

Статистический анализ спектра акустического сигнала электромеханического устройства

Будалян Я. С., 212 гр.
научный руководитель — к.т.н. доц. Грачев Е. А.

2015 г.

В моей работе рассматривается задача сегментации акустического сигнала, снимаемого с некоторого устройства. Сегментация - это процесс классификации отсчетов сигнала по некоторому набору категорий, то есть нахождение моментов изменения состояния устройства, что отражается на свойствах снимаемого сигнала. Практические приложения этой задачи возникают в различных областях:

- Обнаружение неисправностей и диагностика
- Анализ человеческой речи
- Прогнозирование природных катастроф(землетрясения, цунами, и т.д.)
- Мониторинг в биомедицине, и т.д.

В качестве математической модели таких акустических сигналов будет использоваться авторегрессионный процесс. Такой подход часто применяется при моделировании различных случайных сигналов, от человеческой речи и шума от устройств [1, 2] до биологических сигналов [3].

Постановка задачи

Дан авторегрессионный процесс $x = (x_t)_{t \geq 0}$

$$x_t = \phi_0(h_t) + \sum_{i=1}^n \phi_i(h_t)x_{t-i} + B(h_t)\xi_t,$$

где ξ_t - стандартный белый шум.

Параметры процесса могут скачкообразно изменяться, принимая в каждый момент времени $t \geq 0$ один из m известных наборов значений

$[\phi_0(h_t), \dots, \phi_n(h_t), B(h_t)]; h_t \in \{0, \dots, m-1\}$. Также задана матрица $Q = \{q(i, j)\}$ вероятностей переходов между наборами параметров (классами):

$$P(h_t | h_{t-1}) = q(h_{t-1}, h_t)$$

Под задачей сегментации понимается задача отношения каждого отсчета процесса x_t некоторому классу h_t , т.е. восстановление ненаблюдаемой последовательности "переключений" состояний $h_i \rightarrow h_j$.

Рассмотренные алгоритмы

Для решения этой задачи были рассмотрены 2 алгоритма:

- Алгоритм сегментации на основе метода динамического программирования [4]
- Алгоритм на основе статистики кумулятивных сумм (CUSUM)

Оба алгоритма были реализованы в программном коде на языке Python. Было проведено экспериментальное исследование алгоритмов на модельном сигнале и на реальных данных с электромеханического устройства - вентилятора.

Алгоритм на основе метода динамического программирования

Ключевая идея алгоритма заключается в последовательной обработке всех отсчетов сигнала путем рекуррентного пересчета вектора

$$d_t(h_t) = \min[d_{t-1}(h_{t-1}) + \beta_t(h_{t-1}, h_t)],$$

и построении матрицы K по правилу

$$k_t(h_t) = \arg \min[d_{t-1}(h_{t-1}) + \beta_t(h_{t-1}, h_t)],$$

где

$$\beta_t(h_{t-1}, h_t) = \frac{1}{2B(h_t)} [x_t - \phi_0(h_t) - \sum_{i=1}^n \phi_i(h_t)x_{t-i}]^2 - \ln q(h_{t-1}, h_t)$$

Построенная матрица K позволяет найти оптимальную сегментацию по рекуррентной формуле

$$\hat{h}_{s-1} = k_s(\hat{h}_s), \quad s = N, N-1, \dots, 1$$

Алгоритм на основе статистики кумулятивных сумм

Ключевая идея алгоритма CUSUM заключается в том, что условное распределение величин x_t $p(x_t|x_{t-1}, \dots, x_1)$ до и после момента переключения отличается математическим ожиданием. Зафиксируем некоторый переход $h_i \rightarrow h_j$. Возьмем логарифм отношения правдоподобия

$$L_t = \ln \frac{p_j(x_t|x_{t-1}, \dots, x_1)}{p_i(x_t|x_{t-1}, \dots, x_1)} = (x_t - \phi_0(h_i) - \sum_{k=1}^n \phi_k(h_i)x_{t-k})^2 - (x_t - \phi_0(h_j) - \sum_{k=1}^n \phi_k(h_j)x_{t-k})^2$$

и построим кумулятивные суммы по правилу

$$\begin{cases} z_1 = 0, \\ z_t = \max(0, z_{t-1} + L_t), \quad t = 2, 3, \dots \end{cases}$$

При отклонении от ожидаемого среднего кумулятивная сумма растет и в какой-то момент превысит заданный порог T_{ij} ; после этого считается, что произошел переход $h_i \rightarrow h_j$.

Поскольку аналитический способ построения пороговых значений для различных переходов неизвестен, во всех экспериментах примерные значения порогов были выбраны вручную.

Результаты

Модельный сигнал

Ниже приведены графики, на которых изображены сегментации модельного сигнала, построенные обеими программами, наложенные на оригинальную сегментацию.

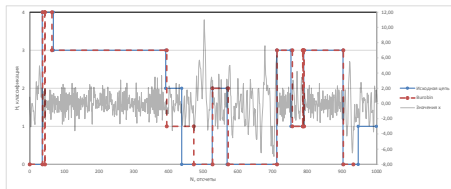


Рис.: Сигнал 1, $t = 1000$

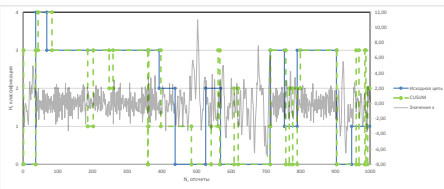


Рис.: Сигнал 1, $t = 1000$

Результаты

Модельный сигнал

Таблица: Параметры модельного сигнала

h	ϕ_0	ϕ_1	ϕ_2	B
0	0	1.36	-0.49	1
1	0	1.02	-0.40	1
2	0	0.82	-0.49	1
3	0	0	-0.49	1
4	0	-0.82	-0.49	1
$q_{ii} = 0.99$				
$q_{ij} = 0.0025, i \neq j$				

Было проведено исследование качества обоих алгоритмов. В качестве меры точности алгоритма было взято среднее время совпадения построенной сегментации с оригинальной, усредненное по $N = 20000$ запусков. Были получены следующие результаты:

$$t = 1000, Q_B = 0.86795, Q_{CUSUM} = 0.76240$$

$$t = 2000, Q_B = 0.86408, Q_{CUSUM} = 0.75599$$

$$t = 3000, Q_B = 0.86235, Q_{CUSUM} = 0.75323$$

Результаты

Сигнал с вентилятора

Были произведены 3 записи шума вентилятора в нормальном режиме работы и с разладкой(в работающий вентилятор засовывалась бумажка) в формате [норма|разладка|норма] по 20 секунд на каждый сегмент. Из каждой сегмента записи было вырезано по секундному отрезку и составлены сигналы длиной по 3 секунды. Общее количество отсчетов в каждом сигнале составило $3с \cdot 44100\text{Гц} = 132300$.

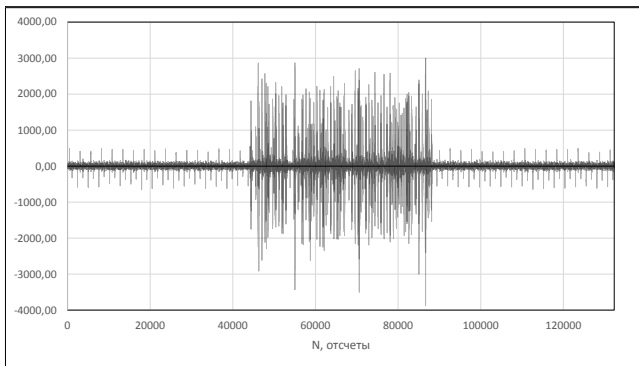
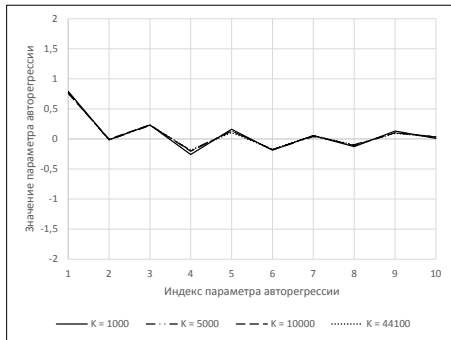


Рис.: Пример анализируемого сигнала

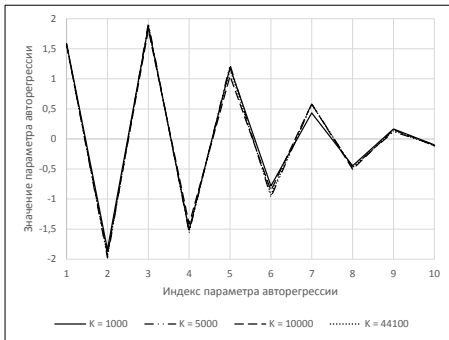
Результаты

Сигнал с вентилятора

Один из сигналов был использован для подбора параметров для алгоритмов. С помощью МНК были определены коэффициенты авторегрессии с глубиной модели $P = 10$. На графиках показаны зависимости коэффициентов авторегрессии для этого сигнала а) до разладки, б) во время разладки от количества отсчетов, анализируемых МНК.



а)



б)

Рис.: Значения коэффициентов авторегрессии до и во время разладки

Результаты

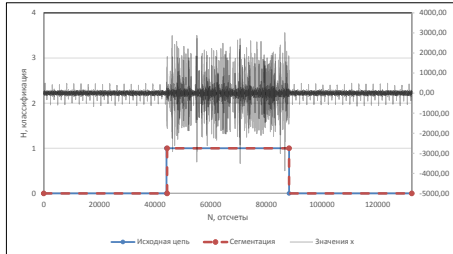
Сигнал с вентилятора

Таблица: Оценки коэффициентов авторегрессий, полученные МНК

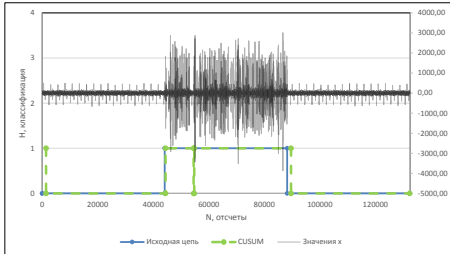
h	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}
0	0	0.78	0	0.23	-0.19	0.11	-0.17	0.06	-0.1	0.1	0
1	0	1.58	-1.94	1.88	-1.53	1.16	-0.93	0.58	-0.47	0.14	-0.11

В таблице h_0 соответствует нормальной работе вентилятора, а h_1 - разладке. Подобранные параметры для алгоритмов составили

$$B(h_i) = 10^5, \quad q_{ii} = 0.99999, \quad T_{ij} = 4 \cdot 10^6.$$



а) алгоритм Буробина



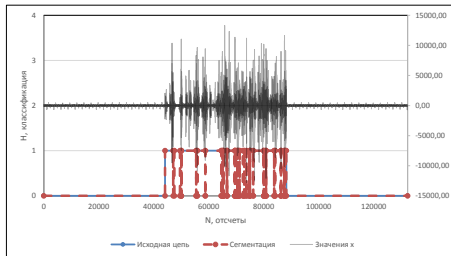
б) CUSUM

Рис.: Сегментация первого сигнала с вентилятора обоими алгоритмами

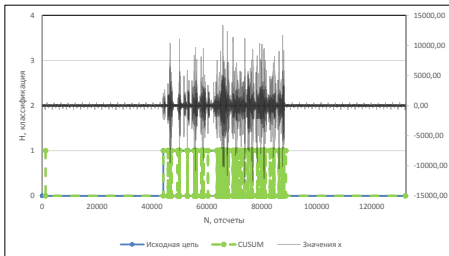
Результаты

Сигнал с вентилятора

Найденные для сигнала 1 параметры были использованы для сегментации сигналов 2 и 3:



а) алгоритм Бурбина

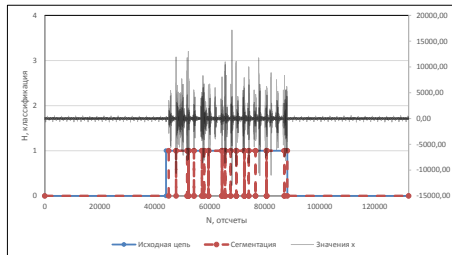


б) CUSUM

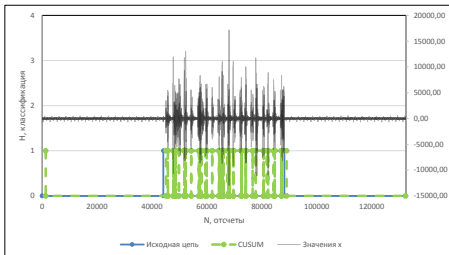
Рис.: Сегментация сигнала 2

Результаты

Сигнал с вентилятора



а) алгоритм Буробина



б) CUSUM

Рис.: Сегментация сигнала 3

Результаты, полученные из экспериментов с модельным сигналом показывают, что

- Алгоритм на основе метода динамического программирования
 - ▶ имеет в среднем на 10% большую точность по сравнению с алгоритмом на основе статистики CUSUM
 - ▶ учитывает вероятности переходов, заданные в матрице Q
- Алгоритм на основе статистики CUSUM:
 - ▶ качество сегментации сильно зависит от выбора пороговых значений T_{ij} для каждой пары переходов
 - ▶ легче в реализации
 - ▶ более производителен

Эксперименты на реальном сигнале с электромеханического устройства (вентилятора) показали, что рассматриваемый подход применим к реальным сигналам, однако требует некоторой доработки.

Список литературы



Chukiet Sodsri, "Time-varying autoregressive modelling for nonstationary acoustic signal and its frequency analysis", 2003



Kie B. Eom, "Analysis of Acoustic Signatures from Moving Vehicles Using Time-Varying Autoregressive Models", Multidimensional Systems and Signal Processing 10, pp. 357-378, 1999



Akay, Y.M., "Noninvasive acoustical detection of coronary artery disease: a comparative study of signal processing methods", Biomedical Engineering 40, pp. 571-578, 1993



Николай Буробин, Вадим Моттль, Илья Мучник, "Алгоритм определения моментов многократного изменения свойств случайного процесса на основе метода динамического программирования", Статистические проблемы управления 65, стр. 49-57, 1984.



Michèle Basseville, Igor V. Nikiforov, "Detection of Abrupt Changes: Theory and Application", pp. 35-43, 1998