

РАСПОЗНАВАНИЕ ДИКТОРОВ ПО ЧАСТОТАМ ПАРАМЕТРИЧЕСКИХ СПЕКТРОВ В МОДЕЛИ РЕЧИ В ВИДЕ СОСТАВНЫХ ВЕКТОРНЫХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Тихонов В.А., Безрук В.М., Хмеларжова (Кудрявцева) Н.В., Хмеларж П.

Введение

Сложность решения задачи распознавания речи и идентификации говорящего, связана с тем, что речь относится к одному из наиболее сложных видов сигналов. В речевом сигнале существенна интеллектуальная составляющая, позволяющая передавать не только огромное количество разнообразной информации, выработанной человеком, но и делать это в художественной форме, эмоционально, применять интонацию. Человек в своем историческом и индивидуальном развитии долго обучался искусству общения и передачи сообщений посредством речевых сигналов. Это повлияло на сложность и разнообразие процесса речеобразования, тонкость слухового анализа речи, большой объем эталонов речи, хранящихся в памяти. Понимание распознаваемой речи также способствует ее точному распознаванию.

Для генерации речевых сигналов человек использует тонкую манипуляцию речеобразующим трактом. Процесс формирования сигнала речеобразующим трактом описывается системой линейных и нелинейных стохастических уравнений. Сформированный коррелированный сигнал обладает как гауссовыми, так и негауссовыми свойствами. Человеческая речь имеет как стационарные, так и нестационарные участки. Фонемы, являющиеся элементами речи, характеризуются многообразием спектральных форм, отличающихся между собой. Нестационарность спектральных характеристик фонем, наличие переходных участков от одной фонемы к другой приводят к перепутыванию фонем при их распознавании.

Спектральные характеристики, являющиеся основой признаков элементов речи, неустойчивы, что приводит к сложностям при определении признаков речевых сигналов. Применяемые методы спектрального оценивания часто дают заметные погрешности в определении спектров анализируемых сигналов. На статистические характеристики речевых сигналов накладываются национальные особенности произнесения звуков, на них влияет возраст говорящего, состояние здоровья речеобразующего тракта, особенности его структуры. Поэтому задача распознавания речи и тесно связанная с ней проблема понимания речи, несмотря на многочисленные исследования, еще не получила удовлетворительного решения. Преимущества вычислительной техники в обработке сигналов заключаются в быстродействии вычислений и их точности, объеме памяти и эффективности ее использования. Компьютер уже достиг превосходства над человеком в игре в шахматы и в игре го. Однако человек все еще превосходит автоматические системы в точности распознавания образов, в частности, речевых сигналов.

Задачи распознавания речи характеризуются разнообразием целей. Они включают: распознавание слитной речи, распознавание слов, распознавание

дикторов, распознавание команд и др. На качество распознавания оказывают существенное влияние условия постановки задачи распознавания, объем словаря слов. При ограниченном словаре слов удалось достичь хороших результатов и поднять эти системы распознавания до уровня их коммерческого использования. Существенно повысить качество распознавания можно при работе в идеальных условиях, в изолированных помещениях с хорошей акустикой.

Достигнутые результаты экспериментов по распознаванию иногда вносят рекламный оттенок, направленный на привлечение финансирования таких исследований для достижения их коммерческого использования. На необходимость дальнейших исследований в области распознавания речи, указывает полученное качество распознавания речи, его близость к точности восприятия речи человеком в естественных условиях окружающей акустической среды общения.

В работе рассматривается задача распознавания дикторов по их речевым сигналам [1-6]. Она подразделяется на две подзадачи: верификация и идентификация диктора. Верификация — это процесс распознавания диктора путем сравнения произнесенного образца речи с хранящимся в библиотеке эталоном [7-11]. При этом производится подтверждение личности говорящего или его опровержение по его идентификатору. Это напоминает речевую подпись говорящего.

Процесс идентификации заключается в определении диктора по голосу среди нескольких возможных [12-15]. Для этого признаки речи говорящего сравниваются с набором эталонных признаков, полученных на этапе обучения. Т.к. говорящий стремится к своей идентификации, то при произнесении идентификатора, он должен избегать ряда факторов, которые способствуют возникновению ошибок принятия решений. К этим факторам можно отнести: неправильное или необычное произнесение идентификатора, исключение влияния на речь эмоционального состояния диктора, отсутствие неблагоприятного акустического фона в виде шума и помех, использование идентичных каналов связи.

Современные системы идентификации диктора в основном используют кратковременные спектральные особенности речевого сигнала говорящего [16, 17]. Мел-частотные кепстральные коэффициенты, перцептивное линейное предсказание и линейное предсказание - широко используемые методы извлечения признаков из-за их значительной производительности и более низкой вычислительной сложности. Во многих работах рассматриваются способы распознавания говорящего по коротким ключевым словам [18-20].

К системам классификации дикторов предъявляются противоречивые требования. С одной стороны, ими производится анализ одинаковых слов-идентификаторов, что обуславливает схожесть их статистических характеристик. С другой стороны система распознавания должна выделять различия в произнесении одинаковых слов разными дикторами. Эталонные признаки определяются на этапе обучения и хранятся в памяти системы распознавания. Заметим, что в системах распознавания слитной речи,

желательно использовать признаки инвариантные для разных дикторов, а словарь слов неограничен.

Современные методы анализа и обработки сигналов и процессов оперируют с понятиями отсчетов или величинами функции в точках, определяемых значениями аргументов. Информационные параметры или статистические характеристики системы характеризуют динамику изменения процесса. Многие стационарные и нестационарные случайные процессы характеризуются не только краткосрочными изменениями (колебаниями или флуктуациями), но и среднесрочными и долгосрочными изменениями. Современные методы статистического анализа не позволяют эффективно исследовать характеристики таких разномасштабных изменений.

Предложенная нами модель представления процессов в виде совокупности процессов меньшей длины применялась для решения задачи идентификации дикторов. Речевой сигнал представлялся в виде последовательности непересекающихся подвекторов. Для такого представления рассчитывалась корреляционная функция и строилась мультипликативная модель авторегрессии идентификационного слова. Сложный параметрический спектр слова, путем предложенной факторизации, представлялся в виде совокупности одномодовых спектров. Признаками идентификатора, произнесенного диктором, служили частоты пиков одномодовых составляющих параметрического спектра.

Проведенный спектральный анализ идентификаторов тридцати пяти дикторов, на основе предложенной модели, показал отличие спектров для различных дикторов. В тоже время наблюдалось сходство спектров и частот пиков идентификаторов, произнесенных одним диктором. Экспериментальные исследования по идентификации дикторов на основе предложенного метода, показали его эффективность.

Авторегрессионная модель линейного предсказания речевых сигналов

Рассмотрим теоретические основы построения модели, лежащей в основе анализа и распознавания речевых сигналов при идентификации дикторов. Дискретные случайные процессы с узкополосным спектром можно описывать с помощью моделей авторегрессии (АР). Корреляция отсчета случайного процесса в текущий момент времени с некоторым конечным или бесконечным числом отсчетов в предыдущие моменты времени положена в основу модели АР. Корреляционные связи позволяют осуществить регрессию текущего отсчета на предшествующие отсчеты. Такая регрессия называется авторегрессией. Для модели АР текущий отсчет $x[t]$ стационарного случайного процесса с нулевым средним выражается через предыдущие $x[t - j]$ с некоторыми коэффициентами веса [21]

$$x[t] = \sum_{i=1}^p \Phi[i]x[t - i] + a[t], \quad (1)$$

где $\Phi[j]$ – коэффициенты АР, p – порядок модели АР, $a[t]$ – некоррелированные случайные отсчеты. Коэффициенты АР и дисперсия ошибки предсказания определяются с помощью системы линейных уравнений

$$R[i] - \sum_{j=1}^p \Phi[j]R[j-i] = 0, \quad i = 1 \div p, \quad (2a)$$

$$R[0] - \sum_{j=1}^p \Phi[j]R[j] = D_a, \quad (2б)$$

где $R[j] = E\{x[t]x[t-j]\}$ – значения функции корреляции случайного процесса, $D_a = \sigma_a^2$ – дисперсия ошибок предсказания модели АР, $R[0]$ – дисперсия случайного процесса $x[t]$. Набор уравнений (1.2а) и (1.2б) – это полная система уравнений Юла-Уокера. Они могут быть выражены в матричном виде с использованием корреляционной матрицы. Коэффициенты АР и дисперсия ошибок предсказания являются решением этой системы уравнений.

В [22] были предложены выражения, связывающие коэффициенты АР и параметры спектральной плотности мощности (СПМ). Для этого полагают, что разностное уравнение АР описывает систему с несколькими собственными частотами и соответствующими декрементами затухания. Решения разностного уравнения выражаются через корни характеристического уравнения, которое в данном случае совпадает с уравнением (1). Представим характеристическое уравнение для разностного уравнения (1.1) в виде

$$\prod_{i=1}^p (c - c[i]) = 0. \quad (3)$$

Из (3) и (1) следует, что характеристическое уравнение можно записать следующим образом

$$c^p - \Phi[1]c^{p-1} - \dots - \Phi[p] = \prod_{i=1}^p (c - c[i]) = 0. \quad (4)$$

Сравнивая в (4) коэффициенты при одинаковых степенях c , получают связь между коэффициентами АР и корнями $c[i]$. Приведем соответствующие формулы для $p = 1, 2$:

$$\begin{aligned} \Phi[1,1] &= c[1],, \\ \Phi[2,1] &= c[2,1] + c[2,2];, \\ \Phi[2,2] &= -c[2,1]c[2,2],, \end{aligned} \quad (5)$$

В (5) первый индекс в квадратных скобках указывает на соответствующий порядок модели, а второй индекс указывает на номер параметра модели или корня. Полученные формулы оказываются весьма полезными для определения коэффициентов АР по заданным характеристикам пиков СПМ случайного процесса.

Отметим, что корни характеристического уравнения полностью описывают модель АР. Свойства модели зависят параметров, через которые они выражаются. Если корни действительные, то их можно представить в виде экспоненциальной функции [22]

$$c[i] = e^{-h_i T},$$

где h_i – коэффициент демпфирования равный $h_i = \pi \Delta f_i$, где Δf_i – ширина полосы i -го пика СПМ. Тогда действительные корни характеристического уравнения зависят от ширины полосы выражением вида

$$c[i] = e^{-\pi \Delta f_i T}. \quad (6)$$

Комплексные корни характеристического уравнения описываются выражениями

$$c[i] = e^{-\pi \Delta f_i T - j \omega_{hi} T}, \quad c[i+1] = e^{-\pi \Delta f_i T + j \omega_{hi} T}, \quad (7)$$

где ω_{hi} – собственная частота модели АР с поправкой на демпфирование, соответствующая i -тому пику СПМ. В зависимости от порядка модели, подставляя (6) или (7) в формулы (5) определяют коэффициенты АР. Показано, что рассчитанные таким образом коэффициенты АР, можно использовать для генерации процессов АР с заданными характеристиками СПМ, а также для получения параметров спектров полипроцессов линейного предсказания.

Формула для нахождения спектра модели АР лежит в основе параметрического спектрального оценивания. Для ее вывода будем рассматривать процесс АР как реакцию фильтра $x[t]$, на вход которого подаются некоррелированные отсчеты $a[t]$. Используя z -преобразование, можно показать, что z -преобразование передаточной функции АР фильтра имеет вид

$$H(z) = 1 / \Phi(z),$$

где

$$\Phi(z) = -\sum_{i=1}^p \Phi[i] z^{-i}, \quad \Phi[0] = -1.$$

z -преобразования СПМ выходного и входного процессов связаны соотношением

$$P(z) = |H(z)|^2 P_a[z] = P_a[z] / |\Phi(z)|^2. \quad (8)$$

Чтобы найти СПМ выходного АР процесса необходимо в (8) сделать замену $z = e^{j2\pi fT}$ и положить, что для белого шума $P_a[z] = D_a$ постоянная величина.

Тогда из (1.8) следует

$$P(f) = D_a / \left| 1 - \sum_{i=1}^p \Phi[i] e^{-j2\pi fT} \right|^2. \quad (9)$$

Выражение (9) широко используется в параметрическом методе спектрального оценивания случайных процессов [23]. В качестве параметров, полностью характеризующих спектральную оценку случайного процесса, выступают коэффициенты АР и порядок модели. Параметрическое

спектральное оценивание обладает рядом достоинств по сравнению с традиционными методами спектрального оценивания. К ним относятся: более высокое спектральное разрешение при использовании коротких выборок, отсутствие боковых лепестков.

Мультипликативные модели авторегрессии речевых сигналов

Для произвольного числа моделей АР, составляющих мультипликативную модель линейного предсказания $AP_1 \times AP_2 \times \dots \times AP_k$, выражение (1) представляется произведением k системных функций АР [24]

$$x[t] = \Phi_k^{-1}(z)\Phi_{k-1}^{-1}(z)\dots\Phi_1^{-1}(z)a[t]. \quad (10)$$

Представив это уравнение в виде

$$\Phi_k(z)\Phi_{k-1}(z)\dots\Phi_1(z)x[t] = a[t], \quad (11)$$

получим систему уравнений для расчета коэффициентов АР. Умножим левую и правую части (11) на $x[t-j]$ и усредним. Учитывая условие оптимальности, получим

$$\Phi_k(z)\Phi_{k-1}(z)\dots\Phi_1(z)R[j] = 0. \quad (12)$$

Эффективный способ решения этого уравнения продемонстрирован на примерах приведенных ниже.

Остановимся на примерах построения мультипликативных АР моделей. Рассмотрим простейший мультипликативный процесс $AP_1 \times AP_2$, состоящий из присоединенных моделей АР. Такие модели полезны, когда выходной процесс получен в результате последовательного преобразования белого шума двумя узкополосными АР формирующими фильтрами.

Для этой модели уравнение (11) выражается через системные функции следующим образом

$$\Phi_2(z)\Phi_1(z)x[t] = a[t]. \quad (13)$$

Из (13) следует разностное уравнение модели $AP_1 \times AP_2$

$$x[t] = \sum_{i=1}^{p_2} \Phi_2[i]x[t-i] + \sum_{n=1}^{p_1} \Phi_1[n]x[t-n] - \sum_{i=1}^{p_2} \sum_{n=1}^{p_1} \Phi_2[i]\Phi_1[n]x[t-i-n] + a[t] \quad (14)$$

Уравнение (14) нелинейное по коэффициентам АР. Условие оптимальности мультипликативной модели $AP_1 \times AP_2$ состоит в статистической независимости $a[t]$. В случае модели $AP_1 \times AP_2$ второго ранга ошибки $a[t]$ некоррелированные

$$E\{a[t]a[t-j]\} = 0, \quad j > 0. \quad (15)$$

Покажем, что условие оптимальности позволяет получить систему уравнений для расчета коэффициентов АР мультипликативной модели $AP_1 \times AP_2$. Умножив правую и левую части уравнения (13) на $x[t-j]$ и усреднив, получим

$$\Phi_2(z)\Phi_1(z)R[j] = 0. \quad (16)$$

При выводе (16) было использовано свойство оптимальности модели (15). Уравнение (16) удобно преобразовать к виду

$$(1 - \sum_{i=1}^{p_2} \Phi_2[i]z^{-i})(1 - \sum_{n=1}^{p_1} \Phi_1[n]z^{-n})R[j] = 0. \quad (17)$$

Поддействовав операторами сдвига на $R[j]$ получим уравнение для расчета параметров модели $AP_{1 \times AP_2}$

$$R[j] = \sum_{i=1}^{p_2} \Phi_2[i]R[j-i] + \sum_{n=1}^{p_1} \Phi_1[n]R[j-n] - \sum_{i=1}^{p_2} \sum_{n=1}^{p_1} \Phi_2[i]\Phi_1[n]R[j-i-n]. \quad (18)$$

Неизвестные коэффициенты AP входят в уравнение (18) нелинейно. Поэтому для его решения необходимо решать систему нелинейных уравнений. Выражение (18) является более общим случаем уравнения Юла-Уокера. Это выражение сводится к уравнению Юла-Уокера при $p_2 = 0$.

Параметрическую спектральную оценку для модели $AP_{1 \times AP_2}$ получают аналогично (9) [24]

$$P(f) = \frac{D_a}{\left| 1 - \sum_{i=1}^{p_2} \Phi_2[i]e^{-j2\pi ftT} \right|^2 \left| 1 - \sum_{n=1}^{p_1} \Phi_1[n]e^{-j2\pi fnT} \right|^2}. \quad (19)$$

Рассмотрим пример построения мультипликативной модели $AP_1(1) \times AP_2(1)$. Из (16) после преобразований получим рекуррентное уравнение

$$R[j] = (\Phi_1[1] + \Phi_2[1])R[j-1] - (\Phi_1[1]\Phi_2[1])R[j-2], \quad (20)$$

или

$$R[j] = \Phi[1]R[j-1] + \Phi[2]R[j-2]. \quad (21)$$

В (21) введено обозначение

$$\Phi[1] = \Phi_1[1] + \Phi_2[1],$$

$$\Phi[2] = -\Phi_1[1]\Phi_2[1].$$

При $j = 1, 2$, из (20) имеем систему нелинейных уравнений для расчета параметров $\Phi_1[1]$ и $\Phi_2[1]$

$$\begin{aligned} R[1] &= (\Phi_1[1] + \Phi_2[1])R[0] - \Phi_1[1]\Phi_2[1]R[1], \\ R[2] &= (\Phi_1[1] + \Phi_2[1])R[1] - \Phi_1[1]\Phi_2[1]R[0]. \end{aligned} \quad (22)$$

Заметим, что уравнение (20) удобно записать в матричном виде

$$\vec{R} = R\vec{\Phi}, \quad (23)$$

где

$$\begin{aligned} \vec{R} &= \begin{bmatrix} R[1] \\ R[2] \end{bmatrix}, & R &= \begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix}, \\ \vec{\Phi} &= \begin{bmatrix} \Phi_1[1] + \Phi_2[1] \\ -\Phi_1[1]\Phi_2[1] \end{bmatrix} = \begin{bmatrix} \Phi[1] \\ \Phi[2] \end{bmatrix}. \end{aligned} \quad (24)$$

Решение уравнения (23) получают в два этапа. Вначале решают матричное уравнение (23)

$$\vec{\Phi} = R^{-1}\vec{R}, \quad (25)$$

и находят $\Phi[1]$ и $\Phi[2]$. Затем решают систему нелинейных уравнений, которую получают из (24)

$$\begin{aligned} \Phi[1] &= \Phi_1[1] + \Phi_2[1], \\ \Phi[2] &= -\Phi_1[1]\Phi_2[1], \end{aligned}$$

и находят $\Phi_1[1]$ и $\Phi_2[1]$. Заметим, что при $\Phi_2[1] = 0$, из (22) имеем

$$R[1] = \Phi_1[1]R[0], \quad R[2] = \Phi_1[1]R[1],$$

или в общем виде $R[j] = \Phi_1[1]R[j-1]$. Это известное уравнение Юла-Уокера для первого порядка модели АР. Т.о. классическая модель АР является частным случаем предложенной нами мультипликативной модели $AP_1 \times AP_2$.

Рассмотрим модель $AP_1(1) \times AP_2(2)$ [24]. Из (1.18) имеем

$$\begin{aligned} R[k] &= (\Phi_2[1] + \Phi_1[1])R[k-1] + (\Phi_2[2] - \Phi_2[1]\Phi_1[1])R[k-2] - \\ &- \Phi_2[2]\Phi_1[1]R[k-3], \end{aligned} \quad (26)$$

которое можно записать в виде

$$R[k] = \Phi[1]R[k-1] + \Phi[2]R[k-2] - \Phi[3]R[k-3]. \quad (27)$$

Это рекуррентное выражение позволяет сначала найти коэффициенты при $R[k-1]$, $R[k-2]$ и $R[k-3]$, а затем, решив систему нелинейных уравнений, вычислить $\Phi_1[1]$, $\Phi_2[1]$ и $\Phi_2[2]$. При $\Phi_1[1] = 0$, получаем уравнение Юла-Уокера для модели АР второго порядка.

Уравнение (1.26) можно записать в матричном виде (22), в котором

$$\begin{aligned} \vec{R} &= \begin{bmatrix} R[1] \\ R[2] \\ R[3] \end{bmatrix}, \quad R = \begin{bmatrix} R[0] & R[1] & R[2] \\ R[1] & R[0] & R[1] \\ R[2] & R[1] & R[0] \end{bmatrix}, \\ \vec{\Phi} &= \begin{bmatrix} \Phi_2[1] + \Phi_1[1] \\ \Phi_2[2] - \Phi_2[1]\Phi_1[1] \\ -\Phi_2[2]\Phi_1[1] \end{bmatrix} = \begin{bmatrix} \Phi[1] \\ \Phi[2] \\ \Phi[3] \end{bmatrix}. \end{aligned} \quad (28)$$

Тогда решение можно представить в виде (25). Вычислив коэффициенты $\Phi[1]$, $\Phi[2]$, $\Phi[3]$, рассчитывают коэффициенты АР составляющих моделей, используя матричное уравнение (28).

Модель авторегрессии речевых сигналов в представлении составных векторных случайных процессов

Дискретный речевой сигнал будем рассматривать в виде векторного случайного процесса. Случайный процесс можно представить в виде последовательности подвекторов \vec{x}_i , одинаковой длины n с однородными статистическими свойствами [25]. Здесь введено понятие «подвектора» \vec{x}_i вектора процесса $\vec{x}[t]$. Назовем такой стационарный случайный процесс «составным векторным случайным процессом» (СВСП). Формула для оценки корреляционной функции сигнала в СВСП представлении имеет вид

$$R^n[k] = \frac{1}{m-k} \sum_{i=0}^{m-k} \sum_{j=0}^{n-1} (x[in+j]x[in+j+kn]).$$

Это выражение можно упростить

$$R^n[k] = \frac{1}{N-kn} \sum_{i=0}^{N-kn} (x[i]x[i+kn]).$$

При $n = 1$ из этого выражения может быть получена известная формула оценки корреляционной функции для стационарного случайного процесса.

Для процесса в СВСП представлении также можно построить модели линейного предсказания. В таких моделях предсказываются не коррелированные отсчеты, а коррелированные подвекторы. Уравнение для модели АР случайного процесса в представлении СВСП имеет вид

$$\vec{x}_i = \sum_{s=1}^p \Phi^n[s] \vec{x}_{i-s} + \vec{a}_i, \quad (29)$$

где $\Phi^n[s]$ – коэффициенты АР СВСП, p – порядок модели АР СВСП, \vec{a}_i – векторы длиной n отсчетов белого шума. Условие оптимальности модели АР СВСП состоит в статистической независимости подвекторов \vec{a}_i . Для модели АР СВСП ошибки \vec{a}_i должны быть некоррелированными, т.е. $E\{\vec{a}_i \vec{a}_{i-k}\} = 0$, при $k \neq 0$. Это условие эквивалентно минимуму дисперсии ошибки предсказания СВСП D_a^n .

АР СВСП состоит из коррелированных подвекторов, сформированных фильтром, но с разными подвекторами порождающего процесса типа белого шума. Естественно, что процессы АР и АР СВСП не совпадают. При $n = 1$ модель АР СВСП вырождается в обычную модель АР случайного процесса.

Для нахождения коэффициентов АР векторов процесса умножим (1.29) на \vec{x}_{i-j} и усредним. После несложных преобразований получим уравнения типа Юла – Уокера для расчета параметров модели АР СВСП

$$R^n[j] = \sum_{i=1}^p \Phi^n[i] R^n[i-j], \quad j = 1, 2, \dots, p.$$

При $j = 0$ получим выражение, связывающее дисперсию подвекторов процесса и векторов процесса и векторов ошибки предсказания

$$R^n[0] = \sum_{i=1}^p \Phi^n[i] R^n[i] + D_a^n.$$

Очевидно, что все свойства модели АР случайных процессов справедливы и для модели АР СВСП. Так, выражение для параметрической оценки СПМ имеет вид

$$P^n(f) = \frac{D_a^n}{\left| 1 - \sum_{i=1}^p \Phi^n[i] e^{-j2\pi f k T} \right|^2}. \quad (30)$$

Факторизация спектров случайных сигналов

На основе приведенной выше теории мультипликативной модели в данном подразделе предложен алгоритм факторизации СПМ случайных процессов [26, 27]. Разложение многомодовой СПМ на составляющие с одной или несколькими модами позволяет проанализировать сложный процесс как совокупность простых. Из многомодового спектра можно выделить представляющие наибольший интерес составляющие и использовать их для решения задач распознавания. Способ факторизации продемонстрирован на нескольких примерах спектрального анализа реальных случайных процессов.

Применение мультипликативной модели позволяет разложить параметрический двухмодовый спектр АР на составляющие. Спектры речевых сигналов могут содержать большее число пиков. Поэтому для их описания используются модели АР 8-го порядка. В зависимости от структуры таких спектров и целей исследования, их можно разложить на 4-8 составляющих моделей 1-го или 2-го порядков.

Рассмотрим, в качестве примера, модель АР речевого сигнала 8-го порядка. Она представляется мультипликативной моделью $AP_1 \times AP_2 \times AP_3 \times AP_4$. Уравнение для расчета параметров четырех моделей АР имеет вид

$$\begin{aligned} & \left(1 - \sum_{v=1}^{p_4} \Phi_4[v] z^{-v} \right) (-R[j] + \sum_{i=1}^{p_1} \Phi_1[i] R[j-i] + \sum_{n=1}^{p_2} \Phi_2[n] R[j-n] + \\ & \sum_{u=1}^{p_3} \Phi_3[u] R[j-u] - \sum_{i=1}^{p_1} \sum_{n=1}^{p_2} \Phi_1[i] \Phi_2[n] R[j-i-n] - \\ & - \sum_{i=1}^{p_1} \sum_{u=1}^{p_3} \Phi_1[i] \Phi_3[u] R[j-i-u] - \sum_{n=1}^{p_2} \sum_{u=1}^{p_3} \Phi_2[n] \Phi_3[u] R[j-n-u] + \\ & \sum_{i=1}^{p_1} \sum_{n=1}^{p_2} \sum_{u=1}^{p_3} \Phi_1[i] \Phi_2[n] \Phi_3[u] R[j-i-n-u]) = 0. \end{aligned} \quad (31)$$

Для решения этого уравнения необходимо задать порядки составляющих моделей. Их сумма должна равняться 8. Затем необходимо привести (31) к разностному уравнению. Приведя подобные члены для одинаковых значений $R[\bullet]$, получим нелинейную систему восьми уравнений с восемью неизвестными коэффициентами модели $AP_1 \times AP_2 \times AP_3 \times AP_4$. Решение несколько упрощается, если сначала найти коэффициенты модели АР(8), используя уравнение Юла-Уокера. Затем необходимо решить систему из восьми нелинейных уравнений. Этот способ нахождения параметров модели $AP_1 \times AP_2 \times AP_3 \times AP_4$ требует большого объема вычислений и анализа

десятков и даже сотен полученных корней системы уравнений. Поэтому был предложен способ факторизации системной функции модели АР

$$\frac{1}{1 - \Phi[1]z^{-1} - \Phi[2]z^{-2} - \dots - \Phi[p]z^{-p}} = \frac{1}{1 - c[1]z^{-1}} \frac{1}{1 - c[2]z^{-1}} \dots \frac{1}{1 - c[p]z^{-1}}.$$

Системная функция определяется характеристическим уравнением АР. Это уравнение удобно представить в виде

$$c^p - \Phi_1 c^{p-1} - \Phi_2 c^{p-2} - \dots - \Phi_p = (c - c[1])(c - c[2]) \dots (c - c[p]). \quad (32)$$

В (32) неизвестные $c[i]$ являются корнями характеристического уравнения (4), которое легко решается в общем случае численными методами. По найденным корням, используя (2а, б), легко найти коэффициенты модели АР первого или второго порядка. В этом случае корни уравнения либо действительные, либо комплексно-сопряженные. Найденные коэффициенты модели используются в (9), для расчета параметрического спектра. При первом и втором порядках модели получается факторизация спектра на одномодовые составляющие. При третьем и четвертом порядках получаем факторизацию на двухмодовые составляющие. Для факторизации применяется также комбинация двух этих случаев. Таким образом, использование корней характеристического уравнения, существенно упрощает процедуру факторизации многомодовых параметрических спектров. Предложенный метод факторизации спектров применялся для анализа спектров речевых сигналов при получении признаков идентификации дикторов.

Ниже рассмотрен ряд примеров разложения многомодовых спектров реальных процессов на составляющие. При получении параметров составляющих моделей используется их однозначная связь с параметрами классических моделей линейного предсказания. Способ расчета коэффициентов АР моделей составляющих по АР модели процесса, описан в предшествующем подразделе.

Речевой сигнал – сложный коррелированный процесс с многомодовым спектром. В задачах распознавания речи, идентификации говорящего и др. необходимо анализировать структуру сигнала, а также форму спектра. В задачах идентификации речи особое внимание уделяется исследованиям гласных звуков, имеющих высокую информационную продуктивность [28]. Для них характерна многомодовость СПМ. На рис. 1 и рис. 2 показана факторизация параметрической АР оценки СПМ фонем “е” и “о” на основе мультипликативной модели.

Как показывает анализ графиков, составляющие спектров имеют больше отличий, чем полный спектр. Эти отличия могут быть использованы для повышения точности идентификации дикторов. Подобные расчеты проводились и для фонем согласных. В этой работе для решения задач идентификации говорящего применялась факторизация спектров речевых сигналов в СВСП представлении.

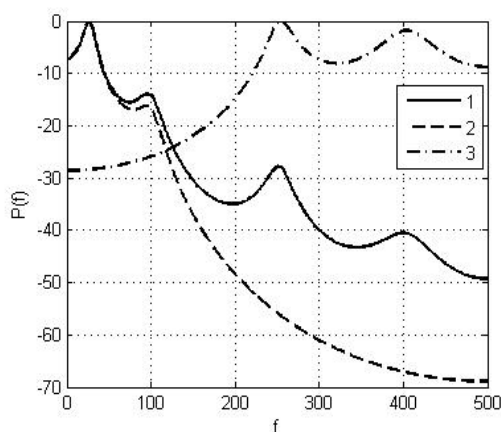


Рис. 1. Факторизация параметрических оценок СПМ фонемы “e”: 1 – СПМ по модели $AR(8)$, 2 – составляющая СПМ для модели $AR_1(4)$, 3 – составляющая СПМ для модели $AR_2(4)$

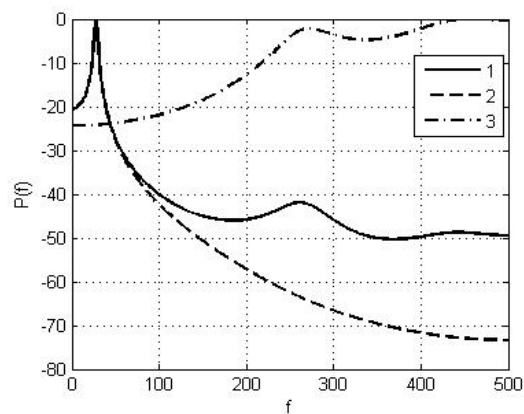


Рис. 2. Факторизация параметрических оценок СПМ фонемы “o”: 1 – СПМ по модели $AR(8)$, 2 – составляющая СПМ для модели $AR_1(4)$, 3 – составляющая СПМ для модели $AR_2(4)$

Идентификация говорящего с помощью авторегрессионного спектра речевого сигнала в СВСП представлении

Примеры АР СПМ в СВСП представлении некоторых слов показаны на рис. 3. Слова, произнесенные одним диктором, имеют различные спектры и могут служить признаками слов. СПМ были получены по модели АР 12-го порядка при длине подвектора 20 отсчетов.

В работе, для идентификации говорящего, предложено рассматривать ключевое слово в виде набора последовательности подвекторов длиной 20 отсчетов. Как показали исследования, АР многомодовые спектры слова в СВСП представлении имеют заметные отличия для разных дикторов. Важным достоинством спектров в СВСП представлении является возможность получать моды на частоте колебаний основного тона каждого диктора, что является важным признаком верификации говорящего. Чтобы использовать в качестве признаков не отсчеты спектров, а параметры спектральных мод речевых сигналов и тем самым сократить размерность пространства признаков, строится мультипликативная модель АР в виде $AR_1 \times AR_2 \times \dots \times AR_k$. Параметры составляющих моделей AR_i вычисляются по корням характеристического уравнения модели АР. Тогда многомодовый спектр p -го порядка (1.30) представляется (факторизуется) набором одномодовых спектров первого или второго порядков. Т.о. многомодовый спектр (1.30) представляется выражением

$$P^n(\omega) = \frac{D_a^n}{\left| \sum_{i=0}^{p_1} \Phi_1^n[i] e^{-j\omega i T} \right|^2 \left| \sum_{i=0}^{p_2} \Phi_2^n[i] e^{-j\omega i T} \right|^2 \times \dots \times \left| \sum_{i=0}^{p_k} \Phi_k^n[i] e^{-j\omega i T} \right|^2}$$

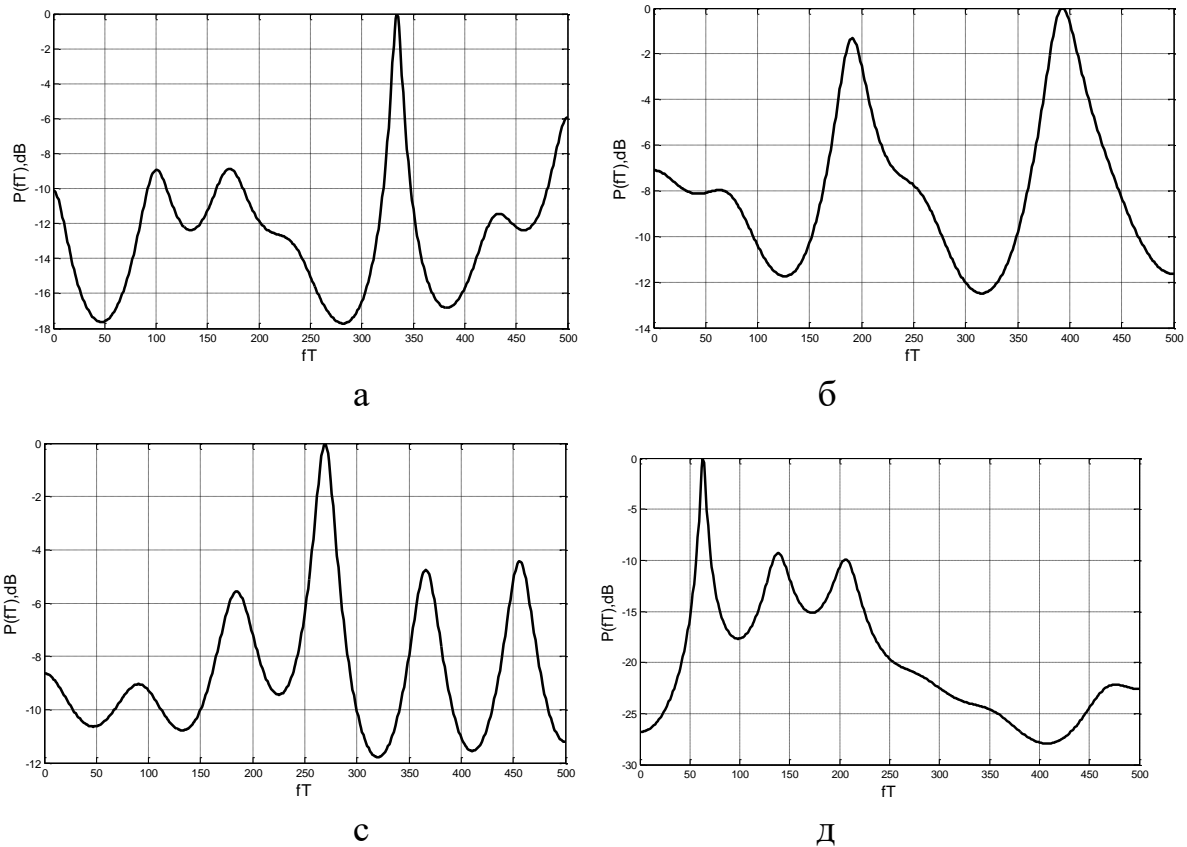


Рис. 3. СПМ по модели AP(12) в СВСП представлении: а – слова “езда”; б – слова “литр”; в – слова “село”; г – слова “hello”, произнесенное вторым диктором

Если корни характеристического уравнения комплексные, то из (5) и (7) следует, что частота пика определяется по формуле

$$f_m^n = \arccos\{\Phi_m^n[1] / (\sqrt{-\Phi_m^n[2]})\} / 2\pi T, \quad (33)$$

а ширина полосы пика равна

$$\Delta f_m^n = \ln(-\Phi_m^n[2]) / 2\pi T.$$

Если корни действительные, то модель AP имеет первый порядок. При этом, частоты пика $f=0$, если $\Phi_m^n[1] > 0$ или $f=1/2T$, если $\Phi_m^n[1] < 0$. Из (5) и (6) следует, что ширина полосы пика в этом случае равна

$$\Delta f_m^n = \ln(|\Phi_m^n[1]|) / \pi T.$$

В качестве признаков идентификаторов-слов предложено использовать частоты пиков параметрического спектра. Исследования показали, что ширина полосы пика спектра менее устойчива и ее использование в качестве дополнительного признака речи говорящего, не дает улучшения качества идентификации. У каждого диктора параметры спектра ключевого слова лежат в некоторых пределах, определяемых на этапе обучения. Для распознавания можно использовать не все параметры спектров, а наиболее устойчивые, например, частоты пиков.

На рис. 4 показаны примеры СПМ слова “hello” для третьего диктора женщины. Как видно из графиков, СПМ для каждого произнесенного слова диктором близки между собой. Это является причиной выбора частотных характеристик АР СПМ речевых сигналов в качестве признаков дикторов.

Для расчета частот пиков применялась факторизация СПМ. Одномодовые составляющие СПМ слова “hello” третьего диктора, спектр которого показан на рис. 4в, представлены на рис. 5. Сравнение графиков на рис. 4в и рис. 5 показывает, что факторизация дает точное совпадение соответствующих частот пиков СПМ. Это особенно заметно на частотах узкополосных составляющих СПМ. Как видно из графиков на рис. 5, слабые по мощности составляющие СПМ, показанного на рис. 4в, имеют значительную ширину полосы пика. Они не всегда устойчивы и поэтому часто не использовались в качестве признаков речевых сигналов дикторов.

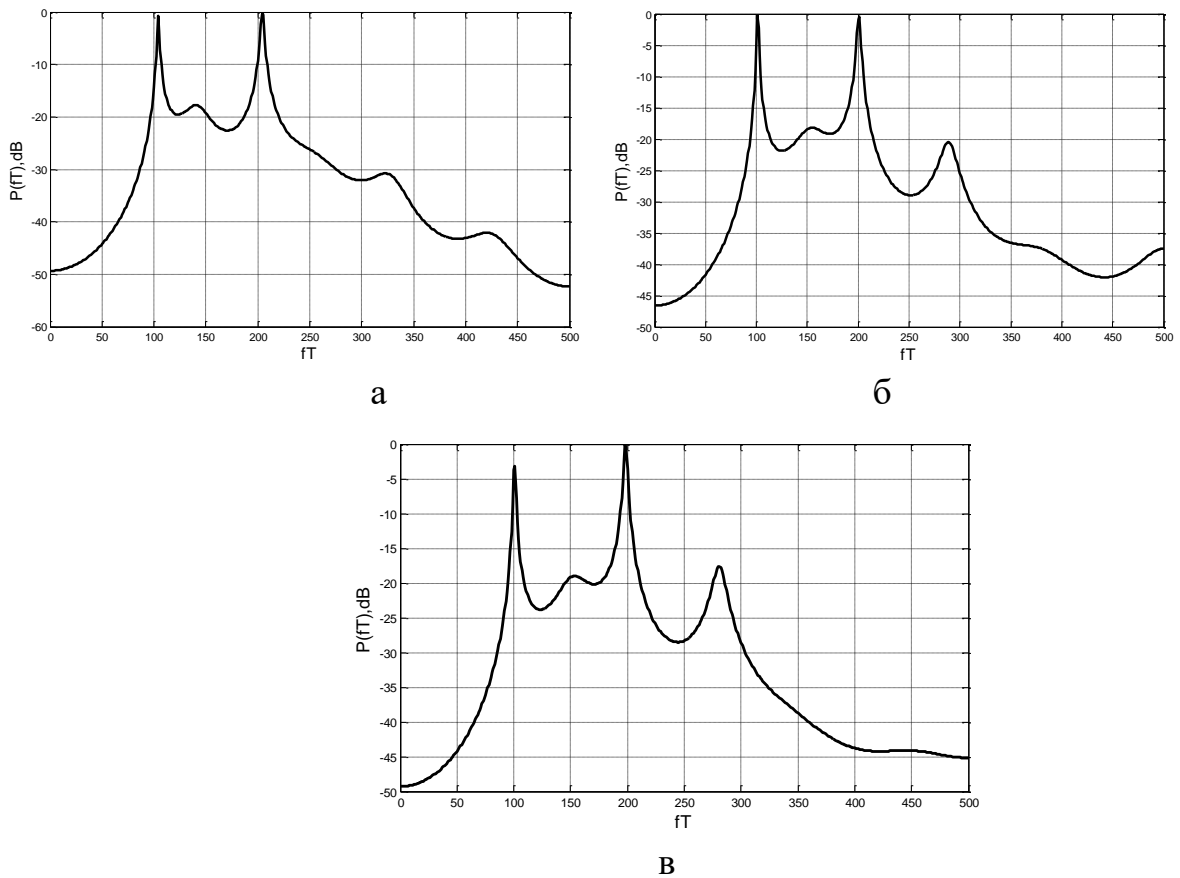


Рис. 4. АР СПМ для 3-го диктора (женщина) слова “hello” (а – произношение 1; б – произношение 2; в – произношение 3)

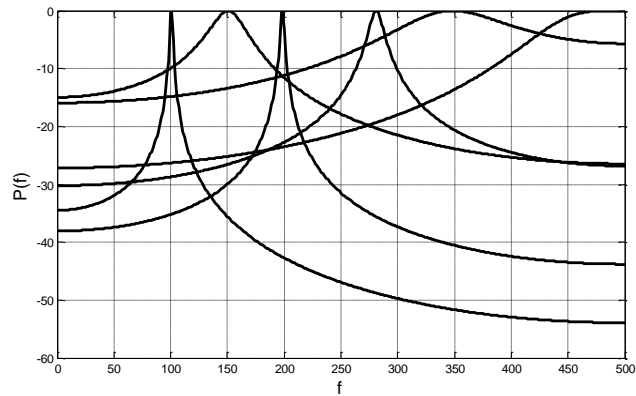


Рис. 5. Одномодовые составляющие СПМ слова “hello” для 3-го диктора, произношение 3 (рис. 1.4в)

В табл. 1 приведены частоты пиков СПМ слова “hello” и их ширины полосы по уровню 0,5, показанного на рис. 5, полученные после факторизации многомодовой СПМ, и рассчитанные по (33). Как показали исследования, частоты пиков более устойчивы по сравнению с ширинами полос. Поэтому именно они использовались в качестве признаков слова при идентификации дикторов.

Таблица 1.

Частоты пиков и их ширины полос составляющих СПМ, представленных на рис. 5

f [Гц]	1	2	3	4	5	6
f_s	453	198	151	101	281	343
Δf_s	93.9	2.76	29.5	1.72	11.94	96.2

Для сравнения, на рис. 6 показаны примеры АР СПМ речевого сигнала слова “hello” для 35-го диктора (мужчины). Форма СПМ в этом случае имеет заметные отличия, по сравнению с СПМ третьего диктора. Факторизация СПМ спектра, представленного на рис. 6в, показана на рис. 7. Частоты пиков и их ширины полос представлены в табл. 2.

Таблица 2.

Частоты пиков и их ширины полос составляющих СПМ, представленных на рис. 7

f [Гц]	1	2	3	4	5	6	7
f_s	500	0	100	244	197	281	400
Δf_s	219.7	152.7	21.0	31.8	73.1	9.1	21.7

Экспериментально исследовалась эффективность использования частот пиков СПМ в качестве признаков речи дикторов. Для этого проводилась идентификация дикторов. При проведении эксперимента 35 дикторов произносили слово “hello” по 10 раз каждый.

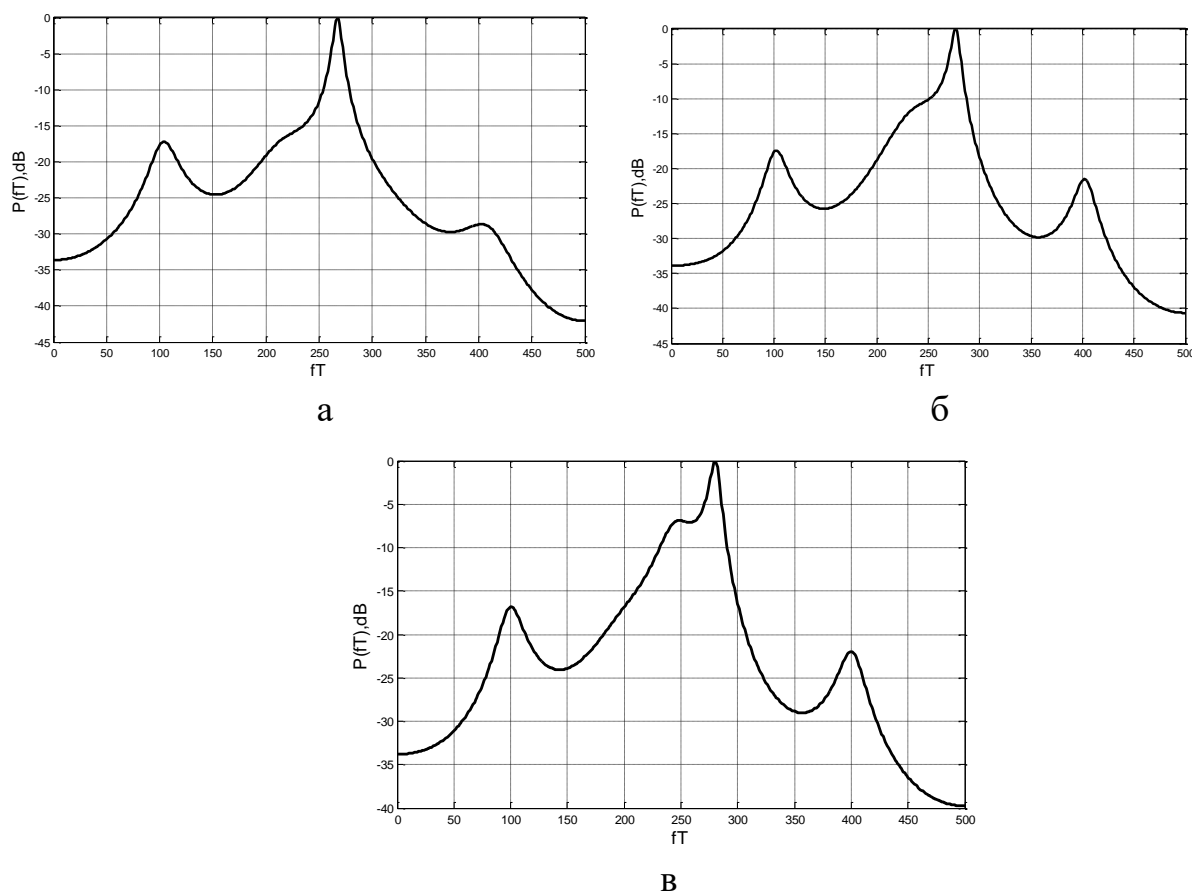


Рис. 6. Пример АР СПМ для 35-го диктора (мужчина) слова “hello” (а – произношение 1; б – произношение 2; в – произношение 30)

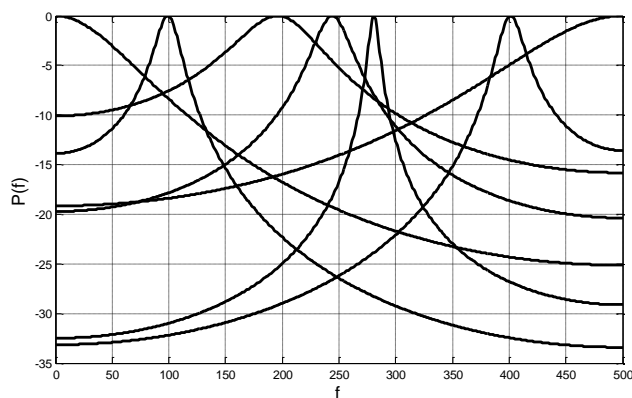


Рис. 7. Составляющие СПМ слова “hello” для 35-го диктора, произношение 3 (рис. 6в)

Речевой сигнал, произнесенного слова, не разбивался на составляющие. Полагалось, что слово — это единая выборка. Сигнал речи представлялся последовательностью подвекторов длиной 20 отсчетов каждый. СВСП метод позволяет определить долговременную характеристику слова. Для всех 10 слов каждого диктора были рассчитаны АР СПМ для 12-го порядка сигнала в СВСП представлении. Для этого порядка модели АР и при длине подвектора 20 отсчетов, спектры дикторов имеют высокие моды и заметно отличаются между собой. В ходе экспериментов определялся наиболее подходящий по вероятности правильного распознавания дикторов порядок.

На этапе обучения для всех слов, произнесенных дикторами, были найдены 3 или 6 наиболее устойчивых частот пиков. Их средние, полученные для каждого диктора, служили эталонными признаками диктора. Для контрольного слова рассчитывались частоты пиков и поочередно сравнивались со всеми эталонами. Решение принималось в пользу того диктора, у которого было минимальное суммарное евклидово расстояние между рассчитанными частотами пиков и эталонными.

Анализ спектров слов, произнесенных разными дикторами, показал, что для 12-го порядка модели AP в спектре присутствуют обычно 6 пиков. В табл. 3 представлены 6 частот пиков составляющих СПМ для 1-го диктора. Как правило, частоты пиков устойчивы и не имеют значительного отклонения от их средних значений. Однако, для некоторых дикторов наблюдается отсутствие некоторых пиков. Например, как видно из табл. 4, в некоторых произнесенных словах 2-м диктором, отсутствуют четвертые пики. Для других дикторов, могут отсутствовать 4-й, 5-й и 6-й пики. Поэтому если в идентифицируемом слове отсутствуют эти пики, распознавание производится по первым трем частотам. Они присутствуют в СПМ у всех произнесенных слов всеми дикторами. В табл. 5 приведены частоты пиков для слов диктора у которого отсутствует 6-й пик. Т.о. если в идентифицируемом слове присутствуют все 6 пиков, в решающем правиле используется эталон для 6 частот. Если в идентифицируемом слове отсутствуют некоторые пики, то в решающем правиле используются 3 частоты слова и эталонов.

Таблица 3.

Частоты пиков составляющих СПМ для 1-го диктора

f1	f2	f3	f4	f5	f6
950	1502	1879	2325	2935	3411
978	1447	1849	2274	3019	3406
988	1495	1925	2269	3041	3400
1000	1539	1816	2277	3111	3457
985	1494	1835	2281	3066	3408
995	1505	1785	2275	3189	3434
989	1534	1874	2255	3206	3439
993	1538	1779	2236	3170	3399
996	1536	1609	2191	3190	3384
981	1425	1763	2230	3309	3354

Таблица 4.

Частоты пиков составляющих СПМ для 2-го диктора

f1	f2	f3	f4	f5	f6
520	1089	1661	2124	2864	3624
481	1075	1677	2131	2911	3645
521	1185	1932	1940	2505	3567
515	1104	1868	2022	2637	3563

662	1227	1857		2437	3437
551	1264	1999		2586	3565
663	1225	1968		2493	3542
665	1204	1890		2454	3390
405	1039	1554	2091	2901	3587
510	1110	1694	2040	2894	3559

Таблица 5.

Частоты пиков составляющих СПМ для 6-го диктора

f1	f2	f3	f4	f5
886	1722	1906	2698	3296
876	1735	2094	2743	3465
877	1718	2007	2424	3233
858	1697	2166	2694	3127
854	1693	2144	2758	3624
884	1736	1886	2762	3484
887	1784	1903	2817	3478
874	1714	1991	2679	3358
877	1730	1923	2813	3344
879	1700	1864	2717	3185

В табл. 6 приведены примеры трех или шести значений эталонов частот для десяти дикторов. Приведенные частоты эталонов поясняют выбор частот в качестве признаков распознавания, приведенных в табл. 3 - табл. 5.

В табл. 7 приведены вероятности правильного распознавания каждого диктора и среднее значение для всех дикторов. Результаты исследования показывают высокую вероятность правильного распознавания. Среднее значение вероятности правильного распознавания для всех дикторов составило 90,92%. После отбрасывания хуже всего распознаваемых пяти дикторов, вероятность правильной идентификации дикторов повысилась до 94,15%. К плохо распознаваемым относятся слова, произнесенные с разной интонацией, а также в присутствии шума.

Таблица 6.

Примеры трех или шести значений эталонов частот для десяти дикторов

Spk.	f [kHz]	1	2	3	4	5	6
Spk. 1	N _{frq. 3}	-	-	-	-	-	-
	N _{frq. 6}	0.99	1.5	1.81	2.26	3.12	3.41
Spk. 2	N _{frq. 3}	0.64	1.23	1.93	-	-	-
	N _{frq. 6}	0.49	1.1	1.73	2.06	2.79	3.59
Spk. 3	N _{frq. 3}	0.82	1.17	1.61	-	-	-
	N _{frq. 6}	0.83	1.15	1.63	2.12	2.64	3.5
Spk. 4	N _{frq. 3}	-	-	-	-	-	-

	N _{frq. 6}	0.52	1.11	1.67	2.19	2.3	3.64
Spk. 5	N _{frq. 3}	0.62	1.16	1.76	-	-	-
	N _{frq. 6}	0.58	1.03	1.59	2.26	2.92	3.61
Spk. 6	N _{frq. 3}	0.82	1.57	1.97	-	-	-
	N _{frq. 6}	-	-	-	-	-	-
Spk. 7	N _{frq. 3}	0.77	1.36	2.03	-	-	-
	N _{frq. 6}	-	-	-	-	-	-
Spk. 8	N _{frq. 3}	-	-	-	-	-	-
	N _{frq. 6}	5.13	1.26	1.49	2.03	2.79	3.55
Spk. 9	N _{frq. 3}	-	-	-	-	-	-
	N _{frq. 6}	4.78	1.14	1.62	2.02	2.78	3.51
Spk.10	N _{frq. 3}	4.3	1.1	1.68	-	-	-
	N _{frq. 6}	4.2	1.11	1.66	2.02	2.95	3.56

Таблица 7.

Вероятности правильной идентификации для десяти дикторов

Spk.	Frequencies	Results "true"	Spk.	Frequencies	Results "true"
Spk.1	N _{frq. 3}	-	Spk. 7	N _{frq. 3}	%71
	N _{frq. 6}	100%		N _{frq. 6}	-
Spk. 2	N _{frq. 3}	96.43%	Spk. 8	N _{frq. 3}	-
	N _{frq. 6}	92.59%		N _{frq. 6}	%99.17
Spk. 3	N _{frq. 3}	100%	Spk. 9	N _{frq. 3}	-
	N _{frq. 6}	100%		N _{frq. 6}	%75.17
Spk. 4	N _{frq. 3}	-	Spk. 10	N _{frq. 3}	%93.8
	N _{frq. 6}	100%		N _{frq. 6}	%100
Spk. 5	N _{frq. 3}	85.71%			
	N _{frq. 6}	99.54%			
Spk. 6	N _{frq. 3}	100%			
	N _{frq. 6}	-			
All 35 speakers				90.92%	
30 speakers in normal conditions				94.15%	

Выводы

Представлен новый способ распознавания дикторов на основе СВСП представления процесса. Модель позволяет получить характерные различия в спектрах речи различных дикторов, произносивших одно и то же слово. Для распознавания использовались частоты пиков в параметрическом спектре сигналов в СВСП представлении произнесенных слов. Для этого речевой сигнал слова представлялся в виде последовательности подвекторов длиной 20 отсчетов. При построении параметрической модели АР, корреляционная функция рассчитывалась по подвекторам, а не как обычно, по отсчетам сигнала. Таким способом строилась долговременная модель слова.

При оценке частот пиков многомодовый спектр слова представлялся эквивалентным набором одномодовых спектров. Для этого использовалась

факторизация спектра на основе мультипликативной модели AP. Частоты пиков спектра рассчитывались по формулам, связывающим корни характеристического уравнения AP и частоту пика одномодового спектра.

Экспериментальная идентификация дикторов осуществлялась по ключевому слову “hello”, произнесенному каждым диктором 10 раз. В качестве эталонов использовались средние значения соответствующих частот, усредненных по 10 словам. Вероятность правильного распознавания рассчитывалась для 35-ти дикторов и составляла 91%. После отбрасывания 5-ти наихудших распознаваемых дикторов, вероятность правильного распознавания дикторов выросла до 94%.

Литература

1. Kinnunen T., Li H. An overview of text-independent speaker recognition: from features to supervectors// *Speech Commun.* - 2010.- **52** (1).- P. 12–40.
2. Hansen J.H.L., Hasan T. Speaker recognition by machines and humans: a tutorial review// *IEEE Signal Process. Mag.*- 2015.- **32** (6). - P. 74–99.
3. Jayanna H., Prasanna S.M. Analysis, feature extraction, modeling and testing techniques for speaker recognition// *IETE Tech. Rev.*- 2009.- **26** (3).- P. 181–190.
4. Ідентифікація людини за параметрами мовних сигналів (проблеми та шляхи їх вирішення)/ Бондаренко М.Ф., Дрюченко О.Я., Коряк С.Ф., Ю.П. Шабанов-Кушнарченко. – Харків: ТОВ «Компанія СМІТ». - 2006. – 260 с.
5. Методы автоматического распознавания речи: В 2-х книгах.: Пер. с англ. / Под ред. У. Ли. - М.: Мир, 1983. - Кн.1. - 328 с.
6. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу. М: Радио и Связь. - 1981. - 222 с.
7. Das R.K., Mahadeva Prasanna S. Exploring different attributes of source information for speaker verification with limited test data// *J. Acoust. Soc. Am.*- 2016. - **140** (1). - P. 184–190.
8. Bimbot F., Bonastre J.F, Fredouille C., Gravier G., Magrin-Chagnolleau I., Meignier S., Merlin T., Ortega-Garcia J., Petrovska-Delacretaz D., Reynolds D. A tutorial on text-independent speaker verification// *EURASIP J. Appl. Signal Process.* - v.4.- 2004.- P. 430–451.
9. Hautamäki V., Kinnunen T., Nosratighods M., Lee K.-A., Ma B., Li H. Approaching Human Listener Accuracy with Modern Speaker Verification// *Interspeech.* - 2010. - P. 1473-1476.
10. Schmidt-Nielsen A., Crystal Th. H. Speaker verification by human listeners //Experiments comparing human and machine performance using the NIST 1998 speaker evaluation data. *Digital Signal Processing.* - v. 10.- 2000.- P.- 249–266.
11. Lamel L.F., Gauvain J.L. Speaker verification over the telephone// *Speech Communication.*- v.31.- 2000.- P.- 141-154.
12. Wu T., Duchateau J., Martens J.-P., Compennolle D.V. Feature subset selection for improved native accent identification // *Speech Communication.*- v.52.-2010.- P.-83–98.
13. Lavner Y., Gath I., Rosenhouse J. The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels// *Speech Communication.*- v.30.-2000.- P.- 9-26.
14. Rosenberg A., Siohan O., Parthasarathy S. Small group speaker identification with common password phrases// *Speech Communication.*- v. 31.-2000.-P.-131-140.
15. Bartkova K., Gac D.L., Charlet D., Jouvét D. Prosodic parameter for speaker identification//*Proc. Internat. Conf. on Spoken Language Processing (ICSLP).*- 2002.-P.-1197–1200.
16. Sahidullah M., Kinnunen T. Local spectral variability features for speaker verification//*Digital Signal Processing.* -v. 50.-March 2016.-P.-1–11.

17. Li Z.Y., Zhang W.Q., Liu J. Multi-resolution time frequency feature and complementary combination for short utterance speaker recognition//Multimedia Tools Appl.-v.74, (3).-2015.- pp.-937–953 .
18. Chen Y., Tang Z.M. The speaker recognition of noisy short utterance//ICISBDE.- 2013.-pp.-666–671.
19. Soldi G., Bozonnet S., Alegre F. et al. Short-duration speaker modelling with phone adaptive training// Proc. Odyssey.- 2014.-pp. -208-215.
20. Li L., Wang D., Zhang C. et al. Improving short utterance speaker recognition by modeling speech unit classes// IEEE Trans. Audio Speech Lang. Process.-v. 24, (6).-pp.-1129–1139 .
21. Бокс Дж., Дженкинс Г. Анализ временных рядов: Пер. с. англ. – М.: Мир, 1974. – Вып.1. – 406с.
22. Тихонов В.А., Русановский Д.Е., Тихонов Д.В. Генерирование узкополосных имитационных случайных процессов // Радиотехника и информатика. – 1999. – №4. – С. 83–85.
23. Марпл.–мл. С. Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.
24. Тихонов В.А. Синтез и анализ мультипликативных процессов линейного предсказания / В.А. Тихонов, Н.В. Кудрявцева // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2010. – № 162. – С. 140 – 143.
25. Тихонов В.А. Математические модели составных векторных случайных процессов / В.А. Тихонов, Н.В. Кудрявцева, И.О. Филь // Восточно-Европейский журнал передовых технологий. – 2011. – № 2/4(50). – С. 17 – 20.
26. N. V. Kudriavtseva. Factorization of processes parametric spectra on the base of multiplicative linear prediction polymodels,” RADIOELEKTRONIKA 2014 24th International Conference,” pp.1-4, 2014.
27. N. V. Kudriavtseva, and I. O. Fil, “Factorization of Rhythmograms Parametric Spectra on the Base of Multiplicative Linear Prediction Models,” IEEE East-West Design & Test Symposium (EWDTS’2012), Kharkiv, Ukraine, pp. 538-540, 2012.
28. Бондаренко М.Ф., Дрюченко А.Я., Шабанов-Кушнаренко Ю.П. Гласные звуки в теории и эксперименте. Х.: Харьк. нац. ун-т радиотехники, 2002. 348 с.