

## **РАСПОЗНАВАНИЕ ПАТТЕРНОВ СУБВОКАЛИЗАЦИИ И ИССЛЕДОВАНИЕ ИХ ПРИМЕНЕНИЯ В ЗАДАЧЕ УПРАВЛЕНИЯ МОБИЛЬНЫМ РОБОТОМ**

Б.О. Лавров

студент группы М22-512 НИЯУ МИФИ, lavrov.bogdan@list.ru

*Аннотация.* В статье рассмотрено распознавание паттернов субвокализации (тихой речи) и исследование их применения в задаче управления мобильным роботом. Процесс распознавания паттернов ЭМГ сигнала обычно состоит из трех этапов: предварительная обработка сигнала, т.е. снижение влияния внешних шумов и улучшение соотношения сигнал/шум; извлечение полезной информации; классификация. Система состоит из набора электродов, крепящихся к горлу и подключенных к компьютеру, который анализирует сигнал и преобразует его в текст или голос. Представлены результаты эксперимента, который показал, что модель робота не полностью адаптирована к условиям реального использования и требует дополнительной настройки или обучения.

*Ключевые слова:* паттерны, классификация ЭМГ сигналов, метрики качества модели, метод выбора признаков, модель машинного обучения.

### **Введение**

При любом процессе возбуждения живой ткани возникает биоэлектрический потенциал. В определенных пределах существует прямая зависимость между напряжением, создаваемым мышцами, и уровнем биоэлектрического потенциала. В связи с этим альтернативой произносимой (вокализованной) речи может быть произносимая речь, т.е. внутренняя речь – процесс мысленного представления слова без его проговаривания. В процессе формирования речи помимо акустических сигналов организм генерирует и другие биологические сигналы, которые возникают в организме во время произнесения речи и могут варьироваться от движений артикуляционных органов (гортань, голосовые связки, язык, небо, зубы, губы, носоглотка) до активности нейронов в мозге. В зависимости от конкретных расстройств, поражающих человека, некоторые этапы процесса формирования речи могут быть нарушены, в то время как другие остаются неизменными. Поэтому биологические сигналы, исходящие от неповрежденных частей речевого аппарата, могут быть зарегистрированы, обработаны и преобразованы либо в звуковой сигнал, либо в текст. Иными словами, путем регистрации биосигналов, поступающих от элементов речевого аппарата, можно распознать речь без ее произнесения, что является чем-то большим, чем мысли, но меньшим, чем фактическая речь, и представляет собой желание что-то сказать. Наиболее известными субвокальными системами являются:

- Система субвокального распознавания NASA для астронавтов, которая позволяет им общаться в условиях высокого шума или отсутствия звука; система состоит из набора электродов, крепящихся к горлу и подключенных к компьютеру, который анализирует сигнал и преобразует его в текст или голос [1];
- AlterEgo – система, разработанная в Массачусетском технологическом институте, которая позволяет человеку общаться с компьютером или другими людьми без открытия рта; устройство надевается на голову и имеет четыре электрода, сигнал передается на беспроводные наушники, которые воспроизводят ответ компьютера или собеседника [2].
- Subvocal – система, разработанная в Вустерском политехническом институте, которая позволяет человеку управлять музыкальными инструментами с помощью субвокализации [3].

Основной целью изучения ЭМГ сигналов, является извлечение полезной информации о мышечной активности. Эта задача носит название распознавания паттернов ЭМГ сигнала, которая обычно состоит из трех этапов [4]:

- 1) предварительная обработка сигнала, т.е. снижение влияния внешних шумов и улучшение соотношения сигнал/шум;
- 2) извлечение полезной информации;
- 3) классификация.

Ключевые технические параметры для исследования субвокализации с помощью ЭМГ могут быть следующими:

- расположение электродов,
- количество каналов данных – обычно используются от двух до восьми каналов данных.
- частота целевого сигнала – обычно целевой сигнал имеет частоту от 0 до 500 Гц.
- частота считывания данных – обычно частота считывания данных составляет от 1000 до 5000 Гц.

### **Способы обработки сигнала**

Несмотря на удобство поверхностной ЭМГ, существуют также и недостатки, которые негативным образом влияют на качество исследований. Когда электромиографический датчик установлен непосредственно на коже, то он неизбежно подвергается действию как внешних, так и физиологических факторов, влияющих на качество обнаружения мышечной активности. Помимо шумов мышц, которые располагаются в зоне чувствительности электрода и вносят паразитные помехи посредством собственной электрической активности, на качество регистрации влияет целый ряд факторов [5-7]:

- Шум, вызванный электронным оборудованием;
- Шум окружающей среды;
- Артефакт, связанный с движением;
- Вход в режим насыщения;
- Естественная нестабильность сигнала.

Способы обработки сигнала для распознавания паттернов субвокализации могут включать следующие этапы:

- Фильтрация: применение различных фильтров для удаления шума и артефактов из сигнала, таких как фильтры высоких и низких частот, полосовые фильтры, вейвлет-фильтры и т.д.;
- Извлечение признаков: это процесс преобразования входного сигнала в набор числовых характеристик, которые отражают его свойства и позволяют отличать различные паттерны субвокализации; эти характеристики называются признаками и используются для обучения моделей распознавания речи; среди основных особенностей ЭМГ сигнала можно выделить временные и частотные характеристики.

### **Временные характеристики сигнала**

Суммирование ЭМГ сигнала представляет собой суммирование сигналов ЭМГ в окне определенного размера, определяемого количеством отсчетов  $N$ . Впоследствии данная сумма сравнивается с ранее определенным порогом и принимается решение об активности мышцы.

$$G = \sum_{i=1}^N |x_i|,$$

где  $x_i$  — это измеренный сигнал ЭМГ,  $N$  — размер исследуемого окна (число отсчетов).

Среднее абсолютное значение (MAV) является одной из самых популярных характеристик, используемых при анализе сигналов ЭМГ. MAV имеет несколько модификаций. Условно их можно разделить на два типа. Первый тип — это модификация с добавлением новой переменной  $w_i$ , которая является своего рода весом промежутка внутри окна. Второй тип — это модификация с использованием нестационарной переменной  $w_i$ .

В [8] для анализа предлагается энергия ЭМГ сигнала, рассчитанная как сумма квадратов значений амплитуды ЭМГ сигнала. Другими характеристиками для анализа ЭМГ сигнала являются дисперсия (VAR) и среднее абсолютное значение степенной функции считываемого сигнала. Среднеквадратичное значение (RMS) также является популярной характеристикой для анализа сигнала ЭМГ. Длина волны (WL) — это мера сложности ЭМГ сигнала. Этот параметр определен как совокупная величина изменения амплитуды сигнала ЭМГ за временной сегмент. Изменение средней амплитуды

(AAC) почти эквивалентно WL, за исключением того, что длина волны усреднена. Пересечение нуля (ZC) – это мера частотной информации сигнала ЭМГ, определенная во временной области. Значение амплитуды сигнала ЭМГ может несколько раз пересечь нулевой уровень амплитуды. Изменение знака наклона (SSC) – это ещё один способ представления частотной информации сигнала ЭМГ. SSC определяется количеством раз, когда наклон сигнала ЭМГ меняет знак.

### **Частотные характеристики**

Частотные характеристики сигнала вычисляются на основе преобразования Фурье. Спектральная плотность мощности ( $P$ ) – это мера распределения энергии сигнала по частотам. Она показывает, какая часть общей энергии сигнала приходится на каждую частоту.  $P$  вычисляется с помощью преобразования Фурье от временного представления сигнала.

Средняя частота (MNF) рассчитывается как сумма произведений спектра мощности ЭМГ на частоту, разделенная на суммарную мощность спектра. MNF также называют центральной частотой  $f_c$  или спектральным центром тяжести.

### **Исследование ЭМГ сигналов**

Для измерения активности мышц в работе были использованы датчики ЭМГ Glove от Seeed Studio. Датчик получает сигнал от мышц, затем обрабатывается с двукратным усилением и подается на Arduino.

Сигналы ЭМГ регистрировались с кожи горла с частотой дискретизации 2.8 кГц в комнате с постоянной температурой. Использовались только одноразовые ЭМГ электроды. Вокруг участника эксперимента была установлена разная оргтехника: принтер, ноутбук, блоки питания, сетевые переключатели и т.д.; таким образом, естественные магнитные поля, окружающие человека в повседневной жизни, не были искусственно занижены. Затем испытуемый последовательно произносил шёпотом и обычной речью слова «вперед», «направо», «назад», которые записывались в соответствующие файлы. Запись набора данных осуществлялась несколько раз в разное время.

### **Модели для работы с данными**

Для каждого файла данных (240 файлов) были рассчитаны все временно-частотные признаки (31 параметр), описанные выше. Для выявления наиболее важных признаков, была построена ковариационная матрица, которая предоставляет информацию о взаимосвязи (корреляции) между признаками. Чем ближе к нулю ковариационное значение, тем лучше (это означает, что признаки независимы).

Анализ ковариационной матрицы показал, что к сильно коррелирующим признакам относятся MAV, MAV\_1, MAV\_2, TM3, TM4, TM5, LOG, SSC. Их решено исключить из анализа.

Используя метод главных компонент (PCA), было решено снизить размерность пространства признаков. Путем определения кумулятивной выборочной дисперсии признаков выясняется, какие признаки определяют большую часть дисперсии данных. Методом главных компонент было выбрано 13 признаков из 31, которые определяют 98% дисперсии данных.

Для классификации паттернов были исследованы следующие модели: SCV, DecisionTreeClassifier, GradientBoostingClassifier, LogisticRegression, RandomForestClassifier. Наиболее высокую точность распознавания 83.5% показал RandomForestClassifier, для которого в дальнейшем были подобраны наилучшие гиперпараметры. В результате анализа установлены следующие закономерности влияния гиперпараметров на точность модели:

- `n_estimators`: значения 300, 500, 700, показывают наилучшие средние результаты;
- `min_samples_split`: небольшие значения, вроде 2 и 7, показывают наилучшие результаты; хорошо выглядит и значение 23; можно исследовать несколько значений этого гиперпараметра, превышающих 2, а также – несколько значений около 23;
- `min_samples_leaf`: возникает такое ощущение, что маленькие значения этого гиперпараметра дают более высокие результаты; а это значит, что мы можем испытать значения между 2 и 7;
- `max_features`: вариант `sqrt` даёт самый высокий средний результат;
- `max_depth`: тут чёткой зависимости между значением гиперпараметра и результатом работы модели не видно, но есть ощущение, что значения 2, 3, 7, 11, 15 выглядят неплохо;
- `bootstrap`: значение `False` показывает наилучший средний результат.

Используя эти результаты, с помощью алгоритма GridSearchCV был выполнен более тонкий поиск наилучшей комбинации гиперпараметров: `'bootstrap': False, 'max_depth': 13, 'max_features': 'sqrt', 'min_samples_leaf': 7, 'min_samples_split': 2, 'n_estimators': 700`. Средняя точность классификации паттернов составила 86%.

Тестирование предварительно обученной модели в режиме реального времени проводилось на мобильной роботизированной тележке. Произносились последовательно слова-команды: «направо», «налево» «назад», «вперед» и «стоп». Всего было выполнено 20 циклов. Для каждого слова-команды была составлена матрица ошибок и рассчитаны: полнота (*recall*) –

способность алгоритма обнаруживать данное слово, точность результата измерений (*precision*), точность измерений (*accuracy*) и *f*-мера – конечный показатель эффективности модели.

Полученные результаты показали, что модель не полностью адаптирована к условиям реального использования и требует дополнительной настройки и обучения.

### Заключение

В статье рассмотрены методы распознавания паттернов субвокализации ЭМГ сигнала. Проведен обзор способов обработки сигнала для распознавания паттернов субвокализации. Разработана модель классификации ЭМГ сигнала на основе `RandomForestClassifier`, которая на тестовых данных имеет точность предсказания 86%. Полученный классификатор в процессе проверки продемонстрировал низкое качество распознавания мышечной активности в режиме реального времени. Тестирование модели в режиме реального времени показало, что она не может точно распознавать команды пользователя. Для улучшения результатов необходимо провести дополнительную тренировку модели на более широком наборе данных. Полученные выводы могут служить основой для дальнейших исследований и развития технологии распознавания субвокализации.

### Список литературы

1. Bluck J. NASA Develops System To Computerize Silent, 'Subvocal Speech'. NASA Ames Research Center, Moffett Field, Calif., March 2004. [Электронный ресурс]. [https://www.nasa.gov/centers/ames/news/releases/2004/04\\_18AR.html](https://www.nasa.gov/centers/ames/news/releases/2004/04_18AR.html). (Дата обращения 04.05.2023).
2. Kapur A., Kapur S., Maes P. AlterEgo: A Personalized Wearable Silent Speech Interface. 2018, pp. 43-53.
3. A Review on Electromyography Decoding and Pattern Recognition for Human-Machine Interaction / M. Simao, N. Mendes, O. Gibaru, P. Neto // IEEE 460 Access, 2019, V. 7, pp. 39564-39582.
4. Resolving the 425 limb position effect in myoelectric pattern recognition / A. Fougner, E. Scheme, A. Chan, K. Englehart, O. Stavadahl // IEEE Trans. Neural Rehabil. Syst. Eng., 2011, V. 19, N. 6, pp. 644-651.
5. Reaz M., Hussain M., Mohd-Yasin F. Techniques of EMG signal analysis: detection, processing, classification and applications. // Biol Proced Online, 2006, V. 8, pp. 11-35.
6. Examining the adverse effects of limb position on pattern recognition based myoelectric control / E. Scheme, A. Fougner, A. Chan, K. Englehart // IEEE inginering in Medicine and Biology Soc., 2010, pp. 6337-6340.
7. Du S., Vuskovic M. Temporal vs. spectral approach to feature extraction from prehensile EMG signals. // In Proceedings of IEEE International Conference on Information Reuse and Integration. 2004, pp. 344–350.
8. Phinyomark A., Phukpattaranont P., Limsakul C. Feature reduction and selection for EMG signal classification. // Expert Systems with Applications, 2012, V. 39, № 8, pp. 7420-7431.