

Я.Ю. КУЛЬКОВ

**Основные подходы к повышению
разборчивости речи в системах
диспетчеризации и управления
производством**

УДК 654.078

Муромский институт
(филиал) ФГБОУ ВПО
«Владимирский
государственный
университет имени
А.Г. и Н.Г. Столетовых»,
г.Муром

Для эффективного использования в управлении систем оперативно-технологической связи необходимо обеспечить максимальную разборчивость речи. В статье рассматриваются факторы, снижающие разборчивость речи в таких системах. Приводятся результаты исследования существующий систем оперативной связи и возможные пути решения.

Введение.

Применение специальных коммуникационных систем в области связи на объектах промышленности и транспорта позволяет существенно повысить эффективность управления ими и обеспечить надежную связь на технологических участках. От работы этих систем во многом зависит безопасность на объектах, особенно в местах с повышенной опасностью возникновения аварий, пожаров, взрывов и химических выбросов.

Системы громкоговорящей связи - это отдельный специализированный вид оборудования, в первую очередь предназначенный для решения задач информационного звукового обеспечения на объекте. Громкоговорящая связь - электросвязь на объекте или в открытом пространстве, в которой воспроизведение информации осуществляется посредством громкоговорителя или акустической системы [1]. По требованиям международных стандартов подобные системы должны размещаться везде, где присутствует большое количество людей [2]. Требования к системам оповещения отличаются от требований к бытовому или профессиональному звуковому

оборудованию. В первую очередь необходимо достижение максимальной разборчивости речевого сигнала.

Использование громкоговорящей связи не только необходимо с точки зрения эффективности и оперативности работы производства, но и регламентировано постановлением Госгортехнадзора РФ от 5 мая 2003 г. № 29 «Об утверждении Общих правил взрывобезопасности для взрывопожароопасных химических, нефтехимических и нефтеперерабатывающих производств» ПБ 09-540-03.

Одно из наиболее важных требований, которым должно отвечать оборудование технологической связи, – работа в условиях шума выше 100 Дб, когда стандартную телефонию использовать невозможно.

Основной причиной гибели людей на пожарах является сопутствующая паника. Техническая основа для связи с персоналом - системы речевого оповещения. основополагающим документом при проектировании таких систем являются «Нормы пожарной безопасности НПБ 104-103».

Действующий в настоящее время документ лишь в общих чертах затрагивает требования к системе. Так, требуется зональность (поочередность) оповещения в многоэтажных зданиях, а также предварительное оповещение персонала здания. Отсутствуют требования к качеству оповещения (разборчивости речи), а ведь в экстремальных условиях внешний фоновый уровень шума значительно возрастает.

Своевременное оповещение людей о пожаре является одним из главных факторов обеспечения их безопасной эвакуации. Необходимость в устройствах оповещения и типы систем оповещения регламентированы СНБ 2.02.02-01 «Эвакуация людей из зданий и сооружений при пожаре».

Неудовлетворительная разборчивость речи в оповещении не просто свидетельствует о низком качественном уровне системы, но и в определенных обстоятельствах ставит под угрозу жизнь людей. На сегодняшний день надежно оценить степень разборчивости можно лишь эмпирическим путем, то есть после того, как система инсталлирована на объекте. Существующие объективные методы расчета разборчивости (например PESQ) очень часто дают результаты, значительно расходящиеся с реальной картиной. Это связано

с тем, что разборчивость речи является результатом взаимодействия системы звукоусиления с акустическими условиями озвучиваемого объекта [3].

Теоретический анализ.

В системах громкоговорящего оповещения передаваемая речь подвергается воздействию акустических шумов. Разборчивость речи в системах ГГС является результатом взаимодействия системы звукоусиления с акустическими условиями озвучиваемого объекта. Основными задачами систем громкоговорящей связи являются: обеспечение хорошей слышимости в различных точках прослушивания при воздействии акустических шумов, обеспечение разборчивости речи.

Разборчивость является некоторой интегральной оценкой речевого сигнала и в соответствии с международным стандартом ISO/TR 4870 определяется как «степень, с которой речь может быть понята (расшифрована) слушателями». Под этим понимается степень, с которой слушатели могут понять смысл фразы, идентифицировать слова, слоги и фонемы.

Разборчивость речи в системах оповещения должна оцениваться как «хорошая» или «отличная». Применение различных алгоритмов шумочистки и постобработки сигналов в подобных случаях малоэффективны так как в разных зонах оповещения на сигнал действуют различные акустические искажения.

Акустические шумы - это шумы в пределах озвучиваемой поверхности, создаваемые публикой, различными агрегатами и т. п., и шумы, попадающие извне, например от транспорта.

По уровню акустические шумы в непроизводственных помещениях имеют диапазон от 55 до 80 дБ. Уровни в производственных помещениях имеют широкий диапазон – от 65 до 95 дБ и более.

Рассмотрим более подробно влияние различных акустических шумов на разборчивость речевого сигнала. Негативное влияние проявляется в виде эффекта маскирования [4]. Маскирование другими звуками, в том числе шумами в реверберирующем помещении и др. Шумы могут создаваться вентиляцией, внешними проникновениями, шумами аппаратуры или оборудования, публикой.

Процент потери разборчивости зависит, прежде всего, от отношения уровня речевого сигнала к уровню шума, которое должно

быть выше определенного уровня, чтобы можно было понять смысловое содержание речи. Степень маскировки шумом будет зависеть от отношения сигнал/шум SNR и от спектрального состава шума. Для широкополосного шума (20-4000 Гц) анализ индекса разборчивости RASTI от показателя SNR показывает, что индекс разборчивости будет больше 0.8 только при SNR более 12 дБ [5].

Влияние шумов на разборчивость речи зависит также от направления их прихода: если направления речевого сигнала и шума совпадают, то степень маскировки и, соответственно, процент потери разборчивости будет наибольшим. Слуховой системе трудно провести их разделение, но чем больше расстояние между ними, тем выше разборчивость.

Сильное воздействие на разборчивость речи оказывает шум от других голосов (шум толпы). Поскольку этот шум сходен с речью по спектральному составу, то уровень словесной разборчивости резко снижается, особенно при увеличении числа мешающих голосов [5].

Основная энергия речи сосредоточена в полосе до 2 кГц. Более 80% звуков речи имеют уровень меньше 50 дБ, и легко могут маскироваться шумами. Среди этих звуков могут оказаться согласные звуки - самые информативные. Гласные звуки имеют основную частоту фонации в пределах 80-250 Гц, и значительная часть их энергии сосредоточена в формантных областях в пределах 450-4000 Гц. По распределению формантных областей в спектре и происходит распознавание гласных звуков. Хотя гласные звуки имеют длительность 30-300 мс, и именно в них сосредоточена основная энергия речевого сигнала, основной вклад в разборчивость вносят согласные звуки, которые имеют значительно меньшую длительность, от 10 до 100 мс. Они ниже по уровню на 27 дБ и их спектр - особенно у шумовых (С, З) и взрывных (Д, Т) согласных - расположен в основном в высокочастотной области 2-10 кГц. Ключевую роль в распознавании речи играют октавные полосы в области 1, 2, 4 кГц. Они содержат до 75% речевой информации. Важную роль играет октавная полоса в области 2 кГц - до 33% речевой информации.[5]

Эффект адаптации при маскировании имеет ряд негативных моментов по отношению к разборчивости речи. Так, если среди акустических шумов имеется интенсивный сосредоточенный выброс, и его частота совпадает с частотой основного тона человека, то вто-

рая f_1 и третья f_2 форманты речевого сигнала будут в зоне маскировки. Как показали эксперименты, это приводит к падению разборчивости на 0.3 пункта по показателю RASTI.

Значения показателя RASTI соотносятся с субъективной оценкой разборчивости речи в соответствии со стандартизованной зависимостью, представленной в таблице 1.

Таблица 1

Соотношение показателя RASTI и субъективной оценкой разборчивости

Значения RASTI	Разборчивость речи
0.75-1.00	Отличная
0.6-0.75	Хорошая
0.45-0.60	Удовлетворительная
0.30-0.45	Плохая
0.00-0.30	Недопустимо плохая

Рассмотренный эффект маскировки позволяет сделать вывод, что наличие во внешнем акустическом шум всплесков высокой интенсивности приводят к появлению эффекта маскировки. Это приводит к потере человеком некоторых звуков слышимой речи, что неминуемо ведет к снижению разборчивости. При использовании систем громкоговорящей связи на промышленных объектах появление таких узкополосных шумов неизбежно.

При таком снижении разборчивости мозгу человека требуется определенное время для разбора смысловой информации, произнесенной по системе оповещения. Работа систем громкоговорящей связи и оповещения связана с появлением критических ситуаций. От скорости реакции человека на информацию зависит жизнь не только его самого, но и жизни других людей. Поэтому ослабление эффекта маскирования, и как следствие, повышение разборчивости речи, является одной из наиболее актуальных при построении систем громкоговорящей связи.

В условиях внешних акустических шумов на приемной стороне важную роль приобретает обеспечение комфортных условий прослушивания речевого сообщения. Перспективными подходами в данном направлении являются разработка адаптивных к внешнему шуму устройств регулировки выходной громкости и частотной характеристики речи [6].

Задачей данных устройств является снижение эффекта маскирования прослушиваемой речи мешающим внешним шумом, путем частотно-взвешенного повышения ее громкости для формирования достаточной глубины артикуляции речевого сигнала на фоне шума во всем частотном диапазоне. Для оценки внешнего шума используются сигнал, получаемый от местного микрофона [6].

Описание известных систем оперативно-технологической связи.

В настоящее время создано большое количество систем громкоговорящей связи, как отдельных, так и в составе комплексов оперативно-технологической связи. Практически все современные системы созданы с использованием цифровых технологий.

Устройство тревожной сигнализации и оповещения УТСО-200 компании Armtel применяется для передачи сообщений по громкоговорящей связи и подачи сигналов для личного состава. В своей структуре не имеет средств повышения разборчивости речи.

Заявлено улучшение разборчивости за счет использования одновременно 2 голосовых каналов при связи между абонентами, что позволяет расширить полосу передаваемого звукового сигнала с 3,4 кГц до 6,8 кГц. На производственных участках с повышенным уровнем шума это практически не влияет положительным образом на разборчивость речи.

Более распространена система на базе цифрового коммутационного оборудования DVS-21 производства немецкой компании Procom. Полудуплексные режимы используются в большинстве существующих на рынке ГГС в качестве основных. Однако данные режимы оправданы и распространены только в бытовых системах ГГС. При этом компания обещает хорошую разборчивость речи в сложных условиях цехов и на больших площадях озвучивания, благодаря специальным акустическим характеристикам. Под этим подразумевается лишь компрессия динамического диапазона.

Среди отечественных разработок можно отметить приборы ПГС-16, ПА-1, П-160, П-164, П-166ВАУ.

Прибор ПА-1 обеспечивает оперативную дуплексную громкоговорящую связь с аналогичным прибором (парная связь), так и с аппаратурой диспетчерской связи АДС-М. При работе с однотипным прибором или с аппаратурой АДС-М обеспечивается разборчивость

речи не менее 92 % при приеме информации на расстоянии не более 1 м от прибора при уровнях шума в местах приема и передачи - не более 60 дБ. Имеют встроенный компандер аудиосигнала для улучшения разборчивости речи [7]. Заявленные характеристики не позволяют использовать его в местах с высоким уровнем акустического шума.

Аналогичные характеристики у системы цифровой оперативно-технологической избирательной связи ПГСИ-90/180.

Приборы громкоговорящей связи ПГС16-xxx-х.11 IPxx ООО «ТехноСвязь» со специальной системой компрессии передаваемого сигнала обеспечивают сжатие динамического диапазона сигнала передаваемого в линию общей связи. Сжатие динамического диапазона передаваемого сигнала при той же мощности дает эффект увеличения громкости, повышает звуковое давление в слышимом диапазоне. Как было показано выше, данный метод не позволяет избежать проблем, связанных с эффектом маскирования.

В настоящее время, в большинстве систем громкоговорящей связи не применяется никаких методов повышения разборчивости речи. При этом как было показано, это один из важнейших показателей качества и эффективности использования систем ГГС. В некоторых системах применяется компрессия динамического диапазона. Это несколько улучшает разборчивость на фоне шумов, но никак не помогает бороться с эффектом маскирования, который ответственен за снижение разборчивости. В большинстве случаев производители пытаются поднять показатель «сигнал/шум» за счет увеличения количества и мощности громкоговорителей. Такой метод имеет ограниченную эффективность, т.к. приведет к сдвигу кривой чувствительности слуха.

Одним из методов повышения разборчивости речи является снижение уровней шумов и помех [8]. В условиях промышленного применения систем ГГС подобные методы являются не только неэффективными, но и нереализуемыми.

Результаты исследования

Анализ подобных систем и существующих методов борьбы с шумами показывает необходимость разработки методов, эффективных именно при воздействии сильных внешних акустических шумов, с которыми невозможно бороться существующими методами.

Одним из важнейших параметров сигнала возбуждения колебаний в речевом тракте является частота основного тона (ОТ) речи, характеризующая высоту голоса при произнесении вокализованных звуков речи. Человек легко определяет частоту ОТ на слух или на глаз при анализе осциллограмм или спектрограмм речи, однако, построить устройство, автоматически определяющее частоту ОТ с малой ошибкой и малой задержкой во времени, даже при относительно низком уровне помех, довольно трудно [9].

В русском языке различают сорок один звук речи (фонем) [10]. По спектральному составу звуки речи различаются друг от друга числом формант и их расположением в частотном спектре. Следовательно, разборчивость речи зависит, прежде всего, от того, какая часть формант дошла до уха слушающего без искажений и какая - исказилась.[11]

Речь содержит в себе форманты, прием которых определяет ее разборчивость и неформантные составляющие, к которым относятся основные тоны, области частот между формантами и составляющие, зависящие от индивидуальных особенностей говорящих [10]

Известно, что от правильного определения основного тона существенно зависит разборчивость речевого сигнала, так как именно основной тон определяет высоту голоса, интонацию говорящего, логическое ударение, а иногда и смысл слов [5]. Также известно, что слоговая разборчивость речи значительно снижается при ошибочном определении ОТ [5].

Наибольший показатель разборчивости речи, искаженной шумами и помехами достигается за счет восстановления нарушенной формантной структуры или фонетической функции речевого сигнала, которая может быть восстановлена по известной гармонической структуре, определяемой, в свою очередь, за счет поиска и реконструкции в спектрограмме зашумленной речи «следов» гармоник полезного речевого сигнала.[11]

Предлагаемым способом снижения маскирующего влияния акустических шумов на разборчивость речевого сигнала является повышение относительной средней мощности излучаемого громкоговорителями сигнала. Причем простое повышение мощности, то есть увеличение максимальной амплитуды не приведет к желаемому эффекту [12]. Увеличение громкости выше некоторого порога при-

водит наоборот к снижению разборчивости речи, так как порождает сдвиг кривой чувствительности в сторону уменьшения. При этом эффект маскирования становится еще более выраженным.

В спектральной области голос каждого человека уникален. В одном голосе присутствует большое количество низкочастотных компонентов, голос другого более насыщен высокими частотами. Соответственно каждый голос потребует различной коррекции его спектра.

Автоматические устройства для коррекции частотного спектра, позволяют корректировать частотный спектр без применения специального тестового сигнала.

Предлагается использовать многополосную фильтрацию в заданной полосе и адаптивную коррекцию коэффициентов передачи каналов в зависимости от индивидуальных свойств спектра речи. Добиться подобного результата можно путем усиления только формантных частот излучаемого речевого сигнала. Это приведет к снижению эффекта маскирования и выделению голоса на фоне акустических шумов. В описываемом устройстве происходит выделение частоты основного тона и усиление первых трех формант.

Эталонные коэффициенты для каждой полосы частот подобраны в соответствии с психоакустическими особенностями строения слуха человека. Обеспечиваемая таким образом коррекция амплитудно-частотной характеристики речевого сигнала приводит к усилению тех частот, на которых чувствительность слухового аппарата ниже, что приводит к улучшению разборчивости передаваемой в системах громкоговорящей связи речи. Это позволяет учитывать индивидуальные свойства каждого конкретного голоса, что приводит к значительному улучшению разборчивости речи диктора.

Заключение.

В ходе работы были проведены экспериментальные исследования, в которых определялась F-картина исследуемого речевого сигнала. Далее выборочно усиливались частоты спектра, соответствующие первым трем формантам исследуемого голоса. Измерялась разборчивость речи, обработанной указанным алгоритмом. В среднем повышение разборчивости по показателю RASTI составило 0.2.

Результаты исследования позволяют говорить о высокой эффективности разработанного алгоритма в системах громкоговорящей связи, и могут быть реализованы в уже имеющихся или установленных системах громкоговорящей связи.

Литература:

1. ГОСТ 24214-80 Связь громкоговорящая. Термины и определения
2. *Мишенков С.Л.* Системы звукового вещания и оповещения: Учебное пособие / А.М. Копылов, А.П. Ефимов. – М.: МТУСИ, 1995, 78 с.
3. *Лукач А.А.* Проектирование систем оповещения о пожаре // Техника без опасности, 2004 - №1 – с.4-10.
4. *Ковалгин Ю.А.* Цифровое кодирование звуковых сигналов / Ю.А. Ковалгин, Э.И. Вологдин. - СПб.: КОРОНА-принт, 2004, 240 с.
5. *Алдошина И.* Основы психоакустики. Часть 17. Речь и слух. Часть 1. // Звукорежиссер. 2002 г., № 9, с.10-15.
6. *Бабкин В.В.* Проблемы построения современных систем цифровой речевой связи. / 9-я Межд. Конф. и выставка Цифровая обработка сигналов и ее применение (DSPА-2007), 2007, г.Москва, с.23-27.
7. *Борисов Г.Н.* Системы связи на металлургическом предприятии // Журнал «Connect! Мир Связи», 2008 - №6 – с.15-17.
8. *Сапожков М. А.* Электроакустика. Учебник для вузов. М., «Связь», 1978, 272 с.
9. Вокодерная телефония. / Под. ред. А. А. Пирогова. - М: Связь, 1974, 536 с.
10. *Покровский Н. Б.* Расчет и измерение разборчивости речи. - М.: Гос. изд-во литературы по вопросам связи и радио, 1962. - 392 с.
11. *Дворянкин С.В., Макаров Ю.К., Хореев А.А.* Обоснование критериев эффективности защиты речевой информации от утечки по техническим каналам // Защита информации. Инсайд, - 2007. № 2 - с.18-25.
12. Звуковое вещание: Справочник / Под ред. Ю.А. Ковалгина. – М.: Радио и связь, 1993, 464 с.