

Текущий уровень качества условий доступа к ресурсам: $QoS_i = 0.65$. Текущее состояние не соответствует желаемому на: $\Delta_2 = 1.65$ - <очень значительно>. Желаемый уровень качества условий доступа к ресурсам: $QoS_d = -1$. В следующий период контроля ожидается изменение состояния условий на: $\Delta_1 = 1 \sim$ <значительное>. Ожидаемый уровень качества доступа к ресурсам на следующем шаге будет равен: $QoS_p = -0.35 \sim$ <среднее>. Это состояние не соответствует желаемому на: $\Delta_3 = 0.65$ - <среднее>.

8. Выводы

В настоящей работе рассмотрены вопросы оценивания качества работы пользователя в компьютерной сети и качества функционирования самой сети. К основным результатам относятся метод оценивания качества доступа пользователя к ресурсам компьютерной сети. Результаты могут быть отнесены к научным, получены впервые.

Сравнение с аналогами. Отличие предлагаемого метода состоит в том, что для большинства подходов оцениваются только количественные характеристики [8]. Использование количественных характеристик снижает возможности интерпретации состояния объектов и оперативности принятия решений.

Общее описание *научного результата* можно представить как: впервые предложен метод оценивания качества доступа к ресурсам для пользователя компьютерной сети. Метод основан на комбинированном оценивании состояния уровня доступа к ресурсам, состоящий в том, что сначала оцениваются количественные показатели качества доступа, далее формируются экспертные оценки, и потом осуществляется переход от количественных оценок к качественным. Такой подход позволит получить простые дискретные оценки, позволяющие реализовать оперативный контроль состояния компьютерной сети в целом. *Научная новизна* заключается в том, что был разработан метод оценки качества условий доступа к ресурсам с учетом индивидуальных характеристик конкретного пользователя.

Практическая значимость состоит в том, что предложенные методы можно успешно использовать в системах непрерывного мониторинга состояния оценок качества условий работы пользователя в сети. Это в свою очередь способствует своевременному выявлению отказов сети, снижению затрат от наличия необнаруженных отказов, созданию предпосылок повышения доходности сети за счет повышения качества обслуживания пользователей.

Пути дальнейших исследований. Предполагается рассмотреть вопросы оперативного оценивания задержек и отказов в сети.

Литература: 1. *Dinesh Verma C.* Simplifying Network Administration Using Policy-Based Management» IBM Thomas J Watson Research Center, IEEE Network, March/April 2002. P. 3-5. 2. *Rich Ptak.* Enterprise Network Configuration Management: A practitioner's guide. 2003 Network World, Inc. and Ptak & Associates, Inc. P. 5-7. 3. *Don Jones.* Tips and Tricks Guide To Network Configuration Management. 2003 Realtimedpublishers.com, Inc. P.5. 4. *Саенко В.И., Сизова О.В.* Модель пользователя компьютерной сети // АСУ и приборы автоматики. 2004. Вып. 129. С. 86-93. 5. *Larsen K.* The role of service level agreements in IT service delivery. // Information Management & computer security. 6/3, 1998. P.128-132. 6. 3GPP Technical Specification 22.121 v4.0.0: "The Virtual Home Environment (Release 4)", October 2000. P.2-4. 7. UMTS Forum Report 1: "A Regulatory Framework for UMTS", June 1997. P.2-7. 8. *Бацамут В.Н.* Метод управления входной нагрузкой сети передачи данных для повышения оперативности доставки информации // АСУ и приборы автоматики. 2004. Вып. 128. С.67-76.

Поступила в редколлегию 09.11.2006

Рецензент: д-р техн. наук, проф. Поповский В.В.

Саенко Владимир Иванович, канд. техн. наук, проф. кафедры информационных управляющих систем ХНУРЭ. Научные интересы: менеджмент компьютерных сетей. Увлечения и хобби: садоводство. Адрес: Украина, 61166, Харьков, пр. Ленина, 14.

Кривонос Татьяна Ивановна, магистр каф. информационных управляющих систем ХНУРЭ. Научные интересы: менеджмент компьютерных сетей. Увлечения и хобби: занятие спортом. Адрес: Украина, 61166, Харьков, пр. Ленина, 14.

УДК004.93'1:004.942:519.23

СИНТЕЗ И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ДИКТОРОВ ПО ХАРАКТЕРИСТИКАМ ОСТАТКОВ ЛИНЕЙНОГО ПРЕДСКАЗАНИЯ

ФЕДОРОВ А.В., ОМЕЛЬЧЕНКО А.В

Описывается синтез и исследование текстонезависимых адаптивных алгоритмов идентификации дикторов по характеристикам остатков линейного предсказания. В качестве информативных признаков используются коэффициенты преобразования Фурье автокорреляционной фун-

кции остатков линейного предсказания, пропущенных через различные временные окна. Выполняется экспериментальная проверка предложенных алгоритмов идентификации.

1. Введение

Задаче идентификации личности по голосу уделяется значительное внимание как у нас в стране [1, 2], так и за рубежом [3-6]. Этот интерес практически не ослабевает на протяжении последних 30 лет. Несмотря на это, к настоящему моменту окончательного решения проблемы не найдено. Тем не менее, современный уровень развития систем связи предполагает внедрение новых услуг, включающих в себя управление своим счетом по телефону, удаленный доступ к информации, call центры и т.д. В связи с этим *актуальной задачей* является разработка алгоритмов иден-

тификации абонентов как мобильной, так и классической телефонной связи.

Целью настоящей работы является создание текстонезависимой системы идентификации дикторов по характеристикам остатков линейного предсказания. Сформулированная задача решается в три этапа. На первом этапе формируется система информативных признаков. На втором – синтезируется правило принятия решения и на третьем – разрабатываются алгоритмы вычисления решающих статистик.

2. Постановка задачи

В настоящее время для анализа речевых сигналов широко используется метод линейного предсказания. Преимущество его состоит в том, что он способствует быстрому и точному выполнению необходимых преобразований. Показав высокую эффективность в задачах распознавания речи [5, 6], метод линейного предсказания оказался приемлемым также и для автоматического опознавания дикторов [3, 5, 6], поскольку в результате анализа речевого сигнала таким методом мы обычно получаем комбинированную информацию о весьма важных параметрах речевой активности конкретного человека. Такими параметрами являются в первую очередь импульсы возбуждения голосовой щели, а также формантные частоты и полосы формантных областей [1]. Кроме того, методом линейного предсказания могут быть оценены параметры основного тона голоса [5, 6].

В основу модели линейного предсказания положено предположение, что любое значение речевого сигнала может быть представлено как линейная комбинация p предыдущих отсчетов [5]. Если непрерывный сигнал $s(t)$ дискретизируется во времени с интервалом Δ , в результате получается последовательность $s_n = s(n\Delta)$, где n – целочисленная переменная, то значение предсказываемой величины может быть записано как

$$\hat{s}_n = \sum_{k=1}^p a_k s_{n-k}, \quad (1)$$

где a_k – коэффициенты предсказания; p – порядок модели линейного предсказания. Определим ошибку (остаток) предсказания как разность между предсказываемым и действительным значением n -го отсчета временного ряда

$$e_n = s_n - \hat{s}_n = s_n - \sum_{k=1}^p a_k s_{n-k}. \quad (2)$$

В рамках модели линейного предсказания синтез речевого сигнала может быть выполнен с использованием рекурсивного линейного цифрового фильтра согласно рис. 1.

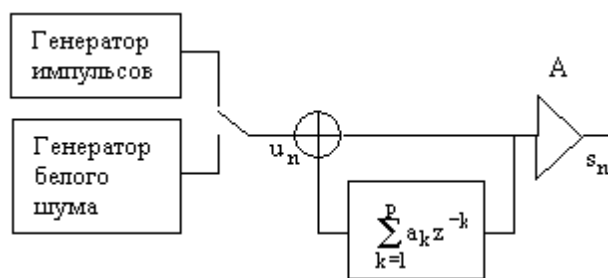


Рис. 1. Цифровая модель речеобразования

Сигналом возбуждения u_n служит либо последовательность импульсов, формируемая генератором импульсов, либо шум, формируемый генератором белого шума. Эти два источника сигнала моделируют вокализованные и невокализованные фрагменты речевого сигнала. Ошибка предсказания, определяемая уравнением (2), в рамках данной модели совпадает с функцией возбуждения u_n . Исходя из этого, сигнал ошибки предсказания e_n может быть использован для решения задач идентификации дикторов, а также для определения периода основного тона и признака тон/шум. В [7, 8] показано, что речевой тракт и источник возбуждения связаны слабо и изменения конфигурации тракта мало влияют на колебания голосовых связок. Таким образом, в пределах справедливости этого положения источник и тракт с точки зрения анализа и построения эквивалентных схем могут рассматриваться независимо друг от друга.

Представим модель речеобразования в виде дискретного линейного полюсного фильтра с p полюсами. Передаточная функция этого фильтра, отражающая как особенности голосового тракта, так и излучение и форму речевой волны, описывается функцией вида

$$H(z) = A / (1 - \sum_{k=1}^p a_k z^{-k}), \quad (3)$$

где A – коэффициент усиления; z – переменная z – преобразования.

Рассматриваемая задача синтеза и исследования алгоритмов идентификации дикторов в рамках модели линейного предсказания может быть сформулирована так. Реализации речи эталонных и идентифицируемых дикторов представляют собой сигналы со следующими характеристиками: частота дискретизации 8 кГц, ИКМ, 16 битов на отсчет. Формирование остатков линейного предсказания выполняется с помощью решетчатого фильтра [5] 10-го порядка по сегментам длительностью 30 мс. Классификация речи на вокализованные и невокализованные сегменты выполняется автокорреляционным методом. При этом сегмент считается вокализованным, если отношение пика нормированной автокорреляционной функции, лежащего в диапазоне 80-200 Гц, к значению этой функции в нуле не менее 0,25. Необходимо синтезировать адаптивные алгоритмы идентификации дикторов, устойчивые к возможным частотным искажениям исходного речевого сигнала в тракте передачи.

3. Информативные признаки

В задачах идентификации дикторов в качестве информативных признаков используются характеристики речевого сигнала, отражающие индивидуальность голоса. Помимо полноты представления информации об особенностях голоса и речи диктора информативные признаки должны обладать рядом других свойств. Во-первых, они должны быть легко измеряемы и мало зависимы от мешающих факторов окружающей среды (шумов и помех), во-вторых, они должны обладать стабильностью во времени и, в-третьих, не должны поддаваться имитации. Создание системы признаков, удовлетворяющей всем этим требованиям, является трудно решаемой задачей. Поэтому, в первую очередь, производят отбор признаков по их способности наиболее адекватно и эффективно произвести описание индивидуального голоса, уделяя меньше внимания другим факторам.

Согласно [1] основные физические проявления индивидуальности следует искать в спектральных и формантных, а также временных и амплитудных характеристиках сигнала.

В результате выполненного анализа речевых фрагментов дикторов, содержащих гласные звуки (фонемы) «а», «и», «о», «у», «э», установлено, что форманты, представляющие собою резонансы голосового тракта, в первую очередь определяют характер произносимого звука и слабо зависят от диктора. В качестве примера на рис. 2, а, б показаны диаграммы формантных траекторий для звука «э», произнесенного двумя различными дикторами мужского пола. Оценивание формантных траекторий выполнялось согласно процедуре, изложенной в [5].

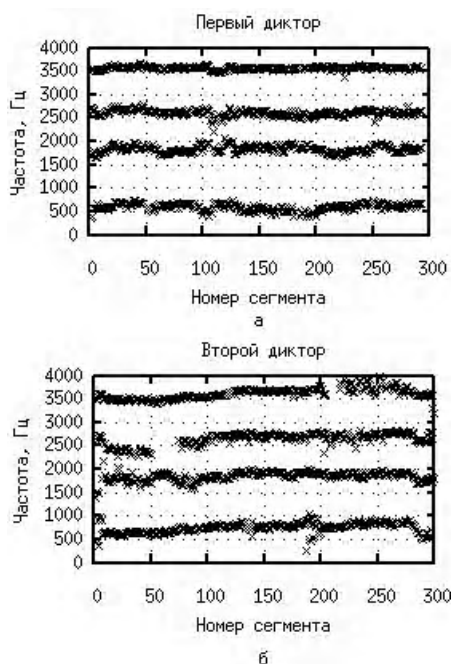


Рис. 2. Диаграммы формантных траекторий для звука «э», произнесенного двумя дикторами мужского пола

Как известно [1], форма и характер следования голосовых импульсов являются индивидуально дифференцирующими свойствами и непосредственно участвуют в образовании характерного для диктора голоса. В свою очередь, энергетический спектр $|S(\omega)|^2$ звукового сигнала равен [8]

$$|S(\omega)|^2 = |V(\omega)|^2 \cdot |T(\omega)|^2, \quad (4)$$

где $|V(\omega)|^2$ – энергетический спектр сигнала голосового возбуждения; $T(\omega)$ – передаточная функция речевого тракта. Таким образом, информация об импульсах голосового возбуждения может быть получена из анализа спектральных характеристик как остатков линейного предсказания, так и исходного речевого сигнала. На рис. 3 представлены спектрограммы исходных речевых сигналов (а, б) и сигналов остатков линейного предсказания (в, г) тех же дикторов. Из анализа приведенных зависимостей легко видеть, что использование оценок энергетических спектров позволяет уверенно различить дикторов по реализациям достаточно большой протяженности.

Процедура формирования системы информативных признаков включает в себя следующие этапы: 1) формирование согласно (2) сигнала остатков линейного предсказания; 2) интегрирование исходного речевого сигнала и остатков линейного предсказания в целях компенсации эффекта дифференцирования, осуществляемого губами говорящего. Передаточная функция цифрового интегрирующего фильтра имеет вид $I(z) = 1/(1 - \alpha z^{-1})$, где α – параметр, значение которого меньше единицы и должно быть близко к единице [5]; 3) выделение информативных признаков из преобразованных сигналов исходной речи и остатков линейного предсказания, полученных на втором этапе. Кроме того, для исследования влияния частотных искажений речи, возникающих в тракте передачи, речевой сигнал на этапе 2 пропускается через цифровой дифференцирующий фильтр с передаточной функцией $D(z) = 1 - \beta z^{-1}$, $\beta = 0,9$.

В настоящей работе в качестве информативных признаков использованы коэффициенты преобразования Фурье автокорреляционной функции преобразованных речевого сигнала и сигнала остатков линейного предсказания, пропущенных через различные временные окна. Если через x_k , $k = 0, n-1$ обозначить координаты вектора информативных признаков $\bar{x} = (x_0, x_1, \dots, x_{n-1})$ размерности n , то

$$x_k = \text{Re} \left(\sum_{m=0}^{n-1} r(m) w(m) e^{-\frac{2\pi}{n} km} \right), \quad (5)$$

где $r(m) = \sum_{l=0}^{n-1-m} s(l) s(l+m)$ – автокорреляционная функция сигнала $s(l)$, полученного на втором этапе формирования системы информативных признаков; $w(m)$ – временное окно, форма которого определяется одним из следующих соотношений:

$$w(m) = 1 - \frac{m}{n-1} \quad (\text{треугольное окно});$$

$$w(m) = 0,5[1 + \cos(\frac{\pi}{n-1}m)] \quad (\text{окно Хеннинга});$$

$$w(m) = 0,54 + 0,46 \cos(\frac{\pi}{n-1}m) \quad (\text{окно Хэмминга}).$$

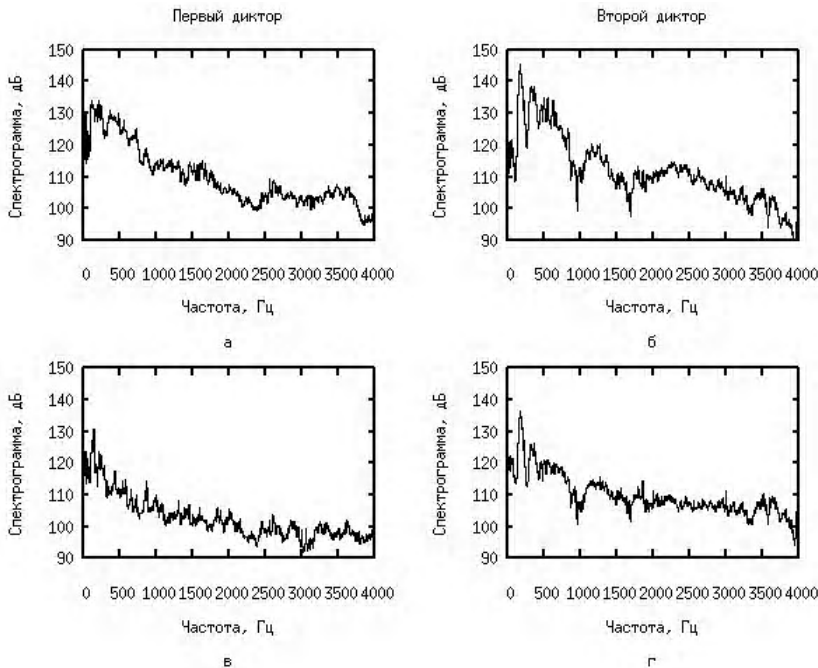


Рис. 3. Спектрограммы исходных речевых сигналов (а, б) и сигналов остатков линейного предсказания (в, г)

4. Алгоритмы идентификации

Правило принятия решения, полученное в рамках адаптивного подхода, имеет вид

$$j = \arg \max_{i=1, M} \max_a \{ \ln W(\vec{x} | i, a) \}, \quad (6)$$

где $j \in \{1, 2, \dots, M\}$ – принятое решение об идентифицируемом дикторе; M – число дикторов; $W(\vec{x} | i, a)$ – функция правдоподобия вектора информативных признаков \vec{x} с координатами, определяемыми соотношением (5); a – параметр, учитывающий неопределенность уровня сигнала. Основываясь на решающем правиле (6), можно предложить ряд алгоритмов идентификации.

Используем предположение о стационарности сигнала в пределах анализируемого отрезка. Тогда, поскольку отсчеты спектрограмм стационарных временных рядов асимптотически независимы и имеют χ_2^2 распределение с двумя степенями свободы [9], выражение для функции правдоподобия может быть записано в виде

$$W(\vec{x} | i, a) = \prod_{k=0}^{n-1} \frac{1}{a \cdot x_k^{(i)}} \cdot \exp\left\{-\frac{x_k}{a \cdot x_k^{(i)}}\right\}, \quad (7)$$

где x_k – k -е значение вектора информативных признаков \vec{x} идентифицируемого диктора; $x_k^{(i)}$ – k -е значение вектора информативных признаков \vec{x} i -го эталонного диктора.

Логарифмируя (7), получаем

$$\ln W(\vec{x} | i, a) \sim \sum_{k=0}^{n-1} \left\{ \ln(a \cdot x_k^{(i)}) + \frac{x_k}{a \cdot x_k^{(i)}} \right\}. \quad (8)$$

Дифференцируя и приравнявая к нулю правую часть соотношения (8), находим значение параметра a , минимизирующего выражение.

$$a = \frac{1}{n} \sum_{k=0}^{n-1} \frac{x_k}{x_k^{(i)}}. \quad (9)$$

Подставляя (9) в (8), получаем

$$\ln W(\vec{x} | i, a) \sim \ln \sum_{k=0}^{n-1} \frac{x_k}{x_k^{(i)}} + \frac{1}{n} \sum_{k=0}^{n-1} \ln x_k^{(i)}. \quad (10)$$

С учетом (10) правило (6) преобразуется в следующий алгоритм:

$$A_1: j = \min_{i=1, M} \left(\ln \sum_{k=0}^{n-1} \frac{x_k}{x_k^{(i)}} + \frac{1}{n} \sum_{k=0}^{n-1} \ln x_k^{(i)} \right). \quad (11)$$

Считая вектор информативных признаков \vec{x} принадлежащим n -му евклидовому пространству \mathbf{R}^n с метрикой

$$d(\vec{x} | i, a) = \sqrt{\sum_{k=0}^{n-1} (\sqrt{x_k} - a \cdot \sqrt{x_k^{(i)}})^2},$$

записываем функцию правдоподобия как

$$W(\vec{x} | i, a) = (2\pi\sigma^2)^{-n/2} \cdot \exp\left(-\frac{d^2(\vec{x} | i, a)}{2\sigma^2}\right), \quad (12)$$

где σ^2 – параметр масштаба распределения, одинаковый для всех дикторов.

Выполнив преобразования, аналогичные предыдущим, получим следующий алгоритм:

$$A_2: j = \min_{i=1, M} \sum_{k=0}^{n-1} (\sqrt{x_k} - a \cdot \sqrt{x_k^{(i)}})^2, \quad (13)$$

$$\text{здесь } a = \frac{\sum_{k=0}^{n-1} \sqrt{x_k x_k^{(i)}}}{\sum_{k=0}^{n-1} x_k^{(i)}}.$$

5. Результаты экспериментальных исследований

В настоящей работе выполнена экспериментальная проверка алгоритмов идентификации (11) и (13) для различных систем информативных признаков с учетом частотных искажений речи, возникающих в тракте передачи. В эксперименте использовалось 44 реализации 11 дикторов мужчин длительностью 18 с каждая. 11 из 44 реализаций использовались на этапе обучения, остальные 33 – на этапе идентификации. Таким образом, каждый из 11 дикторов был представлен 4 реализациями (по две записи двух текстов различного содержания).

Исследования показали, что регуляризация нормированного по мощности к единице речевого сигнала белым шумом с нулевым средним и дисперсией $\sigma_{ш}^2 = 0,1$ позволяет повысить вероятность правильной идентификации дикторов по остаткам линейного предсказания в среднем с 0,7 – 0,91 до 0,94 – 0,97. В таблице приведены оптимизированные по временному окну данные о вероятности правильной идентификации дикторов, соответствующие указанным условиям проведения эксперимента.

Из анализа данных таблицы легко видеть, что точность идентификации дикторов по спектральным характеристикам остатков линейного предсказания выше точности идентификации дикторов по аналогичным характеристикам исходного речевого сигнала. Кроме того, использование остатков линейного предсказания позволяет в значительной степени скомпенсировать частотные искажения речевого сигнала, вносимые трактом передачи с неравномерной частотной характеристикой.

Алгоритм	Анализируемый сигнал			
	Без частотных искажений речи в тракте передачи		С частотными искажениями речи в тракте передачи	
	Исходный речевой сигнал	Сигнал остатков линейного предсказания	Исходный речевой сигнал	Сигнал остатков линейного предсказания
(11)	0,94	0,97	0,82	0,88
(13)	0,94	0,97	0,79	0,94

6. Заключение

Научная новизна работы определяется тем, что в ней предложены и исследованы системы информативных признаков, содержащих информацию о параметрах источника голосового возбуждения. Выполнен синтез алгоритмов идентификации дикторов по указанным системам признаков. Использована процедура предварительного интегрирования анализируемого сигнала в целях компенсации эффекта дифференцирования, осуществляемого губами говорящего, что позволило повысить вероятность правильной идентификации дикторов.

На основании выполненных исследований сделано заключение, что данные об импульсах голосового возбуждения, заключенные в спектральных характеристиках как исходного речевого сигнала, так и остатков линейного предсказания, в отличие от данных о значениях первых формант позволяют уверенно разделить дикторов по реализациям достаточно большой протяженности, в то время как форманты в большей степени характеризуют произносимые звуки, слабо меняясь для разных дикторов.

Осуществлена экспериментальная проверка предложенных алгоритмов. Исследования показали, что регуляризация речевого сигнала белым шумом позволяет повысить вероятность правильной идентификации дикторов по остаткам линейного предсказания в среднем с 0,7 – 0,91 до 0,94 – 0,97. Кроме того, точность идентификации дикторов по остаткам линейного предсказания выше аналогичных результатов для исходного речевого сигнала. Применение остатков линейного предсказания позволяет также в значительной степени скомпенсировать частотные искажения речевого сигнала, вносимые трактом передачи с неравномерной частотной характеристикой.

Практическая значимость работы определяется тем, что разработанные алгоритмы идентификации могут быть использованы в системах ограничения доступа, а также в криминалистике.

Дальнейшая перспектива исследований состоит в переходе от использования блочных алгоритмов обработки речи к непрерывным, а также в более детальном исследовании устойчивости предложенных алгоритмов к возможным искажениям сигнала в тракте передачи.

Литература: 1. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу. М.: Радио и связь, 1981. 224 с. 2. Бондаренко М.Ф., Дрюченко А.Я., Шабанов-Кушнаренко Ю.П. Гласные звуки в теории и эксперименте. Х.: ХНУРЭ, 2002. 348с. 3. Атал Б.С. Автоматическое опознавание дикторов по голосам // ТИИЭР, 1976. Т. 64, № 4. С. 48-66. 4. Розенберг А.Е. Автоматическая идентификация диктора. Обзор // ТИИЭР, 1976. Т. 64, № 4. С. 66-79. 5. Маркел Дж. Д., Грей А.Х. Линейное предсказание речи: Пер. с англ. / Под ред. Ю.Н. Прохорова и В.С. Звездина. М.: Связь 1980. 308 с. 6. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов: пер. с англ. / Под ред. М.В. Назарова и Ю.Н. Прохорова. М.: Радио и связь. 1981. 496 с. 7. Фланаган Д. Л. Анализ, синтез и восприятие речи: Пер. с англ. / Под ред. А.А. Пирогова. М.: Связь, 1968. 391 с. 8. Фант Г. Акустическая теория речеобразования: Пер. с англ. / Под ред. В.С. Григорьева. М.: Наука, 1964. 284 с. 9. Бриллинджер Д. Р. Временные ряды. Обработка данных и теория: пер. с англ. / Под ред. А.Н. Колмогорова. М.: Мир, 1980. 536 с.

Поступила в редколлегию 06.11.2006

Рецензент: д-р физ.-мат. наук Прокопов А.В.

Федоров Алексей Валерьевич, ассистент кафедры «Сети связи» ХНУРЭ. Научные интересы: математическое моделирование в статистической теории распознавания образов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. (057)7021-429.

Омельченко Анатолий Васильевич, канд. техн. наук, доцент кафедры «Сети связи» ХНУРЭ. Научные интересы: методы обработки сигналов и распознавания образов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. (057)7021-429.