

Список литературы

1. Интернет сообщение: www.securitylab.ru
2. Интернет сообщение: www.cnews.ru
3. Kerckhoffs A. La cryptographie militaire // Journal des sciences militaires. - Vol. IX. - Jan. 1883, P. 5-38, (P. 161-191, Feb. 1883).
4. Шеннон К.Э. Работы по теории информации и кибернетике // М.: И.Л., 1963.
5. Жебка С.В. Сжатие звуковой информации // Захист інформації. Спеціальний випуск(40). - 2008. - С. 128 -129; Герасин С.Н., Калиниченко О.В., Лоцман В.П., Трошило А.С. О способе кодирования сигналов различной природы. // Бионика интеллекта. - 2008. - №1 (68). - С. 177-180.
6. Климентов В.В., Трошило А.С., Дыс Л.И., Бурлаков В.М. Проблемы и перспективы применения криптографических систем // Інформаційна Безпека. Матеріали науково-практичної конференції. Україна, Київ, 26-27 березня 2009 року. - С. 270-274.

Поступила 25.11.09

УДК 621.39:534.782

Демьян Н.И., Осмаловский В.А., Хорошко В.А.

**ЛИНЕЙНЫЕ МОДЕЛИ РЕЧЕВОГО СИГНАЛА  
С ЛОКАЛЬНО-ПОСТОЯННЫМИ ПАРАМЕТРАМИ**

**Вступление**

Звуки при формировании которых голосовые связки осуществляют колебательные движения называют вокализованными. Все остальные звуки относятся к невокализованным. Более точно: среди последних различают фрикативные звуки, возникающие при образовании турбулентного широкополосного шума, и взрывные звуки, формируемые путем создания в тракте смычки с последующими внезапным высвобождением сжатого в области за смычкой воздуха.

Волны распространяющиеся в речеобразующей системе, могут быть описаны двумя функциями пространственных  $x, y, z$  и временной (непрерывной)  $t$  координат: звуковым давлением  $p(x, y, z, t)$ . Для акустических колебаний с длиной волны  $\lambda$ , превышающей размеры голосового тракта, можно считать, что вдоль продольной оси тракта распространяется плоская волна. Такое допущение оправдано в частотном диапазоне ниже 5000 Гц [1]. Звуковое давление и объемная скорость тогда являются функциями только двух переменных:  $p(x, t), v(x, t)$ . Если голосовой тракт аппроксимировать цилиндрической трубой с сечением  $S(x, t)$ , то распространение колебаний по трубе можно описать волновым уравнением Вебстера:

$$\frac{1}{S(x,t)} \cdot \frac{\partial}{\partial x} S(x,t) \frac{\partial \bar{\Phi}(x,t)}{\partial x} = \frac{1}{c^2} \frac{\partial^2 \bar{\Phi}(x,t)}{\partial t^2}; \quad (1)$$

где  $\Phi(x,t)$  - потенциал скорости акустических колебаний,

$$\frac{\partial \Phi(x,t)}{\partial x} = \frac{1}{\rho} p(x,t), \quad \frac{\partial \bar{\Phi}(x,t)}{\partial t} = -v(x,t);$$

где  $\rho$  - плотность воздуха в трубе;  $c$  – скорость распространения звука в воздухе.

**Основная часть**

Выражение (1) описывает свободные колебания в тракте. При произнесении вокальных звуков речи на вход тракта воздействует волна, поступающая от голосовых связок, объемная

скорость которой зависит от перепада звукового давления  $\Delta p$  на связках,  $v = (2\Delta p / 0,875\rho)^{1/2}$  [2], а следовательно, от  $p(x,t)$  или  $\Phi(x,t)$  в тракте. Если обозначить «воздействие» на тракт через  $F(\Phi,t)$ , то волновое уравнение, описывающее распространение акустической волны в системе «связки-тракт», запишем в виде:

$$\frac{1}{S(x,t)} \cdot \frac{\partial}{\partial x} S(x,t) \frac{\partial \Phi(x,t)}{\partial x} - \frac{1}{c^2} \frac{\partial^2 \Phi(x,t)}{\partial t^2} = F(\Phi,t) . \quad (2)$$

На основании этого выражения можно сделать вывод: артикуляционный аппарат является нелинейной параметрической системой с распределенными параметрами. Нелинейный характер этой системы определяется зависимостью  $F$  от  $\Phi$ , а параметрический – наличием зависимости коэффициентов от времени. Влияние тракта на работу голосовых связок часто считают пренебрежимо малым, поскольку оно проявляется только на низких частотах. Это позволяет представить артикуляционный аппарат традиционной для инженерной практики схемой: источник возбуждения - линейная система, формирующая на выходе речевое колебание. Такое заключение, будучи вполне приемлемым в рамках линейной системы с постоянными параметрами, становится весьма приближенным в нелинейных параметрических системах, где малым изменениям параметров могут соответствовать существенные изменения выходного сигнала.

Получить аналитическое решение выражения (2) трудно. Можно лишь воспользоваться численными методами решения, поэтому использовать общее выражение (2) для построения моделей речевых сигналов в системах с сосредоточенными параметрами пока затруднительно. Важные для построения моделей речевых сигналов результаты удалось получить из аппроксимации голосового тракта секциями цилиндрических труб с постоянными площадями поперечного сечения [1]. Для одной секции длиной  $l_k$  с площадью  $S_k$  можно записать два уравнения:

$$\frac{\partial p_k(x,t)}{\partial x} = - \frac{\rho}{S_k} \frac{\partial v_k(x,t)}{\partial t} , \quad (3)$$

$$\frac{\partial v_k(x,t)}{\partial x} = - \frac{S_k}{\rho c^2} \frac{\partial p_k(x,t)}{\partial t} . \quad (4)$$

Объединяя эти уравнения в одно, получим уравнение Вебстера относительно одной из функций  $p_k(x,t)$ ,  $v_k(x,t)$ . Решая уравнения (3) и (4), они принимают вид линейной комбинации объемных скоростей прямой  $v_k^+(t - \frac{x}{c})$  и отражений от стыков между секциями  $v_k^-(t + \frac{x}{c})$  волн

$$v_k(x,t) = v_k^+(t - \frac{x}{c}) - v_k^-(t + \frac{x}{c}) , \quad (5)$$

$$p_k(x,t) = \frac{\rho c}{S_k} [v_k^+(t - \frac{x}{c}) + v_k^-(t + \frac{x}{c})] . \quad (6)$$

Проверить эти соотношения, характерные для решений волновых уравнений, можно подстановкой (5) и (6) в (3) и (4), выполняемой в предположении, что  $v_k^+$  и  $v_k^-$  – дважды дифференцируемые функции. Здесь  $x$  – расстояние, измеряемое вдоль продольной оси голосового тракта. В месте стыка возникают отражения, однако звуковое давление и объемная скорость должны оставаться непрерывными функциями. Условия их непрерывности в месте стыка таковы:

$$p_k(l_k, t) = p_{k+1}(0, t), \quad v_k(l_k, t) = v_{k+1}(0, t). \quad (7)$$

Подставляя в соотношения (5) и (6) выражения (7), получаем при  $\tau_k = \tau$

$$v_k^+(t - \tau) + v_k^-(t + \tau) = \frac{S_k}{S_{k+1}} [v_{k+1}^+(t) + v_{k+1}^-(t)],$$

$$v_k^+(t - \tau) - v_k^-(t + \tau) = v_{k+1}^+(t) - v_{k+1}^-(t), \quad (8)$$

где  $\tau$  – время распространения волны вдоль  $k$ -й секции.

Суммируя и вычитая эти уравнения, приходим к соотношениям

$$v_k^+(t - \tau) = \frac{1}{1 + r_k} [v_k^+(t) - r_k v_k^-(t)],$$

$$v_k^-(t - \tau) = \frac{1}{1 + r_k} [v_k^-(t) - r_k v_{k+1}^+(t)], \quad (9)$$

где  $r_k$  - коэффициент отражения.

$$\text{Введем } c_k = \prod_{i=k}^m (1 + r_i) = c_{k+1} (1 + r_k). \quad (10)$$

Умножая обе части (9) на  $c_k$ , имеем

$$c_k v_k^+(t - \tau) = c_{k+1} v_{k+1}^+(t) - r_k c_{k+1} v_{k+1}^-(t),$$

$$c_k v_k^-(t + \tau) = c_{k+1} v_{k+1}^-(t) - r_k c_{k+1} v_{k+1}^+(t). \quad (11)$$

Произведем в (11) замену переменной  $z = t - \tau$ . Тогда  $t = z + \tau$ ,  $t + \tau = z + 2\tau$  и, обозначая  $z$  вновь через  $t$ , получаем

$$c_k v_k^+(t) = c_{k+1} v_{k+1}^+(t + \tau) - r_k c_{k+1} v_{k+1}^-(t + \tau),$$

$$c_k v_k^-(t + 2\tau) = c_{k+1} v_{k+1}^-(t + \tau) - r_k c_{k+1} v_{k+1}^+(t + \tau). \quad (12)$$

Введем время распространения  $t_k$  акустической волны от левого края секции с номером  $k$  до правого края последней  $m$ -й секции. Тогда  $t_k = t_{k+1} + \tau$ ,  $k = 1, 2, \dots, m$ ;  $t_m = \tau$ ;  $t_{m+1} = 0$ . Исходя из этого перепишем выражения (12) с учетом  $t_k$  как:

$$c_k v_k^+(t - t_k) = c_{k+1} v_{k+1}^+(t - t_{k+1}) - r_k c_{k+1} v_{k+1}^-(t - t_{k+1}),$$

$$c_k v_k^-(t - t_k + 2\tau) = c_{k+1} v_{k+1}^-(t - t_{k+1}) - r_k c_{k+1} v_{k+1}^+(t - t_{k+1}). \quad (13)$$

Объясним физический смысл выражений (13). Переменные  $v_k^+(t - t_k)$ ,  $v_k^-(t - t_k + 2\tau)$  - это объемные скорости прямой и отраженной от  $k$ -го стыка волн, наблюдаемых на выходе всей системы. Сдвиг  $t_k$  учитывает конечное время распространения волны от стыка до выхода. Сдвиг  $2\tau$  показывает, что отраженная волна, находясь слева от  $k$ -го стыка,

достигает сначала  $k-1$ -го стыка за время  $\tau$ , отражается, затем  $\tau$  распространяется вдоль  $k$ -й секции (суммарный сдвиг  $2\tau$ ) и далее проходит по остальным  $m-k$  секциям до выхода. Коэффициенты  $c_k$  учитывают отражения волн при их продвижению к выходу. Таким образом выражение (13) так же, как и выражения (5) и (6) описывают взаимосвязь прямой и отраженной волн в  $k$ -й секции, но наблюдаются эти волны на выходе системы. Введем для объемных скоростей наблюдаемых волн новые обозначения:

$$\zeta_k(t) = c_k v_k^+(t - t_k), \tag{14}$$

$$\xi_k(t) = c_k v_k^-(t - t_k + 2\tau).$$

Теперь выражение (13) можно переписать для новых переменных

$$\zeta_k(t) = \zeta_{k+1}(t) - r_k \xi_{k+1}(t - 2\tau), \tag{15}$$

$$\xi_k(t) = \xi_{k+1}(t - 2\tau) - r_k \zeta_{k+1}(t),$$

$$k = 1, 2, \dots, m.$$

Очевидно, что  $\zeta_{m+1}(t) = \xi_{m+1}(t) = x(t)$  - речевое колебание на выходе, а  $\zeta_k(t)$  и  $\xi_k(t)$  - его составляющие, формирующиеся в отдельных секциях. Уравнения (15) показывают как  $x(t)$  разложить на эти составляющие, если известны  $r_k$ . Для этого необходимо, начиная с  $k = m$  выполнять расчет по уравнениям (15), понижая  $k$  вплоть до  $k = 1$ . Переменная  $\zeta_1(t) = \xi_1$  - это объемная скорость волны, поступающей от голосовой щели в тракт.

Если (15) переписать в другом виде, а именно:

$$\zeta_{k+1}(t) = r_k \xi_{k+1}(t - 2\tau) + \zeta_k(t), \tag{16}$$

$$\xi_{k+1}(t - 2\tau) = r_k \zeta_{k+1}(t) + \xi_k(t),$$

$$k = 1, 2, \dots, m, \zeta_1(t) = \xi_1,$$

то задавая  $\xi_1$ , можно, пользуясь (16), рассчитать последовательно, повышая  $k$ , все составляющие вплоть до  $x(t) = \zeta_{m+1}(t) = \xi_{m+1}(t)$ . Вычисления по первому уравнению выражения (16) при  $k = 1, 2$  выполняются во фрагменте схемы приведенной на рис.1. Для вычисления  $\xi_2(t)$  из второго уравнения (15) запишем  $\xi_2(t) = \xi_3(t - 2\tau) - r_2 \zeta_3(t)$ . Вычисления по этому соотношению следует добавить к схеме, показанной на рис.1.

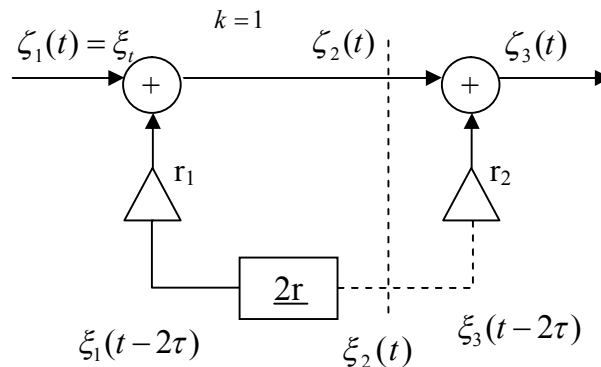


Рис.1. Фрагмент лестничного фильтра с частотными корреляциями

В результате получим более полный фрагмент общей схемы (рис.2). Продолжая такие рассуждения, приходим к известной схеме лестничного фильтра, приведенного на рис.3.

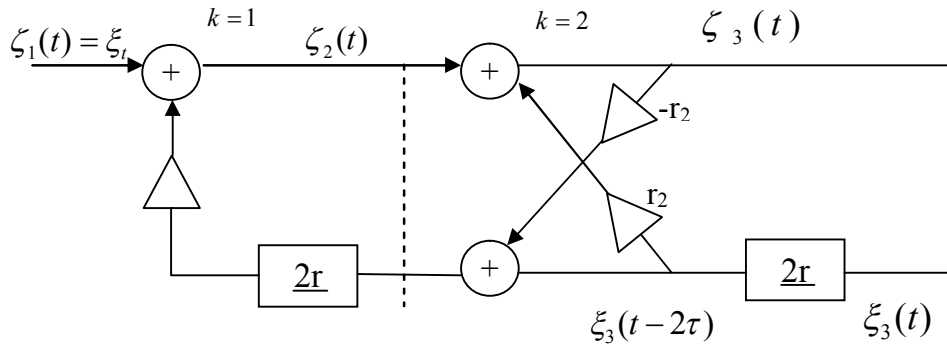


Рис.2. Изображение двух каскадов лестничного фильтра с частотными корреляциями

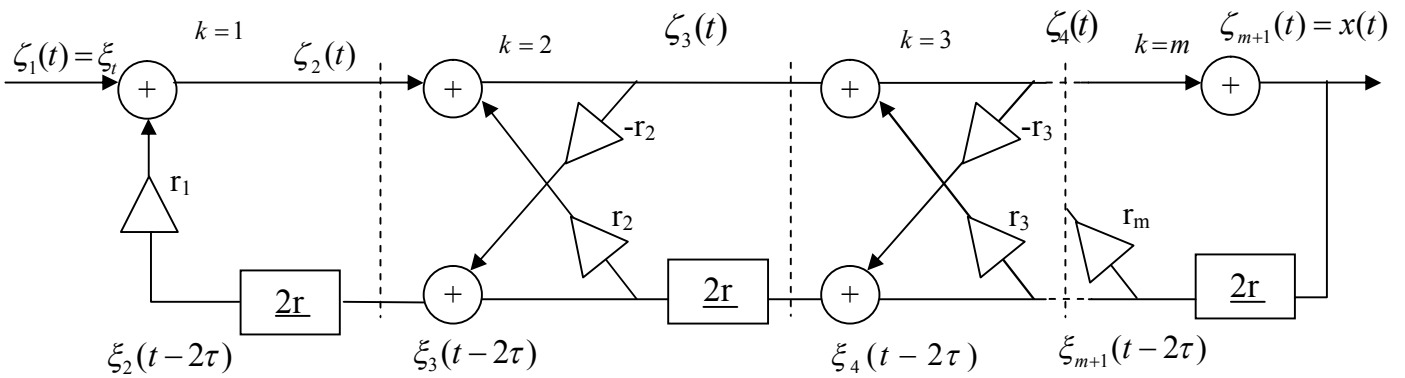


Рис.3. Структурная схема лестничного фильтра

Если теперь перейти к дискретному нормированному времени  $t = 0, 1, 2, \dots$ , приняв  $\Delta t = 2\tau$ , получаем соответствующие уравнения в дискретном времени

$$\zeta_k(t) = \zeta_{k+1}(t) - r_k \xi_{k+1}(t-1), \quad (17)$$

$$\xi_k(t) = \xi_{k+1}(t-1) - r_k \zeta_{k+1}(t),$$

$$k = 1, 2, \dots, m; \quad \zeta_{m+1}(t) = \xi_{m+1}(t) = x(t); \quad t = 1, 2, \dots,$$

и

$$\zeta_{k+1}(t) = r_k \xi_{k+1}(t-1) + \zeta_k(t), \quad k = 1, 2, \dots, m, \quad (18)$$

$$\xi_{k+1}(t-1) = r_k \zeta_{k+1}(t) + \xi_k(t), \quad k = 2, 3, \dots, m,$$

$$\zeta_1(t) = \xi_t.$$

Уравнения (17) были выведены в [3], а (18) – в [4]. Следуя исследованиям проведенным в [4] примем  $m = 2$  и по (17) и (18) запишем

$$\zeta_2(t) = r_1 \xi_2(t-1) + \xi_t, \quad (19)$$

$$\zeta_2(t) = \zeta_3(t) - r_2 \xi_3(t-1) = x_t - r_2 x_{t-1}, \quad (20)$$

$$\xi(t-1) = \xi_3(t-2) - r_2 \zeta_3(t-1) = x_{t-2} - r_2 x_{t-1}.$$

Подставляя (20) в (19), получаем

$$x_t - r_2 x_{t-1} = r_1 [x_{t-2} - r_2 x_{t-1}] + \xi_t, \quad (21)$$

или

$$x_t = r_2(1-r_1)x_{t-1} + r_1 x_{t-2} + \xi_t. \quad (22)$$

Уравнение (22) является линейной моделью авторегрессии второго порядка с параметрами авторегрессии  $a_1 = r_2(1-r_1)$ ,  $a_2 = r_1$ . Причем формулы пересчета  $\{r_1, r_2\}$  в  $\{a_1, a_2\}$  совпадают с известными соотношениями Дарбина [5]. Выполняя такие подстановки для произвольного порядка  $m$  можно получить уравнение авторегрессии

$$x_t = \sum_{i=1}^m a_i x_{t-i} + \xi_t, \quad (23)$$

где  $\{a_i\}$  - параметры авторегрессии. Формулы пересчета  $\{r_i\}$  в  $\{a_i\}$  совпадают с общими соотношениями Дарбина, связывающими так называемые частные корреляции и параметры авторегрессии [5]. Этот факт позволяет интерпретировать  $r_k$  как частные корреляции временного ряда  $x_t$ .

Уравнения (17), (18), (23) задают широко распространенные модели речевых сигналов с локально постоянными параметрами  $\{r_i\}$  или  $\{a_i\}$ . Их можно получить и по-другому, используя уравнение (23), метод обновляющего процесса и осуществляя обычную в теории временных рядов «подгонку» моделей авторегрессии возрастающего порядка  $p = 1, 2, \dots, m$ . Такие рассуждения использованы в [4] и тесно связаны уже и с анализом сигнала  $x_t$ . Сумма в правой части выражения (23) может быть интерпретирована как предсказанное значение  $x_t$ , которое задается параметрами  $\{a_i\}$  или  $\{r_i\}$ . Поэтому модели (17), (18) и (23) называют моделями линейного предсказания.

Если  $\xi_t$  - случайная некоррелированная гауссовская последовательность, то приведенные уравнения описывают сигнал на участках звучания согласных. На участках звучания гласных (и звонких согласных)  $\xi_1$  следует рассматривать как модель квазипериодической волны, формирующейся в голосовой щели. Ее можно описать последовательностью импульсов известной формы  $u_t$ :

$$\xi_t = \sum_i u_{t-t_i} + \omega_t; \quad t_i - t_{i-1} = T_0, \quad (24)$$

где  $t_i$  - временные координаты импульсов;

$\omega_t$  - сопутствующая некоррелированная последовательность слабой интенсивности;

$T_0$  - период основного тока.

Форма импульса  $u_i$  задается обычно треугольной, но может изменяться в широких пределах, так как это всегда можно компенсировать, изменив значения параметров  $a_i$  или  $r_i$ . Важно чтобы  $\xi_i$  обладала широким спектром и, действуя на входе формирующей системы, формировала все ее собственные колебания.

### **Выводы**

Эта модель наиболее полно отражает свойства речевого сигнала в дискретном времени. Причем в цифровых системах передачи разборчивость обычно высока и в понятие «качество звучания» вкладываются такие субъективные характеристики как натуральность, естественность звучания, возможность узнавания и т.п. Предложенная модель позволяет описать все эти субъективные характеристики, а следовательно и обеспечить их.

### **Список литературы**

1. Рабинер Л.Р. Цифровая обработка речевых сигналов: Пер. с англ./ Рабинер Л.Р., Шафер Р.В. // Под ред. М.В. Назарова, Ю.Н. Прохорова. –М.: Радио и связь, 1981. -495 с.
2. Величкин А.И. Передача аналоговых сообщений по цифровым каналам связи. / Величкин А.И. –М.: Радио и связь, 1983. -240 с.
3. Гуревич В.Э. ИКМ в многоканальной связи. / Гуревич В.Э., Лопушнян Ю.Г., Рабинович Г.В. – М.: Связь, 1973. -336 с.
4. Binal M.E. On the Derivation of Certain Statistical Properties of Speech. / Acustica, vol. 40, №3, 1978, P. 137-212.
5. Лупцер Р.Ш. Статистика случайных процессов / Лупцер Р.Ш., Ширяев А.Н. – М.: Наука, 1974. – 696 с.
6. Карпов О.Н. Компьютерные технологии распознавания речевых сигналов / Карпов О.Н., Габович А.Г., Марченко Б.Г., Хорошко В.А., Щербак Л.Н. – К.: ООО «ПолиграфКонсалтинг», 2005. – 138 с.

*Поступила 4.11.09*