

УДК 004.021

М.А. Степаненко, Е.А. Фунтиков

ОЦИФРОВКА И АНАЛИЗ АКУСТИЧЕСКИХ СИГНАЛОВ НА ПЛАТФОРМЕ ANDROID

Нижегородский государственный технический университет им. Р.Е. Алексеева

Посвящена спектральному анализу акустических сигналов и определению нот музыкального произведения в реальном времени при помощи мобильных устройств на платформе Android.

Ключевые слова: захват звука, дискретизация, окно Гаусса, антиалиасинг, гармонический сигнал.

Введение

Вопрос спектрального анализа оцифрованных звуков всегда стоял довольно остро, так как он применяется во многих сферах жизнедеятельности (например, обработка музыкального произведения или определение посторонних шумов у механизмов на производстве). В настоящее время темпы роста производительности мобильных устройств набирают обороты, что позволяет производить сложные и трудоёмкие математические операции прямо на портативном устройстве.

Результаты исследований будут реализованы в виде мобильного приложения, способного в реальном времени оцифровывать и анализировать игру на фортепиано и сравнивать полученные в процессе анализа музыкальные ноты с оригинальным произведением на предмет качества игры. Процесс обработки звука можно разделить на пять этапов: оцифровка (дискретизация аналогового сигнала) поступающей через микрофон устройства музыки, обработка оцифрованных данных, получение частотного и фазового спектров, их анализ и выделение музыкальных нот.

Оцифровка аналогового акустического сигнала

Для того чтобы в дальнейшем обрабатывать звуковой сигнал, его нужно оцифровать. Это делается путём дискретизации (рис. 1) при помощи встроенного в мобильное устройство АЦП (аналогово-цифровой преобразователь) с определёнными параметрами. Платформа Android предоставляет необходимый API с возможностью конфигурирования следующих параметров: частота дискретизации, тип аудиоканала (моно/стерео), аудио формат (PCM 8 bit / PCM 16 bit), размер буфера (фрейма) данных. Одним из наиболее важных параметров является частота дискретизации (рис. 1).

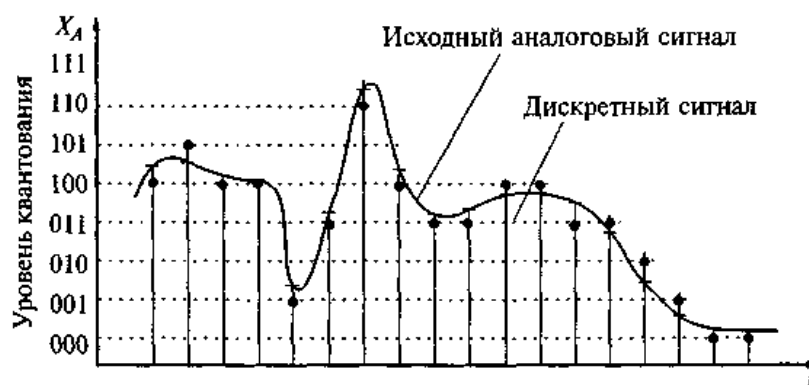


Рис. 1. Дискретизация аналогового сигнала

Она выбирается исходя из диапазона частот музыкальных нот. По теореме Шеннона-Найквиста, частота дискретизации должна быть, как минимум, в два раза больше частоты обрабатываемого звука. В табл. 1 отображены частоты всех музыкальных нот, начиная с «Ля субконтроктавы» и заканчивая самой высокой – «Ми пятой октавы».

Таблица 1

Музыкальные ноты и соответствующие им музыкальные частоты.

Нота	Название	Частота (Hz)
A0	Ля субконтроктавы	27.50
A0 [#]	Ля-диез субконтроктавы	29.13
B0	Си субконтроктавы	30.87
~~~~	~~~~~	~~~~~
D8	Ре пятой октавы	4698.40
D8 [#]	Ре-диез пятой октавы	4978.00
E8	Ми пятой октавы	5274.00

Исходя из описанной теоремы, частота дискретизации выбрана следующим образом:

$$\omega \geq 5274 * 2 \approx 11000 \text{ Hz.}$$

Ближайшее к этой частоте значение из пакета API android.media – 11025 Hz.

В качестве аудиоформата был выбран PCM 16 bit, а тип аудиоканала – моно, соответственно, при такой конфигурации, одна секунда записи будет занимать

$$\frac{11025 * 16}{8} \approx 21.484 \text{ Кб}$$

памяти.

Размер буфера выбирается кратным степени 2, так как для работы алгоритма БПФ (о котором речь идет далее) требуется чётное количество отсчётов.

После задания необходимых параметров начинается цикл сбора и обработки заранее оцифрованного звука. На каждой итерации цикла буфер заполняется дискретными значениями, которые поступают на дальнейший анализ.

### Дискретное преобразование Фурье

Так как вычисления производятся на устройстве с ограниченными ресурсами и в реальном времени будет производиться быстрое преобразование Фурье (БПФ) каждого отдельного фрейма, это позволит сократить количество вычислений ( $N$ ) с  $N^2$  до  $N * \log_2 N$ . БПФ основано на том, что среди множителей (синусов) есть много повторяющихся значений (в силу периодичности синуса). Его алгоритм группирует слагаемые с одинаковыми множителями, существенно сокращая число умножений. В результате быстрое действие БПФ может во множество раз превосходить быстрое действие стандартного алгоритма (зависит от  $N$ ). При этом следует подчеркнуть, что алгоритм БПФ точнее, по сравнению со стандартным, так как сокращение числа операций уменьшает ошибки округления.

Выделяется короткий кадр (фрейм) композиции, состоящий из дискретных отсчётов, который условно считаем периодическим и применяем к нему преобразование Фурье. В результате преобразования получаем массив комплексных чисел, содержащий информацию об

амплитудном и фазовом спектрах анализируемого фрейма. Причём спектры также являются дискретными с шагом равным отношению частоты дискретизации к количеству отсчётов.

$$\delta = \frac{\omega}{n}.$$

Таким образом, чем больше делается отсчётов, тем более точное разрешение мы получаем по частоте. Однако если мы при постоянной частоте дискретизации увеличиваем число отсчётов, то увеличивается анализируемый временной интервал. В реальных музыкальных произведениях ноты звучат с различной длительностью и могут быстро сменять друг друга, а значит, возможно их наложение, поэтому амплитуда длительных нот будет перекрывать собой амплитуду коротких. С другой стороны, такой способ увеличения разрешения по частоте хорошо подходит для гитарных тюнеров, поскольку нота, как правило, звучит одна и достаточно долго.

Также существует относительно простой подход увеличения разрешения по частоте – исходный дискретный сигнал заполняется нулями между отсчётами. Однако применение такого подхода сильно искажает фазовый спектр, но при этом увеличивает разрешение амплитудного спектра. В этом случае возможно применение фильтров Фарроу и искусственное увеличение частоты дискретизации, однако это также вносит искажения в спектры.

Длительность фрейма обычно составляет приблизительно от  $3 \cdot 10^{-3}$  до 1 с. При этом чем фрейм короче, тем лучше разрешение по времени, но хуже по частоте, и наоборот. Это очень напоминает принцип неопределённости Гейзенберга из квантовой механики: чем точнее измеряется одна характеристика частицы, тем менее точно можно измерить вторую. В результате анализа фрейма одиночного синусоидального сигнала амплитудный спектр сильно напоминает дифракционную картинку.

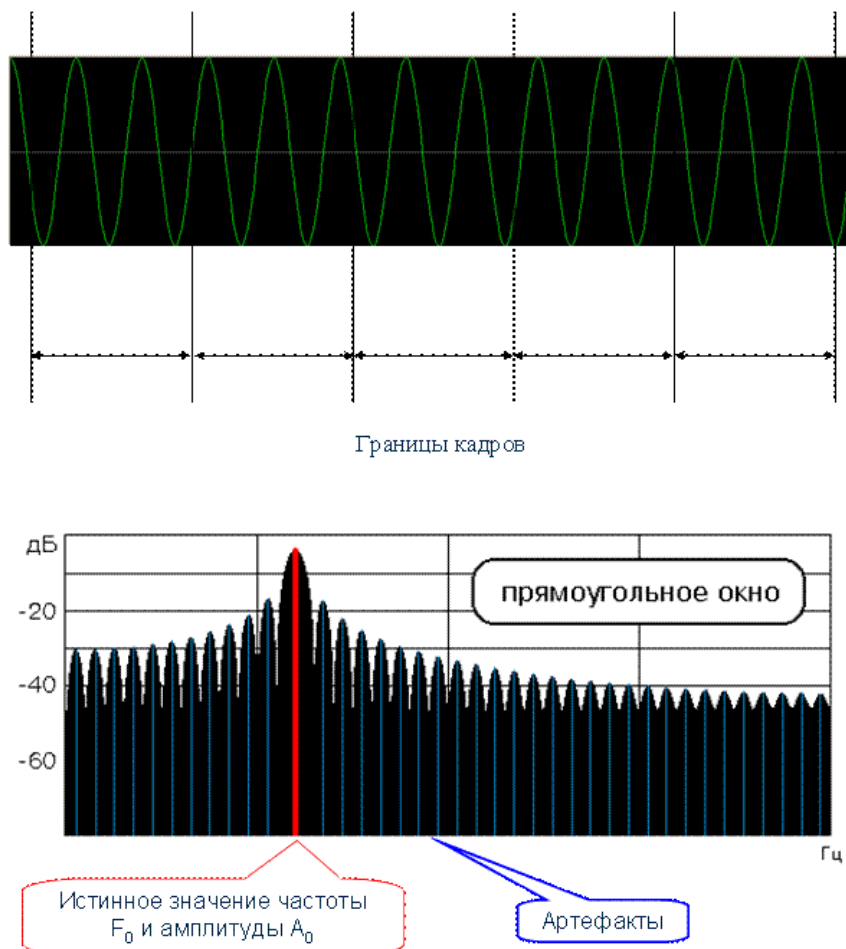


Рис. 2. Амплитудный спектр одного кадра гармонического сигнала

На практике это даёт нежелательный эффект, затрудняющий анализ сигналов, поэтому его стараются понизить путём применения оконных функций (рис. 3). На сегодняшний момент существуют следующие оконные функции: окно Ханна, окно Хэмминга, окно Наттол-ла, окно Гаусса и прямоугольное окно (сигнал без изменения). В качестве оконной функции будет использоваться окно Гаусса, так как его боковые лепестки в двойном логарифмическом масштабе не стремятся к прямой линии, но спадают намного быстрее, чем у остальных окон.

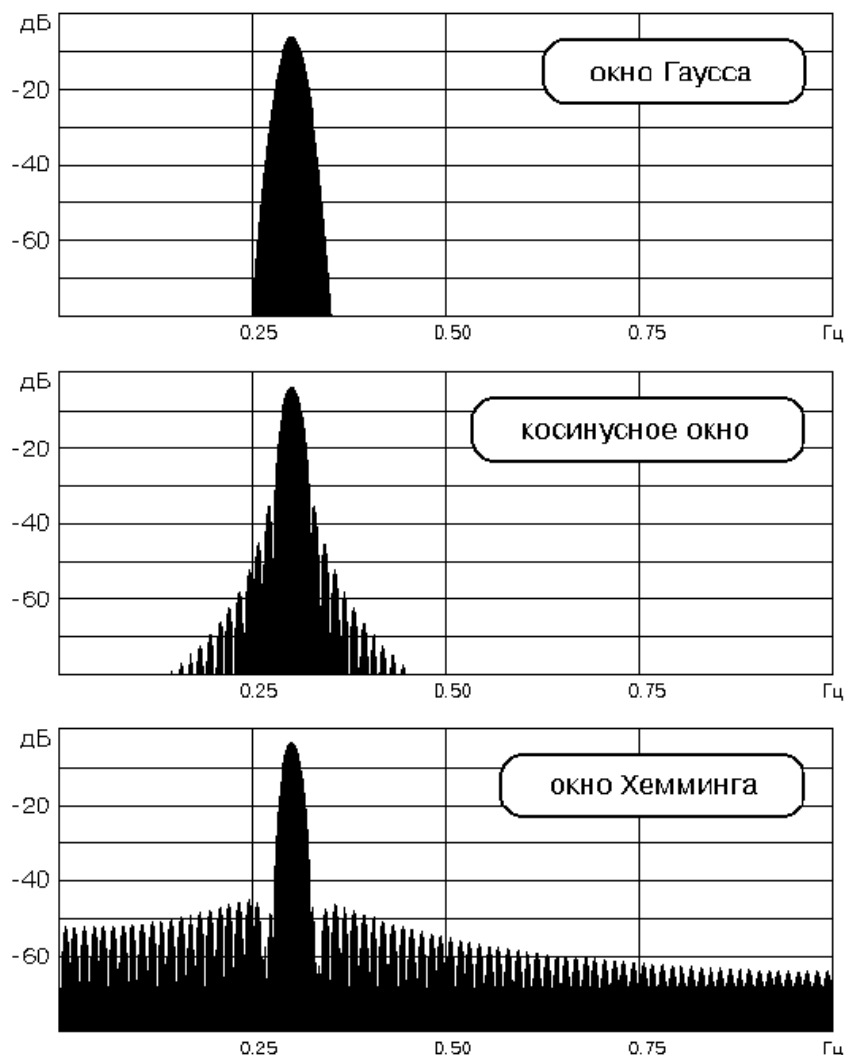


Рис. 3. Различные варианты оконных функций

Существует две разновидности алгоритма БПФ: с прореживанием по времени и частоте. Применение обоих даёт идентичные результаты. На входе функции принимают массив комплексных чисел, состоящий из реальных значений амплитуд сигнала во временной области, а на выходе дают массив комплексных чисел, содержащий информацию об амплитудном и фазовом спектрах.

Конечный массив комплексных чисел содержит в себе ровно половину полезной информации, другая половина является, на самом деле, зеркальным отражением первой и может быть исключена из рассмотрения. Этот момент подтверждает теорему Найквиста-Шеннона.

После преобразования Фурье амплитудный спектр следует нормализовать. Это приведёт к тому, что величины значений амплитуд будут одного порядка независимо от размеров фрейма.

Получив амплитудный и частотный спектры, можно произвести анализ сигнала, например, применить частотную фильтрацию или произвести сжатие. К примеру, таким об-

разом можно сделать эквалайзер: выполнив прямое преобразование Фурье, можно увеличить или уменьшить амплитуду конкретной области частот, после чего выполнить обратное преобразование Фурье (хотя работа настоящих эквалайзеров обычно основана на другом принципе – фазовом сдвиге сигнала) и получить обработанный акустический сигнал путём цифро-аналогового преобразования. К тому же сжать сигнал достаточно просто: необходимо сформировать словарь, где ключ – это частота, а значение – соответствующее ей комплексное число. В словарь нужно добавить только те частоты, амплитуда сигнала на которых превышает заданный минимальный порог. Данные о «тихих» частотах с малой амплитудой, не слышимых человеческим ухом, будут потеряны, но при этом получится ощутимое сжатие с сохранением надлежащего качества звучания. Частично этот принцип является основой для многих кодеков.

### Точное определение частоты

Дискретное преобразование Фурье предоставляет дискретный спектр, где значения амплитуд находятся на равных промежутках по частоте друг от друга. И если при этом частота в сигнале кратна шагу равному отношению частоты дискретизации к количеству отсчётов

$$\partial = \frac{\omega}{n},$$

то, в результате, мы получим выраженный остроконечный пик. Если же частота сигнала находится между границами шага, ближе к середине, то в этом случае образуется пик со «срезанной» вершиной и определение значения частоты сильно затруднится. Существует большая вероятность того, что в сигнале присутствуют две или более частот, находящиеся рядом друг с другом, – в этом и заключается ограничение разрешения по частоте. Подобно тому как на фотоснимке с малым разрешением мелкие предметы склеиваются и становятся неразличимы, тонкие детали спектра тоже могут теряться.

Но частоты музыкальных нот часто не совпадают с сеткой шагов преобразования Фурье, а для повседневных задач, таких как настройка музыкальных инструментов и потоковое распознавание музыкальных нот, необходимо знать точную частоту. К тому же, на низких октавах при разрешении от  $10,24 * 10^3$  отсчётов и ниже, сетка частот Фурье становится настолько редкой, что на одном шаге могут поместиться несколько нот, и определить, какая из них играет в данный момент становится фактически невозможно. Чтобы избежать это ограничение часто применяют аппроксимирующие функции, например, параболические. Но подобные меры являются искусственными. Они улучшают одни показатели и дают искажения в других.

Для уточнения частоты сигнала можно использовать его фазовый спектр. Этот подход основан на вычислении задержки фаз у спектров двух, наложенных друг на друга, фреймов, но, при этом, немного сдвинутых во времени.

В конечном итоге, мы получим словарь частота-амплитуда, где значения частот будут достаточно близки к реальным. Однако «срезанные пики» всё ещё будут проявляться, хоть и менее явно. Чтобы устранить этот недостаток, можно доопределить их путём антиалиасинга.

### Перспективы

Нотный анализ музыкальных произведений открывает ряд интересных возможностей. С помощью готового нотного рисунка можно производить поиск других музыкальных композиций с похожим рисунком.

Например, одно и то же произведение может быть исполнено на другом инструменте, с другой скоростью и/или с другим тембром, либо транспонировано по октавам, при этом нотный рисунок останется похожим, что позволит найти и определить различные варианты исполнения одного и того же произведения.

В некоторых конкретных случаях такой анализ поможет выявить плагиат в музыкальных произведениях. Также по нотному рисунку, теоретически, можно искать произведения определённого настроения или жанра, что поднимает поиск на новый уровень.

### Выводы

В этой статье отображены основные принципы точного определения частот акустических сигналов и выделения нот. А также показана тонкая связь дискретного преобразования Фурье с квантовой физикой.

### Библиографический список

1. Лукин, А. Введение в цифровую обработку сигналов (математические основы) / Лукин А. - Лаборатория компьютерной графики и мультимедиа, МГУ 2007. – С. 15–32.
2. Android API Reference and Guides <http://developer.android.com>
3. Основы спектрального анализа звуков. - Интернет ресурс: <http://pandia.org/text/77/481/644.php>

*Дата поступления  
в редакцию 15.10.2016*

**М.А. Stepanenko, Е.А. Funtikov**

### DIGITIZING AND ANALYSIS OF ACOUSTIC SIGNALS ON ANDROID PLATFORM

Nizhny Novgorod state technical university n.a. R.E. Alexeyev

**Purpose:** This article is devoted to spectral analysis of acoustic signals and identify a piece of music notes in real time using mobile devices on the Android platform.

**Design/methodology/approach:** This article observes the methods of acoustic signals digitizing and time-to-frequency translation.

**Research limitations/implications:** The present study provides an introduction to mathematical base of acoustic signals analysis based on Fourier Transform algorithm.

**Originality/value:** In scope of this study, we described an amount of methods and algorithms for runtime music recognizing.

*Key words:* sound gathering, sampling, Gauss window, antialiasing, harmonic signal.