

## К ВОПРОСУ О ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ АППАРАТОВ «ИСКУССТВЕННАЯ ГОРТАНЬ» ДЛЯ РЕЧЕВЫХ ИССЛЕДОВАНИЙ

Н. Ф. КУЧЕР, В. М. РАЗИН

(Представлена кафедрой вычислительной техники)

Аппараты «Искусственная гортань» (ИГ) были разработаны в качестве протезов для людей, потерявших по какой-либо причине голосовые связки, например, в связи с хирургической операцией. Они позволяют таким людям восстановить возможность общения посредством речевой связи.

В настоящее время имеется несколько разновидностей аппаратов ИГ, которые подразделяются на два класса: вибрационные и звукоизлучающие.

Основным элементом вибрационных аппаратов является механический вибратор, колеблющийся с заданной частотой основного тона, который прикладывается к определенному участку шеи и таким образом возбуждает речеобразующий тракт.

В звукоизлучающих аппаратах колебания основного тона генерируются в виде импульсного акустического давления, которое создается специальным звукоизлучающим элементом типа электродинамического или электромагнитного телефона, включенного на выходе импульсного генератора, и по специальному звукопроводу вводятся в речевую полость говорящего.

Создаваемая с помощью аппаратов ИГ искусственная речь по качеству и громкости приближается к естественной.

В Советском Союзе серийно выпускается только один голосообразующий аппарат вибрационного типа — АГ-61, построенный по принципу зуммера. Несмотря на сравнительную простоту аппарата, речь, получаемая с его помощью, обладает хорошим качеством.

Из зарубежных аппаратов вибрационного типа известен аппарат фирмы Wescon (США), описанный в [1]. Вибратор в этом аппарате приводится в действие электронным генератором прямоугольных колебаний, выполненным на транзисторах.

Звукоизлучающих аппаратов Советский Союз серийно в настоящее время не выпускает. Однако имеются разработанные и подготовленные к серийному выпуску аппараты ИГ типа «Голос-15», «Голос-18», «Голос-21», которые характеризуются следующими данными:

1. Частота основного тона — 80—150 *гц*.
2. Длительность импульса основного тона — 0,3—0,6 *мсек*.
3. Средний уровень громкости — 70 *дб* на расстоянии 1 *м*.

Кроме того, аппарат «Голос-21» позволяет генерировать шумовые звуки со средним уровнем громкости 50—60 *дб*.

Рассмотрим процесс образования речи с помощью аппаратов «Искусственная гортань». Согласно акустической теории речеобразования, наиболее полно разработанной Г. Фантом [2], процесс образования звуков речи происходит следующим образом (рис. 1 служит для иллюстрации этого процесса):

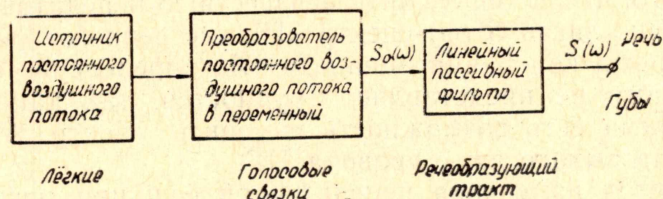


Рис. 1. Блок-схема, иллюстрирующая процесс образования звуков речи

Легкие посылают постоянный воздушный поток к голосовым связкам, которые в зависимости от вида звука преобразуют этот поток или же в почти периодические колебания, или в белый шум. Колебания, излучаемые голосовыми связками, имеют спектр  $S_0(\omega)$ . Далее эти колебания поступают в речеобразующий тракт, который является линейным пассивным четырехполюсником, имеющим частотную передаточную функцию  $K(\omega)$ . На выходе речеобразующего тракта, оканчивающегося губами, возникает речевой сигнал, спектр  $S(\omega)$  которого, согласно теории линейных систем, равен

$$S(\omega) = S_0(\omega) \cdot K(\omega). \quad (1)$$

Процесс образования звуков речи с помощью аппаратов ИГ происходит аналогично описанному (рис. 2).

Колебания «Искусственной гортани» со спектром  $S'_0(\omega)$  воздействуют непосредственно на речеобразующий тракт с передаточной функцией  $K(\omega)$ , на выходе которого образуется «искусственный» речевой сигнал со спектром  $S'(\omega)$ , равным

$$S'(\omega) = S'_0(\omega) \cdot K(\omega). \quad (2)$$

Искусственный речевой сигнал обладает хорошей разборчивостью и громкостью, почти не уступающим естественной речи, и, кроме того, обладает двумя замечательными свойствами: постоянством частоты основного тона и постоянством спектра  $S'_0(\omega)$  сигнала возбуждения речеобразующего тракта. Эти свойства являются чрезвычайно важными в задачах распознавания, синтеза и компрессии речевых сообщений.

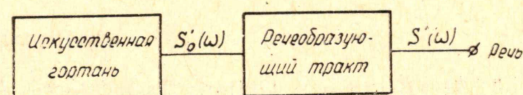


Рис. 2. Образование «искусственной речи»

Таким образом, произнесение речи с помощью аппарата ИГ обеспечивает нормировку речевого сигнала разных дикторов по частоте основного тона и спектру возбуждения, позволяет избавиться при анализе от сложного устройства выделения основного тона, так как сигнал основного тона можно вывести прямо из генератора, входящего в состав ИГ.

Среди большого круга задач, связанных с анализом и синтезом речевых сообщений, выделяются задачи определения передаточных функций речеобразующего тракта человека и определения сигнала возбуждения (импульсов основного тона). Применение аппаратов ИГ позволяет решить эти задачи.

Рассмотрим выражение (2). Передаточная функция речеобразующего тракта  $K(\omega)$  определяется из него следующим образом:

$$K(\omega) = \frac{S'(\omega)}{S'_0(\omega)}, \quad (3)$$

то есть для того, чтобы определить неизвестную передаточную функцию  $K(\omega)$ , достаточно знать  $S'(\omega)$ -спектр «Искусственного» речевого сигнала и  $S'_0(\omega)$ -спектр колебаний аппарата ИГ, определение которых известными методами не представляет трудностей, так как в отличие от естественной речи есть возможность измерить спектр  $S'_0(\omega)$  сигнала возбуждения на выходе звукопровода ИГ.

Другой метод измерения передаточной функции речеобразующего тракта основан на известном свойстве линейных систем, заключающемся в том, что импульсная переходная функция  $g(t)$  линейной системы (функция веса) и частотная передаточная функция  $K(\omega)$  связаны между собой преобразованием Фурье

$$K(\omega) = \int_0^{\infty} g(t) e^{-j\omega t} dt. \quad (4)$$

Считается, что за период  $T$  основного тона переходный процесс (речевой сигнал) в речеобразующем тракте при воздействии на него импульса возбуждения (основного тона) прекращается. Поэтому, если с помощью ИГ сформировать достаточно короткий импульс возбуждения (при периоде основного тона  $T = 10$  мсек достаточно выбрать длительность сигнала возбуждения, равной 100—150 мксек), полученный «искусственный» речевой сигнал с достаточной степенью точности можно считать функцией веса речеобразующего тракта, по которой с помощью выражения (4) легко получить искомую функцию  $K(\omega)$ .

После определения  $K(\omega)$  из выражения (1) можно определить спектр  $S_0(\omega)$  функции естественных голосовых связок человека по формуле

$$S_0(\omega) = \frac{S(\omega)}{K(\omega)}. \quad (5)$$

С помощью обратного преобразования Фурье по спектру  $S_0(\omega)$  определяется временная функция возбуждения  $U(t)$ .

#### ЛИТЕРАТУРА

1. H. Barney. Unitary Transistored Artificial Larynx. IRE Wescon Convention Record, Part 8. August. 1959.
2. Г. Фант. Акустическая теория речеобразования. «Наука», 1964.