Список цитируемых источников

1. *Шилдт, Г.* Java 8. Полное руководство / Г. Шилдт — М. : Вильямс, 2015. — 1377 с.
2. *Вестра, Э.* Разработка геоприложений на языке Python / Э. Вестра — М. : ДМК-Пресс, 2017. — 446 с.
3. *Dalal, N.* Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — San Diego, 2005. — С. 23—26.
4. Классификация и регрессия с помощью деревьев принятия решений [Электронный ресурс]. — Режим доступа : https://habr.com/post/116385 . — Дата доступа : 20.04.2022.
5. *Шапович, Е. Г.* Методы распознавания отпечатков пальцев и реализация на высокоуровневом языке программирования C# / Е. Г. Шапович, А. В. Шах // Актуальные направления научных исследований XXI века : теория и практика. — 2019. — Т. 7, № 1(44). — С. 477—480.

УДК 004.522

О. И. Наранович, А. В. Шах
Учреждение образования «Барановичский государственный университет», Барановичи, Республика Беларусь

МОДУЛЬ ГОЛОСОВОЙ ИДЕНТИФИКАЦИИ ВОДИТЕЛЕЙ ТРАНСПОРТНЫХ СРЕДСТВ

Введение. На сегодняшний день одной из актуальных проблем, возникающих в ходе мониторинга дорожной обстановки, является проблема идентификации водителей транспортных средств. В основе указанной проблемы лежит необходимость обеспечения непрерывности такой идентификации, что позволит оперативно обнаруживать лиц, неправомерно управляющих транспортными средствами, а также идентифицировать личности участников дорожно-транспортных происшествий.

Биометрическая идентификация личности основана на принципе распознавания и сравнения уникальных характеристик человеческого организма. Уникальность голосовой биометрии состоит в том, что это единственная биометрическая модальность, которая позволяет идентифицировать водителя по микрофону, не прибегая к установке дорогостоящего оборудования. При этом по уровню надежности голосовая биометрия не уступает, а по некоторым характеристикам превосходит характеристики других систем биометрической идентификации.

Основная часть. Целью данного проекта является разработка модуля системы биометрической идентификации водителя по голосу. На рисунке 1 приведена общая схема работы приложения. На рисунке 2 более подробно отображен процесс извлечения признаков входного речевого сигнала.

Рисунок 1 — Общая схема работы приложения

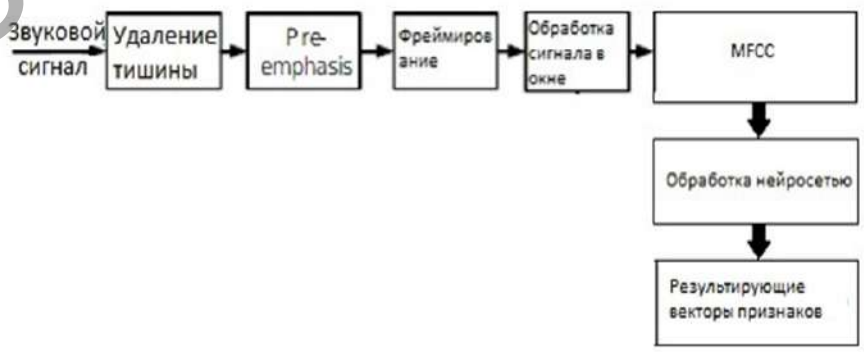

Рисунок 2 — Процесс извлечения признаков из входного речевого сигнала

Захваченный аудиосигнал может содержать тишину в разных позициях, таких как начало сигнала, между словами предложения, конец сигнала и т. д. Если включены бесшумные кадры, ресурсы моделирования расходуются на части сигнала, которые не способствуют идентификации. Настоящее молчание должно быть удалено перед дальнейшей обработкой.

Обычно речевой сигнал предварительно подчеркивается перед любой дальнейшей обработкой, если посмотреть на спектр для вокализованных сегментов, таких как гласные, на более низких частотах больше энергии, чем на более высоких частотах. Это падение энергии на частотах вызвано природой глотательного импульса. Увеличение энергии высоких частот делает информацию от этих высших формантов более информативной для дальнейшей обработки [1].

Для того чтобы получить векторы признаков одинаковой длины, нужно «нарезать» речевой сигнал на равные части, а затем выполнить преобразования внутри каждого сегмента. Обычно сегменты выбирают таким образом, чтобы они перекрывались либо наполовину, либо на 2/3. Перекрытие используется для предотвращения потери информации о сигнале на границе. Чем меньше перекрытие, тем меньшей размерностью в итоге будет обладать вектор свойств, характерный для рассматриваемого участка, поскольку он составляется из кепстральных коэффициентов каждого сегмента в отдельности. Кепстральными коэффициентами называют набор чисел, полученных после спектрального анализа участка звукового сигнала. Обычно выбирается длина участка (сегмента), соответствующая временному интервалу в 20-30 мс [2].

Цель этого этапа обработки — снижение граничных эффектов, возникающих в результате сегментации. Для подавления нежелательных граничных эффектов принято умножать сигнал $s(n)$ на оконную функцию $w(n)$: $x(n) = s(n)w(n)$. В качестве функции $w(n)$ часто используется окно Хэмминга, которое задается следующей формулой:

$$w(n) = \begin{cases} 0,54 - 0,46\cos(\frac{2\pi n}{N-1}), 0 \le n \le N \\ 0, -otherwise \end{cases}.$$

Для спектрального выравнивания речевого сигнала его следует пропустить через низкочастотный фильтр. Цель этого преобразования — снизить влияние локальных искажений на характеристические признаки, которые в дальнейшем будут использоваться для распознавания. Часто низкочастотная фильтрация осуществляется на аппаратном уровне, хотя существуют различные математические методы, которые успешно применяются в задачах работы со звуком. В рассматриваемой далее системе такие методы не использовались. Известно, что наиболее информативные частоты человеческого голоса сосредоточены в интервале 0,1—2КГц, поэтому при решении задач распознавания речи уже на начальном этапе в спектрограмме оставляют только гармоники, частоты которых попадают в этот интервал [3].

В дополнение к проделанным ранее действиям применяется преобразование Фурье и вычисление весовых коэффициентов. Полученный вектор свойств является конечным и позволяет сверять его с уже сохраненными образцами. Так как совпадение из-за различных внешних условий может быть не полным простое сравнение не будет корректным и по этой причине требуется наличие обученной нейронной сети.

На данном этапе полученный вектор свойств сравнивается с существующими данными для каждого пользователя. В случае прохождения порога делается вывод о совпадении говорящего и проверяемой записи. В случае успешной идентификации полученный образец голоса добавляется в БД решая проблему «свежести» данных.

Для того чтобы записать новый образец голоса необходимо заполнить поле «Sample name» и нажать на изображение записи. После нажатия на кнопку записи активируется таймер. Запись останавливается спустя 10 секунд о чем сигнализирует изменившийся цвет таймера. Вид страницы представлен на рисунке 3.

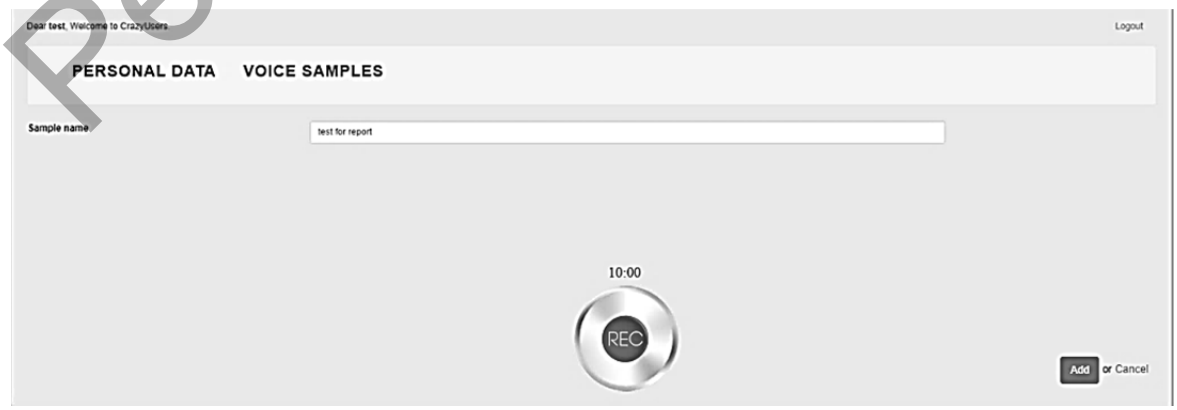

Рисунок 3 — Запись нового образца голоса

После нажатия кнопки «Add» произойдет переход на страницу с образцами голосов включая только что добавленный образец, представленную на рисунке 4.

Рисунок 4 — Результат записи нового образца голоса

Заключение. Уникальность голоса человека обусловлена множеством физиологических особенностей (строением голосовых связок, трахеи, носовых полостей, манерой произношения звуков, расположением зубов). Комбинация этих особенностей индивидуальна, как и отпечатки пальцев. Однако на практике ни одна из унимодальных систем биометрической идентификации, в том числе и голосовая, не может гарантировать 100 %-ной идентификации личности. Основными источниками ошибок при идентификации дикторов являются эффекты среды записи (уровень и тип шума, уровень реверберации), представления (длительность речи, психофизиологическое состояние говорящего (болезнь, эмоциональное состояние и т. п.), язык речевого сообщения, изменение голосового усилия), канала (помехи (импульсные, тональные и т. п.), искажения (амплитудно-частотные характеристики микрофона и канала передачи, вид кодирования в канале и т. д.)).

Основная проблема такого подхода состоит в том, что создание базы данных с необходимым количеством комплексных биометрических образов, невозможно в силу комбинаторного взрыва (резкого роста числа возможных комбинаций биометрических образов). Более того, именно этот комбинаторный взрыв обеспечивает защиту от копирования биометрического образа. Рассматривая приведенный пример, система идентификации может запросить произнести человека произвольную фразу, по которой проведет ее идентификацию, очевидно, что попытка злоумышленником создать базу данных всех возможных фраз (включая не только голос, но и мимику, и ряд других биометрических характеристик) обречена на провал, как и попытка сделать тоже самое специалистами по защите информации. Отсюда можно сделать вывод, что формальные методы слабо подходят для решения задачи биометрической идентификации. При этом, важно понять, что речь идет об использовании формальных методах для объединения результатов комплексной биометрической идентификации, полученных с использованием эвристических методов.

Список цитируемых источников

1. *Граничин, О. Н.* Рандомизированный алгоритм стохастической аппроксимации в задаче самообучения / О. Н. Граничин. — М. : ЕАОИ, 2009. — 232 с.

2. Digital Speech Processing [Электронный ресурс]. — Режим доступа : https://www.academia.edu/15982255/Digital_Speech_Processing . — Дата доступа : 18.04.2022.

3. *Фролов, А. В.* Синтез и распознавание речи. Современные решения / А. В. Фролов. — М. : Вильямс, 2010. — 186 с.

УДК 375.14

О. С. Платун, А. В. Шах

Учреждение образования «Барановичский государственный университет», Барановичи, Республика Беларусь

ПРОВЕДЕНИЕ НЕЛИНЕЙНОГО АНАЛИЗА НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ АВТОМОБИЛЬНОГО АМОРТИЗАТОРА ПРИ ДИНАМИЧЕСКИХ НАГРУЗКАХ

Введение. С развитием тенденции по внедрению ЭВМ, в промышленности появились первые системы автоматизированного проектирования, способные автоматизировать цикл изготовления изделия, начиная от создания технической документации и до процесса изготовления изделия. В настоящее время, в мировой промышленности используется большое количество всевозможных САПР. Примером одной из таких является SolidWorks.