

Критерии выбора полос частотного анализа речевых сигналов

Козлачков С.Б., к.т.н., МГТУ им. Н.Э. Баумана, Бонч-Бруевич А.М., к.т.н., МГТУ им. Н.Э. Баумана

Предложен объективный критерий выбора параметров полос частотного анализа при оценке защищенности речевых сигналов.

Ключевые слова: защита информации, речевые вокализмы, разборчивость речи.

We propose an objective criterion for selecting the parameters of the bands of frequency analysis in assessing the security of voice signals.

Keywords: information security, speech vocalism, speech intelligibility.

Вопросы оценки защищенности речевой информации (РИ), циркулирующей в выделенных помещениях (ВП) от утечки по техническим каналам являются предметом регулярных обсуждений специалистами на конференциях, научных семинарах и в печати [1, 2, 3]. Отдельного рассмотрения заслуживает обоснование выбора полосы частот при анализе речевых сигналов (РС), в т.ч. ширина таких полос.

В настоящее время в действующих нормативно-методических документах ФСТЭК России оценка защищенности РС (определение отношения сигнал-шум – SNR) проводится в октавных полосах [4] без какого-либо обоснования такого выбора. При этом отдельные исследователи [5] справедливо отмечают, что при наличии у злоумышленника (ЗЛ) аппаратуры негласного контроля с возможностями фильтрации шумов в более узких (третьоктавных) полосах, реальное значение SNR и разборчивости речи (РР) окажется больше за счет исключения из анализа частотных полос, в которых сигнал более зашумлен. В связи с этим, когда технические возможности аппаратуры измерения акустических и вибрационных сигналов позволяют

проводить анализ в более узких частотных полосах, чем октавные, необходимо для оценки защищенности РИ определить возможности расчета РР и в более узких (чем октавные) частотных полосах.

Закономерен вопрос о критериях выбора полос анализа РС, и о ширине таких полос (параметрах частотного разрешения) при расчете РР.

В приложениях защиты информации целесообразно рассматривать модели, определяющие семантику РС. Различают три основных типа переносчиков информации: тональный, с помощью которого передаются звонкие звуки речи; шумовой, с помощью которого передаются глухие звуки речи; импульсный, с помощью которого передаются взрывные звуки речи [6, 7]. Передача информации осуществляется путем модуляции определенных параметров каждого из указанных переносчиков [6]. Наибольшим числом видов (параметров) модуляции, и, как следствие, наибольшей информационной емкостью, обладает тональный переносчик, а наименьшей – импульсный.

Одним из наиболее распространенных способов представления РС в частотной области является преобразование Фурье [8, 9, 10, 11]. Во многих программных продуктах Фурье-преобразование играет ведущую роль и активно применяется в задачах исследования, обработки и передачи сигналов, в том числе и речи. Так, для описания вокализованных участков РС на коротких временных интервалах его анализа-синтеза часто используется модель РС [10, 12, 13]:

$$s(t) = \sum_{k=1}^K a_k(t) \cos k\Omega t ,$$

где $a_k(t)$ – амплитуда k -й гармоники основного тона; $\Omega = 2\pi f_{от}$; $f_{от}$ – частота основного тона; k – номер гармоники, обертона.

Вокализованные участки речи (речевые вокализмы – РВ) в таком случае можно представить как звучание отдельных обертонов, входящих в состав исследуемого звукового фрагмента. Параметры таких узкополосных

составляющих РС: амплитуды, частоты и фазы, далее называемые «следами РВ», – содержатся в динамических спектральных развертках, или ДСГ РС [14].

Следы фонообъектов в виде модуляционных параметров амплитуд и фаз узкополосных сигналов их составляющих, проявляются на изображениях динамических спектрограмм в виде совокупности контуров (линий) перепада яркости или треков (цепочек) локальных и глобальных экстремумов цветовой насыщенности в уровнях одного цвета. Пример ДСГ, или графической сонограммы РС, с треками РВ приведен на рисунке 1.

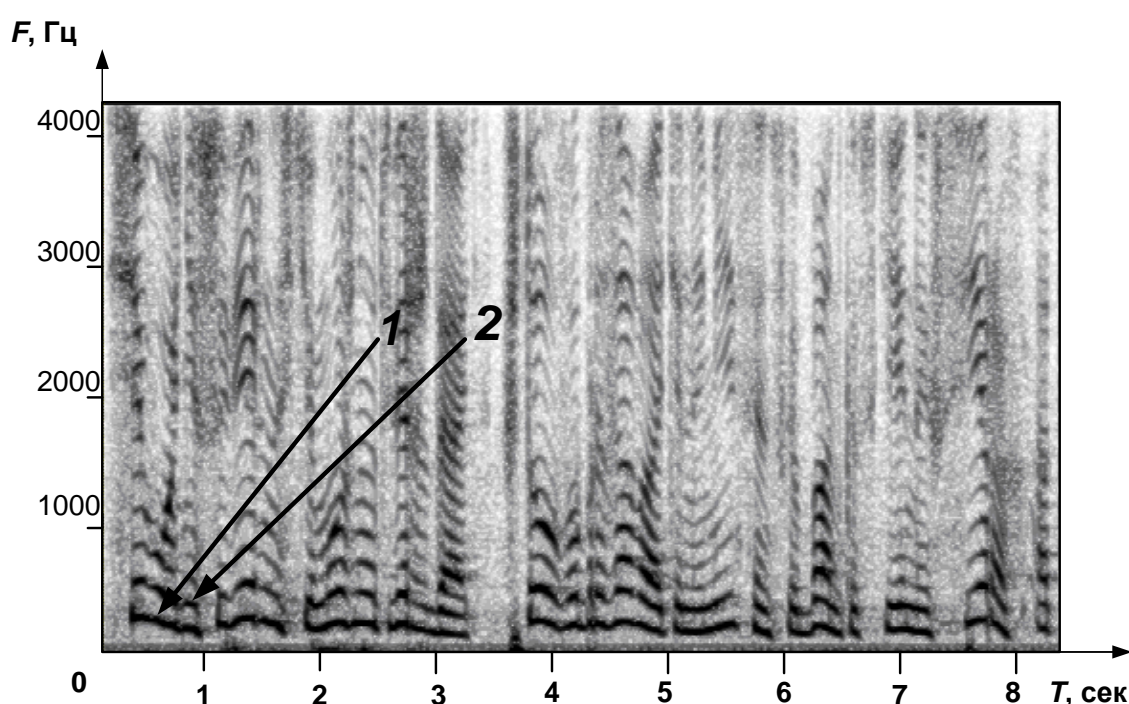


Рис.1. Графическая сонограмма речевого сигнала: 1 – трек первой гармоники РВ, 2 – трек второй гармоники РВ

Здесь, на рис. 1 по оси абсцисс отложено время, по оси ординат – частота. В уровнях серого цвета указана мощность РС на данной частоте в данный момент времени. Наибольшие значения соответствуют – черному цвету, наименьшие – белому.

Как было отмечено выше, мероприятия по защите РИ необходимо проводить с учетом возможностей ЗЛ по фильтрации шумов в более узких полосах, чем октавные. При этом ЗЛ анализируя речь конкретного диктора,

имеет возможность оптимизировать параметры фильтра (в т.ч. полосы разрешения) согласно индивидуальным характеристикам диктора и его РС. В соответствии с технологиями шумоочистки [15, 16, 17] реконструкция искаженных шумами и помехами РС проводится по трекам (следам РС) вокализованных звуков, наблюдаемых (см. рис. 1) на ДСГ. В этом случае полосы анализа не могут быть меньше девиации соответствующих обертонов частоты основного тона (ЧОТ). В противном случае неизбежна потеря существенной части информации и энергии сигнала.

Поэтому представляется закономерным в качестве объективного параметра – полос частотного анализа, соответствующих вышеперечисленным условиям, выбрать диапазон изменения (девиации) треков вокализованных звуков РС, т.е. ЧОТ ($\Delta F_{\text{о.т.}}$) и её обертонов.

По результатам выполненных экспериментов было определено, что для низкочастотных сегментов РС значения девиации обертонов РВ не превышают ширины критических полосок слуха (см. табл. 1).

Таблица 1.

Критические полоски слуха с указанием ширины полосы, Гц (по Цвикеру)

Номер полосы	$F_{\text{ср}}$, Гц	Ширина, Гц	Номер полосы	$F_{\text{ср}}$, Гц	Ширина, Гц	Номер полосы	$F_{\text{ср}}$, Гц	Ширина, Гц
1	51	80	9	1000	160	17	3400	550
2	150	100	10	1170	190	18	4000	700
3	250	100	11	1370	210	19	4800	900
4	350	100	12	1600	240	20	5800	1100
5	450	110	13	1850	280	21	7000	1300
6	570	120	14	2150	320	22	8500	1800
7	700	140	15	2500	380	23	10500	2500
8	840	150	16	2900	450	24	13500	3500

С учетом вышеизложенного и закономерностей формирования РС можно определить, что, например, в области октавы с $f_{\text{ср}i} = 500$ Гц

оптимальная ширина полосы анализа вокализованных сегментов РС составляет не более 100...140 Гц (см. табл. 1), а не 355 Гц как это принято в действующих нормативно-методических документах ФСТЭК России. В таком случае соответствующее значение SNR существенно возрастает – минимум на 3...5 дБ, т.е. не менее чем в 1,5...2 раза.

В заключение следует отметить:

- критерий выбора полос анализа учитывает возможности современных методов и технологий восстановления РС, искаженных шумами и помехами;
- ширина оптимальных полос анализа РС определяется значениями девиации его вокализованных сегментов, и не превышает ширины соответствующих критических полос слуха;
- при наличии априорных данных (образцы голосов дикторов) о параметрах анализируемого РС ширину полосы такого анализа можно дополнительно уменьшить до значений девиации его вокализованных сегментов;
- при оценках защищенности РИ, базирующихся на определении параметров РР с учетом возможностей современных технологий шумопонижения, необходимо вводить дополнительные поправки для соответствующих значений SNR .

Литература.

1. Железняк В.К., Макаров Ю.К., Хорев А.А. Некоторые методические подходы к оценке эффективности защиты речевой информации // Специальная техника. 2000. № 4. С. 39-45.
2. Халяпин Д.Б. Вас подслушивают? Защищайтесь! / Под ред. Л.И. Арефьевой. М.: НОУ ШО «Баярд», 2004. 432с.
3. Хорев А. А., Макаров Ю. К. К оценке эффективности защиты акустической (речевой) информации // Специальная техника. 2000. №5. С.46-56.

4. Герасименко В.Г., Лаврухин Ю.Н., Тупота В.И. Методы защиты акустической речевой информации от утечки по техническим каналам. М.: РЦИБ «Факел», 2008. 256с.
5. Каргашин В. Л. Совершенствование методических принципов оценки защищенности помещений от утечки речевой информации // <http://www.bnti.ru/showart.asp?aid=579&lvl=03.03.04>.
6. Голубинский А.Н. Алгоритм оценки частоты основного тона и обертонов речевого сигнала спектральным методом // Современные проблемы борьбы с преступностью: Сборник материалов международной научно-практической конференции. Воронеж, 2006. С. 41–42.
7. Ролдугин С.В. Голубинский А.Н., Вольская Т.А. Модели речевых сигналов для идентификации личности по голосу // Радиотехника. 2002. № 11. С. 79–81.
8. Дворянкин С.В. Использование вектора приведенных начальных фаз гармоник речевого сигнала как дополнительного признака верификации говорящего // Информатизация правоохранительных систем: Международная конференция. М., 1996. С. 174-177. (Сборник научных трудов Академии управления МВД РФ).
9. Дворянкин С.В. Цифровая шумоочистка аудиоинформации. М.: ИП РадиоСофт, 2011. 208 с.
10. Калинин Ю.К. Разборчивость речи в цифровых вокодерах. М.: Радио и связь, 1991. 220с.
11. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов: Пер. с англ. / Под ред. М.В. Назарова, Ю.Н. Прохорова М.: Радио и связь, 1981. 495с.
12. Дворянкин С.В. Компьютерные технологии защиты речевых сообщений в каналах электросвязи. М.: МТУСИ, 1999. 52 с.
13. Дворянкин С.В. Нормировка слов в системах распознавания речи // Информатизация правоохранительных систем: Международная конференция.

М., 1996. С. 189-191. (Сборник научных трудов Академии управления МВД РФ).

14. Дворянкин С.В. Эксперименты по восстановлению искаженной шумами речи // Управление безопасностью. 2004. Май. С. 42-46.

15. Дворянкин С.В. Взаимосвязь цифры и графики, звука и изображения. // Открытые системы. 2000. №3. С. 34-52.

16. Моделирование и генерация сигналов речевых вокализмов в приложениях оценки защищенности речевой информации / С.В. Дворянкин [и др.] // Проектирование и технология электронных средств. 2007. №4. С. 33-37.

17. Халяпин Д.Б., Рюмин А.А. Коктейль из звуков. Использование динамических спектрограмм для оценки качества зашумления речевого сигнала // Information Security. Информационная безопасность. 2005. №4. С. 28-29.