

Фонетическая функция А.А. Пирогова и помехоустойчивость канала речевой коммуникации

Сергей Борисович Козлачков,
кандидат технических наук, старший преподаватель
МГТУ им. Н.Э. Баумана, Москва

Сергей Владимирович Дворянкин,
профессор, доктор технических наук, заместитель заведующего
кафедрой информационной безопасности Финансового университета, Москва

Надежда Валерьевна Василевская,
сотрудник, ФСТЭК России

Аннотация

В статье рассмотрены некоторые механизмы повышения помехоустойчивости речи. Обращено внимание на определенную согласованность основных характеристик механизма слухового восприятия, параметров и характеристик речевых сигналов, а также физических процессов среды распространения акустических сигналов с сопутствующими искажениями и помехами. Показано, что помехоустойчивость речевых сигналов дополнительно обеспечивается как особенностями амплитудной модуляции, так и сложной широкополосной структурой РС, а также уникальными свойствами фонетической функции А.А. Пирогова. Приведено объяснение эффектам устойчивости разборчивости речи относительно линейных искажений.

Ключевые слова: разборчивость речи, речевой сигнал, искажения, помехоустойчивость, форманты модуляция, фонетическая функция, пороги слухового восприятия.

ВВЕДЕНИЕ

Одной из важнейших характеристик канала связи, в т.ч. канала речевой коммуникации, является его помехоустойчивость, т.е. способность системы выполнять свои функции при наличии помех. В теории оптимального приема помехоустойчивость оценивают интенсивностью помех, при которых нарушение функций устройства ещё не превышает допустимых пределов. Чем сильнее помеха, при которой устройство остаётся работоспособным, тем выше его помехоустойчивость [1-3]. При наличии аддитивных помех в канале передачи информации помехоустойчивость может быть увеличена повышением мощности передаваемых сигналов. Однако при воздействии мультипликативных (неаддитивных) помех простым увеличением мощности сигнала существенного повышения помехоустойчивости достичь нельзя, и требуется радикальное изменение используемых методов, например применение помехоустойчивого кодирования либо самонастраивающегося (адаптивного) приёма [1-3].



В процессе эволюционного развития канал речевой коммуникации также обрел определенные механизмы, способствующие повышению его помехоустойчивости и согласующиеся с положениями теории оптимального приема. В данной статье предпринята попытка анализа некоторых механизмов слухового восприятия по критерию помехоустойчивости.

НЕКОТОРЫЕ МЕХАНИЗМЫ ПОВЫШЕНИЯ ПОМЕХОУСТОЙЧИВОСТИ РЕЧИ

Наиболее наглядно действие некоторых таких механизмов можно пояснить на примере эффекта Ломбарда — форсировании речи при воздействии на диктора аддитивных помех (например, шума) высокого уровня. В таких условиях речь диктора становится более звучной и разборчивой. При этом рост уровня (фонация) речевого сигнала (РС) составляет около 6 дБ. Также значительно возрастают и динамические показатели сигнала в т.ч. глубина модуляция спектральной огибающей РС, что проявляется в более четком выделении формант и антиформант в спектре РС.

Определенный вклад в повышение помехоустойчивости вносит и нелинейный характер ($\log_2 x$) зависимости уровня выходного сигнала от входного значения, сформулированный в законе Вебера-Фехнера, что позволяет значительно расширить динамический диапазон воспринимаемых уровней сигналов, в т.ч. принимать сигналы малого уровня при высоком уровне маскирующих помех.

Существенно влияет на помехоустойчивость и нелинейная зависимость АЧХ механизма слухового восприятия (корректирующая характеристика — фильтр А), благодаря чему усиливается относительный вес среднечастотного диапазона РС. С одной стороны, это позволяет снизить маскирующее влияние разнообразных акустических низкочастотных помех, с другой — подчеркивает значимость диапазона средних частот, в пределах которого находится более 30% информативного объема речи, что обусловлено как статистическими параметрами частоты встречаемости формант, так и преобладающей динамикой дифонных переходов наиболее значимой второй форманты [4].

Отдельного внимания заслуживают иные механизмы повышения помехоустойчивости, реализующие адаптивные свойства слухового восприятия, к числу которых можно отнести акустический рефлекс, предохраняющий слух от повреждений при воздействии длительных сигналов высокого уровня [5]. Важные адаптивные свойства также проявляются и в согласованных изменениях параметров слуховых фильтров (критических полос слуха), что позволяет реализовать один из принципов оптимального приема — согласование характеристик фильтра приемника параметрам РС в значимых спектральных областях (формантах) [6].

В рамках данной статьи планировалось рассмотреть некоторые механизмы повышения помехоустойчивости, связанные с амплитудными характеристиками РС. Также хотелось обратить внимание на то, что пра-

критически все значимые характеристики слухового восприятия описываются нелинейными функциями, что с учетом адаптивных свойств слухового восприятия позволяет решать актуальные задачи оптимизации и таким образом достигать наилучших условий приема в канале речевой коммуникации.

ПОМЕХОУСТОЙЧИВЫЕ СВОЙСТВА ФОНЕТИЧЕСКОЙ ФУНКЦИИ А.А. ПИРОГОВА

При распространении и приеме слуховой системой РС подвергается определенным трансформациям и искажениям, которые могут оказывать значительное влияние на его характеристики. По этой причине передавать информацию можно лишь с помощью каких-то неизменных — инвариантных структур, устойчивых по отношению к искажающим воздействиям. Информация, в т.ч. речевая, передается путем модуляции параметров сигнала — амплитуды, частоты, фазы. В речевом сигнале наиболее важными являются частотная и амплитудная модуляции, поскольку слуховое восприятие не воспринимает вариации значения фазы [7].

Чувствительность слухового восприятия характеризуется абсолютными и дифференциальными порогами. При этом физические значения абсолютных порогов чувствительности слуха почти на порядок выше, чем дифференциальных. Кроме того, слух на порядок лучше воспринимает частотную модуляцию, чем амплитудную (применительно к модуляциям сигналов гармонического типа), что обусловлено более высокой помехоустойчивостью частотной модуляции [7].

Распространение акустических колебаний (и речевых сигналов) в однородной (изотропной) среде из-за наличия различных препятствий сопровождается явлениями рассеяния, затухания и появлением отраженных волн, т.е. интерференцией. В этом случае среду распространения можно представить в виде полосного фильтра с неравномерной АЧХ. Если при этом передаточная характеристика среды распространения носит линейную зависимость, то соответствующие искажения акустического сигнала называют линейными, в ином случае — нелинейными. Из нелинейных наиболее негативное влияние на РР оказывают искажения модуляционного характера, поскольку они искажают саму модулирующую функцию, т.е. основной переносчик информации.

Наиболее распространенными и наименее значимыми (по критерию разборчивости речи) являются линейные искажения. Линейные искажения приводят только к изменению соотношения амплитуд и фаз спектральных составляющих сигнала. В литературе неоднократно отмечалось весьма низкое влияние линейных искажений на восприятие РС в целом [8, 9]. Так даже полное удаление из РС одной либо двух формант влияет только на тембр звука, однако словесная разборчивость остается высокой, что сложно объяснить с позиций формантной теории РР.

Помехоустойчивость амплитудной модуляции РС можно объяснить характером динамики изменения амплитудных и частотных параметров РС, что препятствует как формированию стоячих волн, так и образованию значительных неоднородностей плотности энергий акустического поля. Таким образом, нестационарный характер и относительно широкий диапазон полосы частот РС нивелируют влияние эффектов интерференции звуковых волн на искажение формантной структуры РС. В значительно большей степени эффекты интерференции, фазовых запаздываний и реверберации проявляются в акустике помещений большого объема и с высокой гулкостью, например концертных залах.



С другой стороны, устойчивость амплитудной модуляции РС можно объяснить, прибегнув к иной модели механизма слухового восприятия РС, разработанной А.А. Пироговым [9, 10]. По модели ученого, «каждая фонема отличается главным образом характерным для этой фонемы изменением спектрального распределения, а не самим спектральным распределением, сопутствующим данной фонеме». Исходя из этих соображений, он ввел понятие «фонетической функции речи» (или фонетическая функция А.А. Пирогова — ФФП), согласно которому фонетические элементы речи целиком определяются законом изменения спектров во времени. В качестве оценки спектральных изменений А.А. Пирогов предложил использовать разность логарифмов интенсивностей двух спектральных разрезов, взятых через интервалы, соответствующие разрешающей способности слуха во времени:

$$P(\omega, t) = \ln \left| \frac{S(\omega, t)}{S(\omega, t - \tau)} \right| \quad (1),$$

где:

$S(\omega, t)$ и $S(\omega, t - \tau)$ — интенсивности спектральных отсчетов РС, взятые через интервал τ , учитывающий разрешающую способность слуха во времени.

Рассмотрим характер влияния неравномерной АЧХ среды распространения РС через свойства ФФП.

В общем случае коэффициент передачи тракта распространения сигнала (полосного фильтра) можно представить в виде следующего выражения:

$$K = |K| \cdot e^{j\varphi} \quad (2),$$

где:

$|K|$ — модуль коэффициента передачи, φ — сдвиг фазы, j — комплексная единица.

Зависимость $|K|$ от частоты представляет собой амплитудно-частотную характеристику (АЧХ), а $j\varphi$ — фазочастотную характеристику (ФЧХ).

Спектр сигнала, прошедшего через такой фильтр, определяется выражением:

$$Si(\omega, t) = S(\omega, t) K \quad (3),$$

где:

$S(\omega, t)$ — спектр входного сигнала, K — коэффициент передачи электрического тракта, а $Si(\omega, t)$ — спектр сигнала, прошедшего через электрический тракт.

При допущении, что $|K|$ в период прохождения сигнала не претерпевает нелинейных трансформаций можно выражение (1) представить в следующем виде:

$$P(\omega, t) = \ln \left| \frac{S(\omega, t)K}{S(\omega, t - \tau)K} \right| \quad (4).$$

Из чего следует фактическая инвариантность ФФП относительно неравномерностей АЧХ (K) передаточного тракта.

ЗАКЛЮЧЕНИЕ

На основании изложенного можно сделать вывод о том, что человеческий слух в значительной степени инвариантен в отношении амплитудных линейных искажений, если, конечно, эти искажения не выходят за пределы артикуляторных модуляций и пределов слухового восприятия.

Помехоустойчивость амплитудной модуляции РС обеспечивается как свойствами ФФП, так и сложной широкополосной структурой РС, носящей нестационарный динамический характер, что позволяет нивелировать влияние линейных искажений среды распространения сигнала.

Представляется целесообразным проведение дальнейших исследований механизмов повышения помехоустойчивости речи.

ЛИТЕРАТУРА

1. *Гоноровский И.С.* Радиотехнические цепи и сигналы. М.: Сов. радио, 1977 г. 245 с.
2. *Гуткин Л.С.* Теория оптимальных методов радиоприема при флуктуационных помехах. Изд. 2-е, доп. и перераб. — М.: Сов. радио, 1972. 448 с.
3. *Стратонович Р.Л.* Принципы адаптивного приёма. М.: Советское радио, 1973. 144 с.
4. *Покровский Н.Б.* Расчет и измерение разборчивости речи. М.: Связьиздат, 1962. 392 с.
5. *Алдошина И.А.* Основы психоакустики // Звукорежиссер. 1999. <http://www.625-net.ru>
6. *Moore B.C.J.* An Introduction to the Psychology of Hearing, Sixth Edition. / B.C.J. Moore. — Leiden: Boston Brill. — 2012. 441 с.
7. *Женило В.Р.* Инварианты речевого сигнала и обертона голоса // Материалы Всероссийской научно-практической конференции «Естественнонаучные методы исследований в теории и практике производства судебных экономических и речеведческих экспертиз». Нижний Новгород, 2017. С. 70–81.
8. *Журавлев В.Н., Архипова А.Е.* Анализ противоречий теорий речеобразования и слуха с позиции идентификации информационных параметров и характеристик речевых сигналов // Информационные технологии и компьютерная инженерия. 2007. №2(9). С. 180–185.
9. *Пирогов А.А.* Основы Фонетической теории речи. Фонетическая функция как универсальный природный механизм кодирования-декодирования речевой информации любого происхождения // Научный Журнал Русского Физического Общества. 2001. №1–12. С. 15–28.
10. *Акбулатов А.Ш., Баронин С.П., Куля В.И., Лейтес Р.Д., Муравьев В.Е., Пирогов А.А., Слуцкер Г.С., Соболев В.Н., Трофимов Ю.К.* Вокодерная телефония. Методы и проблемы. М.: Связь, 1974. 536 с.



PHONETIC THE FUNCTION OF A.A. PIROGOV AND NOISE IMMUNITY OF THE CHANNEL SPEECH COMMUNICATION

Sergey B. Kozlachkov,

*candidate of technical Sciences, senior lecturer at MSTU. N. Uh. Bauman,
Moscow*

Sergey V. Dvoryankin,

*Deputy head of Department of information security of the Financial
University, doctor of technical Sciences, Professor, Moscow*

Nadezhda V. Vasilevskaya,

FSTEC employee, Moscow

Abstract

The article discusses some mechanisms to improve robustness of speech. Attention is drawn to a certain consistency the main characteristics of the mechanism of auditory perception, frequency-tours and characteristics of speech signals, as well as physical processes of the medium of propagation of acoustic signals with accompanying distortion and interference. It is shown that noise immunity of speech signals are provided as further features of the amplitudes-frequency modulation and wideband complex structure of RS, as well as unique properties of the phonetic function of A. A. Pirogov. At-Vedeno explanation for the effects of the stability of the intelligibility of speech refers-particularly linear distortion.

Keywords: speech intelligibility, speech signal distortion, stability, modulation formants, phonetic function, thresholds auditory perception.