

Методы и средства обработки сигналов и изображений

УДК 534.782.001:621.39

Л.Б. Дронжевская, В.А. Климков, А.Н. Продеус, канд. техн. наук, Д.А. Шагитова

Формантный и формантно-модуляционный методы оценивания разборчивости речи. Часть 2. Точность и скорость измерений

Сопоставлены, по точности и скорости измерений, формантный и формантно-модуляционный методы измерений разборчивости речи.

In this paper formant and formant-modulation methods are confronted on speed and exactness.

Ключевые слова: разборчивость речи, оценивание, формантный метод, формантно-модуляционный метод, точность, скорость измерений.

Введение

В работе [1] изложена программа исследований, направленных на сопоставление, по точности и скорости измерений, нового формантно-модуляционного (ФМ) метода оценивания разборчивости речи [2] с «классическим» формантным методом [3–5]. Цель данной работы состоит в реализации упомянутой программы.

Перечень сопоставляемых методов и их вариантов представлен в табл. 1. При этом в качестве эталонных значений разборчивости речи приняты результаты, полученные с помощью артикуляционного метода [4].

1. Описание сопоставляемых инструментальных методов

Сущность формантного и ФМ методов состоит в следующем.

При акустической экспертизе помещений, характеризуемых импульсной характеристикой (ИХ) $h(v)$, в определенной точке (местоположение диктора) излучают тестовый сигнал $x(t)$, а в иной точке (местоположение аудитора) принимают сигнал $y(t) = \int_{-\infty}^{\infty} h(v)x(t-v)dv + n(t)$, где $n(t)$ – шумовая помеха.

При использовании ФМ метода [1, 2], как и в формантном методе [3–5], вычисляют артикуляционную разборчивость A , представляющую собой вероятность правильного понимания формант:

$$A = \sum_{k=1}^K p_k \cdot P_k(E_k), \quad (1)$$

где p_k – вероятность пребывания формант в k -й полосе частот; $P_k(E_k)$ – коэффициент восприятия речи; E_k – эффективный уровень ощущения речевого сигнала в k -й полосе частот Δf_k . Заметим, что на практике вместо артикуляционной разборчивости A оперируют слоговой разборчивостью $D = f_1(A)$ или словесной разборчивостью $W = f_2(A)$, представляющими собой вероятности правильного понимания слогов или слов, соответственно. Функциональные зависимости $D = f_1(A)$ и $W = f_2(A)$, найденные экспериментально, представлены в работах [3, 4].

Отличие методов заключается в том, что если в формантном методе используют тестовый сигнал $x(t)$ в виде стационарного случайного процесса (ССП), спектр мощности которого подобен долговременному спектру речи, а затем оценивают E_k в соответствии с соотношением:

$$E_k \approx SNR_k = 10 \lg \frac{D_{sk}}{D_{nk}}, \quad (2)$$

где SNR_k – парциальное отношение сигнал-шум; D_{sk} и D_{nk} – соответственно дисперсии стационарных сигнала и шума в k -й полосе частот; то в ФМ методе для оценивания E_k предложено использовать иной способ, применяющийся в модуляционном методе и позволяющий корректно учесть влияние реверберационной помехи [4, 5].

Таблица 1

Субъективный метод	Инструментальные (объективные) методы				
	Формантный			ФМ	
Артикуляционный	Покровского	Сапожкова	Модернизированный	Полный	Ускоренный

Это косвенный способ измерения отношения сигнал-шум, согласно которому используют тестовый сигнал $x(t)$ в виде нестационарного случайного процесса (НСП), дисперсия которого $D_x(t)$ модулирована по гармоническому закону с частотой модуляции F :

$$x(t) = \xi(t)\sqrt{f(t)}, \quad f(t) = 1 + \cos 2\pi Ft; \quad (3)$$

$$D_x(t) = D_\xi(1 + \cos 2\pi Ft), \quad (4)$$

где $\xi(t)$ – ССП с дисперсией D_ξ и спектром мощности, повторяющим форму долговременного спектра речи.

Дисперсия $D_y(t)$ НСП $y(t)$ будет также модулирована по гармоническому закону с той же частотой F , однако, вследствие влияния реверберационной и шумовой помех, коэффициент модуляции дисперсии $D_y(t)$ оказывается меньше единицы. В полной версии ФМ метода при измерениях используют набор из 14 частот модуляции ($F_i = 0,63; 0,8; 1; 1,25; 1,6; 2; 2,5; 3,15; 4; 5; 6,3; 8; 10; 12,5$ Гц) и измеряют «эквивалентное» отношение сигнал-шум:

$$E_k = \frac{1}{14} \sum_{i=1}^{14} SNR_{ik}; \quad (5)$$

$$SNR_{ik} = 10 \lg \frac{m_k(F_i)}{1 - m_k(F_i)}, \quad (6)$$

где $m_k(F_i)$ – коэффициент модуляции дисперсии $D_y(t)$ сигнала $y(t)$, наблюдаемого в k -й полосе частот [4, 5]. Коэффициент модуляции $m_k(F_i)$ оценивают с использованием преобразования Фурье [1]:

$$\tilde{m}_k(F_i) = \frac{2|A_{ik}(F_i)|}{|A_{ik}(0)|}; \quad (7)$$

$$A_{ik}(f) = \frac{1}{T} \int_0^T y_{ik}^2(t) e^{-j2\pi ft} dt,$$

где $|\cdot|$ – символ модуля; T – длительность реализации процесса

$$y_{ik}(t) = \int_{-\infty}^{\infty} h(v)x_{ik}(t-v)dv + n(t),$$

где $x_{ik}(t) = \xi_k(t)\sqrt{1 + \cos 2\pi F_i t}$ – модулированный полосовой белый шум в k -й полосе частот; $\xi_k(t)$ – результат фильтрации ССП $\xi(t)$ k -м полосовым фильтром.

В работе [8] предложена ускоренная версия ФМ метода, в которой вместо 14 измерительных

опытов с тестовыми сигналами (3) используют единственный тестовый сигнал:

$$x(t) = \xi(t)\sqrt{f_5(t)}, \quad f_5(t) = 1 + 0,32 \sum_{i=1}^5 \sin 2\pi F_i t, \\ F_i = iF, \quad F = 0,7 \text{ Гц.}$$

При этом коэффициент модуляции оценивают в соответствии с соотношением

$$\tilde{m}_k(F_i) = \frac{5|A_{ik}(F_i)|}{|A_{ik}(0)|}, \quad (8)$$

а вместо соотношения (6) используют его аналог:

$$E_k = \frac{1}{5} \sum_{i=1}^5 SNR_{ik}. \quad (9)$$

2. Модернизация формантного метода

В настоящее время российские исследователи, использующие формантный метод с коэффициентами восприятия, предложенными Н.Б. Покровским, проводят измерения либо в пяти октавных полосах частот с центральными частотами 250, 500, 1000, 2000, 4000 Гц, либо в шести октавных полосах, добавляя полосу с центральной частотой 8 кГц [3].

Между тем, в модуляционном методе принято использовать семь октавных полос частот, добавляя, к перечисленным выше, полосу с центральной частотой 125 Гц. Поэтому в рамках этой работы, при сопоставлении формантного и ФМ методов было решено рассматривать семь октавных полос частот.

Как следует из рис. 1, на котором представлены результаты измерений словесной разборчивости формантным методом с использованием коэффициента восприятия Н.Б. Покровского, расширение анализируемого частотного диапазона с пяти до семи октавных полос частот приводит к заметному увеличению оценок разборчивости.

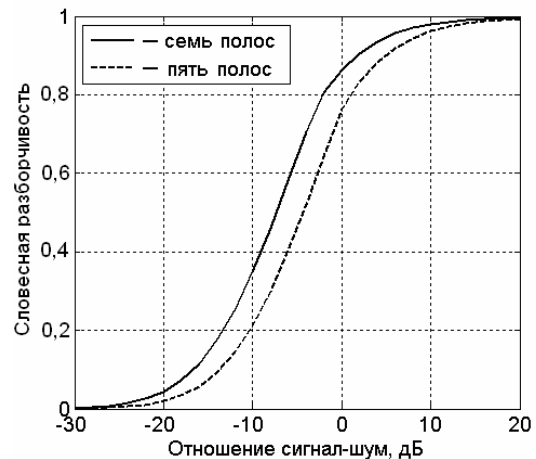


Рис. 1. Результаты измерений словесной разборчивости формантным методом в пяти и семи октавных полосах частот

В работе [3] было предложено модернизировать формантный метод, используя вместо «старых» коэффициентов восприятия Н.Б. Покровского [4] «новые», более точные, коэффициенты восприятия. Поскольку эти коэффициенты существенно отличаются, то и различными получаются оценки разборчивости речи (рис. 2). Графики оценок словесной разборчивости рис. 2, б в дальнейшем будем считать «образцовыми» для модернизированного таким образом формантного метода, сравнивая с ними результаты нового ФМ метода. Интересно, что использование коэффициента восприятия М.А. Сапожкова [5] приводит к результатам, качественно подобным результатам, представленным на рис. 2, б [3], хотя количественно, разумеется, результаты отличаются.

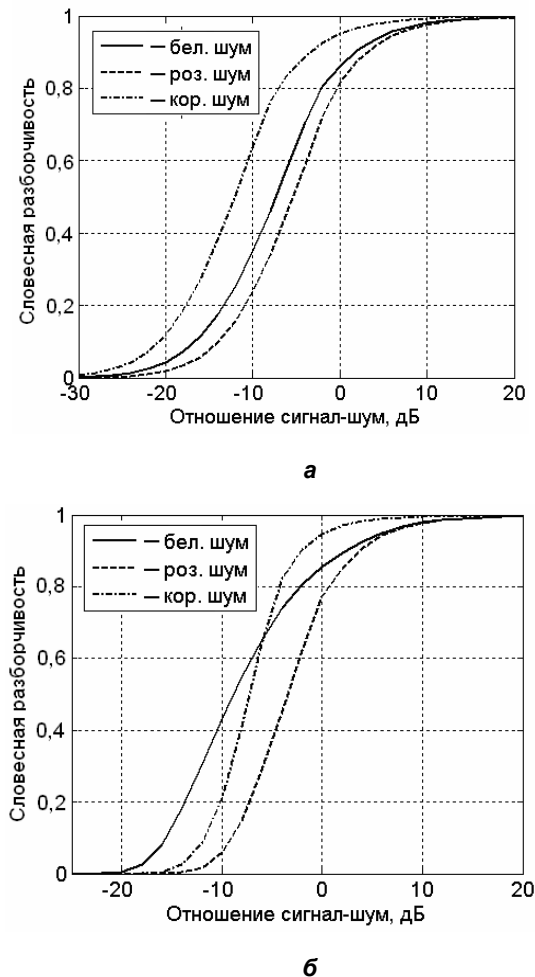


Рис. 2. Результаты измерений словесной разборчивости для «старых» (а) и новых (б) коэффициентов восприятия

3. Точность измерений отношения сигнал-шум в формантном методе

Оценку отношения сигнал-шум, вычисляемую в формантном методе в соответствии с формулой (2), можно считать несмещенной, по-

скольку спектры сигнала и помехи являются практически равномерными в полосах частот Δf_k .

Что касается дисперсии такой оценки отношения сигнал-шум, можно показать (см. Приложение А), что в случае статистически независимых оценок \tilde{D}_{sk} и \tilde{D}_{nk} , а также при выполнении условия $\Delta f T \geq 10^4$ справедливо достаточно точное для инженерных приложений соотношение

$$D[\tilde{S}\tilde{N}R_k] \approx \frac{38}{\Delta f_k T}. \quad (10)$$

Например, для $T = 16$ с погрешность соотношения (10) не превышает 0,1 дБ для всех семи октавных полос частот.

4. Точность измерений отношения сигнал-шум в формантно-модуляционном методе

Предположим, что распределение вероятностей формант по частоте p_k и коэффициент восприятия речи $P_k(E_k)$ известны с высокой точностью. Тогда из формул (5)–(7) следует, что смещение и дисперсия оценки (5) отношения сигнал-шум полностью определяются статистическими свойствами величин $A_{jk}(0)$ и $A_{jk}(F_j)$. В общем случае произвольного соотношения шумовой и реверберационной помех аналитические соотношения для смещения и дисперсии оценки (5) оказываются весьма громоздкими и неудобными для анализа. Поэтому ограничимся рассмотрением двух ситуаций: а) преобладание шумовой помехи; б) преобладание реверберационной помехи.

4.1. Преобладание шумовой помехи

Проведем статистический анализ пары случайных величин $A_{jk}(0)$ и $A_{jk}(F_j)$ в предположении, что шумовая помеха превалирует над реверберационной. В этом случае модель сигнала в k -м частотном канале можно представить так:

$$y_{jk}(t) = \xi_k(t) \sqrt{1 + \cos(2\pi F_j t + \varphi_k)} + n_k(t), \quad (11)$$

где $\xi_k(t)$ и $n_k(t)$ – нормально распределенные полосовые (в k -й полосе частот Δf_k) белые и статистически независимые ССП с нулевым математическим ожиданием; φ_k – неизвестная начальная фаза.

Найдем математическое ожидание (МО) величин $A_{jk}(0)$ и $A_{jk}(F_j)$. Опуская, для упрощения обозначений, индексы в соотношениях (7) и (11), получаем

$$\begin{aligned} \langle A(f) \rangle &= \frac{1}{T} \int_0^T \langle y^2(t) \rangle e^{-j2\pi ft} dt = \\ &= (D_\xi + D_n) e^{-j\pi f T} \text{Sa}(\pi f T) + \\ &+ \frac{D_\xi}{2} e^{j\varphi} e^{-j\pi(f-F_i)T} \text{Sa}[\pi(f-F_i)T] + \\ &+ \frac{D_\xi}{2} e^{-j\varphi} e^{-j\pi(f+F_i)T} \text{Sa}[\pi(f+F_i)T], \end{aligned} \quad (12)$$

где $\langle \cdot \rangle$ – символ МО; $\text{Sa}(x) = \sin(x)/x$. При выполнении условия $T = r/F_i$, где r – произвольное целое положительное число, из формулы (12) следует

$$\langle A(0) \rangle = D_\xi + D_n; \quad (13)$$

$$\langle |A(F_i)| \rangle = \langle |A(F_i)| \rangle = \frac{D_\xi}{2}. \quad (14)$$

Как видим, условие $T = r/F_i$ является условием получения несмещенной оценки отношения сигнал-шум (5).

Чтобы найти дисперсию величин $A_{ik}(0)$ и $A_{ik}(F_i)$, предположим, что коэффициенты корреляции $R_\xi(\tau)$ и $R_n(\tau)$ ССП $\xi(t)$ и $n(t)$, соответственно, определяются соотношением [10]

$$R_\xi(\tau) = R_n(\tau) = \text{Sa}(\pi\Delta f \tau) \cos 2\pi f_0 \tau, \quad (15)$$

где Δf – ширина полосы частот ССП $\xi(t)$ и $n(t)$; f_0 – центральная частота этой полосы. После весьма громоздких выкладок, в предположении $T\Delta f \gg 1$, можно получить следующее соотношение:

$$D[|A(0)|] \approx D[|A(F_i)|] \approx \frac{D_n^2}{\Delta f T} \left(\frac{3q^2}{2} + 4q + 1 \right), \quad (16)$$

где $q = D_\xi/D_n$.

Из соотношения (16), в конечном счете, следует вывод о состоятельности оценки (5), сформированной ФМ методом.

Исследования взаимной коррелированности величин $A_{ik}(0)$ и $A_{ik}(F_i)$ также сопряжены с весьма громоздкими выкладками, поэтому приведем лишь окончательные результаты.

Для $q \ll 1$, при выполнении условий $T\Delta f \gg 1$ и $T = r/F_i$, величины $A_{ik}(0)$ и $A_{ik}(F_i)$ можно считать статистически независимыми, поскольку справедливо приближенное равенство

$$K\{A(0)A^*(F_i)\} \approx K\{A(F_i)A^*(0)\} \approx 0,$$

где $*$ – символ комплексного сопряжения.

Для $q \gg 1$ величины $A_{ik}(0)$ и $A_{ik}(F_i)$ являются статистически зависимыми:

$$K\{A(0)A^*(F_i)\} \approx K\{A(F_i)A^*(0)\} \approx \frac{D_\xi^2}{\Delta f T};$$

коэффициент их взаимной корреляции близок величине 0,66:

$$R\{A(0)A^*(F_i)\} = \frac{K\{A(0)A^*(F_i)\}}{\sigma[A(0)]\sigma[A^*(F_i)]} \approx \frac{2}{3}.$$

Несмотря на коррелированность величин $A_{ik}(0)$ и $A_{ik}(F_i)$, вывод о состоятельности оценки (5), сформированной ФМ методом, остается в силе.

Из-за слабой, вообще говоря, коррелированности величин $A_{ik}(0)$ и $A_{ik}(F_i)$ при выполнении условия $T = r/F_i$ из соотношений (7), (13), (14) следует, что

$$\langle \tilde{m}_k(F_i) \rangle \approx \frac{D_\xi}{D_\xi + D_n} = \frac{q}{1+q} = m, \quad (17)$$

т. е. оценка коэффициента модуляции является несмещенной.

Математическое ожидание квадрата оценки коэффициента модуляции приблизительно равно:

$$\begin{aligned} \langle \tilde{m}_k^2(F_i) \rangle &\approx 4 \frac{\langle |A(F_i)|^2 \rangle}{\langle |A(0)|^2 \rangle} = \\ &= 4 \frac{D[A(F_i)] + \langle |A(F_i)| \rangle^2}{D[A(0)] + \langle |A(0)| \rangle^2}. \end{aligned} \quad (18)$$

Из формул (17) и (18), с учетом соотношений (13)–(16), получаем приближенное выражение для дисперсии оценки коэффициента модуляции:

$$D[\tilde{m}] \approx 4 \frac{\frac{1}{\Delta f T} \left(\frac{3q^2}{2} + 4q + 1 \right) + \frac{q^2}{4}}{\frac{1}{\Delta f T} \left(\frac{3q^2}{2} + 4q + 1 \right) + (1+q)^2} - \frac{q^2}{(1+q)^2},$$

откуда следует

$$D[\tilde{m}_k] \leq \frac{6}{\Delta f_k T}. \quad (19)$$

Далее используем «локальную» линейную аппроксимацию соотношения (6), сущность которой заключается в следующем. Для $\Delta f_k T \gg 1$ дисперсия $D[\tilde{m}_k]$ мала, и можно приближенно

считать справедливыми следующие соотношения:

$$D[\tilde{SNR}_k] \approx \left(\frac{df(m)}{dm} \right)^2 D[\tilde{m}_k];$$

$$f(m) = SNR = 10 \lg \frac{m}{1-m};$$

$$\frac{df(m)}{dm} = \frac{4,34}{m(1-m)} = \frac{4,34(1+10^{-0,1SNR})^2}{10^{-0,1SNR}};$$

$$D[\tilde{SNR}_k] \leq \frac{113(1+10^{-0,1SNR})^4}{10^{-0,2SNR} \Delta f_k T}. \quad (20)$$

График функции

$$c_1(SNR) = \left(\frac{df(m)}{dm} \right)^2 = \frac{18,8(1+10^{-0,1SNR})^4}{10^{-0,2SNR}}$$

показан на рис. 3.

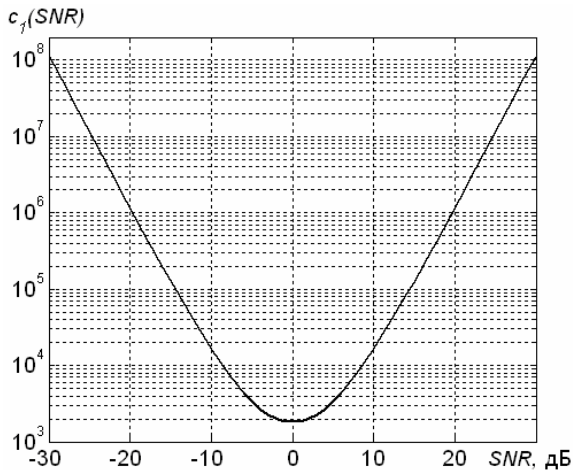


Рис. 3. График функции $c_1(SNR)$

Результаты компьютерного моделирования подтверждают справедливость неравенства (20) в интересном для инженерных приложений диапазоне значений параметра $|SNR| \leq 20$ дБ. Вместе с тем, компьютерное моделирование показало, что верхняя граница дисперсии, определяемая соотношением (20), завышена примерно в два раза, что можно объяснить приближенным характером аналитических выкладок. Компьютерное моделирование также показало, что систематическая составляющая погрешности измерений не превышает 1 дБ для $|SNR| \leq 10$ дБ и возрастает до 4 дБ с приближением $|SNR|$ к ± 20 дБ для полос частот с $f_0 \leq 2$ кГц. Обе составляющие погрешности

измерений можно понизить, увеличивая время анализа T , однако ясно, что при решении практических задач существуют ограничения на увеличение T .

Поскольку оценки \tilde{SNR}_{ik} статистически независимы, в полной версии ФМ метода верхняя граница дисперсии (20) снижается в 14 раз (см. соотношение (5)). Для $SNR = 0$ дисперсия $D[\tilde{SNR}_k]$ ФМ метода минимальна (рис. 3), и из неравенства (20) получаем $D[\tilde{SNR}_k] \leq \frac{130}{\Delta f_k T}$.

Учитывая, что при вычислениях в соответствии с соотношением (20) получаем примерно вдвое завышенные результаты, после коррекции получаем $D[\tilde{SNR}_k] \leq \frac{65}{\Delta f_k T}$, что все равно почти вдвое больше дисперсии (10) формантного метода.

Таким образом, используемый в полной версии ФМ метода косвенный способ измерений отношения сигнал-шум (в ситуации преобладания шумовой помехи) обеспечивает минимальную погрешность измерений при $SNR = 0$, однако, даже в этом случае погрешность измерений примерно в 1,5...2 раза выше, чем для формантного метода. С отклонением от значения $SNR = 0$ погрешность измерений растет.

Анализируя соотношения (5), (7), (8), (9), нетрудно прийти к выводу, что ускоренная версия ФМ метода приводит к еще более высоким погрешностям измерений отношения сигнал-шум, проигрывая полной версии ФМ метода по точности измерений примерно в 4 раза, а формантному методу – в 6...8 раз.

Поскольку для вычислений артикуляционной разборчивости речи используется одно и то же соотношение (1), следует ожидать, что в случае преобладания шумовой помехи наиболее точным окажется формантный метод оценивания, менее точной является полная версия ФМ метода, а ускоренная версия ФМ метода будет наименее точной.

Эти предположения подтверждены результатами компьютерного моделирования процедур измерения словесной разборчивости формантным методом (рис. 4), а также полной (рис. 5) и ускоренной (рис. 6) версиями ФМ метода. Что касается времени измерений, для полной версии ФМ метода оно в 14 раз выше, чем для формантного метода. Время измерений для ускоренной версии ФМ метода и для формантного метода практически одинаково.

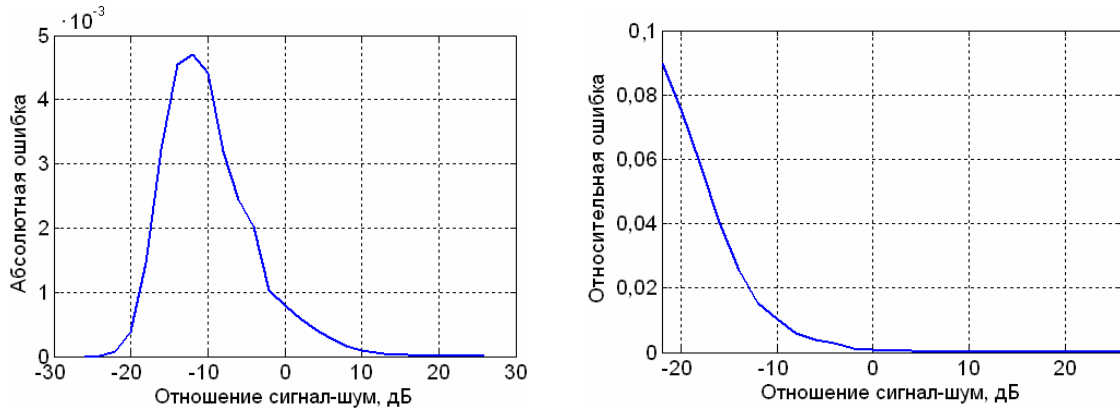


Рис. 4. Ошибка измерений разборчивости формантным методом

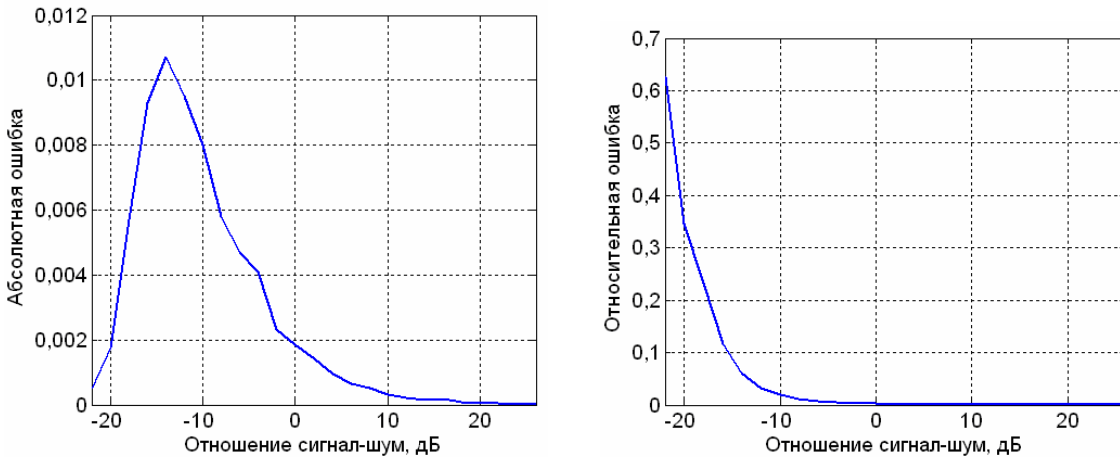


Рис. 5. Ошибка измерений разборчивости ФМ методом (полная версия)

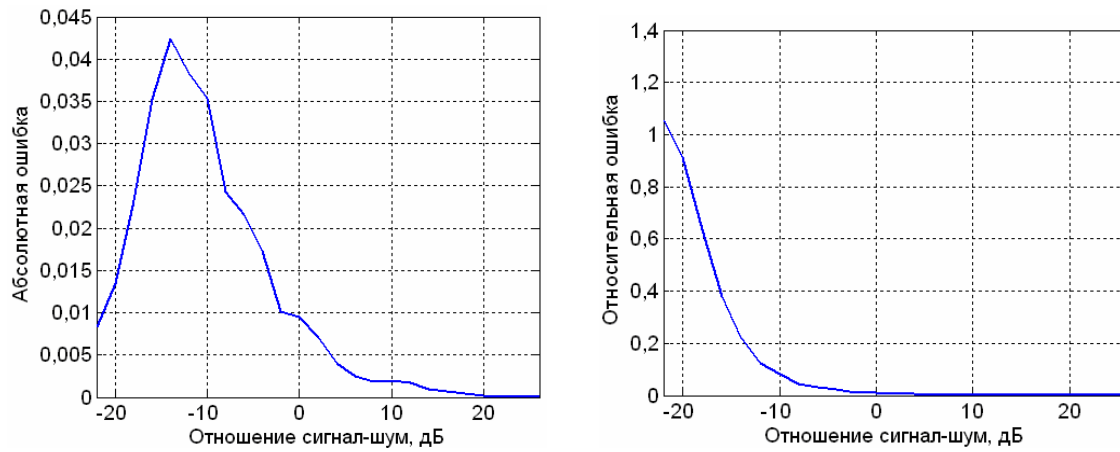


Рис. 6. Ошибка измерений разборчивости ФМ методом (ускоренная версия)

4.2. Преобладание реверберационной помехи

В этом случае модель сигнала в k -м частотном канале удобно представить в виде

$$y_{ik}(t) = \int_{-\infty}^{\infty} h_{ek}(v)x_i(t-v)dv, \tag{21}$$

$$h_{ek}(v) = \int_{-\infty}^{\infty} h(v)h_{k0}(z-v)dv,$$

где $x_i(t) = \xi(t)\sqrt{f_i(t)}$; $h_{k0}(v)$ – ИХ k -го полосового фильтра; $\xi(t)$ – ограниченный по полосе частот ($0 \leq f \leq \Delta f$) белый шум с дисперсией D_ξ .

Тогда, при выполнении условия $T = r/F_i$, можно получить

$$\langle A_k(F_i) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v)e^{-j2\pi F_i v} dv;$$

$$\langle A_k(0) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v) dv.$$

После громоздких выкладок также можно получить выражение для верхней границы дисперсии:

$$D\{A_k(F_i)\} \approx D\{A_k(0)\} \leq \frac{3}{T} \left(\frac{D_\xi}{2\Delta f} \right)^2 \int_{-\infty}^{\infty} B_{h_{ek}}^2(\tau) d\tau;$$

$$B_{h_{ek}}(\tau) = \int_{-\infty}^{\infty} h_{ek}(t) h_{ek}(t+\tau).$$

В предположении слабой коррелированности величин $|A(0)|$ и $|A(F_i)|$, математическое ожидание и дисперсия оценки коэффициента модуляции определяются следующими соотношениями:

$$\langle \tilde{m}_{ki} \rangle \approx \frac{2 \int_0^{\infty} h_{ek}^2(t) e^{-j2\pi F_i t} dt}{\int_0^{\infty} h_{ek}^2(t) dt}; \quad (22)$$

$$D[\tilde{m}_k(F_i)] < \frac{24}{\Delta f_k T}. \quad (23)$$

Соотношение (22) интересно тем, что представляет еще один способ измерений разборчивости речи (предложенный Шредером [7]), сводящийся к измерениям импульсной характеристики помещения.

Сопоставляя соотношения (23) и (19) видим, что в случае преобладания реверберационной помехи выражение для верхней границы дисперсии \tilde{m}_{ki} в четыре раза превышает таковое для случая преобладания шумовой помехи. Причины такого несовпадения нуждаются в дальнейших исследованиях. Одной из возможных причин является приближенный характер аналитических выкладок. Другой причиной может быть различие статистических свойств моделей, описываемых соотношениями (11) и (21).

5. Сопоставление результатов измерений разборчивости речи форматно-модуляционным и артикуляционным методами

Очевидно, в случае преобладания реверберационной помехи, а также в случае сопоставимого влияния реверберационной и шумовой помех, результаты измерений разборчивости речи ФМ методом не могут быть сопоставлены с таковыми для формантного метода, поскольку он изначально предназначен для измерений разборчивости речи в условиях преобладающего влияния шумовой помехи. Поэтому было решено со-

поставить результаты измерений словесной разборчивости речи ФМ методом с таковыми для метода артикуляционных испытаний [4].

Результаты измерений (методом компьютерного моделирования) словесной разборчивости речи ФМ методом в различных помеховых условиях представлены на рис. 7 сплошной (белый шум), штриховой (розовый шум) и штрихпунктирной (коричневый шум) линиями. Результаты артикуляционных испытаний в условиях преобладающего влияния реверберационной помехи представлены звездочками на рис. 7, их среднее значение представлено квадратом. Аналогично результаты артикуляционных испытаний в условиях сопоставимого влияния шумовой и реверберационной помех представлены кружками, их среднее значение представлено треугольником.

Очевидным является достаточно хорошее согласование результатов компьютерного моделирования и артикуляционных испытаний. Имеющиеся расхождения результатов сравнительно невелики для эффективных отношений сигнал-шум близких нулю. В остальных случаях наблюдается некоторое рассогласование, причины которого должны быть изучены в дальнейшем.

Выводы

Проведены аналитические и экспериментальные исследования, позволившие сопоставить форматный и формантно-модуляционный методы измерений разборчивости речи по точности и скорости измерений.

Показано, что в случае преобладания шумовой помехи наиболее точным и быстрым является формантный метод. Полная версия ФМ метода проигрывает ему по точности в 1,5...2 раза, а по времени измерений – в 14 раз. Ускоренная версия ФМ метода требует такого же времени измерений, что и формантный метод, проигрывая по точности в 6...8 раз.

При сопоставимом влиянии шумовой и реверберационной помех вместо формантного метода целесообразно применять ФМ метод, позволяющий корректно учитывать влияние реверберационной помехи на разборчивость речи.

В случае преобладания реверберационной помехи весьма удобен метод измерений разборчивости речи, сводящийся к измерениям импульсной характеристики помещения. Точность и скорость измерений разборчивости речи с использованием такого метода пока исследованы недостаточно.

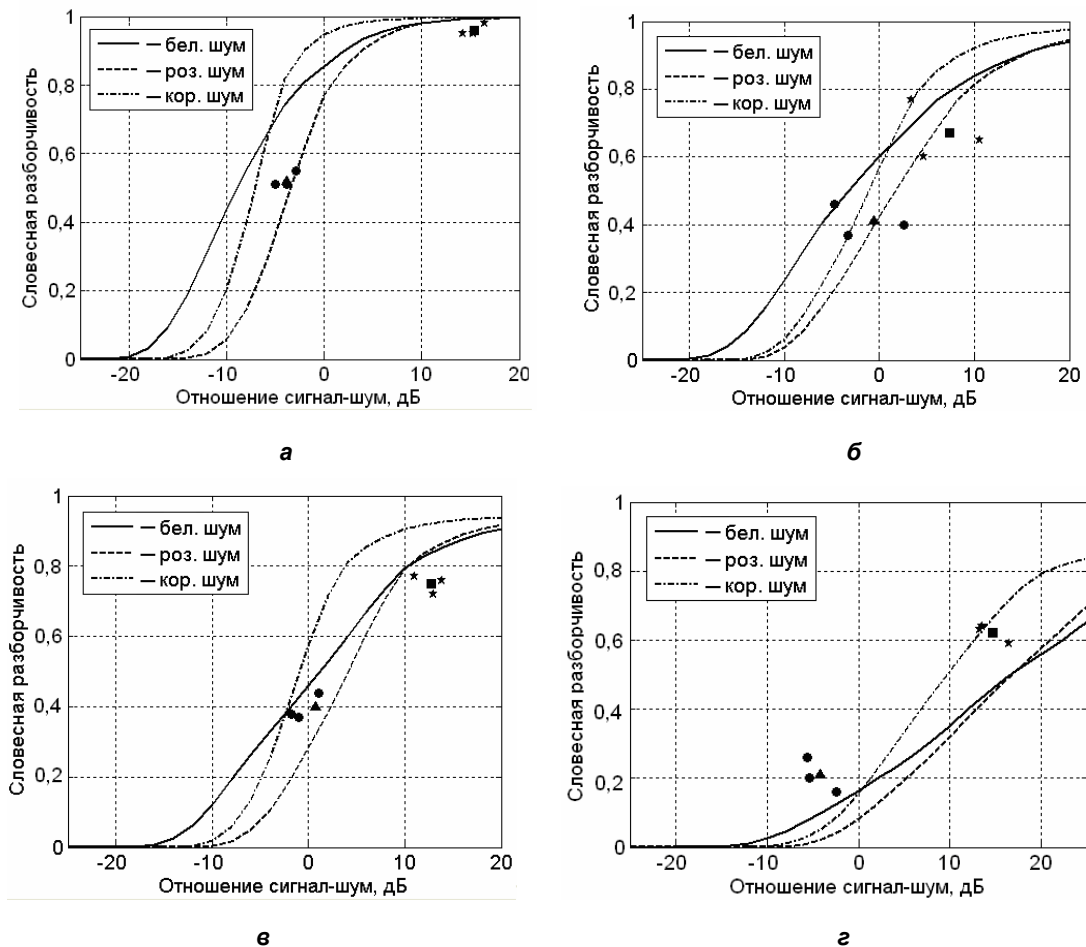


Рис. 7. Результаты измерений словесной разборчивости ФМ методом и методом артикуляционных испытаний: а – $T_{60} = 0,12$ с; б – $T_{60} = 0,6$ с; в – $T_{60} = 1$ с; г – $T_{60} = 2,7$ с

Литература

1. Продеус А.Н., Дронжевская Л.Б., Климов В.А., Шагитова Д.А. Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 1. Унификация алгоритмов // Электроника и связь. – 2010. – № 6(59), ч. 2. – С.117–124.
2. Продеус А.Н. О некоторых особенностях развития объективных методов измерений разборчивости речи // Электроника и связь. Тематический выпуск «Электроника и нанотехнологии». – 2010. – № 2(55). – С. 217–223.
3. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации: Монография. – К.: Имэкс-ЛТД, 2008. – 420 с.
4. Покровский Н.Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – 390 с.
5. Сапожков М.А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. – 472 с.
6. Продеус А.Н., Дронжевская Л.Б., Климов В.А., Шагитова Д.А. Моделирование алгоритмов формантно-модуляционного метода оценивания разборчивости речи // Электроника и связь. Тематический выпуск «Электроника и нанотехнологии». – 2011. – № 2(61). – С. 79–85.
7. Prodeus A. Formant-Modulation Method of Speech Intelligibility Evaluation: Measuring and Exactness // Proceedings of the VII International Conference MEMSTECH 2011. – Lviv, Polyana, 2011. – P. 54–60.
8. Prodeus A. Rapid Version of Formant-Modulation Method of Speech Intelligibility Estimation // Proceedings of the VII International Conference MEMSTECH 2011. – Lviv, Polyana, 2011. – P. 61–63.
9. Двайт Г.Б. Таблицы интегралов и другие математические формулы. – М.: Наука, 1973. – 228 с.
10. Дидковский В.С., Дидковська М.В., Продеус А.М. Комп'ютерна обробка акустичних сигналів. – К.: Імекс-ЛТД, 2010. – 420 с.

Дисперсия оценки отношения сигнал-шум

Оценка отношения сигнал-шум в k -м частотном канале:

$$S\tilde{N}R_k = 10 \lg \frac{\tilde{D}_{sk}}{\tilde{D}_{nk}} = 4,34 \ln \frac{\tilde{D}_{sk}}{\tilde{D}_{nk}} = 4,34 (\ln(\tilde{D}_{sk}) - \ln(\tilde{D}_{nk})). \quad (\text{П.1})$$

Дисперсия оценки (П.1):

$$D[S\tilde{N}R_k] = 18,84 \cdot D(\ln(\tilde{D}_{sk}) - \ln(\tilde{D}_{nk})) = 18,84 [D(\ln(\tilde{D}_{sk})) + D(\ln(\tilde{D}_{nk}))]. \quad (\text{П.2})$$

Рассмотрим $D(\ln(\tilde{D}_{sk}))$:

$$D(\ln(\tilde{D}_{sk})) = \int (\ln(D) - \langle \ln(\tilde{D}_{sk}) \rangle)^2 W(D) dD, \quad (\text{П.3})$$

где $W(D)$ – плотность распределения оценки \tilde{D}_{sk} ; $\langle \ln(\tilde{D}_{sk}) \rangle = \int \ln(D) W(D) dD$.

Если оценка \tilde{D}_{sk} достаточно точная, тогда

$$W(D) \approx \delta(D - D_{sk}),$$

где D_{sk} – дисперсия сигнала в k -й полосе частот, и

$$\langle \ln(\tilde{D}_{sk}) \rangle = \int \ln(D) W(D) dD \approx \ln(D_{sk}). \quad (\text{П.4})$$

Подставляя (П.4) в (П.3), получим

$$D(\ln(\tilde{D}_{sk})) = \int (\ln(D) - \langle \ln(\tilde{D}_{sk}) \rangle)^2 W(D) dD \approx \int (\ln(D) - \ln(D_{sk}))^2 W(D) dD = \int \ln^2\left(\frac{D}{D_{sk}}\right) W(D) dD. \quad (\text{П.5})$$

Применяя к соотношению (П.5) линейную аппроксимацию [9]:

$$\ln x \approx (x - 1), \quad 0 < x \leq 2,$$

получим

$$D(\ln(\tilde{D}_{sk})) \approx \int \ln^2\left(\frac{D}{D_{sk}}\right) W(D) dD \approx \int \left(\frac{D}{D_{sk}} - 1\right)^2 W(D) dD = \frac{1}{D_{sk}^2} \int (D - D_{sk})^2 W(D) dD = \frac{D[\tilde{D}_{sk}]}{D_{sk}^2}.$$

Учитывая, что для полосового белого шума справедливо соотношение [10]

$$D[\tilde{D}_{sk}] \approx \frac{D_{sk}^2}{\Delta f_k T},$$

получаем

$$D(\ln(\tilde{D}_{sk})) \approx \frac{1}{\Delta f_k T}. \quad (\text{П.6})$$

Аналогично для второго слагаемого в квадратных скобках соотношения (П.2) получаем

$$D(\ln(\tilde{D}_{nk})) \approx \frac{1}{\Delta f_k T}. \quad (\text{П.7})$$

Подставляя (П.8) и (П.9) в (П.2), окончательно получаем

$$D[S\tilde{N}R_k] = 18,84 [D(\ln(\tilde{D}_{sk})) + D(\ln(\tilde{D}_{nk}))] \approx \frac{38}{\Delta f T}.$$