УДК 517.17:681.518.54

РАСПОЗНАВАНИЕ ИСКАЖЕННОГО РЕЧЕВОГО СИГНАЛА С ПОМОЩЬЮ ФУНКЦИЙ НЕПРОПОРЦИОНАЛЬНОСТЕЙ

В.В. Авраменко, доц.; Т.Н. Усатенко, ассист.
Сумский государственный университет

В настоящее время известны различные методы автоматического распознавания речи [1, 2, 3, 4]. При передаче речевого сигнала по каналам связи на него обычно накладываются различного рода помехи. В настоящее время эти помехи достаточно хорошо изучены и существует много способов для их фильтрации [5].

Однако при передаче речевого сигнала он также может искажаться. В частности, сигнал искажается при прохождении через нелинейные элементы. Так, например, нелинейные искажения обусловлены нелинейными вольтамперными характеристиками транзисторов и характеристик намагничивания магнитопроводов трансформаторов. В результате, например, усилитель звуковых частот вносит некоторые нелинейные искажения, и его амплитудная характеристика описывается выражением [6]

$$u_{вых} = a_1 u_{вх} + a_2 u_{вх}^2 + a_3 u_{вх}^3 + \ldots + a_m u_{вх}^m , \qquad (1)$$

где

$u_{вх}, u_{вых}$ – амплитуды соответственно входного и выходного сигналов;

$a_1, a_2, a_3, \ldots, a_m$ – коэффициенты, определяемые видом амплитудной характеристики;

m – порядок нелинейности, определяющий число гармоник на выходе усилителя.

Выражением (1) также описываются характеристики модулятора и дискриминатора, используемые при передаче звукового сигнала через радиоканал [7].

При малых значениях амплитуды речевого сигнала на входе нелинейного элемента его распознавание по сигналу на выходе является сложной задачей. В принципе для этого следует знать значения коэффициентов $a_1, a_2, a_3, \ldots, a_m$, входящих в (1). Однако, как правило, эти коэффициенты зависят от многих неконтролируемых факторов и носят случайный характер. Их значения чаще всего неизвестны.

Обычно в системах распознавания речи используется множество эталонных функций, с помощью которых анализируется принимаемый сигнал.

В данной работе в качестве эталонов используется множество числовых функций времени, представляющих отдельные слоги русского языка. Рассматривается случай, когда на удаленном передающем конце, недоступном для контроля, произносится слово. Соответствующий речевой сигнал передается через канал связи с нелинейной характеристикой вида (1). Необходимо по сигналу на выходе канала распознать это слово при неизвестных $a_1, a_2, a_3, \ldots, a_m$ в (1). Считается, что помехи отсутствуют, например, вследствие их фильтрации.

МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

Дано множество эталонных функций $f_i(t)$ времени t, где $t \in [0, T_i]$, $i = 1, 2, \ldots, q$. Известны текущие значения анализируемого сигнала $y(t)$, $t \in [0, T]$, описываемого выражением

$$y(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \ldots + a_m x^m(t), \tag{2}$$

где

$$x(t) = k_i f_i(t), i = 1, 2, \ldots, q. \tag{3}$$

Значения постоянных коэффициентов $a_1, a_2, a_3, \ldots, a_m$ и k_i ($i = 1, 2, \ldots, q$) неизвестны. Порядок полинома m известен. Функции $y(t)$ и $f_i(t)$ ($i = 1, 2, \ldots, q$) гладкие и имеют необходимые производные. Необходимо в момент времени t по текущим значениям функции $y(t)$ и ее производных распознать эталонную функцию на входе нелинейного элемента (2), (3) и получить восстановленный сигнал $v(t)$ в виде последовательности значений распознанных эталонных функций.

РЕШЕНИЕ ЗАДАЧИ

Для решения задачи предлагается использовать разработанные в [8, 9] характеристики непропорциональностей числовых функций. Конкретно будет использоваться непропорциональность по производной первого порядка для функций, заданных параметрически. Эта непропорциональность функции $y(t)$ по $x(t)$ имеет вид

$$@ d_{x(t)}^{(1)} y(t) = \frac{y(t)}{x(t)} - \frac{dy/dt}{dx/dt}. \tag{4}$$

Здесь @ – символ вычисления непропорциональности; d – от англ. derivative – производная.

Если

$$y(t) = kx(t), \tag{5}$$

то непропорциональность (4) равняется нулю независимо от значения коэффициента k в (5).

Кроме непропорциональности (4), будет использована m-непропорциональность по производной первого порядка. m-непропорциональность – это последовательное (m-раз) определение непропорциональности по производной первого порядка для $y(t)$ по $x(t)$. Например, $@(3) @ d_{x(t)}^{(1)} y(t)$ означает $@ d_{x(t)}^{(1)} [@ d_{x(t)}^{(1)} [@ d_{x(t)}^{(1)}]]$.

Доказано [8], что в случае, если зависимость $y = f(x, t)$ имеет вид полинома степени m (2), то m – непропорциональность по производной первого порядка функции $y(t)$ по $x(t)$ равняется нулю независимо от значений, входящих в (2) коэффициентов. Именно это свойство m-непропорциональности предлагается использовать для решения задачи.

АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ

1 Считывается текущее значение $y(t)$ и определяется ее производная.

2 Рассматривается первая эталонная функция $i = 1$.

3 Устанавливается сдвиг во времени $\tau = 0$, эталонной функции $f_i(t)$ по отношению к $y(t)$.

4 Вычисляется m-непропорциональность $z(t)$ функции $y(t)$ по $f_i(t+\tau)$.

5 Проверяется условие $z(t) = 0$. Если оно выполняется, то восстанавливаемому сигналу $v(t)$ присвоить $f_i(t+\tau)$ и перейти к пункту 8. Иначе перейти к пункту 6.

6 Увеличить τ на $\Delta\tau$.

7 Проверить условие $\tau \le T_i$. Если оно выполняется, перейти к пункту 4, иначе – к пункту 8.

8 Перейти к следующей эталонной функции ($i = i + 1$).

9 Проверить условие $i \le q$. Если условие выполняется, перейти к пункту 3. Иначе распознавание не состоялось. В этом случае присвоить $v(t)$ =0 и перейти к пункту 10.

10 Увеличить текущее значение времени t на Δt, т.е. $t = t + \Delta t$

11 Проверить, достигнут ли конец сигнала $y(t)$. Если конец достигнут, выполнить переход к пункту 12, иначе – к пункту 2 .

12 Вывод значений восстановленного сигнала $v(t)$. Остановить процесс.

Рассмотрим пример. Рассматривается случай, когда имеются две эталонных функции: $f_1(t)$, которая соответствует слогу «*при*» (рис. 1), и $f_2(t)$, которая соответствует слогу «*вет*» (рис. 2).

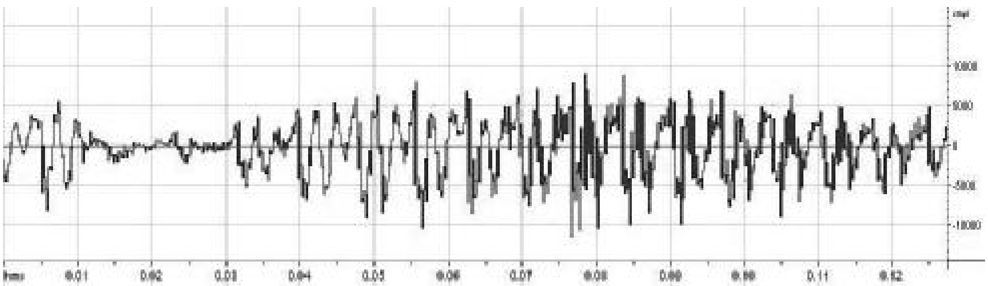

Рисунок 1 – Эталонная функция $f_1(t)$, слог «при»

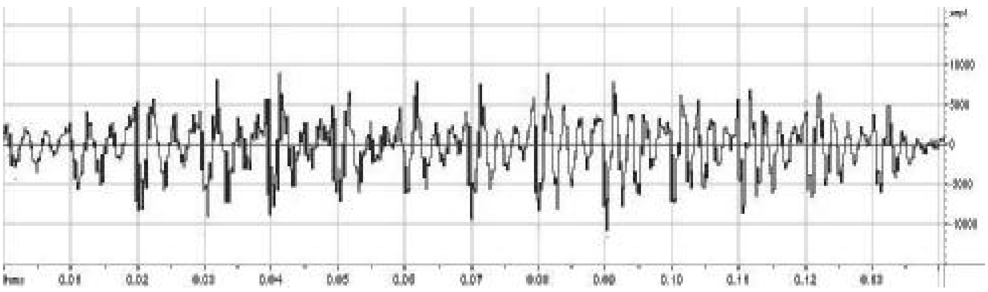

Рисунок 2 – Эталонная функция $f_2(t)$, слог «вет»

Характеристика нелинейного элемента описывается полиномом второго порядка. Сигнал на его выходе описывается выражением

$$y(t) = a_2 x^2(t) + a_1 x(t)\,, \tag{6}$$

где $x(t)$ имеет вид (3); a_2 =0.002; a_1 = 4.

На рисунках 3 и 4 представлены графики анализируемого искаженного $y(t)$ и неискаженного сигналов соответственно. Между ними наблюдаются значительные различия.

Для восстановления сигнала представленные в дискретные моменты времени функции $f_1(t)$, $f_2(t)$ и $y(t)$ предварительно были сглажены с помощью кубических сплайнов.

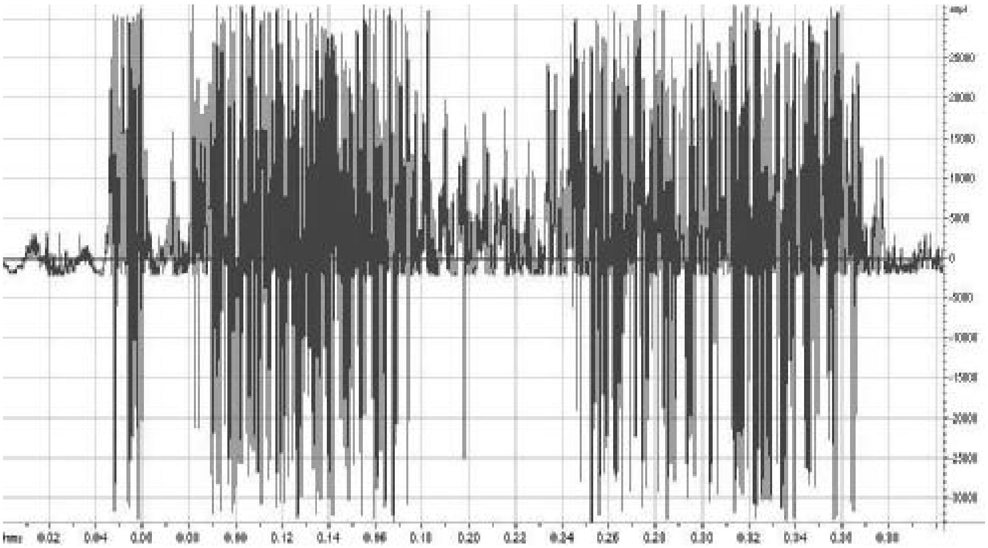

Рисунок 3 – Анализируемый искаженный сигнал $y(t)$

Для каждого момента времени t вычислялась m-непропорциональность $@(2)@d^{(1)}_{f_i(t)}y(t)$ по производной первого порядка функции $y(t)$ по $f_i(t+\tau)$ при $\tau = 0$ ($i = 1, 2, \dots$). Для этого производились следующие вычисления:

$$z_1(t) = @d^{(1)}_{f_i(t+\tau)}y(t) = \frac{y(t)}{f_i(t+\tau)} - \frac{\frac{d}{dt}[y(t)]}{\frac{d}{dt}[f_i(t+\tau)]}, \tag{7}$$

$$z(t) = @d^{(1)}_{f_i(t+\tau)}z_1(t) = \frac{z_1(t)}{f_i(t+\tau)} - \frac{\frac{d}{dt}[z_1(t)]}{\frac{d}{dt}[f_i(t+\tau)]} \tag{8}$$

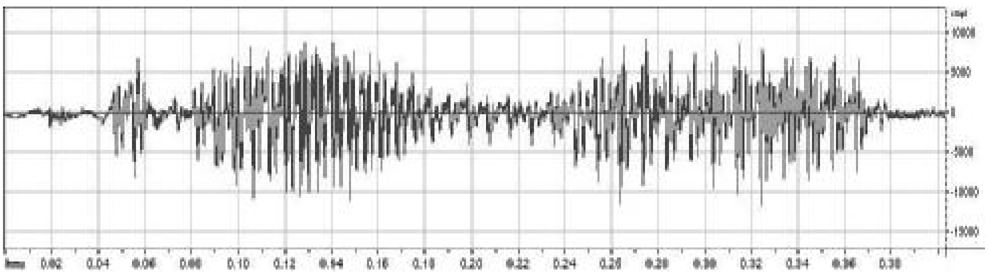

Рисунок 4 – Неискаженный сигнал «привет»

На рисунке 5 приведен график восстановленного $v(t)$ -сигнала.

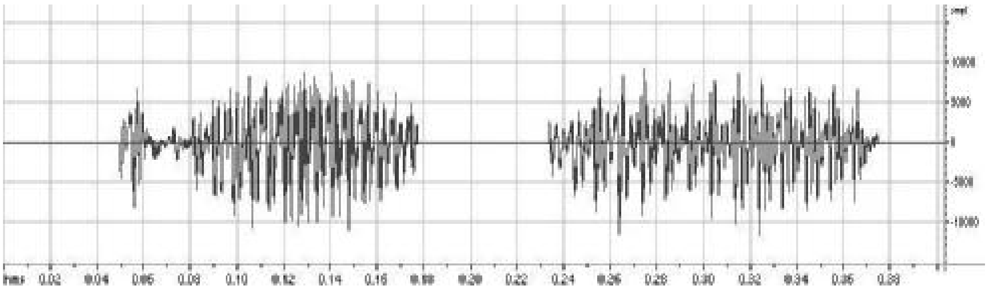

Рисунок 5 –Восстановленный сигнал $v(t)$

Для лучшего сопоставления восстановленного $v(t)$ и исходного $x(t)$ сигналов на рисунке 6 приведены совмещенные фрагменты этих сигналов на отрезке времени $t \in [0.375,\ 0.1825]$, ($300 \le i \le 1460$). Видно, что в моменты времени $t \in [0.05,\ 0.1775]$, ($400 \le i \le 1420$), когда эталон $f_1(t)$ входит в исследуемый сигнал $y(t)$, восстановленный и исходный сигнал полностью совпадают. На интервалах времени $t \in [0.375,\ 0.05)$, ($300 \le i < 400$) и $t \in (0.1775,\ 0.1825]$, ($1420 < i \le 1460$) (рис. 6) $v(t) = 0$, так как сигнал не был распознан.

Совпадение сигналов $v(t)$ и $x(t)$ так же происходит и в моменты времени $t \in [0.2338,\ 0.3745]$, ($1870 \le i \le 2996$), когда в исследуемый сигнал $y(t)$ входит эталонный сигнал $f_2(t)$.

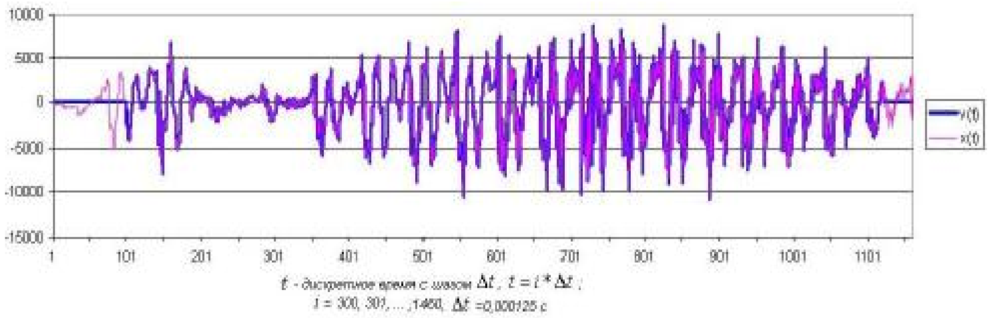

Рисунок 6 – Фрагменты восстановленного $v(t)$ и исходного сигналов $x(t)$

Следует также отметить, что при прослушивании искаженного сигнала $y(t)$ раздается просто треск. В то же время при прослушивании восстановленного $v(t)$ -сигнала отчетливо слышно слово «*привет*». Можно сделать вывод о том, что состоялось практически полное распознавание анализируемого сигнала и его восстановление.

Таким образом, при отсутствии помех и одинаковой скорости записи эталонных функций и анализируемого сигнала $y(t)$ предлагаемый алгоритм позволяет осуществить распознавание сигнала, подаваемого на недоступный для контроля вход, и восстановить его.

SUMMARY

The article concerns problem of the recognition of the speech signal, which distorted by nonlinear element with polynomial feature. Algorithm of the decision of this problem is offered by means of functions of the disproportions on derived first-order for function given parametric.

СПИСОК ЛИТЕРАТУРЫ

1. Васильев В.И. Распознающие системы. –К.: Наукова думка, 1983.
2. Садовой Н.Н., Калмыков О.С. Экспериментальное обоснование возможности выделения фонем гласных звуков в слитной речи // Вестник ДГТУ. - 2002. - Т.1, №2(8).
3. Вінцюк Т. Усний словник-перекладач. / Матеріали III Всеукраїнської Міжнародної конференції у Києві. - Київ, 1996. - С.121-123.
4. Анализ, распознавание и интерпретация речевых сигналов/ Т.К. Винцюк –К.: Наукова думка, 1987.
5. Дьяконов В., Абраменкова И. MATLAB. Обработка сигналов и изображений: Специальный справочник. – СПб.: Питер, 2002.
6. Терещук Р.М., Терещук К.М., Седов С.А. Полупроводниковые приемно-усилительные устройства (справочник радиолюбителя). –К.: Наукова думка, 1989.
7. Тепляков И.М. Радиотелеметрия – М.: Советсткое радио, 1966
8. Авраменко В.В. Характеристики непропорциональности числовых функций - Деп. В ГНТБ Украины, 19.01.98 №59 – Ук98.
9. Авраменко В.В. Характеристики непропорциональности числовых функций и их применения при решении задач диагностики // Вісник Сумського державного університету. - 2000. –№16.

Поступила в редакцию 15 декабря 2006 г.???