый диапазон звуковых частот: от 65,406 Гц до 3951,1 Гц, начало 2-й октавы и конец 7-й октавы, соответственно, этот диапазон включает 72 ноты. Пока речь шла только о частотах, но в контексте вейвлет-преобразования оперируют масштабами. Соответствие частоты масштабу будем понимать в таком ключе: если рассмотреть масштаб и соответствующий ему дочерний вейвлет, то его можно рассматривать как полосовой фильтр, имеющий максимальный коэффициент пропускания на некой частоте, эту частоту и будем понимать как соответствующую данному масштабу.

При вычислении вейвлет преобразования масштабы дискретизируются так, что каждый масштаб соответствует частотам нот, входящих в анализируемый диапазон. Пример:

$$a_1 \rightarrow 65.406 \text{ Гц}, a_2 \rightarrow 65.406 * \sqrt[12]{2} \text{ Гц} \dots$$

Частота дискретизации по времени должна быть больше частоты Найквиста, в данном случае, больше удвоенной частоты верхней границы анализируемого диапазона. Пример скейлограммы для аккорда Am, сыгранного на гитаре, приведен на рисунке 2.

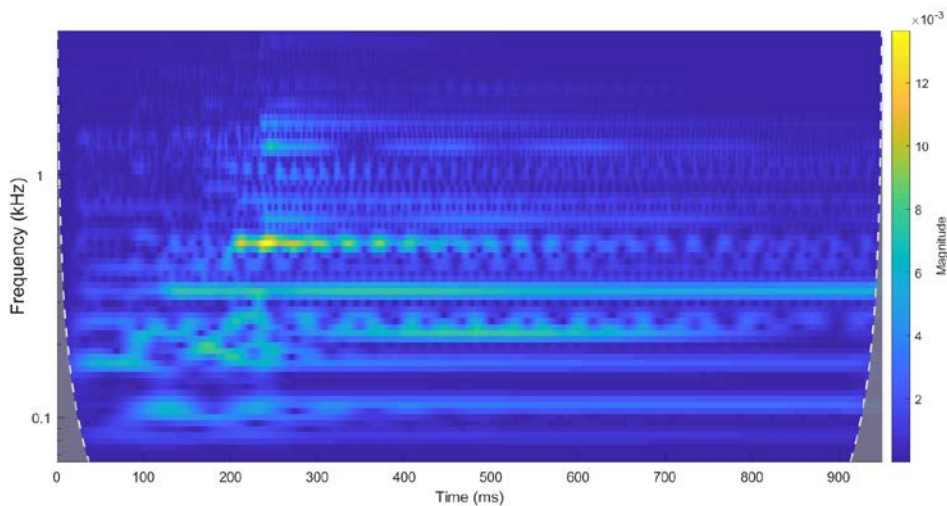


Рисунок 2. - Пример скейлограммы для аккорда Am, сыгранного на гитаре

**Pitch Class Profile.** PCP хорошо себя зарекомендовал в задаче распознавания аккордов [2]. PC (Pitch class) – это множество нот, интервал между которыми составляет целое число октав, всего таких множеств 12 (так как в пределах одной октавы находится 12 нот).

Примеры: {C1, C2, C3, C4...}, {D1, D2, D3, D4...}, {E1, E2, E3, E4...}, {F1, F2, F3, F4...}.

Результатом 1 этапа алгоритма является скейлограмма:

$$X(f_i, t_j); i = \overline{1, 72}; j = \overline{1, N};$$

$N$  – количество отсчетов времени

После этого происходит суммирование по определенному временному отрезку (Frame), использовалась длительность 0,95 секунд, обозначим длину в отсчетах как Length (зависит от частоты дискретизации).

$$Frame(f_i) = \sum_{j=1}^{Length} |X(f_i, t_j)|^2; i = \overline{1, 72}$$

Тогда PCP определяется как

$$PitchClassProfile(k) = \sum_{i=1}^{72} Frame(f_i) \delta(k, PitchClass(f_i))$$

$$PitchClass(f) = 12 \text{round}(\log_2(\frac{f}{f_{ref}})) \text{ mod } 12$$

**Выбор вейвлет-функции.** В работе рассмотрены следующие вейвлеты:

- bump wavelet
- Morlet wavelet (вейвлет Морле)
- Morse wavelet (вейвлет Морзе), с различными параметрами

Экспериментально выбран вейвлет Морзе, поскольку в PCP-векторах, получаемых с помощью этого вейвлета, ноты аккордов локализованы лучше по сравнению с другими вейвлетами, что показано на рисунке 3.

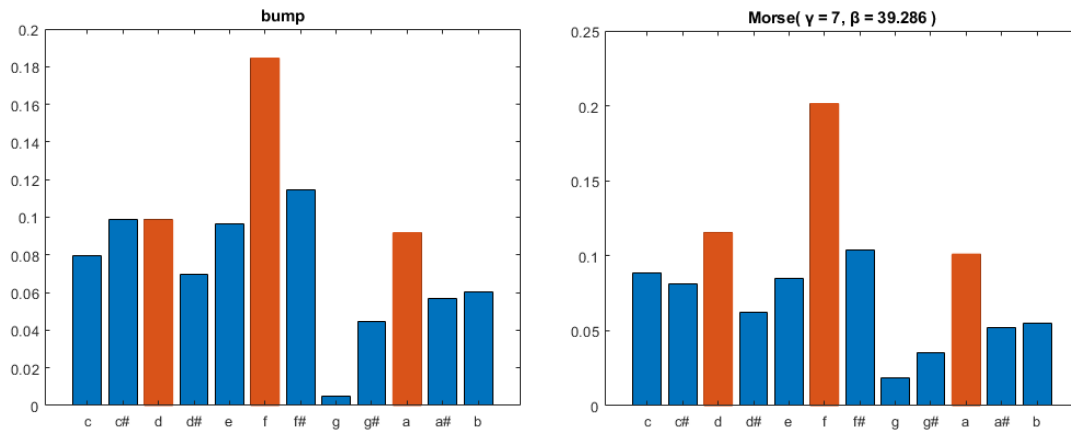


Рисунок 3. – Сравнение вейвлетов

**Классификация на основе нейронной сети.** Экспериментально была выбрана следующая конфигурация нейронной сети: 12-35-12, функция активация скрытого слоя “logsig”, тип выходного слоя “softmax layer”.

Конфигурация нейронной сети зависит от классифицируемых аккордов. Для классификации выбран набор данных [3], который содержит экземпляры 10 минорных и мажорных аккордов популярных в восточной музыке: A, Am, Vm, C, D, Dm, E, Em, F, G. Структура набора:

- 2000 экземпляров сыгранных гитарой, 200 на каждый тип
- 100 экземпляров сыгранных пианино, 10 на каждый тип
- 100 экземпляров сыгранных аккордеоном, 10 на каждый тип

Для обучения нейронной сети использовались экземпляры гитарных аккордов. Аккорды в количестве 2000 экземпляров были разделены на обучающую выборку (LS или learn set) и тестовую выборку (TS или test set), где LS – 70%, TS – 30%. Также нейронная сеть тестировалась на экземплярах аккордов, сыгранных пианино и аккордеоном. Для обучения нейронной сети экспериментально выбран метод регуляризации Байеса. В таблице 1 приведены параметры обучения выбранной нейронной сети, в таблице 2 – результаты тестирования обученной сети.

Таблица 1. – Параметры обучения нейронной сети

epochs	200	mu	0.00001	mu_inc	1.05
perfgoal	0.0005	mu_dec	0.3	min_grad	1e-12

Таблица 2. – Результаты тестирования

Инструмент, размер тестовой выборки	Ошибка
Гитара, 600	2.7 %
Аккордеон, 100	10 %
Пианино, 100	15 %

Величина ошибки обученной сети на обучающей выборке составляла 0,1 %. Величина ошибки при тестировании обученной сети составила от 2,7 % для гитары и до 10 – 15 % для других, что можно объяснить: тренировка происходила только с экземплярами гитарных аккордов, РСР вектора для одного и того же типа аккорда, но разных инструментов будут выглядеть по-разному, вследствие разного распределения энергии по гармоникам у каждого инструмента, что и определяет тембр музыкального инструмента.

#### Литература

1. Н. – G. Stark. Wavelets and Signal Processing. / by Н. – G. Stark, Springer Science & Business Media, Inc., 2005.
2. T. Fujishima. Realtime chord recognition of musical sound: a system using common lisp music // In Proceedings of the International Computer Music Conference (ICMC), 1999. – P. 464–467.
3. J. Osmalskyj, J-J. Embrechts, S. Piérard, M. Van Droogenbroeck. Neural networks for musical chords recognition // In Proceedings of the Journées d'Informatique Musicale (JIM 2012), 2012.
4. Sauer, C. Design and Evaluation of a Simple Chord Detection Algorithm. / C. Sauer // Режим доступа: [https://www.fim.uni-passau.de/fileadmin/files/lehrstuhl/sauer/geyer/BA\\_MA\\_Arbeiten/BA-HausnerChristoph-201409.pdf](https://www.fim.uni-passau.de/fileadmin/files/lehrstuhl/sauer/geyer/BA_MA_Arbeiten/BA-HausnerChristoph-201409.pdf), 2014.