

А. М. Г р и ш и н (Москва, ТВП). **Некоторые соотношения между действительной и мнимой частями спектра Фурье речевого сигнала.**

Представление сигналов или других функций с помощью суммы комплексных экспонент часто приводит к удобным решениям и позволяет глубже понять суть изучаемых явлений. Преобразование Фурье традиционно играет ведущую роль в исследованиях и технике передачи речи. При этом, как правило, рассматриваются различные способы оценки и представления амплитудного и реже фазового спектра речевого сигнала (РС) [1]. Вместе с тем, спектральное представление можно рассматривать как комплекснозначную функцию, состоящую из действительной и мнимой частей. Применительно к РС можно поставить следующие вопросы: насколько действительная и мнимая части спектра сохраняют информацию об исходном сигнале; насколько «близкими» являются мгновенные и средние значения действительных и мнимых частей спектра.

Для ответа на эти и сопутствующие им вопросы были выбраны 4 речевых фрагмента (паузы удалены) на русском, английском и арабском языках общим объемом ≈ 30 мин. Каждая временная функция $x(t)$, в дискретном представлении $\{x[n] = x_n\}$, на основе дискретного преобразования Фурье (ДПФ) переводилась в спектральное представление

$$X_n(e^{iw}) = \text{Re} \{X_n(e^{iw})\} + i \text{Im} \{X_n(e^{iw})\}. \quad (1)$$

Из (1) посредством обратного ДПФ синтезировались временные функции $x_{\text{Re}}[n]$ и $x_{\text{Im}}[n]$, у которых $\text{Im} \{X_n(e^{iw})\}$ и $\text{Re} \{X_n(e^{iw})\}$ равны 0, соответственно.

В качестве статистических характеристик, по значениям которых можно судить о мере сходства и различия функций $x_{\text{Re}}[n]$ и $x_{\text{Im}}[n]$, рассматривались следующие характеристики.

1. Взаимные корреляционные функции

$$\text{Cor}_1[k] = x \text{ cov} (x_{\text{Re}}[n_1], x[n_2]), \quad \text{Cor}_2[k] = x \text{ cov} (x_{\text{Im}}[n_1], x[n_2]),$$

$$\text{Cor}_3[k] = x \text{ cov} (x_{\text{Re}}[n_1], x_{\text{Im}}[n_2]),$$

где $k = n_1 - n_2$ ($x \text{ cov}$ — обозначение используемой для вычислений функции системы MatLAB).

2. Математическое ожидание $\mathbf{M}e$, дисперсия $\mathbf{D}e$ и гистограмма плотности распределения нормированной энергии $e[n]$ функции $x_{\text{Re}}[n]$,

$$e[n] = \sum_{k=n-M/2}^{n+M/2} x_{\text{Re}}^2[k] \bigg/ \sum_{k=n-M/2}^{n+M/2} x^2[k],$$

где M — ширина временного окна при вычислении ДПФ.

3. Математическое ожидание $\mathbf{M}\mu$, дисперсия $\mathbf{D}\mu$ и гистограмма плотности распределения меры близости $\mu[n]$ мгновенных значений спектра функций $x_{\text{Re}}[n]$ и $x_{\text{Im}}[n]$,

$$\mu[n] = \int_0^\Pi \left(\left| X_n^{\text{Re}}(e^{iw}) \right| - \left| X_n^{\text{Im}}(e^{iw}) \right| \right) dw \bigg/ \int_0^\Pi \left(\left| X_n^{\text{Re}}(e^{iw}) \right| + \left| X_n^{\text{Im}}(e^{iw}) \right| \right) dw.$$

Оценки введенных выше величин не зависят от языка. В табл. приведены некоторые числовые данные, которые позволяют сделать следующие выводы.

Таблица.

$\text{Cor}_1[0] \approx \text{Cor}_2 \approx 0,7$	$\mathbf{M}(e) \approx 0,5002$	$\mathbf{M}(\mu) \approx 0,53$
$\max_{-80 \leq k \leq 80} \{\text{Cor}_3[k]\} < 0,1$	$\mathbf{D}(e) \approx 0,0005$	$\mathbf{D}(\mu) \approx 0,03$

1. Качество РС, синтезированного только по действительной (мнимой) части спектра, снижается примерно на 30%, но разборчивость сохраняется.

2. С интегральной точки зрения (энергии) действительная и мнимая части частотного разложения РС являются полностью равноправными, но на различных гармониках могут быть существенные отличия.

СПИСОК ЛИТЕРАТУРЫ

1. *Рабинер Л. Р., Шафер Р. В.* Цифровая обработка речевых сигналов. М.: Радио и связь, 1981, 495 с.