

Автоматический ТРАНСКРИПТОР ГОЛОСА

Обзор

Автоматический транскриптор голоса используется для проведения фонемной транскрипции записи речевого сигнала неизвестного голоса и идентификации его языка к принадлежности к одному из языков существующих в базе системы.

Система состоит из набора модулей:

- Модуль дикторо независимого распознавания фонем;
- Модуль идентификации языка принадлежности речевого сигнала к одному из существующих в базе системы языков.

Модуль распознавания фонем разработан на основе скрытых Марковских моделей (HMMs) и модели длительности фонем. Модуль идентификации языка основан на так называемой "double bi-gram" лингвистической модели языка. "Double bi-gram" модель позволяет отследить в речевом сигнале вероятность переходов между фонемами, с дальнейшим парном ее сравнении с "матрицей языка", хранимой в системе для каждого языка. "Матрица языка" состоит из вероятностей переходов между фонемами данного языка, и для каждого языка она уникальна.

Система эффективно может использоваться:

- Для автоматической идентификации неизвестного голоса по фонограммам телефонных переговоров;
- В системах контроля, где важно достоверно точно идентифицировать язык говорящего диктора и сделать фонетическую транскрипцию его голоса.

В настоящее время модуль идентификации языка обучен для английского языка. В ближайшем будущем планируется обучить для немецкого, французского, китайского, японского и русского языков.

Достоинства

- Операции с низким уровнем SNR;
- Быстрая адаптация к канальным искажениям и внешним шумам;
- Дикторонезависимость;
- Надежность распознавания фонем ~ 75% для речевой базы TIMIT;
- Надежность идентификации языка ~ 95% для речевого сигнала длительностью не менее 10 сек.;
- Простота встраивания в целевое приложение.

Требования к сигналу

- Формат сигнала: 16-bits linear;
- Частота оцифровки: 8 кГц;
- Отношение сигнал-шум (SNR), не менее 10 db;
- Полоса частот сигнала: 300-3500 Гц или лучше.

Доступность

- Float-point C++ код;
- Библиотека для MS Windows;
- PC демо для MS Windows по требованию;
- Портруемость кода на DSP, ARM или RISC платформы.

О компании GritTec

ГритТек лаборатория специализируется на исследованиях и разработках уникальных алгоритмов в области цифровой обработки сигнала (DSP) и аудио-речевых технологий. Основные исследования ГритТек лаборатории сфокусированы в области: технологий шумоподавления и восстановления речевых сигналов, технологий биометрии голоса, анализа и синтеза речевых сигналов. Все наши новые разработки и готовые решения могут быть оптимизированы и встроены в клиентскую систему в соответствии с требованиями заказчика.

Контакты

URL: <http://www.grittec.ru>
E-mail: sales@grittec.ru