АКУСТИЧЕСКИЙ ЖУРНАЛ

Том XXXII 1986 Вып. 4

УДК 612.85

БЕГУЩИЕ ВОЛНЫ В РЕЧЕВОМ ТРАКТЕ

Сорокин В. Н.

Описывается рекурсивная схема для расчета бегущих волн в речевом тракте от легких до губ с учетом податливости стенок и относительно быстрого изменения площади поперечного сечения голосовой щели.

Исследования последних лет обнаружили ряд новых явлений в акустике речеобразования [1]. Так, было установлено, что голосовые складки должны рассматриваться как своего рода артикуляторные органы, поскольку для глухих фрикативных и взрывных звуков площадь голосовой щели сравнима с минимальной площадью поперечного сечения верхней области речевого тракта. Во время колебаний голосовых складок трахея и легкие влияют на частоты и затухание резонансов речевого тракта и вызывают появление дополнительных резонансов. Следовательно, речевой тракт заканчивается не на голосовых складках, а в легких. Податливость стенок речевого тракта приводит к возникновению радиальных акустических колебаний, определяющих характеристики речевого сигнала во время звонкой смычки и в ее окрестности. Моделирование этих явлений, однако, требует больших вычислительных затрат, которые могут быть снижены при использовании метода бегущих волн с аппроксимацией речевого тракта последовательностью цилиндрических секций. В работе [2] для этого метода была предложена эффективная вычислительная схема, но с учетом перечисленных выше явлений она должна быть модифицирована.

Рассмотрим систему уравнений для акустических процессов в речевом тракте:

$$\begin{gathered}
-S\frac{\partial P}{\partial x}=\rho_0\frac{\partial U}{\partial t}+r_1U,\\
-\rho_0c_0^2\frac{\partial U}{\partial x}=S\frac{\partial P}{\partial t}+(r_2+r_3)P+\rho_0c_0^2\frac{\partial S}{\partial t},\\
S=S_0+yL,\\
m\frac{\partial^2 y}{\partial t^2}+b\frac{\partial y}{\partial t}+ky=PL,
\end{gathered}\tag{1}$$

где x — координата вдоль оси речевого тракта, t — время, $P(t, x)$ — акустическое давление, $U(t, x)$ — объемная скорость, $S(t, x)$ — площадь поперечного сечения речевого тракта, $S_0(t, x)$ — площадь поперечного сечения речевого тракта, создаваемая положением артикуляторных органов, yL — приращение площади поперечного сечения в результате податливости стенок речевого тракта, $L(t, x)$ — периметр поперечного сечения, $y(t, x)$ — смещение стенок, ρ_0 — плотность воздуха, $r_1(x)$ — коэффициент вязкого трения, c_0 — скорость звука в воздухе, $r_2(x)$ — коэффициент потерь на теплопроводность, $r_3(x)$ — коэффициент потерь на податливость стенок, $m(x)$ — погонная масса стенок, $b(x)$ — коэффициент вязкого трения в стенках, $k(x)$ — упругость.

Система (1) допускает ряд упрощений. Известно, что резонансные частоты колебаний стенок находятся в диапазоне 150—350 Гц. Для таких частот речевой тракт может рассматриваться как система с сосредоточенными параметрами и, следовательно, зависимостью параметров m, b и k от x можно пренебречь, заменив их эквивалентными постоянными величинами. Скорость изменения S_0 при движении артикуляторных органов го-

раздо меньше скорости изменения площади поперечного сечения при колебаниях стенок. Поэтому приближенно можно принять $\partial S/\partial t \approx \partial(yL)/\partial t$. Площадь поперечного сечения обычно вычисляется по линейным размерам речевого тракта в среднесагиттальной плоскости, причем форма сечения списывается либо кругом с радиусом R, либо прямоугольником с шириной основания d и высотой h. В первом случае периметр $L=2\pi(R+y)$, а во втором — $L=2(d+h+y)$. Соответственно для круглого сечения $\partial(yL)/\partial t=2\pi\partial[y(R+y)]/\partial t$, а для прямоугольного $\partial(yL)/\partial t=2\partial[y(d+h+y)]/\partial t$. При большой площади поперечного сечения ($R \gg y$ или $h \gg y$) $\partial(yL)/\partial t$ пропорционально скорости смещения стенок.

Допустим, что потери в речевом тракте малы и не искажают форму бегущей волны. Тогда для объемной скорости и давления можно записать $U(t, x)=\exp[-g_1(t+\tau)]U^+(t-\tau)-\exp[-g_1(t-\tau)]U^-(t+\tau)$, $P(t, x)=\exp[-g_2(t+\tau)]P^+(t-\tau)+\exp[-g_2(t-\tau)]P^-(t+\tau)$, где $g_1=r_1/S$, $g_2=(r_2+r_3)/S$, $\tau=x/c_0$, индексы «+» и «−» обозначают бегущие волны, распространяющиеся от легких к губам и наоборот.

При аппроксимации формы речевого тракта последовательностью цилиндрических секций их длина должна выбираться, исходя из требований к различению двух ближайших перцептивно значимых мест артикуляции, каковыми являются положения кончика языка для звуков С и Ш. Поэтому длина каждой секции Δx не может быть больше 0,5 см. Если длины всех секций одинаковы, то время пробега волной каждой секции $\Delta t=\Delta x/c_0$, что соответствует тактовой частоте 70 кГц.

Максимальная суммарная скорость кончика языка и нижней челюсти около 20 см/с, так что за интервал Δt произойдет смещение примерно на $3\cdot10^{-4}$ см. Для фрикативных звуков минимальное расстояние между кончиком языка и нёбом составляет около 0,1 см. Следовательно, максимальное относительное изменение площади поперечного сечения составляет 0,3%. Скорость колебаний голосовых складок находится в диапазоне 50—300 см/с для разных частот основного тона при амплитуде колебаний до 0,1—0,2 см. При этом относительное изменение площади голосовой щели за интервал Δt составляет несколько процентов. Скорость колебаний податливых стенок речевого тракта сравнима со скоростью колебаний голосовых складок, но относительное изменение площади поперечного сечения значительно меньше 1% для большинства звуков. Несмотря на такие малые изменения площадей секций за время пробега волны, они могут привести к искажениям речевого сигнала [3, 4]. Это особенно касается модели, в которой голосовая щель находится между трахеей и речевым трактом.

Дополнительная объемная скорость $(\partial S/\partial t)dx$, создаваемая при колебаниях стенок, мала по сравнению со скоростью акустической волны, поэтому можно применить приближенный метод решения системы (1). Обозначим $f=\partial S/\partial t$ и на первом шаге примем, что f не зависит от давления P. Воспользуемся также тем, что для однородной акустической системы для любой дифференцируемой функции $\varphi(t, x)$ справедливо $\partial\varphi(t\pm x/c_0)/\partial x=\pm(1/c_0)\partial\varphi(t\pm x/c_0)/\partial t$. Учитывая это свойство и принимая во внимание малость потерь, для первых двух уравнений системы (1) можно записать так:

$$\frac{\partial P^+(t-x/c_0)}{\partial x}e^{-g_2(t+x/c_0)}+\frac{\partial P^-(t+x/c_0)}{\partial x}e^{-g_2(t-x/c_0)}=$$

$$=\frac{\rho_0 c_0}{S(t,x)}\left[\frac{\partial U^+(t-x/c_0)}{\partial x}e^{-g_1(t+x/c_0)}+\frac{\partial U^-(t+x/c_0)}{\partial x}e^{-g_1(t-x/c_0)}\right],$$

$$\frac{\partial P^+(t-x/c_0)}{\partial x}e^{-g_2(t+x/c_0)}-\frac{\partial P^-(t+x/c_0)}{\partial x}e^{-g_2(t-x/c_0)}=$$

$$=\frac{\rho_0 c_0}{S(t,x)}\left[\frac{\partial U^+(t-x/c_0)}{\partial x}e^{-g_1(t-x/c_0)}-\frac{\partial U^-(t+x/c_0)}{\partial x}e^{-g_1(t-x/c_0)}\right]. \quad (2)$$

Сложим и вычтем почленно уравнения системы (2):

$$\frac{\partial P^+(t-x/c_0)}{\partial x}e^{-g_2(t+x/c_0)}=\frac{\rho_0 c_0}{S(t,x)}\left[\frac{\partial U^+(t-x/c_0)}{\partial x}e^{-g_1(t+x/c_0)}+\frac{1}{2}f(t,x)\right], \tag{3}$$

$$\frac{\partial P^-(t+x/c_0)}{\partial x}e^{-g_2(t-x/c_0)}=\frac{\rho_0 c_0}{S(t,x)}\left[\frac{\partial U^-(t+x/c_0)}{\partial x}e^{-g_1(t+x/c_0)}-\frac{1}{2}f(t,x)\right]. \tag{4}$$

Уравнения (3) и (4) можно проинтегрировать по x для каждой цилиндрической секции, пренебрегая членами с коэффициентами g_1 и g_2:

$$P_i^+\left(t-\frac{x}{c_0}\right)e^{-g_{2i}(t+x/c_0)}=\frac{\rho_0 c_0}{S_i(t)}\left[U_i^+(t-x/c_0)e^{-g_{1i}(t+x/c_0)}+\frac{x}{2}f_i(t)\right], \tag{5}$$

$$P_i^-\left(t+\frac{x}{c_0}\right)e^{-g_{2i}(t-x/c_0)}=\frac{\rho_0 c_0}{S_i(t)}\left[U_i^-(t+x/c_0)e^{-g_{1i}(t+x/c_0)}-\frac{x}{2}f_i(t)\right], \tag{6}$$

Дополнительная объемная скорость, создаваемая колебаниями стенок, имеет разные знаки на правой и левой границах секции, поэтому она суммируется с акустическими волнами, причем только с теми, которые подходят к границам секций. Для правой границы это будет волна, бегущая в положительном направлении, для левой — волна, бегущая в отрицательном направлении.

Возможны различные способы дискретизации $S(t, x)$ во времени. Один из них состоит в сохранении площади сечения секции постоянной в течение всего времени пробега волны, т. е. $S_i(t)=S_i(t-\tau)$, и скачкообразном ее изменении в момент пересечения волной границы между секциями, т. е. $S_i(t)=S_i(t+\tau)$, где $\tau=\Delta t/2$ [3, 4]. Дополнительная объемная скорость вычисляется также в момент $t+\tau$.

На стыке между i-й и $i+1$-й секциями граничные условия для давления и объемной скорости должны быть согласованы также и во времени с учетом принятого способа дискретизации $S(t, x)$:

$$U_i(t,\ \Delta x/2)=U_{i+1}(t,\ -\Delta x/2), \tag{7}$$

$$P_i(t,\ \Delta x/2)=P_{i+1}(t,\ -\Delta x/2). \tag{8}$$

Обозначим $\overline{U}^{\pm}(t-\tau)=U^{\pm}(t-\tau)\exp[-g_1(t+\tau)]+(\Delta x/2)f(t)$, $\overline{U}^{\pm}(t+\tau)=U^{\pm}(t+\tau)\exp[-g_1(t-\tau)]$. Подставив выражения (5) и (6) в (7) и (8), получим

$$\overline{U}_i^+(t-\tau)-\overline{U}_i^-(t+\tau)=\overline{U}_{i+1}^+(t+\tau)-\overline{U}_{i+1}^-(t-\tau), \tag{9}$$

$$\frac{\overline{U}_i^+(t-\tau)}{S_i(t-\tau)}+\frac{\overline{U}_i^-(t+\tau)}{S_i(t+\tau)}=\frac{\overline{U}_{i+1}^+(t+\tau)}{S_{i+1}(t+\tau)}+\frac{\overline{U}_{i+1}^-(t-\tau)}{S_{i+1}(t-\tau)}. \tag{10}$$

Умножим соотношение (10) на $S_i(t+\tau)$ и сложим с (9), а затем умножим (10) на $S_{i+1}(t+\tau)$ и вычтем из него (9). При этом получим рекурсивную схему для трансформации бегущих волн на стыке i-й и $i+1$-й секций в общем случае для произвольно меняющейся площади поперечного сечения речевого тракта во времени:

$$\overline{U}_{i+1}^+=\frac{S_{i+1}(t+\tau)}{S_{i+1}(t-\tau)}\left[\frac{S_i(t+\tau)}{S_i(t-\tau)}\nu_i\overline{U}_i^++\mu_{1i}(\overline{U}_i^++\overline{U}_{i+1}^-)\right], \tag{11}$$

$$\overline{U}_i^-=\frac{S_i(t+\tau)}{S_i(t-\tau)}\left[\frac{S_{i+1}(t+\tau)}{S_{i+1}(t-\tau)}\nu_i\overline{U}_{i+1}^--\mu_{2i}(\overline{U}_i^++\overline{U}_{i+1}^-)\right], \tag{12}$$

где

$$\mu_{1i}=\frac{S_{i+1}(t-\tau)-S_i(t+\tau)}{S_{i+1}(t+\tau)+S_i(t+\tau)}\qquad \mu_{2i}=\frac{S_{i+1}(t+\tau)-S_i(t-\tau)}{S_{i+1}(t+\tau)+S_i(t+\tau)}$$

$$\nu_i=\frac{S_{i+1}(t-\tau)+S_i(t-\tau)}{S_{i+1}(t+\tau)+S_i(t+\tau)}$$

а аргументы при бегущих волнах опущены, причем подразумевается, что волны, прошедшие границу между секциями, имеют аргумент $t+\tau$, а волны, приходящие к границе,— аргумент $t-\tau$. Систему (11), (12) можно упростить, если допустить, что относительное изменение сечения для каждой пары секций одинаково, т. е.

$$\frac{S_i(t+\tau)}{S_i(t-\tau)}=\frac{S_{i+1}(t+\tau)}{S_{i+1}(t-\tau)}.$$

Тогда

$$\overline{U}_{i+1}^{+}=\frac{S_{i+1}(t+\tau)}{S_{i+1}(t-\tau)}\left[\overline{U}_i^{+}+\mu_{1i}(\overline{U}_i^{+}+\overline{U}_{i+1}^{-})\right],$$

$$\overline{U}_i^{-}=\frac{S_i(t+\tau)}{S_i(t-\tau)}\left[\overline{U}_{i+1}^{-}-\mu_{2i}(\overline{U}_i^{+}+\overline{U}_{i+1}^{-})\right].$$

Площадь голосовой щели меняется гораздо быстрее площадей соседних секций, но в любой момент времени она гораздо меньше площади сечения трахеи и морганиевого желудочка. Поэтому справедливо $(S_n+S_{n+1})/S_{n+1}\approx\approx 1$, $(S_n+S_{n-1})/S_{n-1}\approx 1$, где n — номер секции, содержащей голосовую щель. Отсюда получаем аналогичные выражения:

$$\overline{U}_{n+1}^{+}=\frac{S_n(t+\tau)}{S_n(t-\tau)}\left[\overline{U}_n^{+}+\mu_{1n}(\overline{U}_n^{+}+\overline{U}_{n+1}^{-})\right],$$

$$\overline{U}_n^{-}=\frac{S_n(t+\tau)}{S_n(t-\tau)}\left[\overline{U}_{n+1}^{-}-\mu_{2n}(\overline{U}_n^{+}+\overline{U}_{n+1}^{-})\right]$$

и

$$\overline{U}_n^{+}=\frac{S_n(t+\tau)}{S_n(t-\tau)}\left[\overline{U}_{n-1}^{+}+\mu_{1n-1}(\overline{U}_{n-1}^{+}+\overline{U}_n^{-})\right],$$

$$\overline{U}_{n-1}^{-}=\frac{S_n(t+\tau)}{S_n(t-\tau)}\left[\overline{U}_n^{-}-\mu_{2n-1}(\overline{U}_{n-1}^{+}+\overline{U}_n^{-})\right].$$

Таким образом, окончательные выражения для трансформации бегущих волн имеют вид

$$\overline{U}_{i+1}^{+}=\eta_{i+1}\left[U_i^{+}e^{-g_{1i}(t+\tau)}+\frac{\Delta x}{2}f_i(t+\tau)+\mu_{1i}R_{i,i+1}\right]$$

$$\overline{U}_i^{-}=\eta_i\left[U_{i+1}^{-}e^{-g_{1i+1}(t+\tau)}+\frac{\Delta x}{2}f_{i+1}(t+\tau)-\mu_{2i}R_{i,i+1}\right],$$

где

$$R_{i,i+1}=U_i^{+}e^{-g_{1i}(t+\tau)}+U_{i+1}^{-}e^{-g_{1i+1}(t+\tau)}+\frac{\Delta x}{2}\left[f_i(t+\tau)+f_{i+1}(t+\tau)\right],$$

$$f_i(t+\tau)=\frac{S_i(t+\tau)-S_i(t-\tau)}{2\tau}$$

и $\eta_i=S_i(t+\tau)/S_i(t-\tau)$, за исключением случая, когда $i=n$ или $i=n-1$, где n — номер секции, содержащей голосовую щель. В первом случае вместо η_{n+1} используется η_n, а во втором — вместо η_{n-1} используется η_n.

Если площадь поперечного сечения речевого тракта не меняется во времени, то разница между коэффициентами отражения μ_{1i} и μ_{2i} исчезает и $\mu_i=\mu_{1i}=\mu_{2i}=(S_{i+1}-S_i)/(S_{i+1}+S_i)$.

До сих пор функция $f(t)$ считалась независимой от давления. Применяя метод последовательных приближений, на первом шаге положим $f(t)=0$ и определим давление

$$P(t,x)=\frac{\rho_0 c_0}{S(t,x)}\left[U^{+}(t-x/c_0)e^{-g_1(t+x/c_0)}+U^{-}(t+x/c_0)e^{-g_1(t-x/c_0)}\right]e^{g_2(t,x)}.$$

Затем подставим $P(t,\ x)$ в последнее уравнение системы (1) и найдем смещение стенок. По нему вычислим $f(t)$ и вновь подставим ее в систему (1). Выбор значения координаты x не очень важен, поскольку на про-

цессы колебания стенок влияют лишь низкочастотные компоненты акустического давления, а они мало меняются вдоль оси речевого тракта.

Разрешая уравнения (4) и (5) относительно давления и действуя аналогично описанному выше, получим для бегущих волн давления:

$$\overline{P}_{i+1}^{+}=\nu_i\left[P_i^{+}e^{-g_{2i}(t+\tau)}-\frac{\rho_0 c_0\Delta x}{2S_i(t+\tau)}f_i(t+\tau)\right]+\mu_{1i}T_{i,i+1},$$

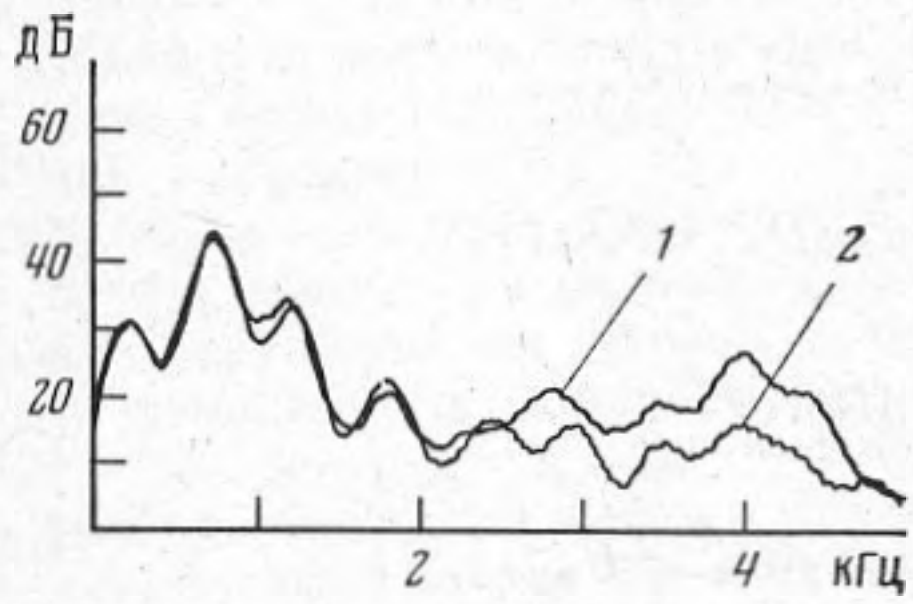
$$\overline{P}_i^{-}=\nu_i\left[P_{i+1}^{-}e^{-g_{2i+1}(t+\tau)}-\frac{\rho_0 c_0\Delta x}{2S_{i+1}(t-\tau)}f_{i+1}(t+\tau)\right]+\mu_{2i}T_{i,i+1},$$

где ν_i, μ_{1i} и μ_{2i} те же, что и раньше, а

$$T_{i,i+1}=P_{i+1}^{-}e^{-g_{2i+1}(t+\tau)}-P_i^{+}e^{-g_{2i}(t+\tau)}+\frac{\rho_0 c_0\Delta x}{2}\left[\frac{f_i(t+\tau)}{S_i(t+\tau)}-\frac{f_{i+1}(t+\tau)}{S_{i+1}(t-\tau)}\right].$$

Как видно, схема для трансформации давления описывается несколько проще. Рекурсивные схемы для давления и объемной скорости имеют разную чувствительность к упрощениям в описании параметров речевого тракта.

Система (1) должна быть дополнена граничными условиями в легких и на губах. Представим упрощенно легкие в виде объема с потерями, имеющего импеданс $Z_л=R_л/S_0^2+1/j\omega C_л$, где $R_л$ — потери ($R_л/S_0^2=1{,}1\div10$), $C_л$ — емкость, $C_л=V_л/\rho_0 c_0^2$, $V_л$ — объем легких ($V_л=3000\div5000$ см^3), S_0 — площадь сечения трахеи. Давление P_0 и объемная скорость U_0 в начале трахеи связаны соотношением $P_0=-Z_л U_0$. Воспользуемся выражением (13), пренебрежем в нем потерями и, принимая во внимание, что умножение на $j\omega$ соответствует однократному дифференцированию во временной области, получим $T_л U_0^{-\prime}+U_0^{-}=G_л U_0^{+\prime}+U_0^{+}$, где $T_л=C_л(\overline{R}_л-\rho_0 c_0/S_0)$, $G_л=C_л(\overline{R}_л+\rho_0 c_0/S_0)$, $\overline{R}_л=R_л/S_0^2$. Переходя к операторной форме, найдем, что передаточная функция легких есть $W_л(p)=U_0^{+}/U_0^{-}=(1+T_л p)/(1+G_л p)$, т. е. представляет параллельное соединение апериодического и реального дифференцирующего звеньев. Эта передаточная функция легко реализуется во временной области.



Спектр синтезированного гласного *А* для статической схемы (*1*) и динамического согласования (*2*)

В работе [5] было показано, что импеданс излучения в области губ $Z_и$ можно представить как параллельное соединение активного и индуктивного сопротивлений $R_и$ и $X_и$: $R_и=(\rho_0 c_0/S_и)128/9\pi^2$, $X_и=(\rho_0 c_0/S_и)8\sqrt{S_и}/3\pi\sqrt{\pi}c_0$, где $S_и$ — площадь излучающего отверстия. Тогда из $P_и=Z_и U_и$, аналогично предыдущему получим $-G_и U_и^{+\prime}+U_и^{+}=-(T_и U_и^{-\prime}+U_и^{-})$, откуда передаточная функция для излучающей секции есть $W_и(p)=U_и^{-}/U_и^{+}=-(1-G_и p)/(1+T_и p)$, где $G_и\approx0{,}5\cdot10^{-5}\sqrt{S_и}$, $T_и\approx2{,}3\cdot10^{-5}\sqrt{S_и}$.

Синтезированные по описанной схеме гласные звуки обладают высокой натуральностью звучания. Эффект динамического согласования граничных условий между секциями показан на фигуре, где видно, что по сравнению со статической системой шумы в высокочастотной области спектра уменьшаются на 10 дБ.

ЛИТЕРАТУРА

1. *Сорокин В. Н.* Теория речеобразования. М.: Радио и связь, 1985.
2. *Kelly J. L., Lochbaum C. C.* Speech sunthesis.— Proc. 4 Intern. Congress on Acoustics, 1962, Paper C42, p. 1–4.
3. *Maeda S.* On a simulation method of dynamically varying vocal tract. Reconsideration of the Kelly – Lochbaum model.— Articulatory Modeling and Phonetics, GALF, Groupe de Communication Parlee, 1977, p. 281–288.
4. *Strube H. W.* Time-varying wave digital filters for modeling analog systems.— IEEE Transactions, ASSP-30, 1982, № 6, p. 864–868.
5. *Фланаган Д.* Анализ, синтез и восприятие речи. М.: Связь, 1968.

Институт проблем передачи информации

Поступила в редакцию 28.11.1985