

Решение данной задачи методом полного перебора требует просмотра вариантов топологических структур, причем просмотр заключается в генерации каждого варианта и проверке соответствующих условий.

УДК 621.391.8

Методы параметризации речевого сигнала на основе анализа синхронизированного с частотой основного тона в системах конверсии голоса

Захарьев В.А., Петровский А.А.

Белорусский государственный университет информатики и радиоэлектроники

Системы конверсии голоса предназначены для преобразования параметров речевого сигнала (РС), характеризующих одного диктора, в параметры другого, без изменения лингвистической составляющей самого сообщения. Первый диктор называется исходным, второй – целевым. Процесс конверсии подразумевает изменение акустических, фонетических и просодических характеристик исходного диктора в характеристики целевого согласно определенному набору правил, представляющему собой модель конверсии голоса.

Для построения эффективных систем конверсии необходимо совершенствование существующих методов анализа сигнала.

Поскольку РС является сложным частотно-модулированным сигналом, классические методы анализа, основывающиеся на предпосылке о его фрагментарной стационарности не всегда адекватны.

Необходимо использование современных методов обработки, учитывающих изменения параметров речевого сигнала в пределах фрейма анализа.

В докладе приведён обзор и сравнение двух методов параметризации речевого сигнала на основе математических моделей представления речевого сигнала.

В частности, обсуждается использование моделей на базе ДПФ, синхронизированной с частотой основного тона (ЧОТ), а также адаптивной интерполяции и спектрального взвешивания (STRAIGHT).

Обе эти модели основаны на представлениях о многокомпонентности речевого сигнала, представляя РС как суперпозицию гармонической составляющей и остаточного сигнала. Разница между двумя моделями заключается в способе описания двух данных составляющих, а также способах поиска параметров модели.

В первом случае анализ ведётся с использованием модифицированного ДПФ учитывающего изменение ЧОТ в пределах фрейма.

Во втором – используются специальные виды временных и частотных окон.

УДК 621.391.8

Построение многоголосого синтезатора речи по тексту на базе системы текстонезависимой конверсии голоса

Захарьев В.А., Петровский А.А.

Белорусский государственный университет информатики
и радиоэлектроники

Повсеместное внедрение средств вычислительной техники в различные сферы человеческой деятельности делает разработку речевого интерфейса для взаимодействия с ЭВМ очень актуальной задачей.

Одной из обязательных составляющих речевого интерфейса является система синтеза речи по тексту (ССРТ). Поэтому от качества её реализации зависит и общее качество работы речевого интерфейса, а, следовательно, и общий уровень удовлетворения потребностей пользователя системы.

На данном этапе развития ССРТ ставится вопрос уже не столько об обеспечении хороших уровней основных характеристик систем этого класса, таких как разборчивость синтезируемой речи, сколько о более сложных и комплексных характеристиках.

Например, естественность синтезируемой речи, возможность поддержки различных языков синтеза и голосов дикторов (так называемые многоголосые ССРТ).

Последний аспект требует особого подхода и внимания, поскольку перенастройка системы на нового диктора требует больших материальных и временных затрат от разработчиков системы.

В докладе рассмотрен возможный путь решения задачи построения многоголосого синтезатора речи с использованием системы конверсии голоса.

Конверсия голоса – это процесс преобразования параметров речевого сигнала, характеризующих исходного диктора, в параметры целевого, без изменения лингвистической составляющей самого сообщения.

Конверсия подразумевает изменение акустических, фонетических и просодических характеристик исходного диктора в характеристики целевого согласно определенному набору правил.

Предложенный авторами подход позволяет рассматривать два этих вида систем не просто как их суперпозицию, а как интеграцию системы конверсии и системы синтеза речи по тексту для максимального