

ИЗМЕРЕНИЕ НЕКОТОРЫХ ХАРАКТЕРИСТИК ГЛАСНЫХ ЗВУКОВ

В. П. БОНДАРЕНКО, В. А. БОНДАРЬ

(Представлена кафедрой вычислительной техники)

Как известно [1, 2, 3], энергия в спектре речевого сигнала гласных звуков в основном концентрируется в области формант. Следовательно, энергетический спектр сигнала можно представить в виде суммы

$$S(\omega) = \sum_{i=1}^n S_i(\omega), \quad (1)$$

где $S_i(\omega)$ — огибающая энергетического спектра i -й форманты, n — число формант.

Можно считать [1, 2, 3], что ширина спектра i -й форманты не превосходит одной октавы. В этом случае выражение (1) для конкретного сигнала можно записать в виде [4]

$$f(t) = \sum_{i=1}^n f_i(t) = \sum_{i=1}^n A_i(t) \sin \int \omega_i(t) dt. \quad (2)$$

Для получения суммы (2) достаточно пропустить сигнал через систему фильтров. Параметры фильтров для неискаженной передачи амплитуды $A_i(t)$ и частоты $\omega_i(t)$ должны удовлетворять условию квазистационарности [5, 6], и в то же время добротность их не должна быть меньше единицы [4].

Таким образом, при анализе речевых сигналов гласных звуков достаточно знать функции $A_i(t)$ и $\omega_i(t)$. Верхняя частотная составляющая функции $A_i(t)$ на выходе фильтра, передаваемая им без искажения, определяется полосой пропускания $\Delta\omega_i$ его на уровне 0,7, а верхняя частота функций определяется из выражения [5, 6]

$$\Delta\omega_i \geq 5 \sqrt{\Delta\omega \Omega_i}, \quad (3)$$

где $\Delta\omega$ — девиация частоты сигнала $f_i(t)$,

Ω_i — верхняя частотная составляющая функция $\omega_i(t)$.

При малой девиации частоты, как следует из (3), фильтр может передать без искажений довольно широкий спектр частотно-модулирующей функции.

Для гласных звуков, как отмечают ряд авторов [1, 2, 4, 7], имеет место изменение частоты формант, т. е. девиация несущей частоты. В ряде работ [1, 7] предполагается, что спектр закона девиации частоты довольно широк, в то же время в работе [4] утверждается, что он ограничивается частотой $50 \div 60$ гц.

Как известно [8], при малом индексе частотной модуляции спектр сигнала по ширине равен удвоенной ширине спектра модулирующей функции, а амплитуда боковых частот пропорциональна половине индекса модуляции и амплитуде несущей. Ширина спектра сигнала, измеренного на уровне 0,7, может оказаться узкой и определяться в основном медленными изменениями в амплитуде несущей частоты. Общая же ширина спектра на более низком уровне будет довольно широка, что имеет место для спектров гласных звуков [3].

Для проверки этого положения были проведены следующие изменения параметров сигналов гласных звуков. Сигнал пропускался через фильтры с резонансными частотами 800 и 1600 гц и добротностью, равной единице. На выходе фильтров сигнал клиппировался и затем измерялся каждый временный интервал между двумя нулями.

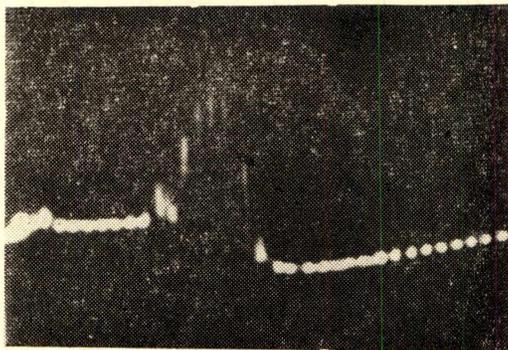


Рис. 1

Число одинаковых временных интервалов запоминалось и затем, после окончания звука, регистрировалось. Таким образом, на выходе регистрирующего прибора получалось распределение периодов по длительности или по частоте.

На рис. 1 представлено типичное распределение мгновенных частот для первой форманты звука «А». Каждой точке соответствует определенная частота. Значение частот по номерам точек слева направо представлено в таблице.

Таким образом, девиация частоты составляет примерно $200 \div 300$ гц, при средней частоте около 700 гц, т. е. девиация частоты составляет $25 \div 30\%$, что подтверждает возможность существования широкого спектра модулирующей функции. Естественно ожидать, что в законе частотной модуляции может содержаться полезная информация о фонеме.

Т а б л и ц а

№	1	2	3	4	5	6	7	8	9	10
<i>f</i> , кгц	2	1,91	1,75	1,6	1,45	1,37	1,26	1,19	1,12	1,06
№	11	12	13	14	15	16	17	18	19	20
<i>f</i> , кгц	0,96	0,87	0,83	0,71	0,65	0,61	0,57	0,55	0,53	0,51

С целью проверки этого предположения были проведены измерения закона частотной модуляции первой форманты.

Временный интервал между двумя нулями фильтрованного и клиппированного сигнала преобразовался в амплитуду пилообразного напряжения. Полученное напряжение и фильтрованный сигнал регистрировались шлейфовым осциллографом на киноплёнке со скоростью 1 м/сек.

На рис. 2 приведены одни из осциллограмм для фонем «А» (верхние кривые) и «О» (нижние кривые). Было замечено, что закон изме-

нения частоты для гласных звуков за период основного тока в общем повторяется и достаточно хорошо совпадает примерно в $50 \div 70\%$ случаев для разных дикторов. Имеется значительное отличие в законах изменения частоты от фонемы к фонеме.

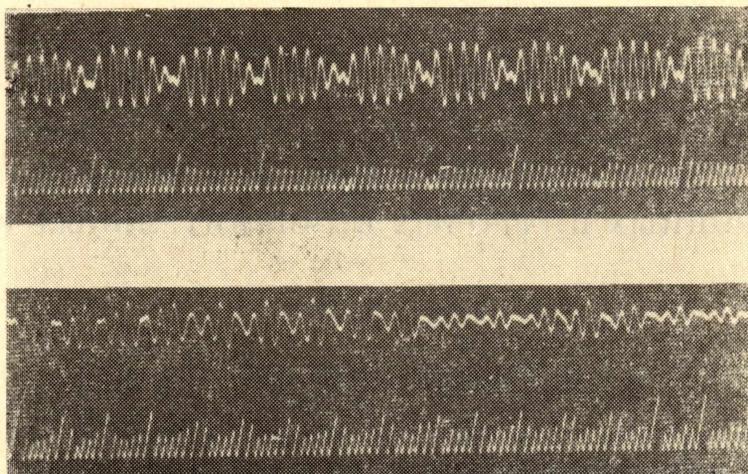


Рис. 2

Невысокая точность измерения, что в основном объясняется дискретностью отсчетов, и малое число ($5 \div 6$ на фонему) дикторов не позволяют сделать окончательного вывода.

Из приведенных исследований видно, что в речевом сигнале частота формант изменяется во времени даже за период основного тона и закон изменения частоты несет некоторую информационную нагрузку о фонеме. Это значит, что при таком подходе к анализу сигнала фильтры должны иметь низкую добротность и постоянный наклон фазовой характеристики, чтобы не исказить законы частотной и амплитудной модуляции. Отмеченные требования к фильтрам, в свою очередь, совпадают с параметрами базилярной мембраны [2].

ЛИТЕРАТУРА

1. М. А. Сапожков. Речевой сигнал в кибернетике и связи. Связьиздат, 1963.
2. Дж. Флэган. Анализ, синтез и восприятие речи. Изд. «Связь», 1968.
3. Е. Т. Семенова. Анализ признаков распознавания речи. Сб. «Бионика», «Наука», 1965.
4. И. Т. Турбович, О. Н. Петров. Об одном методе полного описания одномерных образов совокупностью простых функций (применительно к речевым сигналам). Сб. «Опознавание образов. Теория передачи информации», Изд. АН СССР. Институт проблем передачи информации, 1965.
5. А. С. Винницкий. Модулированные фильтры и следящий прием ЧМ. Изд. «Сов. радио», 1969.
6. С. В. Персон. Нелинейные искажения при усилении ЧМ колебаний. «Радиотехника», 6, № 6, 1951.
7. Calgwell P. Smith, Vokal Respons Synthesizer. The Journal of the Acoustical Society of America. V. 37, № 1, 1965.
8. Н. С. Гоноровский. Радиотехнические цепи и сигналы. Ч. I. Изд. «Сов. радио», 1967.