УДК 517.958:531.33

Академік НАН України **Ю. Г. Кривонос, Ю. В. Крак, І. О. Стеля**

Пряма і обернена задачі моделювання мовного апарату людини

Досліджуються проблеми синтезу мовної інформації для розв'язання задач штучного утворення звуків голосу людини на основі спільного використання фізичних моделей голосового джерела та мовного тракту. Для акустичного рівняння Клейна–Гордона розв'язана обернена задача відновлення параметрів мовного тракту за вимірюваним сигналом на виході.

Запропоновані методи дозволяють синтезувати мовний сигнал шляхом моделювання людського мовного апарату, який включає модель мовного тракту і модель голосових зв'язок. На основі чисельних методів будуються і досліджуються комп'ютерні моделі голосового джерела — для моделювання коливних процесів та моделі мовного тракту — для моделювання поширення акустичних хвиль. Для задачі відновлення параметрів мовного тракту за вимірюваним сигналом на виході використано математичний апарат розв'язання обернених задач.

Модель голосових зв'язок, побудована на основі [1], описує коливання зв'язок двома масами, що зв'язані пружинами як із стінками тракту, так і між собою. Припускається, що зв'язки є двосторонньо симетричними. Маси, що моделюють зв'язки, здійснюють коливання у поперечному до руху повітря напрямку. Система рівнянь для двох мас, що коливаються, записується у вигляді

$$\begin{aligned} m_1\frac{d^2x_1(t)}{dt^2} + r_1\frac{dx_1(t)}{dt} + k_1(x_1(t) - x_{01}) + k_c(x_1(t) - x_2(t)) = l_g d_1 p_{m1}(t),\\ m_2\frac{d^2x_2(t)}{dt^2} + r_2\frac{dx_2(t)}{dt} + k_2(x_2(t) - x_{02}) - k_c(x_1(t) - x_2(t)) = l_g d_2 p_{m2}(t), \end{aligned} \qquad (1)$$

де m_1 і m_2 — маси; $x_1(t)$, $x_2(t)$ — зміщення мас m_1 і m_2; x_{01}, x_{02} — початкове положення мас m_1 і m_2; t — час; r_1 і r_2 — коефіцієнти демпфування; k_1, k_2 — пружність пружин для мас m_1 і m_2; k_c — пружність пружини, що з'єднує маси m_1 і m_2; d_1, d_2 — товщина мас m_1 і m_2; l_g — діюча довжина голосових зв'язок; $l_g d_1$, $l_g d_2$ — поверхні мас m_1 і m_2, на які діють тиски $p_{m1}(t)$ і $p_{m2}(t)$, відповідно.

Розподіл тиску в голосовій щілині апроксимується послідовними дискретними значеннями p_{ij} на кожному j-му кінці кожної i-ї маси, звідки випливає система рівнянь для змін тиску

$$p_s - p_{11}(t) = 0{,}69\rho\frac{u_g^2(t)}{A_{g1}^2(t)} + \int\limits_0^{l_c} \frac{\rho}{A_c(x)}dx\frac{du_g}{dt},$$

$$p_{11}(t) - p_{12}(t) = 12vd_1\frac{l_g^2 u_g(t)}{A_{g1}^3(t)} + \frac{\rho d_1}{A_{g1}}\frac{du_g}{dt},$$

$$p_{12}(t)-p_{21}(t)=\frac{1}{2}\rho u_g^2(t)\left(\frac{1}{A_{g2}^2(t)}-\frac{1}{A_{g1}^2(t)}\right), \tag{2}$$

$$p_{21}(t)-p_{22}(t)=12vd_2\frac{l_g^2u_g(t)}{A_{g2}^3(t)}+\frac{\rho d_2}{A_{g2}}\frac{du_g}{dt},$$

$$p_{22}(t)-p=\frac{1}{2}\rho\frac{u_g^2(t)}{A_{g2}^2(t)}\left[2\frac{A_{g2}(t)}{A_1}\left(1-\frac{A_{g2}(t)}{A_1}\right)\right],$$

де p_s — тиск на вході у голосову щілину; p — атмосферний тиск; ρ — густина повітря; ν — зсувова в'язкість повітря; A_1 — площа голосового тракту на вході; A_{gi} — площа голосової щілини під i-ю масою; $u_g(t)$ — потік повітря; $A_{gi}(t)=(A_{g0i}+2l_gx_i(t))$, $i=1,2$ $(x_1(t)\geqslant x_{01}$, $x_2(t)\geqslant x_{02}$, A_{g01}, A_{g02} — залишкові площі в момент змикання голосових зв'язок).

Для визначення значень тиску $p_{m1}(t)$, $p_{m2}(t)$ використовуються співвідношення

$$p_{m1}(t)=\frac{1}{2}(p_{11}(t)+p_{12}(t)), \qquad p_{m2}(t)=\frac{1}{2}(p_{21}(t)+p_{22}(t)). \tag{3}$$

Шуканим розв'язком системи (1)–(3) є функція $u_g(t)$, яка визначає потік повітря на виході з голосової щілини.

Для реалізації двомасової моделі голосових зв'язок (1)–(3) вперше розроблений удосконалений чисельний метод, в основу якого покладена комбінація методу розв'язання системи рівнянь коливання двох мас і методу розв'язання нелінійної системи для змін тиску. Для перевірки адекваності результатів моделювання проведена серія чисельних експериментів, які показали, що значення вихідних параметрів моделі та розраховані характеристики голосового джерела знаходяться у фізіологічно допустимих межах.

Отриманий за допомогою розробленого методу розв'язок та його похідна використані як голосове джерело для моделей мовного тракту без додаткової обробки [2].

Моделювання мовного тракту людини. Для моделювання поширення акустичних хвиль у мовному тракті як у неоднорідній акустичній трубі, що починається між голосовими зв'язками та закінчується губами, використана система рівнянь акустики в частинних похідних [3]

$$\begin{aligned} -S(x)\frac{\partial p}{\partial x} &= \rho\frac{\partial u}{\partial t}, \\ -\rho c^2\frac{\partial u}{\partial x} &= S(x)\frac{\partial p}{\partial t}, \end{aligned} \tag{4}$$

де $0\leqslant x\leqslant L$, $t>0$, L — довжина мовного тракту; $p(x,t)$ — тиск у тракті в момент часу t; $u(x,t)$ — об'ємна швидкість потоку; ρ — густина повітря в тракті; c — швидкість звуку; $S(x)$ — функція площі поперечного перерізу.

Оскільки тракт має неоднорідний поперечний переріз, він розбивається на циліндричні секції однакової довжини з постійною площею перерізу.

Як крайову умову на вході в тракт вибрано потік $u_g(t)$, знайдений з (1)–(3). На основі цього отримується крайова умова для системи (4): $u(0,t)=u_g(t)$. На протилежному кінці тракту задана умова $p(L,t)=0$.

Різницева задача для апроксимації системи рівнянь (4) побудована на рознесеній сітці. Для розв'язання використовується явний метод "чехарда".

Для моделювання поширення акустичних хвиль у тракті також використане рівняння Вебстера

$$S(x)\frac{\partial^2 P}{\partial t^2} = c^2\frac{\partial}{\partial x}\left(S(x)\frac{\partial P}{\partial x}\right), \tag{5}$$

де x — просторова координата вздовж середньої лінії тракту в середньосагітальній площині; t — момент часу; $p(x,t)$ — шуканий тиск у тракті; $S(x)$ — профіль площі поперечного перерізу вздовж тракту; c — швидкість звуку в тракті.

Крайовою умовою на вході в тракт вибрана похідна від потоку повітря

$$P(0,t) = -\frac{\rho}{S(0)}\frac{du_g(t)}{dt}.$$

Для розв'язання задачі (5) використано скінченнорізницевий метод. Для розв'язання системи різницевих рівнянь застосовано ітераційний метод послідовної верхньої релаксації.

Розглянуто задачу відновлення форми мовного тракту за виміряними акустичними параметрами сигналу на базі акустичного рівняння Клейна–Гордона [4]. Для цього введена нова змінна $\varphi(x,t)$, яка визначається виразом

$$\varphi(x,t) = P(x,t)S(x)^{1/2}. \tag{6}$$

Це дозволяє записати акустичне рівняння у формі Клейна–Гордона:

$$\frac{\partial^2\varphi(x,t)}{\partial t^2} = c^2\frac{\partial^2\varphi(x,t)}{\partial x^2} - c^2U(x)\varphi(x,t), \qquad 0 < x < L, \qquad 0 < t \leqslant T. \tag{7}$$

Рівняння (7) має форму хвильового, де функція $U(x)$ визначена в термінах площі поперечного перерізу мовного тракту як

$$U(x) = \frac{d^2S(x)^{1/2}/dx^2}{S(x)^{1/2}}. \tag{8}$$

Мовна обернена задача — це задача знаходження функції $S(x)$ за виміряними параметрами мовного сигналу на виході з тракту. Математично вона розв'язується як задача пошуку мінімуму функціонала при різного роду обмеженнях.

Нехай на виході з тракту вимірюється тиск $P(L,t)$, пов'язаний з розв'язком рівняння Клейна–Гордона співвідношенням (6). Позначимо $\Phi(t)$ функцію, яка вимірюється на виході з тракту. Тоді обернена мовна задача зводиться до мінімізації функціонала

$$J(U) = \int_0^T (\Phi(t) - \varphi_U(L,t))^2 dt, \tag{9}$$

де $\varphi_U(L,t)$ — розв'язок задачі (7) при заданій функції $U(x)$.

Для мінімізації функціонала (9) використано градієнтний метод. Приріст функціонала записується у вигляді

$$\Delta J(U) = J(U+h) - J(U) = \int_T^{\int} 0(L,t))\Delta\varphi dt + \int_T^{\int} 0^2 dt,$$

де $\Delta\varphi = \varphi_{U+h}(x,t) - \varphi_U(x,t)$. Для визначення градієнта функціонала побудована спряжена задача

$$\frac{\partial^2\Psi(x,t)}{\partial t^2} = c^2\frac{\partial^2\Psi(x,t)}{\partial x^2} - c^2U(x)\Psi(x,t), \qquad 0 < x < L, \qquad 0 < t \leqslant T.$$

Градієнт функціонала визначено через розв'язок спряженої задачі за формулою

$$J^{'} = -\varphi_{\Psi}.$$

Після визначення $U(x)$ функція $S(x)$ знаходиться з (8).

У тестових розрахунках була використана така стратегія перевірки працездатності побудованого алгоритму та створеного програмного забезпечення: а) розв'язувалася пряма задача та визначався сигнал на виході; б) отриманий сигнал використовувався як виміряний при заданому початковому наближенні функції $U(x)$ для розв'язування оберненої задачі.

Точність розв'язку оцінюється за процедурою ресинтезу: синтезований за знайденим розв'язком сигнал має не сильно відрізнятися від вихідного сигналу, за параметрами якого розв'язувалася обернена задача. Результати чисельного моделювання продемонстрували адекватність і ефективність розробленого підходу до розв'язування оберненої мовної задачі.

Таким чином, у даному повідомленні наведені нові результати для розв'язування задачі комп'ютерного відтворення звуків, що утворює людина на основі спільного використання фізичних моделей голосового джерела та мовного тракту. Сформульовано і розв'язано обернену задачу відновлення параметрів мовного тракту за вимірюваним сигналом на виході. Запропонований підхід має практичну цінність для побудови артикуляторного синтезатора мови.

1. *Ishizaka K., Flanagan J. L.* Synthesis of voiced sounds from a two-mass model of vocal cords // Bell Syst. Tech. J. – 1972. – **51/(6)**. – P. 1233–1268.
2. *Крак Ю. В., Стеля І. О.* Чисельне моделювання голосових зв'язок за двомасовою моделлю // Журн. обчислюв. та прикл. математики. – 2007. – № 94. – С. 55–60.
3. *Kinsler L. E., Frey A. E., Coppens A. B., Saunders J. V.* Fundamentals of acoustics. – San Diego: Academic Press, 1982. – 496 p.
4. *Forbes B. J., Pike E. R, Sharp D. B.* The acoustical Klein–Gordon equation: The wave-mechanical step and barrier potential functions // J. Acoust. Soc. Am. – 2003. – **114/(3)**. – P. 1291–1302.

Інститут кібернетики
ім. В. М. Глушкова НАН України, Київ

Надійшло до редакції 21.02.2011

Academician of the NAS of Ukraine **Yu. G. Krivonos, Iu. V. Krak, I. O. Stelia**

The direct and inverse problems of the human speech apparatus modeling

The problems of speech information synthesis for solving the issue of artificial formation of human voice sounds based on the common use of the physical models of voice source and vocal tract are studied. For the acoustic Klein–Gordon's equation, the inverse problem of recovering the parameters of a vocal tract according to the measured output signal is solved.