Физические поля корабля, океана и атмосферы

Научная статья УДК 004.93 https://doi.org/10.24866/2227-6858/2021-4/28-38

Л.Г. Стаценко, Д.А. Кузин

СТАЦЕНКО ЛЮБОВЬ ГРИГОРЬЕВНА – д.ф.-м.-н., профессор, lu-sta@mail.ru, https://orcid.org/0000-0001-7498-9534 КУЗИН ДЕНИС АЛЕКСАНДРОВИЧ – старший преподаватель (автор, ответственный за переписку), kuzindeal@gmail.com, https://orcid.org/0000-0002-4247-5675 Политехнический институт Дальневосточный федеральный университет Владивосток, Россия

Спектральные признаки гидроакустических шумов в решении задач классификации подводных объектов

Аннотация: В настоящее время, во многих сферах успешно применяются системы принятия решения, основанные на анализе большого объема данных и методах машинного обучения. Чтобы верно классифицировать объект (или действие), необходимо определить его признаки, которые позволят добиться высокой точности классификации. В статье аналитически и экспериментально обосновывается выбор спектральных признаков гидроакустических шумов для решения задачи классификации подводных объектов с помощью методов машинного обучения. Обучение модели на большом количестве акустических данных и реализация сенсорной сети позволят создать систему оперативного реагирования на проникновение нежелательных объектов (катеров или ныряльщиков) в акваторию, подлежащую охране.

Ключевые слова: машинное обучение, признаки гидроакустических шумов, классификация акустических шумов, обработка звукового сигнала

Для цитирования: Стаценко Л.Г., Кузин Д.А. Спектральные признаки гидроакустических шумов в решении задач классификации подводных объектов // Вестник Инженерной школы Дальневосточного федерального университета. 2021. № 4(49). С. 28–38. https://doi.org/10.24866/2227-6858/2021-4/28-38

Введение

Методы машинного обучения используют определенные признаки в исходных данных, чтобы добиться лучшего результата классификации. В разных сферах используются разные данные: в банковском секторе применяются данные об устройстве, с которого обычно осуществляются переводы, локация, в которой обычно находится владелец, и время, когда обычно эти переводы совершаются, чтобы обнаружить подозрительную активность; в агрономии используют цвет и размеры лепестков цветов, чтобы автоматизировать классификацию видов; для бизнее, связанного с недвижимостью, важны данные о районе, этажности, годе постройки, цене, чтобы спрогнозировать актуальную цену для жилья.

Цель работы – анализ признаков гидроакустических шумов, которые будут применяться для решения задачи классификации подводных объектов. Приводятся результаты вычислений признаков для четырех классов шумов, результаты вычисления степени значимости признака и графики, обосновывающие выбор признаков. В статье представлены результаты экспериментов по классификации гидроакустических шумов, собранных в базу данных, с использованием реализованного алгоритма машинного обучения на языке программирования Python.

[©] Стаценко Л.Г., Кузин Д.А., 2021

Статья: поступила: 15.11.2021; рецензия: 30.11.2021; исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90084\19.

Признаки гидроакустических шумов

В данной работе рассматриваются спектральные признаки, которые широко используются для анализа и классификации акустических сигналов [4–6, 8, 14]: человеческая речь, музыкальные инструменты, жанры музыкальных произведений. Были исследованы и описаны следующие спектральные признаки для гидроакустических шумов: хромаграмма, мелчастотные кепстральные коэффициенты, среднеквадратичное значение энергии, спектральный центроид, ширина полосы частот, частота пересечения нуля. Сравнение признаков приводится на примере четырех классов гидроакустических шумов: дыхание ныряльщика, грузовое судно, портовый тягач, стук капель дождя по поверхности моря. Расчет и визуализация признаков осуществляется с помощью алгоритма, реализованного на языке программирования Руthon.

Хромаграмма отражает количество энергии, которое приходится на каждый из 12 полутонов звуковой октавы в определенный момент времени, выраженное определенным цветом [9, 10]. Хромаграммы гидроакустических шумов представлены на рис. 1.



а – дыхания ныряльщика; б – грузового судна; в – портового тягача; г – стука капель дождя по поверхности моря

Представленные хромаграммы рассчитаны с помощью библиотеки "Librosa" и выполнены в среде Python для одинаковых промежутков времени, равных 5 секундам. Проводя анализ хромаграмм четырех классов гидроакустических шумов, можно увидеть явные различия между каждым из классов, что позволяет сделать заключение о состоятельности идеи использования данного признака для решения задачи классификации.

Мел-частотные кепстральные коэффициенты — это представление спектра звукового сигнала, основанное на линейном косинусоидальном преобразовании логарифмического спек-

тра на мел-шкале, которая является нелинейной [2, 6]. Мел-шкала выражает зависимость высоты тона от частоты. Процесс получения мел-частотных кепстральных коэффициентов представлен в виде алгоритма (рис. 2).



Рис. 2. Алгоритм вычисления мел-частотных кепстральных коэффициентов

Исходный сигнал делится на сегменты (фреймы) и представляется следующим образом: $x_i(n) - j$ -й сегмент размером n.

Для каждого кадра вычисляется частотный спектр для определения присутствующих частот. Расчет производится с применением преобразования Фурье:

$$X_j(k) = \sum_{n=0}^{N-1} x_j(n) \cdot e^{\frac{-2\pi i}{N}kn}.$$
 (1)

Вычисленный спектр переносится на мел-шкалу с помощью набора треугольных фильтров, рассчитанных по формуле

$$H_{m}(k) = \begin{cases} 0, \ k < f(m-1) \\ \frac{k-f(m-1)}{f(m)-f(m-1)}, \ f(m-1) \le k < f(m), \\ \frac{f(m+1)-k}{f(m+1)-f(m)}, \ f(m) \le k \le f(m+1), \\ 0, \ k > f(m+1), \end{cases}$$
(2)

где т – число фильтров, которые нам необходимо получить.

Частоты для набора фильтров рассчитываются по следующей формуле:

$$f(m) = \left(\frac{N}{F_s}\right) \cdot B^{-1}(B(f_1) + m \cdot \frac{B(f_h) - B(f_1)}{M+1}),\tag{3}$$

где $B^{-1}(b) = 700(e^{\left(\frac{b}{1125}\right)} - 1)$ – преобразование из высоты звука в частоту.

Применяем дискретное косинусоидальное преобразование для получения мел-частотных кепстральных коэффициентов:

$$c_{j}(n) = \sum_{m=0}^{M-1} S_{j}(m) \cos(\frac{\pi n \left(m + \frac{1}{2}\right)}{M}),$$
(4)
rge $S_{j}(m) = \ln\left(\sum_{k=0}^{N-1} |X_{j}(k)|^{2} \cdot H_{m}(k)\right).$

Для четырех классов гидроакустических шумов на рис. 3 представлены мел-частотные кепстральные коэффициенты.

Кепстральный коэффициент, имеющий положительное значение, обозначает, что спектр энергии сигнала располагается в области низких частот. Если кепстральный коэффициент имеет отрицательное значение, это означает, что большая часть спектральной энергии сосредоточена на высоких частотах.

Анализируя изображения мел-частотных кепстральных коэфициентов для различных гидроакустических шумов, можно обнаружить их значительное различие.

ВЕСТНИК ИНЖЕНЕРНОЙ ШКОЛЫ ДВФУ. 2021. № 4(49)



Рис. 3. Мел-частотные кепстральные коэффициенты для гидроакустического шума: а – дыхания ныряльщика; б – грузового судна; в – портового тягача; г – стука капель дождя по поверхности моря

Среднеквадратичное значение энергии описывает общую амплитуду сигнала [7].

$$\boldsymbol{E_{rms}} = \sqrt{\frac{1}{N} \sum_{n} |\boldsymbol{x}(n)|^2},\tag{5}$$

где x(n) – дискретный сигнал.

Производя сравнение рассчитанных среднеквадратичных значений энергий гидроакустических шумов, можно обнаружить сходство данного признака для шума портового тягача и шума стука капель дождя по поверхности моря. Графическое представление среднеквадратичного значения энергии (красная кривая на графиках) для каждого из классов показано на рис. 4.





Рис. 4. Среднеквадратическое значение энергии для гидроакустического шума: а – дыхания ныряльщика; б – грузового судна; в – портового тягача; г – стука капель дождя по поверхности моря

Ширина полосы частот характеризует разброс частот, присутствующих в сигнале, и рассчитывается с помощью формулы [12

$$S_f^2 = \sum_{k \in K_+} (k - C_f)^2 \tilde{X}(k),$$
(6)

где C_f – спектральный центроид,

k – индекс частоты,

К₊ – набор (индексов) частот с положительными коэффициентами,

 $\tilde{X}(k)$ – нормализованный спектр.

Спектральный центроид также является признаком, который используется для классификации акустических сигналов [13], и рассчитывается по формуле

$$C_f = \sum_{k \in K_+} k \tilde{X}(k), \tag{7}$$

где k-индекс частоты.

К₊ – набор (индексов) частот с положительными коэффициентами,

 $\tilde{X}(k)$ – нормализованный спектр.

Нормализованный спектр рассчитывается по формуле

$$\tilde{X}(k) = \frac{|X(k)|}{\sum_{k \in K_{+}} |X(k)|'}$$
(8)

где *X*(*k*) является спектром, полученным с помощью дискретного преобразования Фурье, k – индекс частоты;

К₊ – набор (индексов) частот с положительными коэффициентами.

Рассчитанные графики спектрального центроида с полосой частот нанесены на спектр гидроакустического шума для каждого из рассматриваемых классов и представлены на рис. 5.

ВЕСТНИК ИНЖЕНЕРНОЙ ШКОЛЫ ДВФУ. 2021. № 4(49)





Из рисунков видно, что спектральный центроид не подходит в качестве признака для решения задачи классификации гидроакустических шумов, так как его значение для всех четырех классов является схожим.

Частота пересечения нуля является величиной, отражающей, как часто сигнал пересекает ноль (меняет свой знак с положительного на отрицательный и наоборот) в течение одного кадра [8]. Данный признак вычисляется по формуле:

$$Z(i) = \frac{1}{2N} \sum_{n=0}^{N-1} |sgn[x_i(n)] - sgn[x_i(n-1)]|,$$
(9)

где $x_i(n)$ – семпл і-го кадра,

n = 0,1... N-1,

$$sgn[x_i(n)] = \begin{cases} 1, x_i(n) \ge 0\\ -1, x_i(n) < 0. \end{cases}$$

На рис. 6 изображены частоты пересечения нуля и форма сигналов для четырех классов гидроакустических шумов.



Рис. 6. Частота пересечения нуля и форма сигнала гидроакустического шума: а – дыхания ныряльщика; б – грузового судна; в – портового тягача; г – стука капель дождя по поверхности моря

Далее перейдем к решению задачи бинарной классификации гидроакустических шумов по описанным выше признакам. В обучении и тестировании участвуют два класса гидроакустических шумов: грузового судна и прочие шумы.

Классификация гидроакустических шумов с использованием методов машинного обучения

На языке программирования Python реализована и обучена модель логистической регрессии для решения задачи бинарной классификации. Используя все описываемые признаки гидроакустических шумов, получим матрицу ошибок при использовании 21 признака гидроакустичеких шумов (ИП – истинно положительные точки данных; ИН – истинно негативные точки данных; ЛП – ложно положительные точки данных; ЛН – ложно негативные точки данных):

ИП	ЛП
32	0
ЛН	ИН
0	17

Негативной точкой данных является класс «Другие шумы», положительной точкой данных является класс «Шумы грузового судна».

Время, затраченное на вычисление 21 признака из 249 гидроакустических шумов, составило 47,3 с.

Результаты метрик [3] для предсказания моделью логистической регрессии составили: accuracy: 0.976; precision: 0.958; recall: 0.9605; f1 score: 0.9586.

Рассчитав значимость каждого признака [11], получим следующие значения (см. таблицу).

Признак	Важность
6-й мел-частотный коэффициент	2,20
Хромаграмма	1,78
15-й мел-частотный коэффициент	1,68
Ширина полосы частот	1,50
9-й мел-частотный коэффициент	1,39
Среднеквадратичное значение энергии	1,15
7-й мел-частотный коэффициент	0,99
Частота спада	0,97
5-й мел-частотный коэффициент	0,96
1-й мел-частотный коэффициент	0,80
2-й мел-частотный коэффициент	0,80
12-й мел-частотный коэффициент	0,66
Частота пересечения нуля	0,65
14-й мел-частотный коэффициент	0,64
13-й мел-частотный коэффициент	0,62
Спектральный центроид	0,54
3-й мел-частотный коэффициент	0,50
10-й мел-частотный коэффициент	0,50
8-й мел-частотный коэффициент	0,48
11-й мел-частотный коэффициент	0,37
4-й мел-частотный коэффициент	0,37

Важность признаков для решения задачи бинарной классификации

Оптимизируем алгоритм, используя для обучения и тестирования только признаки с большей значимостью (хромаграмма, среднеквадратичное значение энергии, ширина полосы

частот, шестой, девятый и пятнадцатый мел-частотный кепстральный коэффициенты) получим следующую матрицу ошибок:

ИП	ЛП
31	1
ЛН	ИН
0	17

Результаты метрик для предсказания моделью логистической регрессии составили: accuracy: 0.9519; precision: 0.9151; recall: 0.9292; f1 score: 0.9181.

Время, затраченное на вычисление шести признаков из 249 гидроакустических шумов, составило 42,2 с. Чтобы более очевидно продемонстрировать выигрыш во времени, при вычислении только значимых признаков, воспользуемся базой данных из 3240 шумов сердца. Получение шести признаков из 3240 шумов, обучение и тестирование модели заняло 22,46 мин. Получение 21 признака из 3240 шумов, обучение и тестирование модели заняло 25,8 мин. Выигрыш по времени составляет 13%.

Для визуализации доказательства необходимости использования комбинации более значимых признаков построим графики зависимости класса шумов от выбранных признаков. При нанесении точек данных шумов двух классов на плоскость с осями, представляющими значения признаков с худшими показателями (11-й и 4-й мел-частотные кепстральные коэффициенты) получим следующее изображение графика логистической регрессии (рис. 7).



Рис. 7. График логистической регрессии при использовании 11-го (ось x2) и 4-го (ось x1) мел-частотного кепстрального коэффициентов в качестве признаков гидроакустических шумов

Из рисунка видно, что провести линию логистической регрессии, которая бы разделяла классы шумов с высокой точностью, не получается.

При использовании в качестве осей двух признаков с лучшими показателями полезности (хромаграмма и 6-й мел-частотный коэффициент) получим следующее изображение графика логистической регрессии (рис. 8):



Рис. 8. График логистической регрессии при использовании хромаграммы (ось x2) и 6-го мел-частотного кепстрального коэффициента (ось x1) в качестве признаков гидроакустических шумов

Из рис. 8 видно, что провести линию логистической регрессии, точно разделяющую два класса, возможно. Таким образом, можно сделать вывод, что для решения задачи классификации гидроакустических шумов необходимо применять признаки с большим значением важности.

Выводы

Можно успешно использовать методы машинного обучения для классификации гидроакустических шумов. При этом необходимо применять определенные признаки шума для повышения точности и увеличения скорости обучения и принятия решения. Реализованный на языке Python алгоритм машинного обучения извлекает из набора гидроакустических шумов только значимые спектральные признаки и обучает по ним модель. Далее алгоритм, принимая на вход новые данные, классифицирует их. Общая точность классификации достигает 95%. Алгоритм может работать как с наборами, так и отдельными файлами шумов, записанными в формате "wav". В настоящее время идет работа над созданием устройства на базе микрокомпьютера Raspberry Pi для классификации гидроакустических шумов в режиме реального времени.

Заявленный вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Воронцов К.В. Математические методы обучения по прецедентам (теория обучения машин). Москва, 2011. URL: http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf (дата обращения: 10.11.2021).
- 2. Кузин Д.А., Стаценко Л.Г., Анисимов П.Н., Смирнова М.М. Применение методов машинного обучения для классификации акустических сигналов по спектральным характеристикам // Известия СПбГЭТУ ЛЭТИ. 2021. № 3. С. 48–54.
- 3. Лялин С.Г. Исследование мел-частотных кепстральных коэффициентов // Общество. Наука. Инновации (НПК-2017), Киров, 1–29 апреля 2017 г. Киров: Вятский гос. ун-т, 2017. С. 1499–1503.
- 4. Политай К.Д., Козлова Е.И. Методы машинного обучения для распознавания эмоций в речевом сигнале // Компьютерные технологии и анализ данных (СТDA'2020): материалы II Междунар. науч.-практ. конф., Минск, 23–24 апр. 2020 г. Минск: БГУ, 2020. С. 194–197.
- Baniya B.K., Lee J., Li Z.N. Audio feature reduction and analysis for automatic music genre classification. 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2014. 457–462 p.
- 6. Brunese L., Martinelli F., Mercaldo F., Santone A. Deep learning for heart disease detection through cardiac sounds. Procedia Computer Science. 2020;(176):2202–2211.
- 7. Ganchev T., Mporas I., Fakotakis N. Audio features selection for automatic height estimation from speech. Hellenic Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, 2010. 81–90 p.
- Kamarudin N., Al-Haddad S.A. R., Hashim S.J., Nematollahi M.A. Feature extraction using spectral centroid and mel frequency cepstral coefficient for Quranic accent automatic identification. 2014 IEEE Student Conference on Research and Development. IEEE, 2014. 1–6 p.
- Kattel M., Nepal A., Shah A.K., Shrestha D. Chroma feature extraction. Chroma Feature Extraction using Fourier Transform, Jan., 2019. URL: https://www.researchgate.net/publication/330796993_Chroma_Feature_Extraction – 10.11.2021.
- 10. Kronvall T., Juhlin M., Swärd J., Adalbjörnsson S.I., Jakobsson A. Sparse modeling of chroma features. Signal Processing. 2017;(130):105–117.
- 11. Le P.N., Ambikairajah E., Epps J., Sethu V., Choi E.H. Investigation of spectral centroid features for cognitive load classification. Speech Communication. 2011;53(4):540–551.
- 12. Patil N.M., Nemade M.U. Content-based audio classification and retrieval using segmentation, feature extraction and neural network approach. Advances in computer communication and computational sciences. Springer, Singapore, 2019. 263–281 p.
- 13. Rubin J., Abreu R., Ganguli A., Nelaturi, S., Matei I., Sricharan K. Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients. 2016 Computing in cardiology conference. IEEE, 2016. 813–816 p.

- Yang N., Dey N., Sherratt R.S., Shi F. Recognize basic emotional states in speech by machine learning techniques using mel-frequency cepstral coefficient features. J of Intelligent & Fuzzy Systems. 2020;39(2):1925–1936. doi:10.3233/JIFS179963
- 15. Yang W., Krishnan S. Combining temporal features by local binary pattern for acoustic scene classification. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2017;25(6):1315–1321.

FEFU: SCHOOL of ENGINEERING BULLETIN. 2021. N 4/49 *Physical Fields of Ship, Ocean and Atmosphere*

www.dvfu.ru/en/vestnikis

Original article https://doi.org/10.24866/2227-6858/2021-4/28-38

Statsenko L., Kuzin D.

LUBOV G. STATSENKO, Doctor of Physico-Mathematical Sciences, Professor, lu-sta@mail.ru, https://orcid.org/0000-0001-7498-9534 DENIS A. KUZIN, Senior Lecturer (corresponding author), kuzindeal@gmail.com, https://orcid.org/0000-0002-4247-5675 Polytechnic Institute *Far Eastern Federal University* Vladivostok, Russia

Spectral features of hydroacoustic noises in determination of class of underwater objects

Abstract: Decision-making systems based on the analysis of big data and machine learning methods are successfully applied in many areas. It is necessary for correct classification of an object (or action) to determine its features helping to achieve the necessary classification accuracy. The machine learning method was approved analytically and experimentally for solution of the problem of classification of underwater objects using spectral features of hydroacoustic noises. The training of the model with a bid acoustic data and the development of a sensor network will allow to create a system of quick response to the unauthorized access (boats and persons) to protected water area.

Keywords: underwater object, machine learning, hydroacoustic noise, classification, audiosignal processing

For citation: Statsenko L., Kuzin D. Spectral features of hydroacoustic noises in determination of class of underwater objects. FEFU: School of Engineering Bulletin. 2021;(49):28-38. (In Russ.). https://doi.org/10.24866/2227-6858/2021-4/28-38

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

REFERENCES

- 1. Vorontsov K.V. Mathematical methods of teaching by precedents (theory of machine learning). Moscow, 2011. URL: http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf 10.11.2021.
- 2. Kuzin D.A., Statsenko L.G., Anisimov P.N., Smirnova M.M. Applying machine learning methods to acoustic signal classification using spectrum characteristics. Proceedings of Saint Petersburg Electrotechnical University. 2021;(3):48–54.
- 3. Lyalin S.G. Research of mel-frequency cepstral coefficients. Society. The Science. Innovation. Kirov, Russia, April 1–29, 2017. Kirov, Vyatka State University, 2017. 1499–1503 p.
- 4. Politai K.D., Kozlova E.I. Machine learning methods for recognizing emotions in a speech signal. Computer technologies and data analysis, Proc. II Int Conf, Minsk, Belarus, April 23–24, 2020. Minsk, Belarusian State University, 2020. 194–197 p.
- Baniya B.K., Lee J., Li Z.N. Audio feature reduction and analysis for automatic music genre classification. 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2014. 457–462 p.
- 6. Brunese L., Martinelli F., Mercaldo F., Santone A. Deep learning for heart disease detection through cardiac sounds. Procedia Computer Science. 2020;(176):2202–2211.

- 7. Ganchev T., Mporas I., Fakotakis N. Audio features selection for automatic height estimation from speech. Hellenic Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, 2010. 81–90 p.
- Kamarudin N., Al-Haddad S.A. R., Hashim S.J., Nematollahi M.A. Feature extraction using spectral centroid and mel frequency cepstral coefficient for Quranic accent automatic identification. 2014 IEEE Student Conference on Research and Development. IEEE, 2014. 1–6 p.
- 9. Kattel M., Nepal A., Shah A.K., Shrestha D. Chroma feature extraction. Chroma Feature Extraction using Fourier Transform, Jan., 2019. URL: https://www.researchgate.net/publication/330796993-_Chroma_Feature_Extraction 10.11.2021.
- 10. Kronvall T., Juhlin M., Swärd J., Adalbjörnsson S.I., Jakobsson A. Sparse modeling of chroma features. Signal Processing. 2017;(130):105–117.
- 11. Le P.N., Ambikairajah E., Epps J., Sethu V., Choi E.H. Investigation of spectral centroid features for cognitive load classification. Speech Communication. 2011;53(4):540–551.
- 12. Patil N.M., Nemade M.U. Content-based audio classification and retrieval using segmentation, feature extraction and neural network approach. Advances in computer communication and computational sciences. Springer, Singapore, 2019. 263–281 p.
- Rubin J., Abreu R., Ganguli A., Nelaturi, S., Matei I., Sricharan K. Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients. 2016 Computing in cardiology conference. IEEE, 2016. 813–816 p.
- Yang N., Dey N., Sherratt R.S., Shi F. Recognize basic emotional states in speech by machine learning techniques using mel-frequency cepstral coefficient features. J of Intelligent & Fuzzy Systems. 2020;39(2):1925–1936. doi:10.3233/JIFS179963
- 15. Yang W., Krishnan S. Combining temporal features by local binary pattern for acoustic scene classification. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2017;25(6):1315–1321.