

Метод измерения частоты основного тона речевого сигнала для систем акустического анализа речи

А. В. Савченко¹, В. В. Савченко²

¹Национальный исследовательский университет Высшая школа экономики, Нижний Новгород, Россия, e-mail: avsavchenko@hse.ru

²Нижегородский государственный лингвистический университет, Нижний Новгород, Россия, e-mail: vvsavchenko@yandex.ru

Разработан новый метод измерения частоты основного тона речевого сигнала с повышенной помехоустойчивостью. Проблема защиты от интенсивного фонового шума в нём решена путём частотной селекции вокализованных отрезков речевого сигнала по схеме с гребенчатым фильтром межпериодного накопления. Эффективность метода исследована теоретически и экспериментально с помощью разработанного многоканального частотного измерителя для акустического анализа речи. Показано, что при отношении сигнал–шум 10 дБ и выше погрешность метода не превышает 2 %.

Ключевые слова: акустические измерения, акустика речи, речевой сигнал, основной тон, акустический шум, помехоустойчивость.

Под акустическим анализом речи (ААР) специалисты понимают измерения акустических характеристик речевого сигнала некоторого (контрольного) диктора [1–3]. Частота основного тона (ЧОТ) относится к числу наиболее информативных характеристик и играет важную роль в системах ААР разного назначения [4–6]. В частности, повышение ЧОТ в процессе речеобразования служит надёжным признаком (индикатором) внутреннего волнения контрольного диктора. Поэтому, измеряя ЧОТ в динамике, можно судить о текущем психоэмоциональном состоянии данного диктора, психотипе его личности и поведении в будущем [3]. Проблема состоит в выборе метода измерения ЧОТ, рассчитанного на условия практического применения систем ААР и предназначенного для учёта действия фонового (внешнего) акустического шума.

Большинство известных методов измерения ЧОТ хорошо зарекомендовало себя при отсутствии внешних помех, но их эффективность значительно снижается, когда отношение сигнал–шум составляет 10 дБ и ниже [7]. Между тем, это самые характерные для систем ААР условия функционирования [8]. Ещё больше вопрос обостряется в условиях малых выборок наблюдений на относительно коротких (десятичные доли секунды) интервалах вокализации речевого сигнала [9]. Поэтому измерение ЧОТ в режиме реального времени является актуальной научно-технической задачей, решению которой и посвящена настоящая статья. При решении поставленной задачи будем использовать математический аппарат статистической теории сигналов [10, 11].

Постановка задачи. Принцип действия современных систем ААР предполагает последовательное деление речевого сигнала $x(t)$ на конечные отрезки или фреймы данных $x_1(t)$, $x_2(t)$, ..., где t – текущее время, длительность T

которых согласована с максимальной длительностью T_0 интервалов вокализации сигнала, причём $T \geq T_0$. В пределах каждого фрейма $x_k(t)$, $k=1, 2, \dots$, ЧОТ диктора можно считать постоянной величиной, равной $F_0(k)$, если не учитывать при этом короткие переходные процессы в его речевом тракте. Таким образом, задача состоит в отслеживании значения ЧОТ в динамике, т. е. от одного речевого фрейма $x_{k-1}(t)$ к другому – $x_k(t)$.

Из теории оптимальных статистических оценок известно, что потенциально достижимая точность частотных измерений на фоне белого гауссовского шума характеризуется дисперсией погрешности [10]:

$$\sigma_0^2 = (T_0^2 q_0^2)^{-1},$$

где q_0^2 – отношение сигнал–шум по мощности на интервале вокализации речевого сигнала длительностью T_0 .

В диапазоне рабочих значений ЧОТ речи мужчин $F_0=100\dots 200$ Гц допустимая погрешность частотного измерителя в среднем квадратичном выражении не должна превышать $\sigma_0=5\dots 10$ Гц. При отношении сигнал–шум $q_0^2=10\dots 20$ дБ длительность отрезков речевого сигнала, пригодных для измерений ЧОТ с указанной точностью, должна отвечать условию

$$T_0 = (\sigma_0^2 q_0^2)^{-1/2} \geq (250)^{-1/2} \approx 60 \text{ мс}. \quad (1)$$

Требованию (1) соответствуют, например, гласные звуки речи длительностью $T_0=70$ мс и более [9]. Следовательно, постановка задачи сводится к последовательному выделению из речевого сигнала $x(t)$ в пределах каждого очередного речевого фрейма $x_k(t)$ вокализованных отрезков максимальной длительности T_0 и измерению в его границах текущего или мгновенного значения ЧОТ контрольного диктора [4, 5].

Ниже рассмотрим решение задачи на основе амплитудно-частотной селекции речевого сигнала из его аддитивной смеси с окружающим (фоновым) шумом.

Синтез алгоритма. Согласно акустической теории речеобразования, описанной в работе [12], сигнал $x(t)$ на интервалах действия гласных фонем достаточно большой длительности T_0 имеет линейчатый частотный спектр. Его минимальная (нижняя) частота F_0 – частота первой гармоники – и определяет искомое текущее значение ЧОТ в границах наблюдаемого речевого фрейма $x_k(t)$. Произведение $M=F_0T_0$ определяет в таком случае количество периодов основного тона в пределах одного вокализованного отрезка речевого сигнала. При $T_0=100$ мс и более для дикторов-мужчин будет существовать не менее $100 \text{ мс} \times 100 \text{ Гц}$ периодов основного тона. Это серьёзный стимул для применения межпериодного накопления речевого сигнала как эффективного средства его частотной селекции. В общем случае оно реализуется по схеме гребенчатого фильтра накопления (ГФН), настроенного на частоту F_0 [13]. В условиях априорной неопределённости теория, изложенная в работе [14], рекомендует многоканальную систему обработки сигнала с использованием набора из N параллельно включённых ГФН. Причём частотные характеристики таких фильтров должны перекрывать весь диапазон возможных значений ЧОТ. При $N \gg 1$ в качестве ГФН применяют рециркулятор [14]. Его динамика описывается рекуррентным уравнением первого порядка

$$y_{i,k}(t) = x_k(t) + by_{i,k}(t - T_i), \quad (2)$$

где $T_i = 1/F_i$ – период накопления сигнала $x(t)$; $b < 1$ – коэффициент усиления сигнала в цепи обратной связи.

Чем ближе значение b к единице, тем более выражен в формуле (2) эффект накопления. Однако на практике следует учитывать конечную длительность речевого сигнала на интервалах вокализации. Поэтому существует оптимальное значение b_0 параметра b из условия максимизации достигаемого выигрыша в отношении сигнал (шум) за счёт накопления сигнала. В первом приближении можно записать

$b_0 \approx 1 - 1,27/M$ [13] или в рассматриваемом случае $b_0 \approx 0,95$. При этом период накопления сигнала в каждом канале соответствует определённому варианту ЧОТ $F_i = 1/T_i$ в пределах диапазона ожидаемых значений $[F_{\min}; F_{\max}]$. В режиме дискретного времени $t=0, \Delta T, 2\Delta T, \dots$, при частоте F дискретизации речевого сигнала будем иметь $T_i = T_{i-1} + \Delta T$, где $\Delta T = 1/F$ – период взятия отсчётов. При равенстве $T_0 = 1/F_{\max} - \Delta T$ получим выражение

$$T_i = i\Delta T, \quad i = \overline{1, N}, \quad (3)$$

где $N = (T_N - T_1)/\Delta T = F(F_{\min}^{-1} - F_{\max}^{-1})$ – число каналов.

Например, при частоте дискретизации $F = 8$ кГц, согласованной с полосой пропускания стандартного телефонного канала связи, в расчёте на мужские голоса имеем $N = 8000(100^{-1} - 200^{-1}) = 8000/200 = 40$. Дополним выражение (2) оценкой средней мощности накопленного сигнала на интервале его вокализации конечной длительности T_0 по формуле эмпирической дисперсии $P(y_{i,k})$ из работы [9]:

$$P(y_{i,k}) = (FT_0)^{-1} \sum_{l=1}^{FT_0} y_{i,k}^2(t + l\Delta T). \quad (4)$$

Далее получим квазиоптимальный алгоритм измерения ЧОТ в пределах текущего речевого фрейма

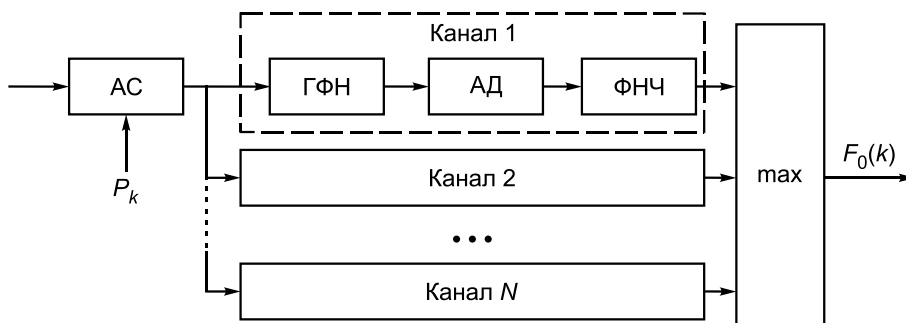
$$x_k(t): \hat{F}_0(k) = 1/T_m, \quad m = \arg \max_{i \leq N} P(y_{i,k}), \quad k = 1, 2, \dots, \quad (5)$$

где $\arg \max$ – значение целочисленного аргумента $i \leq N$, при котором достигается максимум средней мощности накопленного сигнала.

Структурная схема многоканального измерителя ЧОТ представлена на рисунке. Амплитудный селектор вокализованных отрезков речевого сигнала служит для выделения гласных звуков речи, средняя мощность которых превышает среднюю мощность сигнала P_k на интервале в один (текущий) фрейм $x_k(t)$. Согласно этой схеме решение в отношении оптимальной оценки $\hat{F}_0(k)$ ЧОТ принимается по принципу максимума средней мощности $P(y_{i,k})$ сигнала

$y_{i,k}(t)$ на выходе ГФН в составе i -го канала, $i = \overline{1, N}$. Эффект от используемого межпериодного накопления речевого сигнала состоит в существенном увеличении точности измерений ЧОТ в условиях действия случайных помех. При этом вычислительная сложность, а вслед за ней и быстродействие синтезированного алгоритма отвечают требованиям работы измерителя в режиме реального времени.

Анализ эффективности алгоритма. Точность измерения ЧОТ согласно алгоритму (2)–(5) определяется, главным образом, его инструментальной погрешностью. С учётом разрешающей способности ГФН по частоте



Структурная схема многоканального измерителя частоты основного тона:

АС – амплитудный селектор вокализованных отрезков речевого сигнала; ГФН – гребенчатый фильтр накопления; АД – амплитудный (квадратичный) детектор; ФНЧ – фильтр нижних частот

$\Delta f \approx \pi^{-1} F_i (1 - b_0) / (1 + b_0)$ рассчитаем погрешность алгоритма через частотную расстройку двух соседних каналов измерителя ЧОТ:

$$\Delta F_i = 0,5(1/T_i - 1/T_{i+1}) = (0,5 \Delta T) / (T_i T_{i+1}) = 0,5 / (F T_i T_{i+1}).$$

Погрешность алгоритма ограничена сверху предельным значением

$$0,5 / (F T_1^2) = (0,5 F_{\max}^2) / F.$$

Например, в рассматриваемом диапазоне частот можно записать выражение

$$\Delta F_i \leq (0,5 \times 200^2) / 8000 = 2,5 \text{ Гц}.$$

Отметим, что полученный результат на 30–40 % превышает показатели точности современных методов измерения ЧОТ, представленных в работе [8]. Помехоустойчивость предложенного метода можно охарактеризовать его выигрышем по мощности, который достигается за счёт накопления полезного сигнала [14]:

$$W \approx 0,8 M. \quad (6)$$

В рассматриваемой задаче $W \approx 12$ дБ. Данный выигрыш в равной мере относится как к помехам типа белый шум, так и к подобным речевому сигналу помехам с «окрашенными» спектрами мощности [15]. Второй тип помех имеет неотличимую от полезного сигнала внутрипериодную тонкую (формантную) структуру и поэтому представляет неразрешимую на данный момент проблему для большинства известных методов измерения ЧОТ [12]. Однако с точки зрения межпериодной обработки (см. формулу (2)) подобные помехи неотличимы от белого шума, так как не содержат в своём спектре достаточно мощных периодических компонентов в пределах рабочего диапазона частот. Следовательно, вне зависимости от вида действующих помех эффективность предложенного метода в статистическом смысле близка к потенциально достижимой эффективности измерения частоты. В пересчёте к средней квадратичной погрешности измерений через эффективную длительность речевого сигнала, приблизительно равную WT_p , запишем выражение

$$\sigma_0 \approx [q^2 (W T_i)^2]^{-1/2} \leq F_{\max} / (W \sqrt{q^2}).$$

При $q_2 = 20$ дБ получим близкую к потенциально достижимой [16] в рассматриваемых в работе условиях погрешность

$$\sigma_0 \approx 200 / (0,8 \cdot 20 \cdot 10) = 1,25 \text{ Гц}.$$

При уменьшении отношения сигнал–шум до $q^2 = 10$ дБ указанная погрешность возрастает, но лишь до $\sigma_0 = 3,95$ Гц. Это значение меньше 2 % от номинала, соответствующего 200 Гц. И даже при $q^2 = 0$ дБ погрешность остаётся приемлемой для практики: $\sigma_0 = 12,5$ Гц.

Программная реализация. Для подтверждения и развития результатов теоретического исследования авторы настоящей статьи создали и экспериментально исследовали программную реализацию системы ААР – комплекс «Voice Self-Analysis», в котором впервые применили рассмотренный выше метод измерения ЧОТ [17]. Программный комплекс предназначен для автоматической оценки акустического качества речи (АКР) контрольного диктора [3]. При этом показатель АКР вычисляется по формуле [3]:

$$Q(k) = \frac{100}{1 + 10 \delta_F(k)}, \quad (7)$$

где $\delta_F(k) = \begin{cases} \Delta F(k) / \bar{F}_0(k-1), & \Delta F(k) > 0; \\ 0, & \Delta F(k) \leq 0, \end{cases}$ – индикаторная величина; $\Delta F_0(k) = \hat{F}_0(k) - \bar{F}_0(k-1)$ – приращение ЧОТ в динамике; $\bar{F}_0(k-1)$ – выборочное среднее значение ЧОТ к моменту начала очередного речевого фрейма $x_k(t)$.

В процессе дальнейших испытаний длительность фрейма $x_k(t)$ не превышала 2 с. Все вычисления проводились на ноутбуке Asus N61D, 4 Гбайт ОЗУ, ОС Windows 10. Ввод речевого сигнала осуществлялся через встроенный в ноутбук микрофон; частота F его дискретизации была установлена равной 8 кГц.

В идеальном случае, когда ЧОТ стабилизировалась и не увеличивается по ходу речи диктора, его АКР характеризуется максимальным значением показателя $Q(k) = 100$ %. Любая фальшь в речи диктора проявляется в программе в виде снижения показателя из формулы (7). Чем глубже данное снижение, тем фальшь сильнее. В тех случаях, когда выполняется соотношение

$$Q(k) \leq Q_0, \quad (8)$$

где $Q_0 = 20 \dots 30$ % – пороговый уровень, при котором можно говорить о фальши в голосе диктора как признаке его неуверенности в своих словах или о его недостаточно глубоком владении предметом разговора.

В рамках проведённых авторами испытаний пороговый уровень в формуле (8) был установлен равным 20 %, а число каналов обработки речевого сигнала $N = 40$. При этом границы анализируемого диапазона частот $[F_{\min}; F_{\max}]$ по умолчанию составили 100 и 200 Гц соответственно. Отметим, что для них выполнялось соотношение $F_{\max} = 2F_{\min}$, и это имело принципиальное значение. В самом деле, эффект от межпериодного накопления речевого сигнала (2) согласно выражению (6) реализуется не только в основном (рабочем) диапазоне ЧОТ, но и в диапазонах как кратно больших $[2F_{\min}; 2F_{\max}]$, так и кратно меньших $[0,5F_{\min}; 0,5F_{\max}]$ частот повторения. Указанная особенность позволяет использовать разработанный программный комплекс «Voice Self-Analysis» для ААР диктора любого пола.

Результаты экспериментальных исследований. Программой эксперимента было предусмотрено измерение АКР студентов факультета международных отношений, экономики и управления Нижегородского государствен-

Таблица 1

Распределение студентов по успеваемости

Способ группирования	Категория студентов		Всего
	1	2	
Оценка преподавателя	25	11	36
Автоматическая оценка АКР	23	13	36

ного лингвистического университета в процессе экзамена по эконометрике. Цель исследования заключалась в проверке валидности методики (7), (8) в задаче группирования студентов по признаку академической успеваемости. Для этого задействовали две учебные группы второго курса бакалавриата. Объём выборки составил 36 человек. Их успеваемость объективно охарактеризовали постфактум – по итоговым экзаменационным оценкам. В ходе контроля успеваемости студентов на экзамене по эконометрике преподаватели выставили 8,3 % оценок «неудовлетворительно», 22,3 % оценок «удовлетворительно», 58,3 % оценок «хорошо» и 11,1 % оценок «отлично» [18]. В зависимости от оценки преподавателя каждый студент был отнесён авторами статьи к группе с высокой или низкой академической успеваемостью – к категориям 1 и 2 соответственно (табл. 1).

Методика эксперимента предусматривала для каждого студента формирование среднего значения $\bar{Q}(k)$ показателя АКР на интервале в 10 последовательных фреймов речи. При выполнении условия $\bar{Q}(k) > 75\%$ студент характеризовался высоким показателем АКР и относился к категории 1, в противном случае – недостаточно высоким показателем и зачислялся в категорию 2. В условиях небольшой учебной аудитории и окружающего шума отношение сигнал–шум в процессе эксперимента не превышало 10 дБ. Из сравнения оценок, выставленных преподавателями и вычисленных программным путём, можно сделать вывод о существенной степени коррелированности результатов распределения студентов по категориям успеваемости в представлении преподавателя и согласно показателю АКР.

По результатам первого этапа экспериментальных исследований авторы сформировали две связанные между собой выборки бинарных признаков студентов, объёмом 36 отсчётов каждая, т. е. по числу студентов в контрольной группе. Первая из них оттапливалась от показателя академической успеваемости студента (признак *A*), вторая – от его показателя АКР (признак *B*). В итоге каждый студент был представлен в обеих выборках бинарными отсчётами – в зависимости от его принадлежности к одной из двух категорий успеваемости. Результаты анализа сопряжённости признаков *A*, *B* сведены в табл. 2.

По данным табл. 2 с помощью on-line калькулятора по эконометрике вычислили коэффициент $\phi_{AB} \approx 0,8$ контингенции двух признаков в роли характеристики степени их коррелированности в номинальной (относительной) шкале из-

Таблица 2

Анализ сопряжённости показателей успеваемости и АКР

Категории студентов по АКР (признак <i>B</i>)	Категории студентов по оценкам преподавателя (признак <i>A</i>)		Всего студентов
	1	2	
1	22	1	23
2	3	10	13
Всего студентов	25	11	36

мерений. Полученный результат свидетельствует о высокой степени коррелированности признаков *A*, *B* в задаче группирования студентов по категориям успеваемости. Таким образом, измеряя показатель Q_k из формулы (7) по короткому (минута–две) фрагменту речи студентов на экзамене, даже в условиях действия интенсивного фонового шума удалось почти безошибочно спрогнозировать итоговые оценки их успеваемости, выставленные преподавателем.

Также полученное значение коэффициента подтверждает высокую степень валидности рассмотренной методики (7), (8) ААР с применением нового метода измерения ЧОТ в динамике. Обоснованность этого вывода подтверждается и вероятностью совпадения признаков *A*, *B* в пределах контрольной группы студентов:

$$P_{AB} = P(A=1|B=1) + P(A=2|B=2) = (22+10)/36 \approx 0,9.$$

Указанное совпадение – практически достоверное событие. Из работы [19] следует, что полученный результат $P_{AB} \approx 0,9$ с высокой доверительной вероятностью 0,95 гарантирует погрешность оценки вероятности P_{AB} не более $196/\sqrt{360} \approx 10\%$ по выборке объёмом 36.

Заключение. Результаты проведенного в статье исследования имеют не только теоретическое, но и очевидное практическое значение. Между динамикой ЧОТ в речи диктора и его текущим психоэмоциональным состоянием существует взаимозависимость. В нормальном состоянии человек говорит внятно, чётко и в узнаваемом для окружающих ключе. Его ЧОТ при этом стабилизируется на определённом уровне. Однако в результате волнений, стрессов или снижения концентрации внимания состояние диктора объективно ухудшается, и его речь мгновенно утрачивает акустические свойства. Поэтому, отслеживая ЧОТ во времени, можно опосредованно получить часто скрытую от окружающих информацию в отношении текущего психоэмоционального состояния личности. В указанном эффекте и состоит идея предложенного метода измерения ЧОТ. Благодаря своей высокой помехоустойчивости и быстродействию он впервые переводит данную идею в плоскость практического применения. Представленные результаты экспериментальных испытаний в натурных условиях подтверждены авторской программной разработкой комплекса «Voice Self-Analysis», в котором новый метод был впервые реализован и применён в задаче ААР.

Работа выполнена при поддержке Национального исследовательского университета «Высшая школа экономики – Нижний Новгород» (лаборатория алгоритмов и технологий анализа сетевых структур).

Литература

1. **Lech M., He L.** Stress and Emotion Recognition Using Acoustic Speech Analysis / In: Mental Health Informatics. – Springer, Berlin: Heidelberg. 2014. V. 491. P. 163–184.
2. **Dalton J. C., Keegan L. C.** Using Speech Analysis to Unmask Perceptual Bias // Perspectives of the ASHA Special Interest Groups. 2017. V. 2 (SIG 19). P. 9–18.
3. **Савченко В. В.** Метод измерения фонетического качества речи на основе теоретико-информационного подхода // Измерительная техника. 2018. № 1. С. 60–64.
4. **Adalbjornsson S. I., Jakobsson A., Christensen M. G.** Multi-pitch Estimation Exploiting Block Sparsity // Signal Processing. 2015. V. 109. P. 236–247.
5. **Атаянц Б. А., Паршин В. С.** Измерение частоты гармонического сигнала на фоне аддитивного белого шума // Измерительная техника. 2004. № 6. С. 42–45.
6. **Алимурадов А. К.** Алгоритм измерения частоты основного тона речевых сигналов на основе комплементарной множественной декомпозиции на эмпирические моды // Измерительная техника. 2016. № 12. С. 53–57.
7. **Savchenko V. V.** Words Phonetic Decoding Method with the Suppression of Background Noise // Journal of Communications Technology and Electronics. 2017. V. 62. No. 7. P. 788–793.
8. **Hasan M. A.** Pitch Detection Algorithm Based on Windowless Autocorrelation Function and Modified Cepstrum Method in Noisy Environments // International Journal of Computer Science and Network Security. 2017. V. 17. No. 2. P. 106–112.
9. **Savchenko V. V.** The Principle of the Information-Divergence Minimum in the Problem of Spectral Analysis of the Random Time Series Under the Condition of Small Observation Samples // Radiophysics and Quantum Electronics. 2015. V. 58. No. 5. P. 373–379.
10. **Казаринова Ю. М.** Радиотехнические системы. М.: Издательский центр «Академия», 2008.
11. **Savchenko V. V., Savchenko A. V.** Information Theoretic Analysis of Efficiency of the Phonetic Encoding–Decoding Method in Automatic Speech Recognition // Journal of Communications Technology and Electronics. 2016. V. 61. No. 4. P. 430–435.
12. **Фант Г.** Акустическая теория речеобразования. М.: Наука, 1964.
13. **Лезин Ю. С.** Оптимальные фильтры и накопители импульсных сигналов. М.: Советское радио, 1969.
14. **Лезин Ю. С.** Введение в теорию и технику радиотехнических систем. М.: Радио и связь, 1986.
15. **Savchenko V. V.** Enhancement of the Noise Immunity of a Voice-Activated Robotics Control System Based on Phonetic Word Decoding Method // Journal of Communications Technology and Electronics. 2016. V. 61. No. 12. P. 1374–1379.
16. **ГОСТ 12692–67.** Измерители частоты резонансные. Методы и средства поверок.
17. **Программный комплекс Voice Self-Analysis.** [Электрон. ресурс]: <https://sites.google.com/site/frompldcreators/VoiceSelfAnalysisDemo.zip> (дата обращения: 09.09.2018).
18. **Протокол натурных испытаний** [Электрон. ресурс]: https://sites.google.com/site/frompldcreators/VSA_Experimental_results.pdf (дата обращения: 09.09.2018).
19. **Савченко В. В.** Определение объёма контрольной выборки в условиях априорной неопределённости по принципу гарантированного результата // Научные ведомости Белгородского госуниверситета. Серия: Экономика. Информатика. 2015. № 1 (198). С. 74–78.

Дата принятия: 14.11.2018

A method of the speech signal pitch frequency measurement for acoustic speech analysis systems

A. V. Savchenko¹, V. V. Savchenko²

¹National Research University Higher School of Economics, Laboratory of algorithms and technologies for network analysis, Nizhny Novgorod, Russia, e-mail: avsavchenko@hse.ru

²Nizhny Novgorod State Linguistic University, Nizhny Novgorod, Russia, e-mail: vvsavchenko@yandex.ru

In this paper we propose a novel method for measuring the pitch frequency of the increased noise immunity. The problem of protection against intensive background noise is solved in it by the frequency selection of vocalization pieces of a speech signal according to the scheme of the combfilter of inter-periodical accumulation. The efficiency of a new method is theoretically and experimentally examined using the specially developed multi-channel frequency measurement device for the acoustic speech analysis. It is shown that a measurement error in relative terms does not exceed 2 % for the 10 dB signal-to-noise ratio.

Key words: acoustic measurements, speech acoustics, speech signal, fundamental tone, acoustic speech analysis, acoustic noise, noise stability.