

В.Ю. Иванов, С.В. Маркова

V.Y. Ivanov, S.V. Markova

**Моделирование истинных значений локальных максимумов амплитудного спектра
в системах обработки речевой информации**

**Modeling the true values of the local maxima of the amplitude spectrum
in the speech processing systems**

Аннотация, abstract: В статье описывается математический способ моделирования истинных значений локальных максимумов амплитудного спектра звукового сигнала для последующего корректного синтеза звука по их значениям.

The article describes a mathematical method of modeling the true values of the local maxima of the audio signal's amplitude spectrum for subsequent correct sound synthesis by their values.

Ключевые слова, keywords: моделирование, амплитудный спектр, локальные максимумы амплитудного спектра, преобразование Фурье, modelling, amplitude spectrum, local maxima of the amplitude spectrum, Fourier transformation

Авторы, authors: Иванов Вячеслав Юрьевич – Московский Университет МВД России им. В.Я. Кикотя, кандидат технических наук, доцент кафедры специальных информацион-

ных технологий учебно-научного комплекса информационных технологий.

Маркова Светлана Владимировна – Московский Университет МВД России им. В.Я. Кикотя, доцент кафедры специальных информационных технологий учебно-научного комплекса информационных технологий.

Ivanov, Vyacheslav Y. – Moscow University of the Ministry of the Interior of the Russian Federation named after V.Y. Kikot, Russia, Moscow; PhD, associate professor.

Markova, Svetlana V. – Moscow University of the Ministry of the Interior of the Russian Federation named after V.Y. Kikot, Russia, Moscow; PhD, associate professor.

УДК 004.942

Статья поступила в редакцию: 05.08.2014

Статья принята к печати: 03.09.2014

© В.Я. Иванов, С.В. Маркова, 2014

Для систем повышения разборчивости речи и систем шумоочистки достаточно часто приходится сталкиваться с задачами синтеза фонограммы. Первоначальный сигнал представляется в виде сонограммы, а затем, используя принципы работы графического редактора, в нем убираются нежелательные помехи. После окончания процесса шумоочистки, как правило, синтезируют новую фонограмму на основе отредактированного изображения сонограммы.

В основе синтеза фонограммы ключевым моментом является поиск локальных максимумов амплитудного спектра звукового сигнала. Точность их нахождения принципиальна для качественного синтеза фонограммы.

Рассмотрим способ математического моделирования их истинного значения на примере мгновенного амплитудного спектра двух гармоник.

По логарифмическому амплитудному спектру находим частоту локального максимума w_{max} . Так как логарифмический амплитудный спектр является дискретной функцией по частоте и имеет разрешающую способность (Δf) ,

$$\Delta f = \frac{F_{АЦП}}{NFFT},$$

где $F_{АЦП}$ – частота работы АЦП, а $NFFT$ – база Фурье преобразования, то точность вычисления w_{max} будет равна $\Delta f / 2$, что является недостаточным.

Рассмотрим следующий пример. На рисунке представлен мгновенный спектр двух гармоник, отстоящих друг от друга на достаточно большом расстоянии, чтобы не влиять друг на друга.

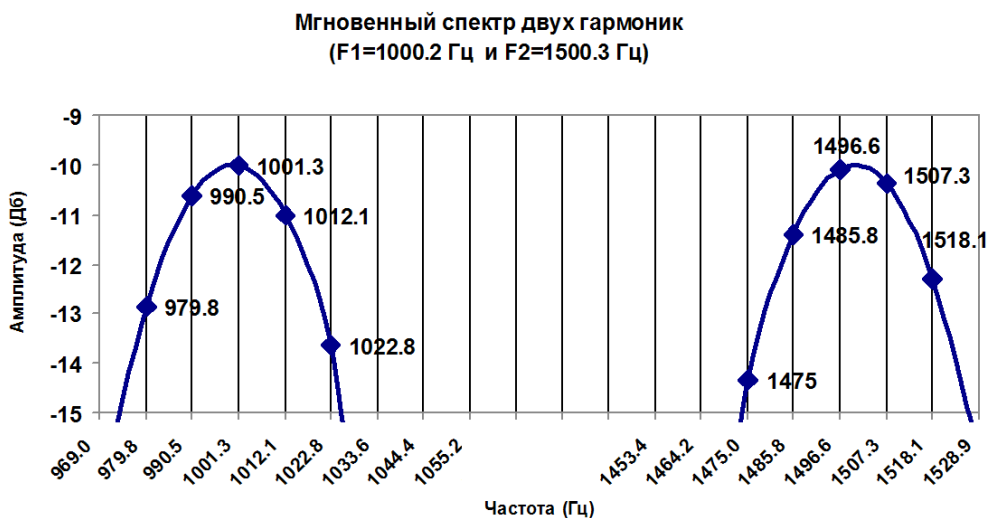


Рисунок 1. Логарифмический амплитудный спектр двух гармоник.

Истинная частота первой гармоники $F1=1000.2$ Гц, а второй гармоники $F2=1500.3$ Гц. Частота работы АЦП $F_{\text{АЦП}} = 11025$ Гц, преобразование Фурье имеет базу $NFFT = 1024$ точки. Следовательно, шаг по частоте логарифмического амплитудного спектра $\Delta f = 10.766$ Гц.

После преобразования Фурье с исходными условиями получаем, что $F1=1001.3$ Гц, а $F2=1496.6$ Гц, что несколько отличается от истинной частоты исследуемых гармоник.

Для точного нахождения частоты локальных максимумов можно предложить следующий способ, основанный на нахождении вершины параболы по три точки в окрестности локального максимума.

Уравнение параболы через ее вершину X_0 выглядит следующим образом:

$$y = a(x - X_0)^2 + c;$$

где a и c – некоторые коэффициенты.

Для трех точек параболы $P_1(X_1, Y_1)$, $P_2(X_2, Y_2)$ и $P_3(X_3, Y_3)$ можно записать следующую систему уравнений:

$$\begin{cases} Y_1 = a(X_1 - X_0)^2 + c; \\ Y_2 = a(X_2 - X_0)^2 + c; \\ Y_3 = a(X_3 - X_0)^2 + c; \end{cases}$$

откуда, путем линейного преобразования найдем значение X_0 , выраженное через a .

$$X_0 = \frac{a(X_2^2 - X_1^2) - Y_2 + Y_1}{2a(X_2 - X_1)},$$

a , в свою очередь, находим следующим образом:

$$a = \frac{(Y_2 - Y_3)(X_2 - X_1) - (X_3 - X_2)(Y_1 - Y_2)}{(X_2^2 - X_3^2)(X_2 - X_1) + (X_3 - X_2)(X_2^2 - X_1^2)}.$$

Используя этот метод можно с высокой степенью точности найти истинное значение частоты локальных максимумов исследуемого сигнала и, следовательно, w_{max} . Эта точность будет ограничена лишь точностью представления чисел с плавающей точкой в компьютере.