

УДК: 519.2, 612.087, 621.319.7

Калашников Д.М., Захаров О.С., Ахметов Б.Б., Пашенко Д.В.

Создание среды моделирования «БиоНейроГолос», ориентированной на обучение больших искусственных нейронных сетей алгоритмом ГОСТ Р 52633.5 при выполнении лабораторных работ русскоязычными студентами университетов России, Беларуси и Казахстана

В настоящее время проблема получения достоверных биометрических данных студентами России, Беларуси и Казахстана решена через использование среды моделирования «БиоНейроАвтограф» [1]. Необходимость в создании этого продукта обусловлена тем, что фирмы производители средств биометрии не обеспечивают бесплатную поддержку своих продуктов.

Например, если ориентироваться на проведение лабораторных работ с линейкой продуктов фирмы Fujitsu (Япония), то каждому университету придется покапать сканер подкожных кровеносных сосудов (стоимость 20 000 рублей), также один из университетов должен приобрести SDK (комплекс средств разработки стоимостью 200 000 рублей). В итоге создание лабораторных работ под эту биометрическую технологию оказывается коммерчески невыгодным.

Аналогичная ситуация возникает при анализе возможности создания лабораторных работ, ориентированных на применение сканеров рисунка отпечатка пальца, например, фирмы BioLink со стоимостью датчика 8 000 рублей и стоимостью SDK 10 000 рублей.

Ситуация меняется, если ориентироваться на использование мыши или тачпада для ввода рукописных слов. В этом случае стоимость сканера биометрии оказывается нулевой, что позволяет выполнять лабораторные работы по биометрии без каких либо материальных затрат в свободно распространяемой среде моделирования «БиоНейроАвтограф» [1]. Если же исследователям университетов требуются биометрические данные более высокого качества, то среда моделирования [1] должна использоваться с любым графическим планшетом. Ориентация на графический планшет фирмы Wacom потребует материальных затрат от 5 000 рублей и выше.

Именно в силу незначительных материальных затрат среда моделирования «БиоНейроАвтограф» [1] оказалась востребована университетами России, Беларуси и Казахстана. Однако единственная технология анализа рукописных биометрических данных не может адекватно отражать перспективные технологии в курсах университетов. Это становится очевидным в силу того, что нейронные сети оказываются удобны для воспроизведения и поддержки эффектов квантовой суперпозиции [2, 3]. То есть достоверные биометрические данные в ближайшем будущем будут востребованы университетами не только по линии кафедр «Информационной безопасности», но и кафедрами «Вычислительной техники» для проведения лабораторных работ по изучению квантовых вычислений.

Возникает необходимость в создании еще одной среды моделирования «БиоНейроГолос», ориентированной на использование подключаемого или встроенного в компьютер микрофона или ларингофона. Предположительно новая среда моделирования будет использовать уже имеющиеся программные модули

среды «БиоНейроАвтограф» и отличаться некоторыми дополнительными функциями. В рамках данного доклада мы попытаемся описать дополнительные функции нового программного продукта.

Одним из отличий нового программного продукта является то, что приходится учитывать возможную рассинхронизацию «пальцев» и «голоса» пользователя. Для ввода голосового образа с помощью микрофона пользователь должен нажать клавишу «Record», при этом не всегда пользователь нажимает клавишу до произнесения парольной фразы. На рисунке 1 приведена ситуация, когда пользователь верно нажал клавишу ввода речи.

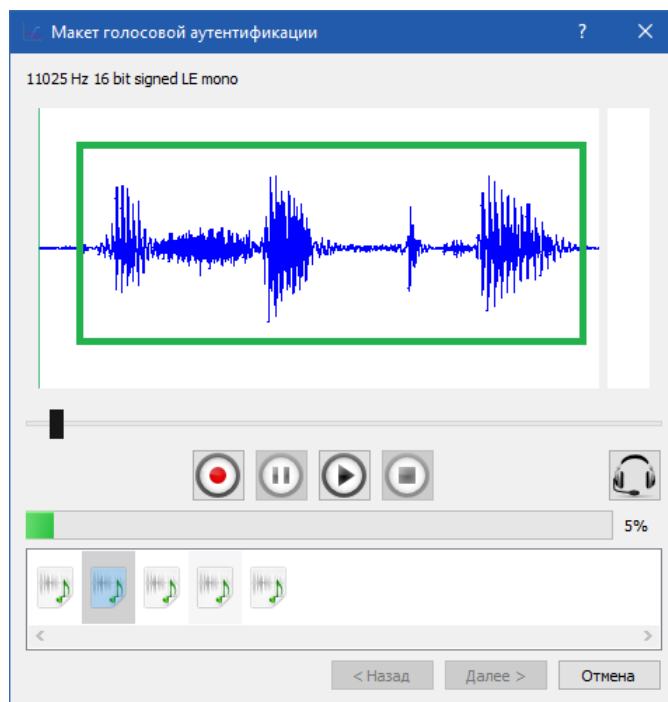


Рис. 1. Пример корректного ввода голосовой фразы при выделении ее начала и конца по тональным компонентам

На рисунке 2 отображена ситуация, когда пользователь нажал клавишу записи речи с опозданием. При этом данные с микрофона/ларингофона начали поступать с опозданием. На рисунке 2 в сравнении с рисунком 1 отсутствует первый тональный звук. В связи с этим автоматически выделенный интервал обработки звука оказывается неверным. Формируемый для обработки звуковой файл оказывается намного короче других файлов в обучающей выборке. Он помечается как «ошибочный» и пользователь должен подтвердить обнаруженную ошибку ввода данных.

Необходимость в подобных дополнительных функциях обусловлена тем, что в режиме работы среды моделирования «БиоНейроГолос» в режиме преобразователя биометрия-код предполагается использовать упрощенный алгоритм выделения голосовой фразы по началу первого из обнаруженных тональных фрагментов и по концу последнего из обнаруженных тональных фрагментов.

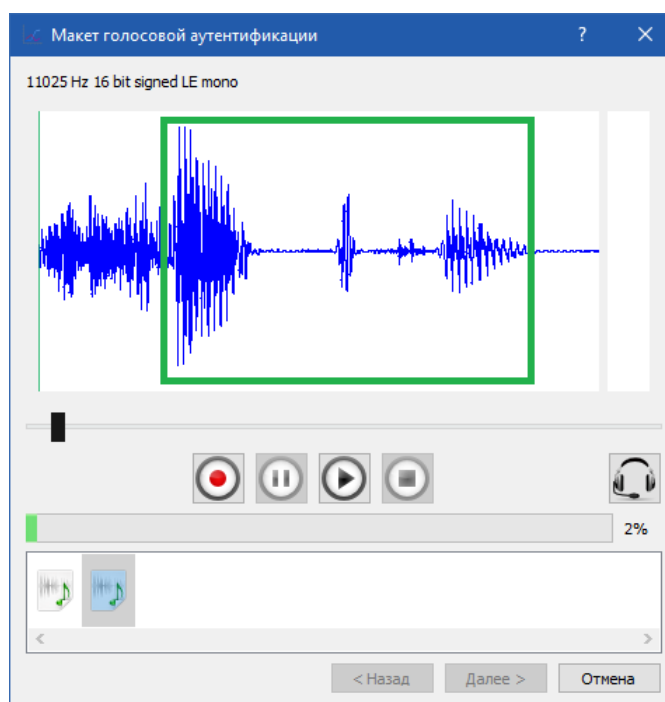


Рис. 2. Пример не корректного ввода голосовой фразы при выделении ее начала и конца по тональным компонентам

В связи с тем, что разрабатываемая среда моделирования «БиоНейроГолос» имеет двойное применение (биометрия и квантовые компьютеры) в ней предполагается иметь два режима входной предобработки данных. Если первый режим выделяет все слово-пароль (см. рисунок 1 и рисунок 2), то второй режим должен быть ориентирован на выделение отдельных звуков, как это показано на рисунке 3.

Отличия первого и второго режимов работы принципиальные. В первом режиме используется анализ длинного звукового фрагмента, состоящего из нескольких слов парольной фразы. На данных этого длинного звукового фрагмента формируется большая нейронная сеть, анализирующая порядка 1000 биометрических параметров с длиной выходного кода 256 бит. Такая сеть обучается алгоритмом ГОСТ Р 52633.5 [4] и имеет малое число ошибок выходного кода, корректируемых сверточным кодом без избыточности [5].

Совершенно иная ситуация возникает, когда происходит дробление слитной речи на звуки (рисунок 3). В этом случае на каждый звук создается собственная нейронная сеть [6]. Нейронных сетей оказывается множество и все они должны образовывать систему выходных кодов, обладающую высокой корректирующей способностью [2]. При этом соседние нейронные сети должны корректировать друг друга, а их групповая корректирующая способность должна быть намного выше, чем корректирующая способность классических кодов с избыточностью (например, БЧХ кодов).

Программный автомат выделения звуков и их комбинаций должен непрерывно подстраиваться под текущий темп речи человека [7, 8], опираясь на знание среднего значения длины звука и знание его допустимого интервала вариаций.

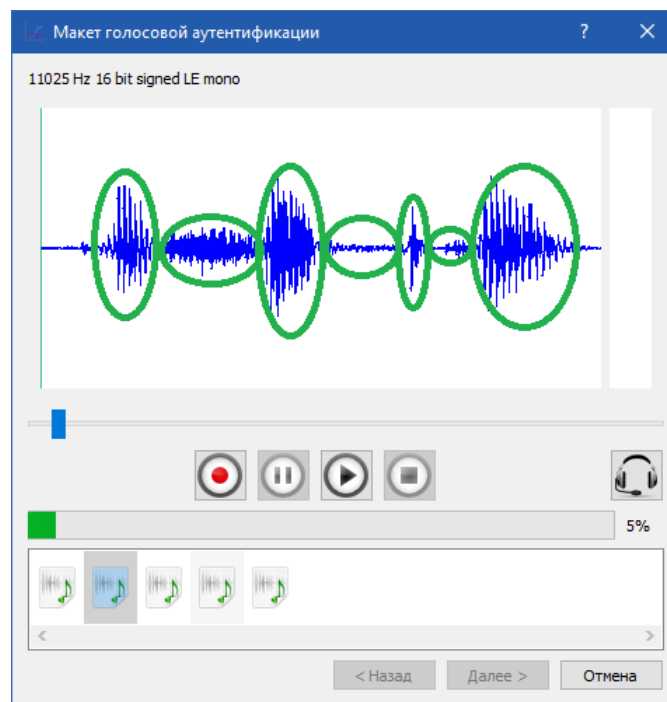


Рис. 3. Пример обработки потока звуковых колебаний с автоматизированным выделением звуковых фрагментов, образующих разные звукоформы

Так как на данный момент создать достаточно качественный фрагментатор речи на звуки никому не удалось, среда моделирования «БиоНейроГолос» должна иметь режим «ручной» корректировки работы этого автомата.

Таким образом, среда моделирования «БиоНейроГолос» будет иметь существенно более сложную предобработку биометрических данных. Это прежде всего связано с компенсацией искажений, вносимых в голосовые данные «неподготовленным» пользователем. Предположительно необходимость в ручной корректировке исходных данных, вводимых студентами будет снижаться по мере выполнения курса лабораторных работ по нейросетевому преобразованию биометрических данных в длинный код парольной фразы или в последовательность более коротких кодов звуков и их сочетаний.

Вопрос шумоподавления внешних источников шума в первой версии разрабатываемого программного обеспечения «БиоНейроГолос» предполагается решить применением ларингофонов, что приведет к росту материальных затрат до 800 рублей на покупку одного ларингофона. Тип используемых ларингофонов и возможное появление программных отличий для входа «ларингофон» уточняются на конечных этапах разработки.

ЛИТЕРАТУРА:

1. Иванов А.И., Захаров О.С. Среда моделирования «БиоНейроАвтограф». Программный продукт создан лабораторией биометрических и нейросетевых технологий, размещен с 2009 г. на сайте АО «ПНИЭИ» <http://пниэи.пф/activity/science/noc.htm> для свободного использования университетами России, Белоруссии, Казахстана.
2. Иванов А.И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции. Издательство АО «ПНИЭИ», Пенза-2016 г., 133 с. Свободный доступ <http://пниэи.пф/activity/science/BOOK16.pdf>

3. Кулагин В., Иванов А., Газин А., Ахметов Б. Циклические континуально-квантовые вычисления: усиление мощности хи-квадрат критерия на малых выборках. /Аналитика № 5, 2016 (30), с. 22-29.
4. ГОСТ Р 52633.5-2011 «Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа».
5. Безяев А.В. Нейросетевой преобразователь в самокорректирующийся код, совершенно не обладающий избыточностью «Нейрокомпьютеры: разработка, применение» №3, 2012 с. 52-55.
6. Иванов А.И., Хальметова А.Н., Захаров О.С., Рыболовлев А.А., Рыжаков А.П. Нейросетевой вокодер-архиватор, сохраняющий биометрические особенности голоса говорящего при высоком уровне сжатия шипящих звуков «Нейрокомпьютеры: разработка, применение» №3, 2012 с. 44-48
7. Калашников, Д.М. Биометрия: предсказание длительности тональных звуков слитной речи /Д.М. Калашников // Инфокоммуникационные технологии. – 2014. – № 3. – С. 72–76.
8. Бойков И.В., Иванов А.И., Калашников Д.М. Алгоритм построения статистического дискретно-континуального описания длительности звуков потока осмысленной речи диктора. Известия высших учебных заведений. Поволжский регион. Технические науки. – Пенза: ПГУ, №4, 2015 с. 45-49

Статья поступила 25.09.2016, опубликована 03.11.2016 по положительной рецензии к.т.н. Зефирова С.Л.