

УДК 004.421:534.6

АЛГОРИТМ ОЦЕНКИ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ МОЩНОСТИ
ВЕТРОВОЙ ПОМЕХИ В АУДИОСИГНАЛАХ

В. Э. ВОРОБЬЕВА

Научный руководитель И. А. БУЛЯКОВА

Государственный университет «Дубна»

Дубна, Россия

При записи звука на открытом воздухе воздушный поток, обтекающий микрофон и корпус регистратора, создаёт пульсации давления у капсуля. В записи они проявляются как нестационарный низкочастотный шум, маскирующий слабые акустические события при биоакустическом, экологическом и техническом мониторинге. Фильтр верхних частот подавляет её грубо и без количественного описания, а в той же полосе присутствуют вибрации корпуса, шаги, техника и низкие компоненты речи. Целью работы была разработка алгоритма, который оценивает спектральную плотность мощности (СПМ) ветровой составляющей и локализует её на частотно-временной сетке.

Исходными данными служили полевые WAV-записи с частотой дискретизации 44 100 Гц. В программе Audacity вручную выделено 1389 фрагментов, приведённых к длительности 1 с. Каждому присвоен экспертный индекс выраженности ветровой помехи I от 0,5 до 5,0. Индекс отражает слышимую и видимую на спектрограмме выраженность помехи и не является оценкой скорости ветра: синхронные анемометрические измерения не выполнялись. После вычитания среднего СПМ фрагмента оценивалась методом Уэлча [1, 2]: окно Ханна длиной 2048 отсчетов ($\Delta f \approx 21,5$ Гц), перекрытие 50 %, усреднение модифицированных периодограмм.

По оценке СПМ вычислялись нижняя f_{low} и верхняя f_{high} границы области, превышающей относительный энергетический порог, ширина полосы, интегральная энергия в диапазоне 0...500 Гц и спектральный центроид. Нижняя граница почти не менялась с ростом I , верхняя росла. Интегральная энергия зависела от усиления тракта и оказалась менее устойчивым признаком. Для агрегированных по уровням индекса значений f_{high} получена квадратичная модель

$$f_{high}(I) = 201,17 + 106,11 \cdot I + 3,31 \cdot I^2 \quad (1)$$

с коэффициентом детерминации $R^2 = 0,9983$. Модель описывает расширение полосы примерно от 255 Гц при $I = 0,5$ до 814 Гц при $I = 5,0$ и применима лишь к использованной аппаратуре и диапазону индекса.

Для локализации помехи матрица мощности $P(m, k)$ по временным фреймам m и частотным бинам k нормируется на максимум фрейма и рассматривается как решётка клеточного автомата [3] с четырьмя состояниями: фон, граница, вероятный ветер, ветер. Клетки с частотой выше $f_{high}(\hat{I})$, где $\hat{I}$ – оценка индекса по низкочастотной энергии фрейма, ограниченная диапазоном 0,5...5,0, переводятся

в состояние «фон». Для остальных клеток на каждой итерации вычисляется оценочная функция

$$u(m, k) = \alpha \cdot q(m, k) + \beta \cdot p(m, k) + \gamma \cdot n(m, k) + \delta \cdot b(m, k), \quad (2)$$

где q – нормированное предыдущее состояние клетки; p – нормированная мощность; n – доля соседей в активных состояниях в окрестности Мура 3×3 ; b – индикатор допустимой полосы; веса $\alpha = 0,35$, $\beta = 0,40$, $\gamma = 0,20$, $\delta = 0,05$.

Новое состояние задается порогами 0,25, 0,45 и 0,70. Выполняется пять итераций: к этому шагу карта состояний стабилизировалась, а дальнейшие итерации лишь расширяли границы области. Маска ветровой области объединяет клетки в состояниях «вероятный ветер» и «ветер». Ее произведение на $P(m, k)$ дает оценку СПМ ветровой составляющей.

Проверка выполнена на размеченной выборке. Согласие автоматической маски M_{auto} с эталонной областью M_{ref} , построенной по ручной разметке и спектральным признакам фрагмента, оценивалось коэффициентом пересечения

$$D = N(M_{ref} \cap M_{auto}) / N(M_{ref} \cup M_{auto}), \quad (3)$$

где $N(\cdot)$ – число клеток области.

Среднее значение D составило 0,82, медианное – 0,84. Доля активных клеток вне допустимой полосы – 0,035 в среднем и 0,028 по медиане. Расхождения сосредоточены на фронтах ветрового участка, где энергия меняется плавно, а ручная метка задает резкую границу. Короткие удары и вибрации попадают в полосу, но не образуют протяженной области и подавляются требованием поддержки соседей. Обработка секундного фрагмента занимает в среднем 0,046 с (примерно в 21,7 раза быстрее реального времени), пиковое потребление памяти – около 168 МБ.

Модуль реализован на C++17 с библиотеками libsndfile и FFTW3 и сборкой CMake. Он работает как консольное приложение и как динамическая библиотека libwindca.so / windca.dll с C-интерфейсом: функция обработки принимает путь к WAV-файлу, сохраняет признаки в CSV и возвращает оценку индекса, границы полосы и энергию. Маска на частотно-временной сетке допускает визуальную проверку и может служить входом для последующей фильтрации. Ограничения: шкала I имеет только акустический смысл; пороги и веса подобраны для одного микрофонного тракта; проверка на независимых записях и калибровка по показаниям анемометра составляют предмет дальнейшей работы.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Welch, P. D.** The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms / P. D. Welch // IEEE Transactions on Audio and Electroacoustics. – 1967. – Vol. AU-15, № 2. – P. 70–73.
2. **Марпл-мл., С. Л.** Цифровой спектральный анализ и его приложения / С. Л. Марпл-мл. – М. : Мир, 1990. – 584 с.
3. **Бандман, О. Л.** Клеточно-автоматные модели естественных процессов и их реализация на современных компьютерах / О. Л. Бандман // Прикладная дискретная математика. – 2017. – № 35. – С. 102–121.