

ВИЗУАЛИЗАЦИЯ ЗВУКОВОГО СПЕКТРА СРЕДСТВАМИ ЯЗЫКА СИ/СИ++

В работе описана реализация программы на языке Си/Си++ для генерации спектрограммы аудиосигнала. Рассмотрены ключевые этапы обработки: парсинг WAV-файла, разбиение на окна по 2048 сэмплов, применение оконной функции Ханна, вычисление ДПФ, нормализация по шкале децибел, цветовое кодирование и запись в формат BMP. Рассмотрено применение спектрального анализа в промышленных системах распознавания музыки на примере сервиса Shazam.

Введение

Спектральный анализ занимает центральное место в цифровой обработке сигналов и находит применение в распознавании речи, медицинской диагностике, сейсмологии и музыкальных технологиях. Спектрограмма – двумерное изображение, у которого горизонтальная ось соответствует времени, вертикальная – частоте, а яркость – интенсивности сигнала. Цель работы – разработать программу на языке Си, реализующую полный конвейер генерации спектрограммы: от чтения WAV-файла до записи BMP-изображения, используя исключительно знания первого семестра [1].

I. Реализация программы генерации спектрограммы

Практическую значимость спектрального анализа демонстрирует сервис Shazam [2]: алгоритм строит спектрограмму короткого фрагмента, выделяет пики амплитуды – «акустические отпечатки» – и хеширует их пары с учётом временного расстояния (метод созвездий). Программа, реализованная в данной работе, воспроизводит начальный этап этого алгоритма.

Программа включает семь этапов. **Парсинг WAV:** модуль wav_parser.cpp считывает чанки формата RIFF функцией fread, из чанка fmt извлекает параметры сигнала, fseek используется для пропуска неизвестных чанков. **Выделение окон:** сигнал разбивается на фреймы по window_size = 2048 сэмплов с перекрытием overlap = 0,5. **Функция Ханна:** коэффициенты $w(n) = 0,5 \cdot (1 - \cos(2\pi n / (N - 1)))$ подавляют спектральные утечки [1]. **ДПФ и матрица:** выделя-

ется массив num_windows × 1025 ячеек; амплитуды вычисляются через суммирование произведений сэмплов на cos и –sin. **Нормализация:** амплитуды переводятся по формуле $20 \cdot \log_{10}(A/A_{\max} + \varepsilon)$, диапазон [–80, 0] дБ нормируется в [0, 1]. **Тепловая карта:** нормированное значение отображается в RGB линейной интерполяцией между восемью опорными цветами – от чёрного до белого. **Запись BMP:** 54-байтный заголовок BITMAPINFOHEADER записывается одним fwrite, пиксели хранятся в формате BGR по 24 бита, строки дополняются до кратности 4 [3].

Заключение

Реализована программа на языке Си, выполняющая полный конвейер генерации спектрограммы: парсинг WAV, оконное разбиение, функция Ханна, ДПФ, нормализация, цветовое кодирование и запись BMP. Весь конвейер опирается исключительно на средства языка Си первого семестра. Полученные спектрограммы наглядно отображают частотно-временную структуру сигнала, а применённые методы лежат в основе промышленных систем, таких как Shazam.

Список литературы

1. Лайонс Р. Цифровая обработка сигналов / Р. Лайонс. – 2-е изд. – М.: Бинном-Пресс, 2006. – 656 с.
2. Wang A. An Industrial Strength Audio Search Algorithm / A. Wang // Proc. of the 4th Int. Conf. on Music Information Retrieval (ISMIR). – Baltimore, 2003. – Pp. 7–13.
3. Murray J. D. Encyclopedia of Graphics File Formats / J. D. Murray, W. vanRyper. – 2nd ed. – O'Reilly Media, 1996. – 1072 p.

Подворный Владислав Максимович, студент факультета информационных технологий и управления БГУИР, vladvelikij555@gmail.com.

Научный руководитель: Шатилова Ольга Олеговна, старший преподаватель кафедры вычислительных методов и программирования БГУИР, o.shatilova@bsuir.by.