

## ФОРМАНТНО-МОДУЛЯЦИОННЫЙ МЕТОД ОЦЕНКИ РАЗБОРЧИВОСТИ РЕЧИ: ТОЧНОСТЬ И СКОРОСТЬ ИЗМЕРЕНИЙ

А. Н. ПРОДЕУС

*Национальный технический университет Украины “КПИ”, Киев*

Произведено сопоставление, по точности и скорости измерений, формантного и формантно-модуляционного методов измерений разборчивости речи. Полученные результаты позволяют обоснованно выбирать метод измерений, учитывая условия измерений.

### ВВЕДЕНИЕ

Среди многообразия объективных (инструментальных) методов акустической экспертизы помещений и линий связи, на сегодняшний день наиболее популярны два метода: формантный и модуляционный [1], каждый из которых обладает своими достоинствами и недостатками. Формантный метод удобно сопоставлять с результатами артикуляционных испытаний, однако неудобно применять в условиях преобладания реверберационной помехи. Модуляционный метод обладает диаметрально противоположными свойствами. Поэтому недавно был предложен новый формантно-модуляционный (ФМ) метод, объединяющий достоинства обоих методов [2].

### 1. ФОРМАНТНО-МОДУЛЯЦИОННЫЙ МЕТОД: КРАТКОЕ ОПИСАНИЕ

При акустической экспертизе помещения любым объективным методом, в определенной точке помещения (местоположение диктора), характеризуемого импульсной характеристикой (ИХ)  $h(v)$ , излучают тестовый сигнал  $x(t)$ , а в иной точке помещения (местоположение аудитора) принимают сигнал  $y(t)$ :

$$y(t) = \int_{-\infty}^{\infty} h(v)x(t-v)dv + n(t),$$

где  $n(t)$  - шумовая помеха.

При использовании ФМ метода [2], как и в формантном методе [3], вычисляют артикуляционную разборчивость  $A$ :

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

где  $p_k$  - вероятность пребывания формант в  $k$ -той полосе частот;  $P_k(E_k)$  - коэффициент восприятия речи;  $E_k$  - эффективный уровень ощущения речевого сигнала в  $k$ -той полосе частот  $\Delta f_k$ . Отличие 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:

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

где  $SNR_k$  - парциальное отношение сигнал-шум,  $D_{sk}$  и  $D_{nk}$  - соответственно дисперсии стационарных сигнала и шума в  $k$ -той полосе частот, в ФМ методе для оценивания  $E_k$  предложено использовать иной способ, применяющийся в модуляционном методе и позволяющий корректно учесть влияние реверберационной помехи [1].

Это косвенный способ измерений 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  $F$ :

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

где  $\xi(t)$  - ССП с дисперсией  $D_\xi$  и со 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 из-за влияния реверберационной и шумовой помех коэффициент модуляции  $m$  дисперсии  $D_y(t)$  будет тем меньше единицы, чем сильнее влияние помех. В полном варианте ФМ метода используют дискретный набор из 14 частот модуляции  $F_i, i = \overline{1,14}$ , измеряя «эквивалентное» отношение сигнал-помеха:

$$\tilde{E}_k = \sum_{i=1}^{14} \tilde{SNR}_{k,i} / 14, \quad (4)$$

$$\tilde{SNR}_{k,i} = 10 \lg \{ \tilde{m}_k(F_i) / [1 - \tilde{m}_k(F_i)] \}. \quad (5)$$

где  $m_k(F_i)$  - коэффициент модуляции дисперсии  $D_y(t)$  сигнала  $y(t)$ , наблюдаемого в  $k$ -той полосе частот [2]. Коэффициенты модуляции  $m_k(F_i)$  оценивают с использованием преобразования Фурье [1]:

$$\tilde{m}_k(F_i) = 2|A_{ik}(F_i)|/|A_{ik}(0)|, \quad A_{ik}(f) = \int_0^T y_{ik}^2(t) e^{-j2\pi ft} dt / T. \quad (6)$$

где  $T$  - длительность реализации процесса  $y_{ik}(t) = \int_{-\infty}^{\infty} h(v)x_{ik}(t-v)dv + n(t)$ ;  $|\cdot|$  - символ модуля;  $x_{ik}(t) = \xi_k(t)\sqrt{1 + \cos 2\pi F_i t}$  - модулированный полосовой белый шум в  $k$ -той полосе частот;  $\xi_k(t)$  - результат фильтрации ССП  $\xi(t)$   $k$ -тым полосовым фильтром.

В работе [4] предложен ускоренный вариант ФМ метода, в котором вместо 14 измерительных опытов с сигналами (3) используют единственный тестовый сигнал:

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

При этом коэффициент модуляции оценивают в соответствии с соотношением:

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

а вместо соотношения (4) используют его аналог:

$$\tilde{E}_k = \sum_{i=1}^5 \tilde{SNR}_{k,i} / 5. \quad (9)$$

## 2. ДИСПЕРСИЯ ОЦЕНКИ SNR ДЛЯ ФОРМАНТНОГО МЕТОДА

Можно показать, что для формантного метода дисперсия оценки отношения сигнал-шум описывается приближенным соотношением:

$$D_\phi[\tilde{SNR}_k] \approx 38/\Delta f_k T, \quad (10)$$

погрешность которого, при  $\Delta f T \geq 10^4$ , не превышает 0,1 дБ в диапазоне значений  $SNR = -20$  дБ... + 20 дБ. Соотношение (10) понадобится в дальнейшем при сопоставлении ФМ и формантного методов по точности измерений.

## 3. ИЗМЕРЕНИЯ В УСЛОВИЯХ ПРЕОБЛАДАНИЯ ШУМОВОЙ ПОМЕХИ

Если распределение вероятностей формант по частоте  $p_k$  и коэффициенты восприятия  $P_k(E_k)$  известны с высокой точностью, тогда из (1), (3), (4)-(6) следует, что смещенность и дисперсия оценки (1) полностью определяются статистическими свойствами величин  $A_{ik}(0)$  и  $A_{ik}(F_i)$ .

В случае, когда шумовая помеха превалирует над реверберационной, модель сигнала в  $k$ -том частотном канале можно представить так:

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

где  $\xi_k(t)$  и  $n_k(t)$  - нормально распределенные полосовые белые и статистически независимые ССП с нулевыми математическими ожиданиями (МО);  $\varphi_k$  - неизвестная начальная фаза.

Можно показать, что при условии  $T = r/F_i$ , где  $r$  - произвольное целое положительное число, МО величин  $A_{ik}(0)$  и  $A_{ik}(F_i)$  равны соответственно:

$$\langle A(0) \rangle = D_\xi + D_n, \quad \langle |A(F_i)| \rangle = D_\xi / 2, \quad (12)$$

где  $\langle \cdot \rangle$  - символ МО. Таким образом, равенство  $T = r/F_i$  есть условие получения несмещенной оценки (1).

Полагая  $R_\xi(\tau) = R_n(\tau) = Sa(\pi \Delta f \tau) \cos 2\pi f_0 \tau$ , где  $R_\xi(\tau)$  и  $R_n(\tau)$  - коэффициенты корреляции ССП  $\xi(t)$  и  $n(t)$ , соответственно,  $\Delta f$  - 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 дисперсии величин  $A_{ik}(0)$  и  $A_{ik}(F_i)$ :

$$D[A(0)] \approx D[|A(F_i)|] \approx (D_n^2 / \Delta f T) (3q^2 / 2 + 4q + 1), \quad (13)$$

где  $q = D_{\xi}/D_n$ . Из соотношения (13) следует вывод о состоятельности оценки (1).

В силу слабой коррелированности величин  $A_{ik}(0)$  и  $A_{ik}(F_i)$ , при  $T = r/F_i$  из соотношений (6), (12) следует:

$$\langle \tilde{m}_k(F_i) \rangle \approx D_{\xi}/(D_{\xi} + D_n) = q/1 + q = m, \quad (14)$$

т.е. оценка коэффициента модуляции является несмещенной. Приближенное выражение для верхней границы дисперсии оценки коэффициента модуляции:

$$D[\tilde{m}_k] \leq 6/\Delta f_k T. \quad (15)$$

Для  $\Delta f T \gg 1$  из (15), с учетом (5), следует:

$$D[\tilde{SNR}_k] \leq 113 \cdot (1 + 10^{-0,1 \cdot SNR})^4 10^{0,2 \cdot SNR} / \Delta f_k T. \quad (16)$$

Результаты компьютерного моделирования свидетельствуют, что верхняя граница дисперсии, определяемая соотношением (16), завышена примерно в 2 раза, что можно объяснить приближенным характером аналитических выкладок.

В силу статистической независимости оценок  $\tilde{SNR}_{k,i}$ , в полной версии ФМ метода верхняя граница дисперсии (16) снижается в 14 раз благодаря усреднению оценок  $\tilde{SNR}_{k,i}$  согласно соотношению (4). Для  $SNR = 0$  дисперсия минимальна:  $D_{ПФМ}[\tilde{SNR}_k] \Big|_{SNR=0} \leq 130/\Delta f_k T$ . Учитывая, что соотношение (16) дает примерно вдвое завышенные результаты, после коррекции получаем  $D_{ПФМ}[\tilde{SNR}_k] \Big|_{SNR=0} \leq 65/\Delta f_k T$ , что все равно почти вдвое больше дисперсии (10) формантного метода. С отклонением от значения  $SNR = 0$  дБ, проигрыш в точности измерений увеличивается.

Анализируя соотношения (4), (6), (8), (9), нетрудно прийти к выводу, что ускоренный вариант ФМ метода приводит к еще более высоким погрешностям измерений отношения сигнал-шум, проигрывая полному варианту ФМ метода, по дисперсии оценки разборчивости, примерно в 3 раза, а формантному методу – примерно в 6 раз. Эти аналитические результаты подтверждаются результатами компьютерного моделирования. Для времени измерений  $T = 16$  с относительная погрешность измерений не более 10% обеспечивается при  $SNR > -25$  дБ для формантного метода, при  $SNR > -15$  дБ - для полного ФМ метода, при  $SNR > -10$  дБ - для ускоренного ФМ метода.

Что касается времени измерений, для полного варианта ФМ метода оно в 14 раз выше, чем для формантного метода. Время измерений для ускоренного варианта ФМ метода и для формантного метода практически одинаково.

#### 4. ИЗМЕРЕНИЯ В УСЛОВИЯХ ПРЕОБЛАДАНИЯ РЕВЕРБЕРАЦИИ

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

$$y_{ik}(t) = \int_{-\infty}^{\infty} h_{ek}(v) x_i(t-v) dv, \quad h_{ek}(v) = \int_{-\infty}^{\infty} h(v) h_{k0}(z-v) dv, \quad (17)$$

где  $x_i(t) = \xi(t)\sqrt{f_i(t)}$ ;  $h_{k0}(v)$  - ИХ  $k$ -того полосового фильтра;  $\xi(t)$  - ограниченный по полосе ( $0 \leq f \leq \Delta f$ ) белый шум с дисперсией  $D_\xi$ . Тогда, при выполнении условия  $T = r/F_i$ , после громоздких выкладок можно получить:

$$\langle A_k(F_i) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v) e^{-j2\pi F_i v} dv, \quad \langle A_k(0) \rangle \approx \frac{D_\xi}{2\Delta f} \int_{-\infty}^{\infty} h_{ek}^2(v) dv.$$

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

Математическое ожидание и дисперсия коэффициента модуляции:

$$\langle \tilde{m}_{ki} \rangle \approx \left| \frac{\int_0^{\infty} h_{ek}^2(t) e^{-j2\pi F_i t} dt}{\int_0^{\infty} h_{ek}^2(t) dt} \right|, \tag{18}$$

$$D[\tilde{m}_k(F_i)] < 24/\Delta f_k T. \tag{19}$$

Заметим, что соотношение (18) описывает способ измерений разборчивости речи, сводящийся к измерениям ИХ помещения и ранее предложенный Шредером [7].

Сопоставляя соотношения (19) и (15) видим, что в случае реверберационной помехи выражение для верхней границы дисперсии  $\tilde{m}_{ki}$  в 4 раза превышает такое для шумовой помехи. Причины такого несовпадения нуждаются в дальнейших исследованиях. Одна из возможных причин – приближенный характер аналитических выкладок. Другая причина – различие статистических свойств моделей, описываемых соотношениями (11) и (17).

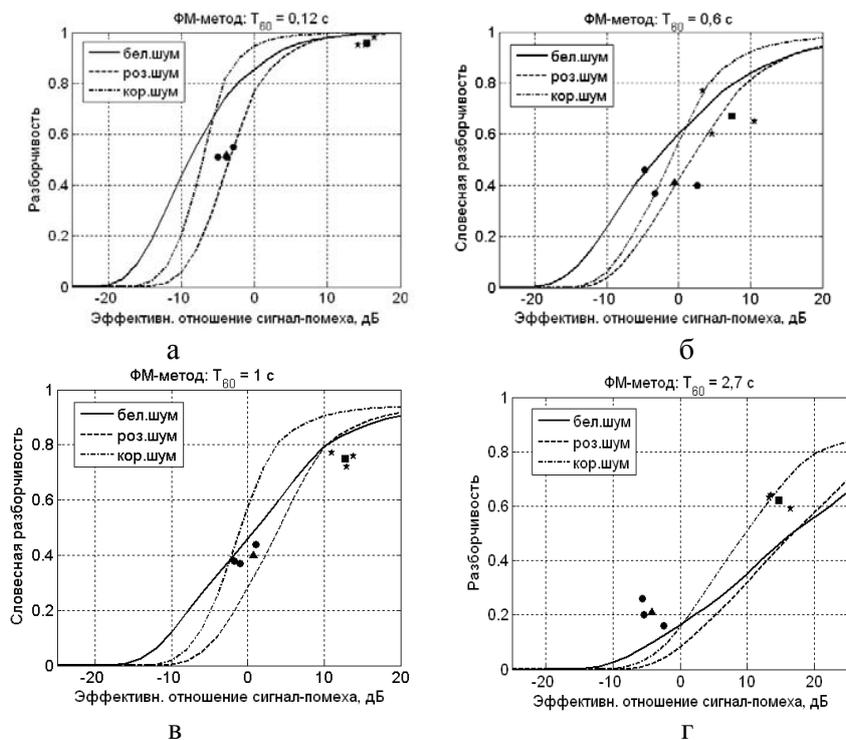


Рис. 1. Оценки разборчивости для ФМ метода и артикуляционных испытаний

На рис. 1 сопоставлены результаты оценивания словесной разборчивости с помощью ФМ метода и артикуляционных испытаний в различных помеховых условиях. Сплошной, пунктирной и штрих-пунктирной линиями представлены результаты для ФМ метода. В случае пренебрежимо малого влияния реверберационной помехи (рис. 1, а) эти кривые практически полностью совпадают с таковыми для формантного метода. На рис. 1, б-г, наблюдается заметное снижение разборчивости с ростом влияния реверберации.

Результаты артикуляционных испытаний в условиях преобладающего влияния реверберационной помехи представлены звездочками, их среднее значение представлено квадратиком. Аналогично результаты артикуляционных испытаний в условиях сопоставимого влияния шумовой и реверберационной помех представлены кружочками, их среднее значение представлено треугольником.

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

## ЗАКЛЮЧЕНИЕ

В условиях заметного влияния реверберационной помехи, ФМ метод предпочтительнее формантного метода, поскольку позволяет корректно учитывать влияние как шумовой, так и реверберационной помех. Хорошее совпадение результатов аналитических и экспериментальных исследований свидетельствует о работоспособности предложенного ФМ метода и достоверности получаемых на его основе результатов.

Вместе с тем, в случае преобладания шумовой помехи наиболее точным и быстрым является формантный метод, что позволяет опровергнуть известный тезис об «устарелости» формантного метода [5].

## ЛИТЕРАТУРА

1. Дидковский В.С., Дидковская М.В., Пролдеус А.Н. Акустическая экспертиза каналов речевой коммуникации: Монография. – К.: Имэкс-ЛТД, 2008. – 420 с.
2. Пролдеус А.Н. О некоторых особенностях развития объективных методов измерений разборчивости речи // Электроника и связь, тематический выпуск «Электроника и нанотехнологии». - 2010. – №2. – С. 217-223.
3. Покровский Н.Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – 390 с.
4. Prodeus A. Rapid Version of Formant-Modulation Method of Speech Intelligibility Estimation. // Proceedings of the VII International Conference MEMSTECH 2011. - Lviv, Polyana, 2011. - Pp.61-63.
5. Алдошина И., Приттс Р. Музыкальная акустика. Учебник. – СПб.: Композитор, 2006. - 720 с.