

УДК 537.531

ПРОГРАММНОЕ СРЕДСТВО РАСПОЗНАВАНИЯ ДИКТОРА ПО РЕЧИ

Куницкий Ю.О., Лунь Н.С., Зайковский В.С.

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Зельманский О. Б. – канд. техн. наук

Обоснована целесообразность применения уникальных особенностей речи при решении задачи верификации пользователя. Предложены программные модули распознавания речевого сигнала и верификации диктора по голосу.

Стремительное развитие и повсеместное распространение информационных систем и технических средств обуславливает необходимость компьютерной обработки речевой информации, поскольку голосовой интерфейс взаимодействия представляется наиболее удобным и востребованным. Таким образом, целесообразно использовать голосовые технологии для биометрики, а именно верификацию говорящего по его голосу. Верификация по голосу представляет собой процедуру подтверждения личности с помощью индивидуальных особенностей речи. Различают текстозависимые и текстонезависимые системы верификации. Текстозависимые требуют произнесения определенной фразы и сравнивают ее с эталоном для каждого пользователя. Верификация в текстонезависимых системах осуществляется на основании любого речевого фрагмента заданной длины.

Первоочередной задачей верификации говорящего является распознавание речевого сигнала. Для ее решения предлагается программный модуль осуществляющий анализ акустической обстановки и расчет следующих параметров сигнала: среднеквадратическое значение, среднее число перехода сигнала через нуль, период основного тона. Далее рассчитанные значения параметров сравниваются с устанавливаемыми пользователем пороговыми значениями. В случае если значения одновременно всех параметров удовлетворяют установленным требованиям, то анализируемый фрагмент сигнала считается речевым. Данный модуль реализован на языке программирования C++ в среде Visual Studio. Предусмотрена возможность вручную задавать пороговые значения параметров с целью адаптации настройки алгоритма и изучения особенностей распознавания речевых сигналов. Структурная схема модуля представлена на рисунке 1.



Рисунок 1 – Структурная схема модуля распознавания речи

Непосредственная верификация пользователя реализована в отдельном модуле, на вход которого подаются фрагменты сигнала, классифицированные как речевые. Модуль верификации также реализован на языке программирования C++ в среде Visual Studio. Работа модуля построена таким образом, что предварительно требуется тестовый ввод логина с целью идентификации пользователя, а следом произнесения пароля. Это позволяет избежать необходимости эмпирического поиска баланса между возможностью возникновения ошибок первого и второго рода, понижая при этом коэффициент когнитивного сопротивления конечного пользователя [1]. В модуле из частотной характеристики сигнала с помощью дискретного косинусного преобразования выполняется расчет вектора мел-кепстральных коэффициентов [2], который сравнивается с базой эталонных записей пользователей. В качестве эталонной записи для каждого пользователя используется усреднённое значение вектора мел-кепстральных коэффициентов трех вариантов

произнесения парольной фразы. Сравнение реализовано при помощи самоорганизующейся нейросети Кохонена.

Список использованных источников:

1. Куницкий, Ю.О. Верификация диктора по голосу на базе метода динамического искажения времени / Ю.О. Куницкий, О.Б. Зельманский // Технические средства защиты информации: материалы XIX Белорус.-российск. науч.-техн. конф., Минск, 8 июня 2021 г. / БГУИР ; редкол.: Т.В. Борботько [и др.]. – Минск, 2021. – С. 59.
2. Запрягаев С.А., Коновалов А.Ю. Распознавание речевых сигналов. Вестник ВГУ, 2009, №2, с. 39-48.