

МОДУЛЬ СИНТЕЗА РЕЧЕПОДОБНОГО СИГНАЛА ДЛЯ ЗАЩИТЫ АКУСТИЧЕСКОЙ ИНФОРМАЦИИ

К.П.Шакин, О.Б.Зельманский

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Ключевые слова: синтез речеподобных сигналов, компиляция сигналов, база данных аллофонов, защита информации.

Стремительное развитие информационных технологий неразрывно связано с защитой информации, что должно являться важнейшим направлением деятельности любой организации. Одной из актуальных проблем информационной безопасности является утечка информации по акустическому каналу. Защита речевой информации является одним из звеньев по обеспечению информационной безопасности и осуществляется с использованием пассивных и активных методов защиты информации. Активные методы защиты речевой информации предполагают создание маскирующих шумовых помех средствам акустической разведки. Для формирования шумовых помех в настоящее время используются генераторы белого, розового и речеподобного шума. Доказано, что речеподобная помеха является эффективнее других применяемых видов шумов, так как по своему спектральному составу близка к речевому сигналу.

В качестве основы модуля, генерирующего речеподобную помеху, предлагается применить компиляционный метод синтеза, заключающийся в соединении готовых минимальных акустических единиц – аллофонов. Данный метод подразумевает формирование базы данных звуковых фрагментов из артикуляционных слоговых и словесных таблиц, соответствующих ГОСТ 16600-72 [1].

Предлагаемый модуль синтеза речеподобных сигналов выполняет следующие функции: формирование фонемного псевдотекста, компиляцию речеподобного сигнала. Структурная схема модуля включает: генератор псевдотекста и базу аллофонов, соединенные с компиляционным синтезатором

речи, а также акустическую систему для воспроизведения формируемого сигнала.

В свою очередь компиляционный синтезатор речи содержит блок обработки естественного языка и блок обработки цифрового сигнала [3, 4, 6].

1. Модуль обработки естественного языка выполняет анализ и обработку входного орфографического текста. Непосредственно в прикладных системах модуль естественного языка делится на три подмодуля: лингвистической, фонетической и просодической обработки.

В подмодуль лингвистической обработки включены такие функции, как очистка текста, расшифровка числительных, даты и времени, аббревиатур, сокращений, Интернет ресурсов, автоматическая расстановка ударений, объединение слов в акцентные группы и членение на синтагмы.

В подмодуле фонетической обработки выполняется автоматическое транскрибирование текста в фонемный вид, а затем фонемного текста в аллофонный.

При этом ударные гласные маркируются индексом 0, предударные индексом 1, заударные индексом 2. В служебных словах ударная гласная маркируется индексом 5, а порядок индексирования заударных и предударных при их наличии остается прежний.

Просодическое оформление текстовой информации, в лингвистической обработке, заключается в сопоставлении интонационных контуров к соответствующим типам синтагм. Для этого необходимо классифицировать акцентные группы и разделить их на составляющие.

2. Модуль обработки цифрового сигнала заключается лишь в акустическом подмодуле. Основная задача акустического модуля — генерация (синтез) речевого сигнала на основе трех типов параметров:

просодических параметров (F_0 — частота основного тона, T — длительность звуков, A — амплитуда звуков), которые поступают от просодического подмодуля;

фонетических параметров, поступающих от фонетического подмодуля (в зависимости от типа фонетического процессора эти параметры могут быть различными: формантными параметрами (F1, F2, A1 ...), параметры сечений речевого тракта, номер аллофона или сегмента и т.д.);

параметры синтезируемого голоса, обеспечивающие желаемую тембровую индивидуальность.

Создание базы аллофонов целесообразно выполнить в следующие этапы:

1. Формирование представительного текстового корпуса (набора текстов) и соответствующих этим текстам фонограмм речи (речевой базы) диктора. Для этого применяются артикуляционные слоговые и словесные таблицы из ГОСТ 16600-72.

2. Обработка созданной речевой базы и выделение аллофонов посредством аудиоредактора Adobe audition, сохранение полученного набора в аллофонно-волновой базе данных.

Выделение аллофонов предлагается выполнить, открывая записанный файл в Adobe audition и нахождении границы аллофонов по внешнему виду сигнала, при этом прослушивая этот файл.

Далее на основании статистики русского языка генерируется псевдотекст, который озвучивается компиляционным синтезатором речи при помощи базы аллофонов.

Псевдотекст предлагается получить при использовании вероятностей появления слогов русского языка. Для формирования таблицы безусловных вероятностей слогов большие объемы текстов были разбиты на слоги в соответствии с предложенными в [2] правилами.

Данный метод формирования псевдотекста заключается в следующем:

- из фраз формируется фоноабзац, при обязательном учете частоты встречаемости фоноабзацев, состоящих из определенного числа фраз;

- фразы из синтагм составляются по тому же принципу, что и в предыдущем пункте;

- формируется синтагма из слов;

- слово составляется из слогов, определяется положение ударного и первого предударного слогов. Для создания очередного слога используется модель, позволяющая осуществлять учет предыдущего слога;

- слова орфографического текста преобразуются в слова фонемного текста;

- последовательности букв фонемного текста преобразуются в последовательность аллофонов и озвучиваются. При озвучивании учитываются паузы, расставляемые в конце синтагмы, фразы и фоноабзаца, соответствующей длительности, а так же просодическое оформление сформированных синтагм и фраз.

Для формирования слогов с учетом предыдущего выполняются следующие действия.

Статистически представительный текст разбивается на слоги в соответствии с правилами:

- в слоге всегда одна гласная буква;
- все согласные до первой гласной в слове принадлежат первому слогу;
- все согласные после последней гласной принадлежат последнему слогу;
- если между двумя соседними гласными находится более одной согласной, то самая первая согласная относится к начальному слогу, а остальные согласные к следующему;
- если между двумя соседними гласными находится одна согласная, то она относится к следующему слогу.

При разбиении текста на слоги учитывается число одинаковых слогов, встречающихся в тексте. После того, как процедура разбиения текста на слоги закончена, формируется словарь наиболее часто встречающихся слогов, покрывающих 90% текста.

По сформированному в соответствии с приведенными правилами словарю слогов текст заново разбивается на слоги. Но теперь задача состоит не в выделении слогов, а в получении статистических данных относительно

условной встречаемости слогов словаря (с учетом предыдущего слога или отсутствием такового). Данные заносятся в таблицу.

Для получения условных вероятностей появления слогов в тексте сформированная таблица нормируется построчно.

Алгоритм формирования последовательности слогов состоит в генерировании случайного числа, равномерно распределенного от 0 до 1, и поиска слога, указываемого этим числом.

Распределение, используемое при поиске слога, определяется предыдущим сформированным слогом, указывающим на соответствующую ему строку таблицы статистических данных.

Сформированные слова орфографического текста преобразуются в слова фонемного текста с использованием правила перевода “Буква-фонема”. Слова, составленные из фонем, преобразуются в последовательность аллофонов с помощью алгоритма “Фонема-Аллофон”.

Просодическое (интонационное) оформление сформированных синтагм и фраз осуществляется в соответствии с рядом правил. Более подробно приведенные процедуры описаны в [5].

При решении задач эффективной защиты речи для составления таблицы вероятностей следует использовать тексты статей, речей, докладов защищаемого лица. Наряду с обработкой текстов, настоятельно рекомендуется применение базы аллофонов этого лица.

В результате формируется аудиосигнал, содержащий речеподобную помеху, который в дальнейшем поступает на вход акустической системы.

Заключение

В данной работе был предложен способ формирования базы аллофонов, не требующий сложных операций и вычислений. Рассмотрен метод генерации псевдотекста, наиболее близкого к естественной речи, как итог обеспечивается высокое качество синтезируемой речи.

Стоит отметить, что данный подход не требует никаких знаний об устройстве речевого тракта и структуре языка.

Литература:

1. Авдеев В.Б., Трушин В.А., Кунгуров М.А. Унифицированная речеподобная помеха для средств активной защиты речевой информации. // Информатика и автоматизация. 2020. Том 19 № 5.
2. Воробьев В.И., Давыдов А.Г., Лобанов Б.М. // XIII сессия Российского акустического общества: Сб. тр. Н. Новгород, 25–29 августа 2003 г. / М., 2003. Т. 3.
3. Зельманский О.Б., Шакин К.П. Сборник НК 2021.
4. Киселев В.В., Лобанов Б.М. Доклады БГУИР, 2004 .
5. Лобанов Б.М. Синтез речи по тексту. Сборник научных трудов 4-й Международной школы-семинара по искусственному интеллекту. Мн.: БГУИР, 2000.
6. Рыбин С.В. Синтез речи.- СПб: Университет ИТМО, 2014 г.