

АЛГОРИТМ НАХОЖДЕНИЯ ПЕРИОДОВ ОСНОВНОГО ТОНА С ПОМОЩЬЮ МНОГОУРОВНЕВОГО АНАЛИЗА ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ СИГНАЛА

Мишунин О. Б., Савинов А. П., Михалёва Е. В., Петровская Т. С.

Национальный исследовательский Томский политехнический университет,

Россия, г. Томск, пр. Ленина, 30, 634050

E-mail: rjawiyyvoz@gmail.com

PITCH DETECTION ALGORITHM USING MULTILEVEL LOCAL EXTREMA ANALYSIS

Mishunin O. B., Savinov A. P., Mikhalyova E. V., Petrovskaya T. S.

National Research Tomsk Polytechnic University

Russia, Tomsk, Lenin str., 30, 634050

E-mail: rjawiyyvoz@gmail.com

***Annotation.** A new method of pitch detection is explored. It is based on finding two levels of local extrema in filtered source signal, second of which is derived from the first. In different parts of signal the sequences of these extrema may represent the pitch period markers, and task of the algorithm shown below is to determine whether they do it or not and build united sequence of markers that accurately match pitch periods.*

Для поиска периодов основного тона был разработан алгоритм, состоящий из следующих шагов:

1. Фильтрация сигнала с помощью фильтра нижних частот Баттерворта второго порядка с частотой среза 500 Гц.
2. Для уменьшения влияния шума установка сигнала равным нулю в местах, где его значение не превышает 0.005 (это значение устанавливается экспериментально и может меняться).
3. Поиск координат всех максимумов в сигнале (Уровень 1).
4. Повторный поиск максимумов в списке, полученном на этапе 3 (Уровень 2).
5. Поиск на втором уровне сигнала достаточно продолжительного участка, имеющего разброс расстояний между соседними максимумами меньше граничного значения. Разброс определяется как $1 - \frac{G(x)}{M(x)}$, где G - среднее геометрическое расстояний между соседними максимумами, M - среднее арифметическое расстояний между соседними максимумами. Граничное значение 0.01, длина интервала - 8 максимумов. Среднее расстояние между соседними максимумами на этом участке можно считать одним из возможных значений периода основного тона в анализируемом сигнале, и все остальные возможные его значения должны отличаться от него не более, чем на некоторую величину «pitch_threshold», например, 50%. Полученное значение назовем «average_pitch». Поиск заканчивается на первом попавшемся участке, соответствующем условию.
6. На каждом из уровней максимумов делаем поиск, аналогичный предыдущему этапу, но ослабляем условия: длина интервала уменьшается до трех максимумов, максимальное значение разброса увеличивается до 0.015. Перебираем все последовательности по три максимума, и те, что соответствуют условию, дополнительно проверяем следующим образом:

Минимальное расстояние между соседними максимумами должно быть больше, чем

$$average_pitch - average_pitch * pitch_threshold.$$

Если условие выполняется, все три максимума текущего интервала помечаем, как соответствующие периодам основного тона.

7. Получаем набор максимумов, соответствующих периодам основного тона, сложив множества отмеченных на предыдущем этапе максимумов так, чтобы значения не дублировались.

8. Восстанавливаем максимумы, расположенные на участках, где амплитуда сигнала плавно падает или возрастает. Они отсутствуют на втором уровне, а на первом уровне не соответствуют условию из шага 6. Для этого применим следующий алгоритм для каждого из максимумов, полученных на предыдущем этапе:

Если расстояние от текущего максимума до следующего не лежит в диапазоне

$$average_pitch - average_pitch * pitch_threshold \leq x \leq average_pitch + average_pitch * pitch_threshold,$$

делаем предположение, что в этом диапазоне может быть потерянный максимум, и восстанавливаем его с первого уровня максимумов, взяв из него точку, ближайшую к координате, соответствующей координате текущего максимума сложенную с расстоянием до предыдущего максимума. Повторяем алгоритм до тех пор, пока он не перестанет находить новые точки.

9. Считаем частоту пересечений нуля на каждом из найденных периодов основного тона и отсеиваем те, где она превышает некоторое значение, например, 5000. Это позволит отсеять лишние периоды, расположенные на согласных вроде «с» или «ш». Частота пересечения нуля рассчитывается на исходном сигнале, не прошедшем фильтрацию.

10. Сдвигаем маркеры основного тона с максимумов до ближайшего левого пересечения нуля на отфильтрованном сигнале и далее до ближайшего слева пересечения нуля на исходном сигнале. Теперь маркеры расположены в начале периодов основного тона.

11. Рассчитываем частоту основного тона для каждого из найденных периодов и к полученному массиву применяем медианное сглаживание.

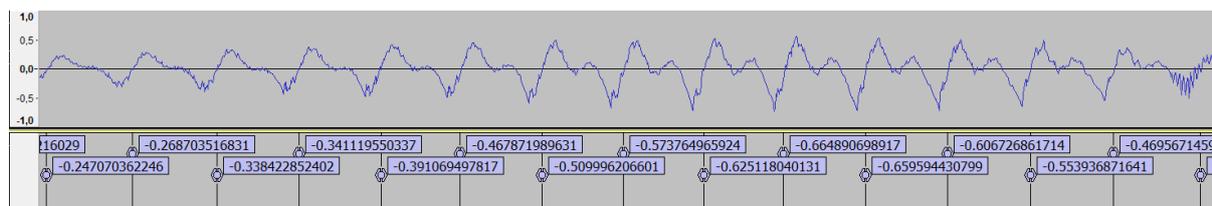


Рис. 1. Пример сегментации сигнала на периоды основного тона

Эксперименты с различными записями звуковых сигналов показали корректную расстановку маркеров основного тона на вокализованных участках за редкими исключениями, которые могут быть исправлены с помощью дополнительной корректировки параметров алгоритма для ослабления или усиления некоторых заданных ограничений. Тишина и невокализованные участки игнорируются.

СПИСОК ЛИТЕРАТУРЫ

1. С.Ю. Клименко, А.П. Савинов. - Математическое модулирование модуляции биений, возникающих при суперпозиции акустических сигналов // Известия томского политехнического университета №2 Том 316. - Томск: Издательство ТПУ, 2010. - С. 135-142.
2. Shlomo Dubnov (2004). "Generalization of Spectral Flatness Measure for Non-Gaussian Linear Processes". Signal Processing Letters 11 (8): 698–701. doi:10.1109/LSP.2004.831663. ISSN 1070-9908.