

Распознавание речевых команд на фоне интенсивных акустических шумов по кросскорреляционным портретам

А.И. Армер^а, Е.Ю. Галицкая^б, Н.А. Крашенинникова^с

^аУльяновский государственный технический университет, 432027, ул. Северный венец, 32, Ульяновск, Россия

^бУльяновское конструкторское бюро приборостроения, 432071, ул. Крымова, 10А, Ульяновск, Россия

^сУльяновский государственный университет, 432970, ул. Льва Толстого, 42, Ульяновск, Россия

Аннотация

Для создания речевых информационных систем в шумных условиях на производстве и транспорте требуются методы распознавания РК в условиях сильных помех. В настоящей работе рассматривается дикторозависимый способ распознавания речевых команд (РК) из ограниченного словаря на фоне интенсивных акустических шумов путём преобразования РК в кросскорреляционные портреты (ККП), то есть особые изображения, в которые преобразуются РК. Распознаваемая РК относится к классу с минимальным расстоянием (метрикой) между ККП этой команды и эталонными ККП класса. Разработаны алгоритмы преобразования команд в ККП, метод уточнения границ команд, способы оптимизации библиотеки эталонных команд и выбора метрики. В результате получено довольно уверенное распознавание РК при сильных помехах.

Ключевые слова: речевая команда; интенсивные шумы; распознавание; кросскорреляционный портрет; метрика; уточнение границ; эталонная команда; оптимизация библиотеки

1. Введение

Растущая интенсивность производства и транспорта ведёт к увеличению нагрузки на операторов. Для уменьшения этих нагрузок применяются речевые информационные системы. Однако эти системы зачастую должны надёжно распознавать РК при особо интенсивных акустических помехах.

К настоящему времени разработано большое количество систем распознавания речи в условиях пониженного шума. К ним относятся, например, Via Voice фирмы IBM, показывающая в среднем 97 % правильного распознавания при словаре до 2000 РК, Naturally Speaking фирмы Dragon Systems, показывающая в среднем 70 % при словаре до 60000 слов, Voice Xpress фирмы Lernout and Hauspie, показывающая 90 % - 98 % при словаре до 1000 слов и др. Существуют и пользовательские системы понимания и обработки слитной речи – VocalIQ, Siri, Google Now и Cortana. При сравнении VocalIQ с Siri, Google Now и Cortana, запросив у систем сложные запросы на естественном языке, получено успешно выполненных запросов [1] у VocalIQ более 90%, в то время как Google Now, Siri и Cortana справились лишь с 20% подобных задач. Из отечественных разработок известна система VoiceCom фирмы ООО «Центр речевых технологий», способная распознавать 100-200 команд в дикторозависимом и 30-50 РК в дикторонезависимом варианте с вероятностью до 98 %. Однако данные системы не выполняют своей функции в условиях даже небольших акустических шумов.

Системы распознавания РК из ограниченного словаря в условиях акустических помех в настоящее время разрабатываются в основном для авиации и используются в приборах речевого управления и контроля полета. Качество работы таких систем на сегодняшний день составляет от 90 до 98% правильного распознавания команд в зависимости от условий испытания и размера словаря. Практически все испытываемые системы – дикторозависимые. По данным лаборатории Wright Laboratory авиабазы Wright-Patterson, проводимые в 1996 г. летные испытания систем распознавания речевых команд VRS-1290 фирмы ИТТ и VAT31 фирмы Verbex, показали следующие результаты: VRS-1290 – 92-98% правильного распознавания при словаре 50 команд, VAT31 – до 97% (данные о размере словаря отсутствуют). В 1997 г. были получены результаты летных испытаний системы распознавания РК канадской организации National Research Council. Система была встроена в навигационную систему Avionics Management System вертолета Bell 412HP и показала в среднем 95% правильного распознавания при словаре в 80 слов, разбитых на 24 группы. По данным о системе Speech Recognition Module фирмы Smiths Industries, встроенной в блок CAMU истребителя Eurofighter, правильность распознавания команд в типовом полете не менее 95% при словаре в 250 слов, 25 из которых могут быть одновременно активны. В настоящее время для истребителей Rafale фирмой Thales Avionics разрабатывается система распознавания РК с требуемыми характеристиками точности распознавания выше 95% при словаре 50-300 слов. Истребитель 5-го поколения F-35 был пробно оснащён системой распознавания РК DynaSpeak, разработанной SRT International. Качество распознавания по заявлению разработчиков составляет 98%. На многоцелевом истребителе 4-го поколения Eurofighter была установлена система речевого управления, разработанная фирмой Logica. Словарь составляет 250 слов, вероятность правильного распознавания не менее 95%. На экспортный вариант самолета Rafale Block 05 фирма Thales Avionics разработала речевую систему управления с надёжностью распознавания, по заявлению разработчика, не менее 95% при словаре до 300 РК, но данные о внедрении отсутствуют. В патенте США US6529866 B1 от 4.03.2003 (The United States Of America As Represented By The Secretary Of The Navy) описаны способ и система преобразования звукового сигнала, содержащего компоненты речевого сигнала и шума в речь, данные об испытаниях и внедрении отсутствуют. В патенте WO 1999040571 A1 от 3.02.99 (Qualcomm Incorporated) на систему для повышения точности распознавания речи в условиях шума также данные об испытаниях и внедрении отсутствуют.

Из отечественных разработок отметим следующие. Система распознавания РК, испытания которой проводились на самолете МиГ-29. Вероятность правильного распознавания составила 56-81%, данные о словаре отсутствуют. Патент РФ

№2267820 С1 от 25.04.2006 г. (Ульяновский государственный технический университет). Вероятность правильного распознавания составила 92% при словаре 23 РК и уровне шума 3дБ. Данные о внедрении отсутствуют. Патент РФ № 2271578 С2 от 10.03.2006 г. (ООО "Центр Речевых Технологий") Изобретение относится к анализу речи в неблагоприятных условиях окружения, например, в условиях движущегося транспортного средства или в условиях механического производства с высоким уровнем шумности. Данные об испытаниях отсутствуют.

Несмотря на имеющиеся разработки, нет данных о реальном применении систем распознавания РК в действующей авиации, так как в реальных полётах разработанные системы показывали существенно меньшую эффективность, чем предполагалось.

Таким образом, задача создания систем распознавания РК в условиях сильных шумов остаётся актуальной.

В настоящей работе рассматривается дикторозависимый метод распознавания РК из ограниченного словаря путём преобразования РК в портреты, то есть в изображения.

2. Методы распознавания РК

Методам распознавания речи и, в частности РК, посвящено большое количество работ. Первые методы автоматического распознавания звуков речи были получены еще в первой половине XX века [2]. Среди методов распознавания речи можно выделить следующие основные классы: спектральные методы [3, 4, 5, 6, 7], методы, основанные на вейвлет-анализе [3], статистические методы [5, 8, 9, 10, 11], использование нейронных сетей [12, 13].

В настоящей работе рассматривается метод распознавания РК, основанный на их преобразовании в портреты, то есть плоские изображения с дальнейшим применением методов обработки изображений [14, 15, 16, 17, 18].

3. Автокорреляционные портреты

Пусть $S = s_0, s_1, s_2, s_3, \dots, s_{N-1}$ цифровые отсчеты РК. Тогда его автокорреляционный портрет (АКП) будет являться двумерное изображение $X(i, k) = \{x_{ik} : i = 1, 2, 3, \dots; k = 1..K\}$, полученное следующим образом. Разобьем РК S на M частей и выполним преобразование

$$X(i, k) = \frac{Cov(S_n, S_{n+k})}{\sigma_n \sigma_{n+k}}, \quad (1)$$

где $Cov(S_n, S_{n+k})$ – выборочная ковариация отрезков S_n, S_{n+k} сигнала S , отстоящих друг от друга на время $k\Delta t$, $\sigma_n^2, \sigma_{n+k}^2$ – выборочные дисперсии соответственно отрезков S_n, S_{n+k} . Таким образом, k -й элемент i -й строки АКП равен коэффициенту корреляции между i -м отрезком S_i и отрезком, сдвинутым относительно S_i на k отсчетов. На рисунке 1 приведены примеры АКП.

Отметим некоторые свойства АКП, делающие их привлекательными для распознавания РК. Портреты РК достаточно индивидуальны, то есть АКП разных РК различны, а одинаковых, но произнесенных в разное время – схожи. Автокорреляционное преобразование выполняет функцию нормирования сигнала, вследствие чего АКП слабо чувствительны к громкости и малочувствительны к медленно изменяющимся аддитивным добавкам. Если шум белый аддитивный с дисперсией σ_θ^2 , то его АКП и отсчеты АКП РС, искаженного шумом, будут отличаться на постоянный множитель.

Однако АКП обладают и некоторыми отрицательными качествами, среди которых зависимость яркостей элементов от различий в тембре произнесения РК, а также геометрические искажения АКП из-за вариации темпа речи. Эти искажения можно ослабить путём применения некоторых модификаций построения АКП, например, учитывая экстремумы громкости.

Распознавание РК по их АКП производится следующим образом. В памяти хранятся АКП эталонных РК. Распознаваемая РК преобразуется в АКП и относится к тому классу, у которого минимально расстояние между его эталонным портретом и АКП распознаваемой команды.

Это расстояние (метрика) между двумя АКП (то есть изображениями) вычисляется следующим образом. Сначала два изображения совмещаются, то есть для каждой строки одного изображения находится соответствующая строка другого изображения. За расстояние между портретами принимается среднее расстояние (например, Евклидово) между соответствующими строками. Это соответствие для АКП одной и той же команды означает близость звука фрагментов РК, поэтому оно сравнительно невелико, так как происходит только от разницы в произнесении и наличия шума. Если же сравниваются АКП разных команд, то это расстояние обычно значительно больше из-за различия звуков. При совмещении команд применялось динамическое программирование по критерию минимума расстояния.

В эксперименте по распознаванию РК, произносимых диктором в реальном времени был использован словарь, содержащий десять групп РК – по 4-23 команды авиационной тематики в каждой группе. Всего в словаре содержалось более 100 РК. В качестве фонового и опорного шума использовался реальный шум авиационного двигателя, записанный в режиме полета, отношение сигнал/шум 5-0 дБ. В испытаниях принимали участие четыре диктора мужского пола.

Предварительно каждым из дикторов были произнесены эталонные РК, при этом каждая РК из словаря произносилась два раза. В ходе эксперимента по распознаванию РК каждый из дикторов произносил все команды из словаря по три раза, всего в эксперименте было произнесено более 1200 РК. Правильно было распознано более 95 % команд.

Однако дальнейшие исследования показали, что с течением времени после начитывания эталонов вероятность правильного распознавания может существенно снижаться. Связано это с эффектом «старения» эталонов, то есть с течением времени тембр голоса диктора может меняться, и ранее произнесённые эталоны недостаточно отражают особенности речи диктора во время настоящего распознавания РК. Поэтому эталоны надо бы время от времени обновлять (например, перед полётом), что создаёт определённые неудобства.

Один эталон РК не отражает возможные разнообразия её произнесения, поэтому было увеличено их количество, то есть диктор произносил каждую команду несколько раз в разное время. Совокупность этих произнесений уже несколько отражала многообразие произнесения. Однако увеличение количества эталонов усложняет и замедляет алгоритм распознавания, что допустимо лишь до определённой степени. Поэтому количество эталонов оказывается ограниченным. Но при этом эти эталоны должны бы максимально отражать вариативность произнесений. Оказалось, что вероятность правильного распознавания очень сильно (на десятки процентов) зависит от того, насколько удачно выбраны эталоны из имеющегося набора. Таким образом, из нескольких произнесений каждой РК следует так выбрать заданное количество в качестве эталонов, чтобы полученная библиотека эталонов обеспечивала наилучшее распознавание РК. Эта задача оптимизации библиотеки эталонов рассматривалась в [19, 20, 21]. Полный перебор вариантов библиотеки осуществить технически невозможно, поэтому был разработан способ направленного перебора, дающий почти оптимальный вариант. Иногда бывает возможно изменять состав самих РК, используя синонимы слов. Эта задача тоже рассматривалась и было найдено её решение на основе анализа различимости имеющихся групп синонимов.

4. Кросскорреляционные портреты

Другим способом ослабления влияния вариативности произнесения РК является использование вместо АКП другого вида портретов.

При формировании АКП находятся коэффициенты корреляции между отрезками одной и той же РК (автокорреляция). При распознавании по АКП находятся расстояния между АКП распознаваемой и АКП эталонов. При нахождении расстояния между АКП распознаваемой команды и АКП её же эталона будут сравниваться АКП двух разных произнесений этой команды, которые могут значительно отличаться друг от друга (расстояние будет большим). Поэтому целесообразно, чтобы при сравнении портретов влияние разницы произнесений было меньше. Для этого нужно, чтобы вариативность произнесений в какой-то мере учитывалась в портретах. Рассмотрим для этой цели кросскорреляционный портрет (ККП), который состоит из коэффициентов корреляции между отрезками двух РК (кросскорреляция) [15, 16, 17, 22].

Пусть имеются две команды $S1$ и $S2$. Разобьем каждую из команд на M отрезков одинаковой длины и вычислим выборочные коэффициенты корреляции x_{ik} между i -м отрезком команды $S1$ и участком команды $S2$, начиная с k -го отсчёта i -го отрезка команды $S2$. В результате получаем двумерный массив (изображение) $X = \{x_{ik}\}$, называемый ККП команд $S1$ и $S2$.

Рассмотрим подробно построение ККП. В качестве примера рассмотрим построение ККП двух произнесений одной команды авиационной тематики, первое из них $S1$ и $S2$ – второе. После разбиения на отрезки одинаковой длины $N1$ – длина каждого отрезка сигнала $S1$, $N2$ – длина каждого отрезка сигнала $S2$, обозначим за $N = \min\{N1, N2\}$ минимальное из этих длин. При задании количества отрезков каждой команды M необходимо учесть, что слишком маленькое его значение не позволит вместить в отрезок значимую часть фонемы, а слишком большое значение наоборот вместит в отрезок слишком много фонем, что отрицательно скажется на коэффициенте корреляции между отдельными фонемами в разных командах при построении ККП.

Вычислим коэффициенты корреляции i -го отрезка сигнала $S1$ и i -го отрезка сигнала $S2$, сдвигаемого на $k = 0..K$ отсчётов вправо.

$$x_{ik} = \frac{\frac{1}{N} \sum_{j=0}^{N-1} S1_{i:N1+j} S2_{i:N2+j+k} - \mu1_i \mu2_{i,k}}{\sigma1_i \sigma2_{i,k}}, \quad (2)$$

$$\mu1_i = \frac{1}{N} \sum_{j=0}^{N-1} S1_{i:N1+j}, \quad (3)$$

$$\mu2_{i,k} = \frac{1}{N} \sum_{j=0}^{N-1} S2_{i:N2+j+k}, \quad (4)$$

$$\sigma1_i^2 = \frac{1}{N} \sum_{j=0}^{N-1} S1_{i:N1+j}^2 - \mu1_i^2, \quad (5)$$

$$\sigma2_{i,k}^2 = \frac{1}{N} \sum_{j=0}^{N-1} S2_{i:N2+j+k}^2 - \mu2_{i,k}^2. \quad (6)$$

При выборе параметра K , необходимо учесть, что с увеличением его значения, величина x_{ik} уменьшается, это связано с ослаблением корреляции отсчетов РК вдоль строки. Данное свойство указывает на нецелесообразность использования больших значений K при построении ККП (под большими K понимается значение $K > N$).

Очевидно, что если сформировать ККП одного и того же произнесения ($S1 = S2 = S$), то получится АКП команды S . Имеет смысл рассматривать ККП двух произнесений одной и той же команды. Он зависит от двух произнесений, поэтому вариативность произнесений влияет на вид портрета.

На рисунке 2 показаны ККП нескольких РК. На этом рисунке, например, «Маневр3 + Маневр4» означает, что этот ККП получен из третьего и четвертого произнесения команды «Маневр».

Отметим, что свойства ККП аналогичны свойствам АКП. Но ККП меньше зависят от особенностей произнесения, так как как учитывают в себе два различных произнесения.

Распознавание РК при использовании ККП производится аналогично распознаванию по АКП. Для каждой РК строится эталонный ККП из двух произнесений этой команды, эти эталонные ККП хранятся в памяти. Для распознаваемой команды строятся её ККП с одним из произнесений каждого класса команд и находится расстояние между этим ККП и эталонным ККП этого класса. Распознаваемая команда относится к классу с наименьшим расстоянием.

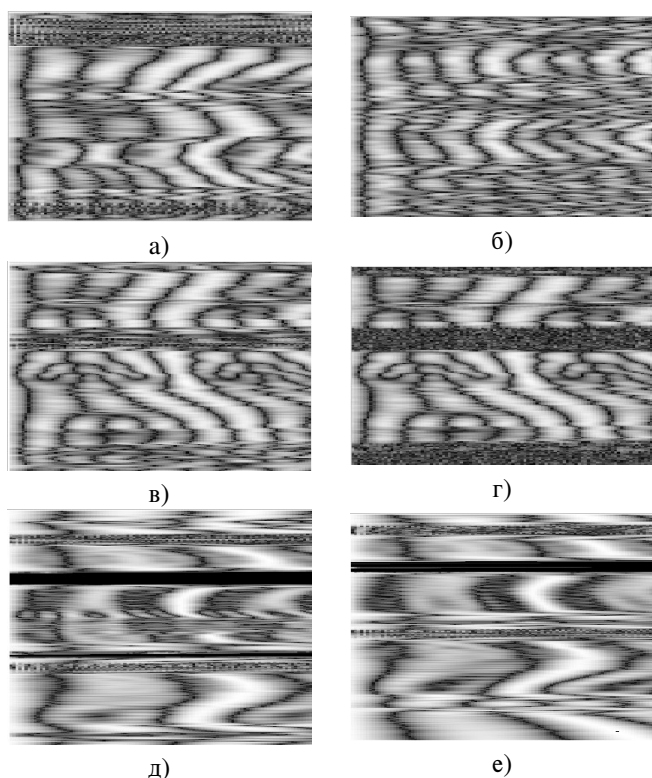


Рис. 1. Автокорреляционные портреты: а) РК «Свет больше», б) РК «Свет больше» на фоне шума авиационного двигателя, в) РК «Кондиционер», г) РК «Кондиционер» на фоне белого шума с нулевым средним и дисперсией равной пяти, д), е) – РК «Высота абсолютная», сказанной в разное время.

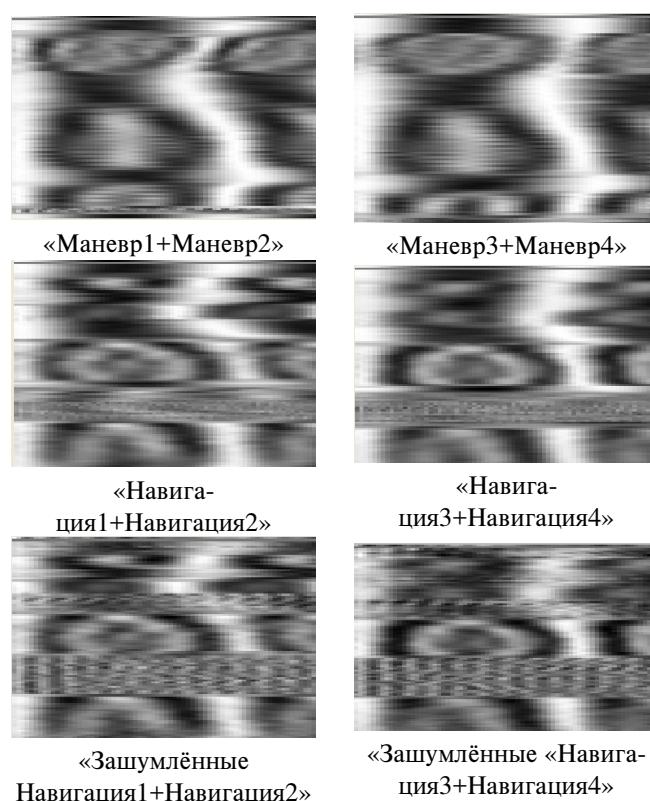


Рис. 2. Кросскорреляционные портреты команд.

5. Оптимизация распознавания команд по их ККП

Применяемые ККП имеют ряд особенностей, следующих как из свойств самих речевых сигналов, так и структуры построения ККП. Рассмотрим некоторые приёмы повышения эффективности распознавания по ККП.

5.1. Зашумление эталонов

Распознаваемые команды содержат сильный шум, что увеличивает расстояние от её ККП до её же эталонного портрета, сформированного на произнесениях без шума. Поэтому применялось «зашумление эталонов», то есть в эталонные произнесения искусственно добавлялся шум, поступающий от микрофона, расположенного вдали от оператора. В результате искажённые эталоны и распознаваемая команда содержали примерно одинаковый шум, что существенно повышало вероятность правильного распознавания.

5.2. Уточнение границ

При формировании портретов желательно, чтобы временные границы команд были определены как можно точнее, тогда достигается более точное совмещение портретов. Из нескольких известных способов обнаружения полезного сигнала на фоне шумов был отобран наилучший в рассматриваемых условиях. Кроме того, после определения границ команд этим способом выполнялись небольшие пробные подвижки границ, что несколько улучшило распознаваемость.

5.3. Удаление пауз

При произнесении некоторых РК имеются микропаузы, в частности, между словами. Эти паузы могут иметь разную продолжительность, но не содержат информации, поэтому был разработан способ их удаления.

5.4. Оптимизация ширины портретов

Ширина портрета, то есть длина его строк может выбираться произвольно, поэтому желательно выбрать оптимальную в смысле наилучшей распознаваемости. Оказалось, что следует длину этих строк в портрете распознаваемой команды брать примерно равной $K = D/(5M)$, где D – длина распознаваемой команды, M – количество строк портрета. Длины же строк эталонных ККП берутся «с запасом», чтобы они были не меньше значений K .

5.5. Выбор метрики

Распознавание РК по их ККП основано на нахождении наиболее близких друг к другу портретов. Отсюда вытекает задача определения расстояния между двумя ККП, то есть метрики, определённой на ККП. Это расстояние принимается равным среднему расстоянию между соответствующими строками ККП, но при этом могут быть использованы любые метрики, определённые на строках, то есть на конечных последовательностях или векторах. Было опробовано 12 известных метрик (Евклидова, Гильбертова, Журавлёва и т.д.) и их вариантов. Оказалось, что наилучшие результаты давали в применении к рассматриваемой задаче пять метрик: мера Журавлева (при $\varepsilon = 10$, $\varepsilon = 20$ и $\varepsilon = 30$), расстояние Ружечки и расстояние Брея-Кертиса. Кроме того, при анализе результатов распознавания с использованием данных метрик, было выявлено, что ошибки распознавания часто происходят на разных командах при использовании различных метрик. Поэтому можно улучшить распознавание, если использовать, например, две метрики. Если результаты распознавания по ним совпадают, то команду считать распознанной, если же не совпадают, то команду считать «не распознанной». В этом случае диктор должен будет повторить команду еще раз.

5.6. Оптимизация библиотеки эталонов

Как и при распознавании РК по АКП, состав библиотеки эталонов существенно влияет на качество распознавания. Поэтому и при распознавании по ККП производилась оптимизация библиотеки эталонных портретов.

5.7. Фурье-анализ

Каждая строка ККП представляет собой последовательность выборочных значений корреляционной функции. В силу квазипериодичности речевых сигналов корреляционная функция оказывается близкой к периодической, что было использовано для повышения качества портретов путём удаления из спектра каждой строки ККП незначимых гармоник. Эта операция производилась с помощью БПФ. Выделение основных гармоник каждой строки ККП уменьшало влияние вариативности различных произнесений речевого сигнала.

6. Результаты экспериментов

Эксперименты, показали, что применение ККП с описанными выше модификациями в значительной мере ослабили влияние вариативности произнесения команд и эффект старения эталонов, приблизив вероятность правильного распознавания к её значению при распознавании по АКП на «свежих» эталонах. В частности, для оценки эффективности предложенного метода был проведён эксперимент. В распознавании участвовало 2 набора РК, каждый из которых состоял из 10 команд. Первый набор включает в себя однословные команды, второй набор РК включает как однословные, так и команды состоящие из двух слов. Каждая команда произнесена по 100 раз женщиной-диктором. Результаты эксперимента указаны в таблице 1. Максимальная вероятность правильного распознавания 95,6 % команд.

Таблица 1. Результаты распознавания РК по ККП.

Команды	Отношение сигнал/шум (дБ)				
	5	4	3	2	1
Набор 1	95,6	90,1	87,4	82,9	61,2
Набор 2	94,6	92,1	87,4	83,5	67,2

7. Заключение

В настоящей работе предложен и исследован дикторозависимый метод распознавания речевых команд из ограниченного словаря в условиях интенсивных акустических помех, например, на фоне шума двигателя самолёта. Этот метод заключается в преобразовании оцифрованных команд в особые изображения с дальнейшим применением методов обработки изображений. Произведены различные модернизации этого метода для повышения вероятности правильного распознавания. Эксперименты на большом количестве речевого материала показали достаточно высокую эффективность предложенного метода.

Список литературы

- [1] Businessinsider. How apples vocaliq ai works [Electronic resource]. — 2016. — URL: <http://uk.businessinsider.com/how-apples-vocaliq-ai-works-2016-5>.
- [2] Рабинер, Л. Цифровая обработка речевых сигналов: Пер. с англ. под ред. М.В. Назарова, Ю.Н. Прохорова / Л.П. Рабинер, Р.В. Шафер. — Москва, Россия. [S. 1.] : Наука, 1981. — Р. 495.
- [3] Бойков, Ф. Применение вейвлет-анализа сигнала в системе распознавания речи / Ф.Г. Бойков, Старожилова Т. К. // Труды международной конференции «Диалог'2003». — Звенигород, Россия. — [Б. м. : б. и.], 2003. — С. 12–19.
- [4] Гудонавичюс, Р. Распознавание речевых сигналов по их структурным свойствам / Р.В. Гудонавичюс, П.П. Кемешис, А.Б. Читавичюс. — Ленинград, СССР. [S. 1.] : Энергия, 1977. — Р. 64.
- [5] Мясникова, Е. Объективное распознавание звуков речи / Е.Н. Мясникова. — Ленинград, СССР. [S. 1.] : Энергия, 1967. — Р. 148.
- [6] Пиконе, Д. Методы моделирования сигнала в распознавании речи / Д. Пиконе. — Кемерово, Россия. [S. 1. : s. n.], 2000. — Р. 79.
- [7] Потапова, Р. Речь: коммуникация, информация, кибернетика / Р.К. Потапова. — Москва, Россия. [S. 1.] : Радио и связь, 1997. — Р. 568.
- [8] Сорокин, В. Скрытые марковские модели в распознавании речи / В.Н. Сорокин, В.А. Суханов // Речевая информатика. Сборник трудов под ред. В.В. Зяблова. — Москва, Россия. — [Б. м. : б. и.], 1989. — С. 104–118.

- [9] Discriminative codebook design using multiple vector quantization in hmm-based speech recognizers / A. Peinado, J. Segura, A. Rubio [et al.] // *IEEE Trans. Speech and Audio Processing*. — 1996. — Vol. IV, No. 2. — P. 89–94.
- [10] Jelinek, F. *Statistical Methods for Speech Recognition* / F. Jelinek. — Cambridge. [S. l.] : MIT Press, 1998.
- [11] Shahshahani, B. A markov random field approach to bayesian speaker adaptation / B. Shahshahani // *IEEE Trans. Speech and Audio Processing*. — 1997. — Vol. V, No. 2. — P. 183–191.
- [12] Федяев, О. Нейросетевой интерпретатор речевых команд для управления программными системами [Текст] / О.И. Федяев, С.А. Гладунов // *Труды 7-й всероссийской конференции «Нейрокомпьютеры и их применение» под редакцией А.И. Галушкина*. — Москва, Россия. — [Б. м. : б. и.], 2001. — С. 298–301.
- [13] Lippmann, R. Neural classifiers useful for speech recognition / R. Lippmann, B. Gold // in: *Proc. IEEE First Int. Conf. Neural Net.* — Vol. IV. — [S. l. : s. n.], 1987. — P. 417–422.
- [14] Распознавание речевых команд на фоне интенсивных шумов с помощью автокорреляционных портретов [speech command recognition on the background of noise using autocorrelation portraits] / V.R. Krasheninnikov, A.I. Armer, N.A. Krasheninnikova, A.V. Khvostov // *Naukoemkie tekhnologii*. — 2007. — Т. 9. — С. 65–76.
- [15] Cross-correlation portraits of voice signals in the problem of recognizing voice commands according to patterns / V.R. Krasheninnikov, A.I. Armer, V.V. Kuznetsov, E.Yu Lebedeva // *Pattern Recognition and Image Analysis*. — 2011. — Vol. 21, No. 2. — P. 192–194.
- [16] Крашенинников, В. Вариация границ речевых команд для улучшения распознавания речевых команд по их кросскорреляционным портретам / В.П. Крашенинников, Е.Ю. Лебедева, В.К. Капырин // *Известия Самарского научного центра РАН*. — 2013. — Vol. 4(4). — P. 928–930.
- [17] Крашенинников, В. Повышение вероятности правильного распознавания сигналов по их кросскорреляционным портретам / В.П. Крашенинников, Н.А. Крашенинникова, Е.Ю. Галицкая // *Радиотехника*. — 2014. — Vol. 7. — P. 107–110.
- [18] Васильев, К. Статистический анализ изображений / К.К. Васильев, В.П. Крашенинников. — Ульяновск, Россия. [S. l.] : УлГТУ, 2014. — P. 216.
- [19] Армер, А. Использование онтологии для формирования наборов эталонов речевых команд в задаче распознавания речевых команд на фоне шумов / А.И. Армер, В.С. Мошкин // *Радиотехника*. — 2016. — Vol. 9. — P. 72–77.
- [20] Армер, А. Подход к формированию наборов эталонов речевых команд с использованием онтологии / А.И. Армер, В.С. Мошкин // *Онтология проектирования*. — 2016. — Vol. 6. — P. 270–277.
- [21] Optimization of dictionary and model library for recognition of speech commands / V.R. Krasheninnikov, N.A. Krasheninnikova, V.V. Kuznetsov, E.Yu Lebedeva // *Pattern Recognition and Image Analysis*. — 2011. — Vol. 21, No. 3. — P. 505–507.
- [22] Optimization of dictionary and model library for recognition of speech commands based on cross-correlation portraits / V.R. Krasheninnikov, N.A. Krasheninnikova, V.V. Kuznetsov, E.Yu Lebedeva // *Pattern Recognition and Image Analysis*. — 2013. — Vol. 23, No. 1. — P. 80–86.