НАУЧНОЕ ИЗДАНИЕ МГТУ ИМ. Н. Э. БАУМАНА

НАУКА и ОБРАЗОВАНИЕ

Эл № ФС77 - 48211. Государственная регистрация №0421200025. ISSN 1994-0408

электронный научно-технический журнал

Идентификация личности по голосу на основе самоорганизующейся карты Кохонена

11, ноябрь 2012

DOI: 10.7463/1112.0500433

Карпенко А. П., Морозов А. Н.

УДК 519.6

Россия, МГТУ им. Н.Э.Баумана <u>apkarpenko@mail.ru</u> <u>alexnikmorozov@gmail.com</u>

Введение

Системы голосовой идентификации личности (СГИЛ) интенсивно развиваются в последние годы. Стимулом для развития СГИЛ является их востребованность в таких областях, как криминалистический учет и биометрический поиск, голосовая верификация водителя и пассажиров, разграничение прав доступа к информации с помощью голосовой биометрии и т. д. [1]. Важным достоинством СГИЛ по сравнению с другими биометрическими системами идентификации является их дешевизна. Важно также, что современные СГИЛ по уровню надежности идентификации не уступают, а иногда и превосходят, например, системы идентификации человека по изображению [2].

Отметим близость задачи идентификации личности по голосу к задаче распознавания речи. Системы распознавания речи, например, Google Voice Search, Dragon Naturally Speaking, Voice Navigator, Microsoft Voice Command, автоматически переводят голос пользователя в текст, освобождая его руки. Эволюция систем распознавания речи привела к созданию интеллектуальных систем, позволяющих не только распознавать, но и автоматически синтезировать человеческую речь. Такие системы широко используются с целью снижения нагрузки на операторов контакт-

центров и секретарей, для организации общения пользователей с голосовыми порталами и т.д.

Несмотря на уникальность голоса человека, обусловленную уникальностью строения его голосовых связок, трахеи и носовых полостей, манерой произношения звуков, расположением зубов и так далее, ни одна из СГИЛ, как и любая другая биометрическая система, не может гарантировать 100 % надежность идентификации. Основными источниками ошибок в СГИЛ являются: окружение (шум, реверберация и т.д.); особенности речи (длительность, тональность, уровень голосового усилия и т.д.); канал связи (искажения микрофона и канала передачи, погрешности кодирования аудиосигнала и т.д.) [2].

В общем случае идентификация личности по голосу требует решения большого числа разнородных задач, основными из которых являются следующие:

- выделение вокализованных участков аудиосигнала путем отбрасывания пауз и участков, содержащих различного рода помехи;
 - разделение речи дикторов (задача диаризации);
 - выделение характерных признаков голоса диктора.

Задаче выделения вокализованных участков аудиосигнала посвящено большое число работ. Известны решения этой задачи, обеспечивающие уровень ошибок, не превышающий 10 % (см., например, работу [3]).

Компоненту СГИЛ, которая решает задачу диаризации, называют детектором речевой активности (*voice activity detector, VAD*). Обзор основных алгоритмов, используемых в этих детекторах, приведен, например, в работе [2].

В нашей работе основное внимание уделено последней, третьей задаче выделения характерных признаков голоса (ХПГ). Вообще говоря, для выделения ХПГ можно использовать спектрально-формантный анализ, статистические характеристики основного тона голоса, параметрическое представление аудиосигнала. В последнем случае ХПГ можно искать в

амплитуде и диапазоне частот звукового сигнала, а также в его спектре. Мы используем метод, предполагающий идентификацию человека по голосу на основе набора основных частот в его спектре. Обширный список публикаций, посвященных перечисленным методам, представлен в работе [4].

Набор ХПГ называем картой ХПГ. После формирования карты ХПГ данного человека возникает задача сравнения этой карты с картами ХПГ, хранящимися в базе данных СГИЛ. Оригинальность нашей работы заключается в использовании с этой целью самоорганизующейся карты Кохонена (Self-Organizing Map, SOM), которая представляет собой соревновательную нейронную сеть с обучением без учителя. Сеть преобразует входные многомерные векторы в одно- или двухмерную дискретную карту [5].

В первом разделе статьи приводим постановку задачи и общую схему предложенного метода идентификации личности по голосу. Второй раздел содержит описание программной реализации метода. В третьем разделе приведены результаты тестирования исследования И эффективности принятых алгоритмических и программных решений. В заключении сформулированы основные результаты работы И перспективы ее развития.

1. Постановка задачи и общая схема метода

Полагаем, что аудиосигнал представляет собой одномерный равноотстоящий ряд $X=(x(t_i),\ t_i\in[0;\hat{T}]),\ x(t_i)=x_i\in R^1,\ i\in[1:T]$ абсолютных значений громкости сигнала в диапазоне 0-90 дБ. В аудиосигнале представлены частоты голоса человека в диапазоне от 40 до 1600 Гц и длительностью \hat{T} от нескольких секунд до нескольких минут. Аудиосигнал содержится в файле формата *.wav.

В общем случае идентификатор C сигналов определяют [6] как оператор, сопоставляющий сигналу X номер класса k, к которому он относится:

$$C(X) = k, k \in \{1, 2, ..., K\}.$$

Очевидной является идея использования величин x_i , $i \in [0:T]$ в качестве $(T \times 1)$ -вектора признаков сигнала. Однако в силу высокой размерности так определенного вектора (длина сигналов может достигать нескольких тысяч отсчетов), а также различной длины сигналов, данную идею на практике не используют. Подавляющее большинство методов идентификации временных рядов предполагает переход с помощью некоторого оператора Ψ от ряда X к некоторому вектору характерных признаков $\mathcal G$ этого ряда (статическому паттерну) размерности M:

$$\Psi(X) = \mathcal{G}, \quad \mathcal{G} \in \mathbb{R}^M.$$

Таким образом, задачу идентификации временного ряда ставим как задачу поиска операторов Ψ, C таких, что

$$\Psi(X) = \mathcal{G}, \quad \mathcal{G} \in \mathbb{R}^M,$$

$$C(\mathcal{Y})=k, \quad k\in \{1,2,....,K\}.$$

Система *VOICE* состоит из обучающей и распознающей подсистем. Первая из подсистем предназначена для формирования базы данных голосов в форме их карт ХПГ. Вторая подсистема призвана обеспечить поиск в указанной базе заданного числа ранжированных голосов, в некотором смысле наиболее близких к анализируемому голосу. Основу обеих подсистем составляет метод, обеспечивающий формирование карт ХПГ, а также сравнение этих карт (рисунок 1).

376

10.7463/1112.0500433

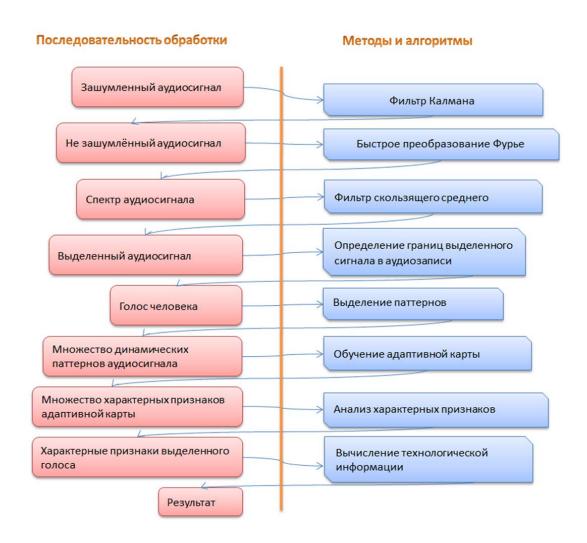


Рисунок 1 – Схема метода формирования карты ХПГ

Таким образом, формирование карты XПГ в программной системе VOICE производим по следующей схеме.

- 1) Калмановская фильтрация. Зашумленный аудиосигнал $X = (x_i, i \in [0:T])$ с записью речи человека обрабатываем фильтром Калмана [7]. Наряду с фильтром Калмана, в вычислительной практике рекурсивные фильтры Чебышева, Бесселя, широко используют эллиптические фильтры и гребенчатые фильтры [8]. Баттерворта, Предпочтение фильтру Калмана вследствие отдано высокой его эффективности.
- 2) Быстрое преобразование Фурье. Полученный на выходе фильтра Калмана сигнал $\overline{X} = (\overline{x}_i, i \in [0:T])$ передаем на вход быстрого преобразователя Фурье, выходом которого является спектр сигнала

- $S = (s_j, j = 1, 2,...)$. Известной альтернативой БПФ является вейвлет-преобразование [8]. Предпочтение БПФ отдано, прежде всего, вследствие простоты его реализации и интерпретации результатов преобразования, высокой скорости вычислений. Важно также, что по своей природе голос человека представляет собой сумму ограниченного числа гармонических сигналов. Поэтому функция x(t) не может содержать, например, импульсов прямоугольной или треугольной формы.
- 3) Фильтрация методом скользящего среднего [9]. Обрабатываем спектр S фильтром скользящего среднего так, чтобы в полученном спектре \overline{S} присутствовал только аудиосигнал.
- 4) *Определение границ сигнала*. Удаляем из аудиозаписи фрагменты, содержащие только шум (когда человек не говорил).
- 5) Выделение паттернов (ХПГ). Осуществляем переход от исходного пространства динамических данных к пространству многомерных статических данных. Выделение паттернов производим методами многомерной классификации.
- 6) Обучение карты Кохонена. Результатом обучения является карта ΧПГ данной личности, позволяющая представить эти признаки в компактной И удобной ДЛЯ распознавания форме. Обучение самоорганизующейся карты Кохонена производим в четыре этапа до тех пор, пока не будут исчерпаны все входные векторы: инициализация сети, подвыборка, поиск максимального соответствия, коррекция весов нейронов [5].
- 7) Анализ характерных признаков (только в распознающей подсистеме) заключается в сравнении полученной карты ХПГ с аналогичными картами других личностей, хранящимися в базе данных системы. Если при этом обнаруживается карта ХПГ, близкая данной карте, то возвращаемся на один из предыдущих этапов обработки. В противном случае признаем карту корректной и сохраняем в базе данных под уникальным именем.

2. Программная организация системы

*IDEF*0-схема системы *VOICE* представлена на рисунке 2.

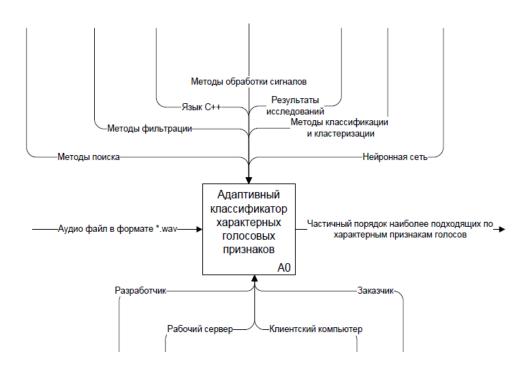


Рисунок 2 - *IDEF*0-схема системы *VOICE*

Структурно система состоит из трёх модулей, которые, вообще говоря, могут выполняться на различных ЭВМ или процессорах (рисунок 3).

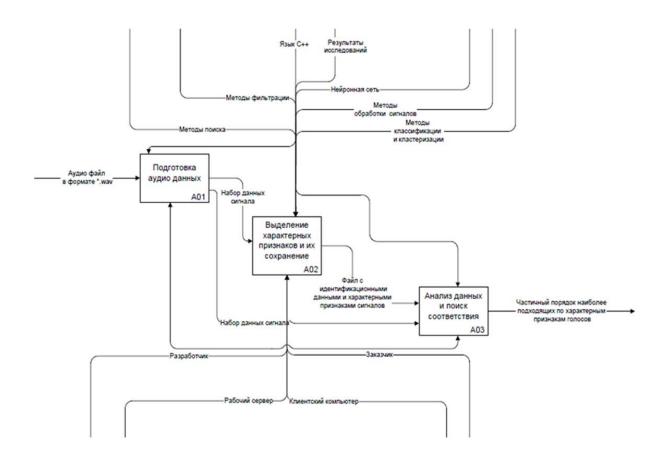


Рисунок 3 – Структура системы *VOICE*

Модуль А01 представляет собой утилиту подготовки исходных аудиоданных. На вход модуль принимает от двух до трёх параметров:

```
imperial@linux-ckfc:~/voice/audio_prep> ./convert
./convert input_file.wav out_putfile.txt <max_length (= wav_len)>
```

Первые два параметра являются обязательными и указывают на аудиофайл в формате *.wav и имя текстового файла, в который будет записана информация об обрабатываемом звуковом сигнале. Третий параметр позволяет менять длину подлежащего обработке сигнала в том случае, когда нас интересуют данные не из всего аудиофайла, а только из его некоторой части. При отсутствии третьего параметра вся информация из аудиофайла поступает в результирующий текстовый файл. Пример результата работы модуля представлен ниже.

imperial@linux-ckfc:~/voice/audio_prep>time ./convert roland.wav roland.txt

Sample rate: 44100

Channels: 1

Bits per sample: 16 Duration: 00:18

subchunk2Size 1611776

blockAlign 2 real 0m1.878s user 0m0.052s sys 0m0.852s

Пример показывает, что перекодирование 1611776 дискретных значений заняло на тестовой виртуальной машине в системе *OpenSUSE* менее двух секунд.

Структура указанного выше текстового файла имеет следующий вид:

imperial@linux-ckfc:~/voice/samples> head -n 5 alex_full.txt length=454656 time=10.3097 0 -185 2.26757e-05 -193 4.53515e-05 -75

Здесь length – число дискретных значений во временном ряде, time - полное время аудиозаписи. Далее следуют пары (t_i, x_i) , i = 1, 2, ...

Модуль А02 предназначен для построения базы данных ХПГ. На вход модуля требуется подать текстовый файл, полученный в результате работы модуля A01:

imperial@linux-ckfc:~/voice/execute> ./buildbase Preparing data Error: wrong parameter ./buildbasefile.txt

Текстовые сообщения, фиксирующие этапы работы модуля A02, выводятся в стандартный поток вывода. В приведенном выше примере таким является сообщение *Preparingdata*, свидетельствующее о том, что модуль приступил к подготовке данных. Ошибки в ходе выполнения модуля выводятся в стандартный поток вывода ошибок. Если ошибка является критической, то модуль завершается аварийно. Пример такой ошибки представлен выше (отсутствие входного текстового файла).

При нормальном завершении модуля выводятся сообщения обо всех пройденных этапах обработки данных, а также сообщение о добавлении записи в базу данных системы. Далее выводится время, затраченное на выполнение всех этапов обработки, освобождается оперативная память и завершается выполнение модуля:

imperial@linux-ckfc:~/voice/execute>time ./buildbase bruce.txt

Preparing data

Processing FFT

Processing filter

Processing FFT inverse

Teaching SOM

E=4.6229

Working with database

7 records in database.txt

Preparing result.html

Work time in ms: 5519

real 0m5.549s

user 0m1.540s

sys 0m3.432s

Характерные признаки сигнала модуль записывает в файл database.txt. Время работы модулей A02, A03 в значительной степени зависит от большого числа свободных параметров метода. В приведенном выше примере полное время работы модуля составило около пяти с половиной секунд.

Модуль А03 выполняет сравнение карты ХПГ, соответствующей аудиосигналу из файла *database.txt*, с картами, хранящимися в базе данных модуля. В качестве входного модуль А03 также использует выходной текстовый файл модуля А01.

Правила вывода информации в стандартные потоки вывода и ошибок в модуле A03 соответствуют аналогичным правилам модуля A02. В качестве результата поиска выводятся идентификаторы тех записей базы данных, которые обеспечивают максимальные значения используемой меры близости ХГП (в приведенном ниже примере это число равно четырем).

imperial@linux-ckfc:~/voice/execute> time ./identify alex.txt

Preparing test signal

Processing FFT for Test Signal

Processing filter for test signal

Teaching SOM

E=0.331418

Loading database.txt

Checking test signal

Classifier check result (sortted by relevance):

- 1) alex (id=0) 72.00%
- 2) bruce (id=1) 22.00%
- 3) kathy (id=3) 4.00%
- 4) zarvox (id=7) 2.00%

Preparing result.html

Work time in ms: 3995

real 0m4.024s

user 0m0.588s sys 0m3.180s

Модули А02, А03 продуцируют также файл result.html, содержащий отчёт об их работе. Отчет включает в себя графики сигнала до и после фильтрации шумов, веса карты Кохонена и спектр сигнала. Также в файле дублируется текстовая информация, поступающая в стандартный вывод указанных модулей. Содержимое файла result.html является вспомогательным и может быть востребовано при отладке системы и интерпретации результатов ее работы. Для сокращения вычислительных затрат запись указанных графиков в файл может быть отменена.

3. Тестирование и исследование эффективности системы

Тестирование системы *VOICE* выполнено по следующей схеме.

- 1) Двадцать человек прочли по два небольших фрагмента текста так, что длительность каждой из 40 полученных аудиозаписей составила 20-30 секунд. Аудиозаписи были получены со встроенного микрофона ноутбука в условиях зашумленности и при отключенной фильтрации шумов, которую реализует используемая операционная система.
- 2) Все аудиозаписи были разделены на два набора таким образом, что голос каждого из 20 дикторов присутствовал в каждом из наборов.

Первый набор записей использован в качестве обучающего, а второй – в качестве тестового.

3) Тестирование считалось успешным, если человек из второго набора был однозначно идентифицирован (оказался на первом месте в числе четырех претендентов, указанных в предыдущем разделе).

Тестирование обеспечило 100 % идентификацию в указанном смысле.

3.1. Оптимизация обучения карты Кохонена. Обучение адаптивной карты Кохонена представляет собой, по сути, задачу глобальной оптимизации – найти такие веса нейронов, которые обеспечивают минимум ошибки обучения *Err*. С целью повышения вероятности локализации глобального минимума этой величины в процессе обучения используем метод мультистарта по случайным начальным весам нейронов. Обозначим число этих стартов N.

Результаты исследования минимальной ошибки обучения карты Кохонена в функции числа мультистартов N представлены в таблицах 1, 2, которые иллюстрируют рисунки 4, 5.

Таблица 1 - Минимальная ошибка обучения карты Кохонена min (Err) в функции числа мультистартов N: этап построения базы данных

Голос	N							
	1	2	5	10	20			
alex	5,23	5,08	4,88	4,80	4,79			
zarvox	16,11	15,68	14,53	13,47	13,23			

Таблица 2 - Минимальная ошибка обучения карты Кохонена min (Err) в функции числа мультистартов N: этап распознавания

Голос	N							
	1	2	5	10	20			
alex	5,86	5,29	4,87	4,74	4,70			
zarvox	13,75	13,62	13,53	13,04	13,02			

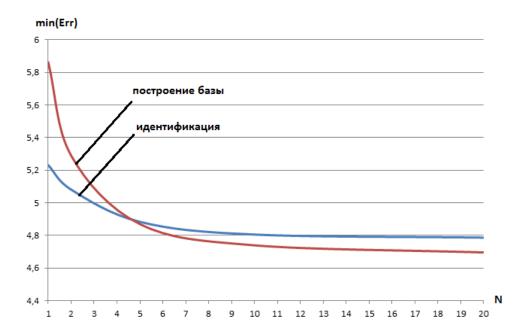


Рисунок 4 - Лучший результат обучения карты Кохонена min (Err) в функции числа мультистартов N: голос alex

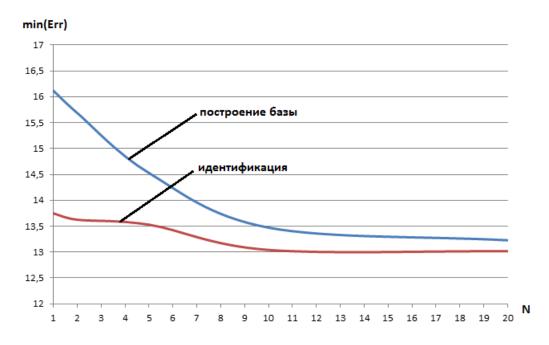


Рисунок 5 - Лучший результат обучения карты Кохонена min (Err) в функции числа мультистартов N: голос zarvox

Представленные результаты показывают, во-первых, что мультистарт может значительно повысить качество обучения карты Кохонена. Во-вторых, из представленных данных следует, что, начиная примерно с 10 стартов, имеет место стагнация ошибки обучения min (Err). На этом основании в качестве умолчательного значения величины N в системе VOICE используем значение 10.

3.2. Оптимизация числа характерных признаков. Важнейшим свободным параметром идентификатора является число характерных голосовых признаков M (равное числу нейронов карты Кохонена). В таблицах 3-6 представлены результаты исследования зависимости временных затрат на идентификацию t и минимальной ошибки обучения карты Кохонена min (Err) в функции этой величины. Указанные таблицы иллюстрируют рисунки 6-9.

Таблица 3 - Время идентификации t (сек) в функции числа ХПГ: этап построения базы данных

Голос	M								
1 0.10 0	4	9	16	25	36	49	64		
alex	5,82	7,27	9,28	11,82	14,89	19,24	25,87		
zarvox	23,33	33,14	46,68	64,00	85,49	116,43	160,30		

Таблица 4 - Минимальная ошибка обучения карты Кохонена *min* (*Err*) в функции числа ХПГ: этап построения базы данных

Голос	M							
	4	9	16	25	36	49	64	
alex	12,30	8,44	4,95	3,58	2,48	2,11	1,12	
zarvox	78,70	37,04	16,15	11,54	8,61	4,71	3,68	

Таблица 5 - Время идентификации t (сек) в функции числа ХПГ: этап распознавания

Голос	M									
	4	9	16	25	36	49	64			
alex	4,51	5,17	6,06	7,10	8,86	10,38	12,92			
zarvox	12,06	16,37	21,97	29,37	39,34	51,65	71,65			

Таблица 6 - Минимальная ошибка обучения карты Кохонена $\min(Err)$ в функции числа XПГ: этап идентификации

Голос	M							
	4	9	16	25	36	49	64	
alex	12,82	8,36	5,54	3,29	2,59	1,67	1,32	
zarvox	77,82	33,99	15,04	12,40	8,30	5,55	4,10	

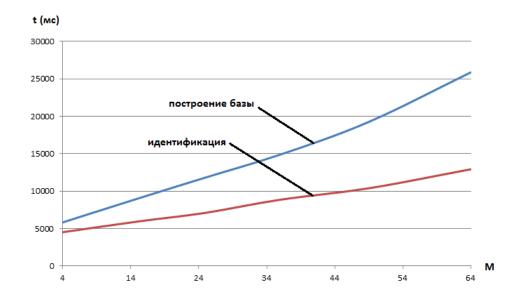


Рисунок 6 - Время идентификации t в функции ХПГ: голос alex

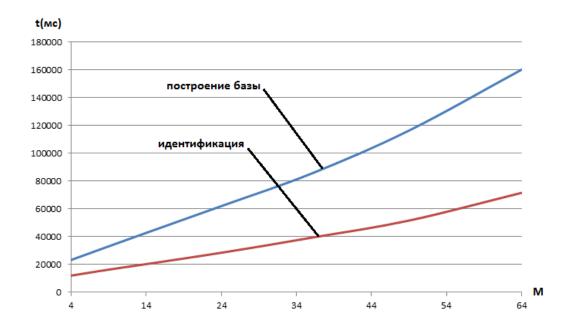


Рисунок 7 - Время идентификации t в функции числа XПГ: голос zarvox

Рисунки 6, 7 показывают, что в широких пределах изменения числа $X\Pi\Gamma$ для обоих рассматриваемых голосов имеет место почти линейная зависимость времени идентификации t в функции этого числа. Данный результат свидетельствует о высокой эффективности принятых алгоритмических и программных решений.

Минимальную ошибку обучения карты Кохонена min(Err) в функции числа ХПГ для рассматриваемых голосов иллюстрируют рисунки 8, 9.

Выбор оптимального значения числа M определяют две противоположные тенденции. С одной стороны, на основании малого числа ХПГ невозможно с высокой надежностью идентифицировать голос. другой стороны, большое число ЭТИХ признаков увеличивает вычислительные затраты на идентификацию и, тем самым, понижает ее скорость (если не говорить об использовании параллельных вычислений). На основании результатов данного исследования в качестве оптимального числа XПГ выбрано M = 25. Если в процессе идентификации в базе данных системы *VOICE* обнаруживается два или более близких голосов, то может оказаться целесообразным увеличение этого числа.

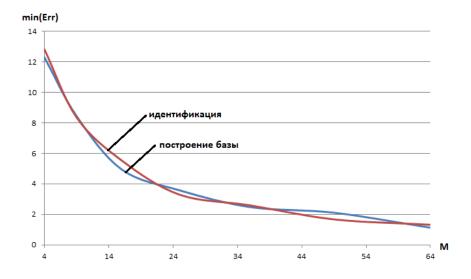


Рисунок 8 - Минимальная ошибка обучения карты Кохонена min(Err) в функции числа $X\Pi\Gamma$: голос alex

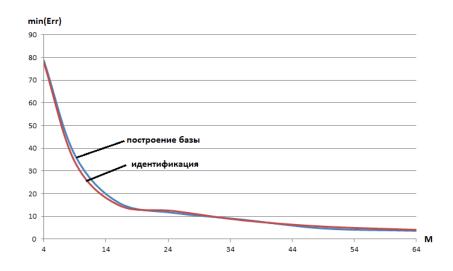


Рисунок 9 - Минимальная ошибка обучения карты Кохонена min(*Err*) в функции числа XПГ: голос *zarvox*

3.3. Исследование надежности идентификации. Выполнено также исследование надежности идентификации голоса в зависимости от величины отношения mym/curhan. В рассматриваемом диапазоне частот шум полагаем белым с нулевым математическим ожиданием. В качестве меры надежности классификации используем относительное число ошибочных решений.

Результаты исследования представлены на рисунке 10. Рисунок показывает 100 % нулевое число ошибок распознавания голоса при

отношении $\frac{\text{шум}}{\text{сигнал}}$ в диапазоне от нуля до 50 %. При увеличении указанного отношения от 50 до 100 % число ошибок повышается с 0 до $\sim 40\%$

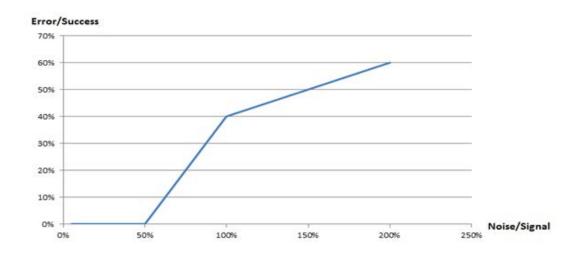


Рисунок 10 – Относительный уровень ошибок идентификации в функции уровня шума

Заключение

В работе предложен метод идентификации личности по голосу. Выполнена программная реализация идентификатора, его тестирование и исследование эффективности. Результаты исследования показывают высокую эффективность принятых алгоритмических и программных решений, а также перспективность развития метода идентификации и его программной реализации.

VOICE Разработанная быть программная система может использована не только для идентификации человека по голосу, но и для распознавания ГОЛОСОВ животных, например, птиц. Представляется перспективным использование системы для распознавания «голосов» различных технических объектов, например, с целью диагностики их состояния. Идентификатор может быть использован в телефонных конференциях для сокращения объема передаваемой информации и борьбы с шумом, обусловленным разговорами участников конференции.

Список литературы

- 1. Центр речевых технологий. Режим доступа: http://www.speechpro.ru/ (дата обращения 25.11.2012).
- 2. Матвеев Ю.Н. Технологии биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н.Э. Баумана. Электронное научно-техническое издание. 2012. № 3(3). Режим доступа: http://vestnik.bmstu.ru/catalog/it/biometric/91.html (дата обращения 25.11.2012).
- 3. Матвеев Ю.Н., Симончик К.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2010 // ГрафиКон'2010: труды 20-й международной конференции по компьютерной графике и зрению. СПб.: СПбГУ ИТМО, 2010. С. 315-319.
- 4. Коваль С.Л. Комплексная методика идентификации дикторов по голосу и речи // Информатизация и информационная безопасность правоохранительных органов (Москва, 24 25 мая 2011 г.) : труды XX международной научной конференции. М.: Академия управления МВД России, 2011. С. 364-370.
- 5. Саймон Хайкин. Нейронные сети. Полный курс: пер. с англ. 2-е изд., испр. М.: Вильямс, 2008. 1103 с.
- 6. Трофимов А.Г., Скругин В.И. Адаптивный классификатор многомерных нестационарных сигналов на основе анализа динамических паттернов // Наука и образование. МГТУ им. Н.Э. 2010. Баумана. Электрон. $N_{\underline{0}}$ 8. Режим журн. доступа: http://technomag.edu.ru/doc/151934.html (дата обращения 25.11.2012).
- 7. Балакришнан А.В. Теория фильтрации Калмана : пер. с англ. М.: Мир, 1988. 168 с.
- 8. Акулов Л.Г., Муха Ю.П. Методы обработки электроэнцефалографических данных // Известия ВолгГТУ. Сер.

Электроника, измерительная техника, радиотехника и связь. 2008. Т. 4, № 2. С. 66-69.

9. Грешилов А.А., Стакун В.А., Стакун А.А. Математические методы построения прогнозов. М.: Радио и связь, 1997. 112 с.

SCIENCE and EDUCATION

EL Nº FS77 - 48211. Nº0421200025. ISSN 1994-0408

electronic scientific and technical journal

Personal identification by voice based on the Kohonen selforganizing map

11, November 2012

DOI: 10.7463/1112.0500433 Karpenko A.P., Morozov A.N.

Russia, Bauman Moscow State Technical University

<u>apkarpenko@mail.ru</u>

alexnikmorozov@gmail.com

The work deals with the problem of identification by voice. The focus is on building an adaptive classifier of specific voice features based on the Kohonen self-organizing map. The authors present the algorithms used for audio processing and classification, as well as the software system VOICE which implements these algorithms. The results of the studies confirming the effectiveness of the algorithmic and software solutions are given in this paper.

Publications with keywords: person identification on a voice, Kochonen self-organizing map

Publications with words: person identification on a voice, Kochonen self-organizing map

References

- 1. *Tsentr rechevykh tekhnologii* [Speech Technology Center]. Available at: http://www.speechpro.ru/, accessed 25.11.2012.
- 2. Matveev Iu.N. Tekhnologii biometricheskoi identifikatsii lichnosti po golosu i drugim modal'nostiam [Biometric technologies of person identification by voice and other modalities]. *Vestnik MGTU im. N.E. Baumana* [Herald of the Bauman MSTU], 2012, no. 3(3). Available at: http://vestnik.bmstu.ru/catalog/it/biometric/91.html, accessed 25.11.2012.
- 3. Matveev Iu.N., Simonchik K.K. Sistema identifikatsii diktorov po golosu dlia konkursa NIST SRE 2010 [The system of identification of announcers by voice for competition NIST SRE 2010]. *GrafiKon'2010: trudy 20-i mezhdunarodnoi konferentsii po komp'iuternoi grafike i zreniiu* [GrafiKon'2010: Proc. of the 20th International Conference on Computer Graphics and Vision]. St. Petersburg, SPbGU ITMO Publ., 2010, pp. 315-319.

- 4. Koval' S.L. Kompleksnaia metodika identifikatsii diktorov po golosu i rechi [Complex method of identification of announcers by voice and enunciation]. *Informatizatsiia i informatsionnaia bezopasnost' pravookhranitel'nykh organov: trudy 20-i mezhdunarodnoi nauchnoi konferentsii* [Informatization and information security of law enforcement agencies: proc. of the 20th international scientific conference], Moscow, 24-25 May, 2011. Moscow, The Academy of Management of the Interior Ministry of Russia Publ., 2011, pp. 364-370.
- 5. Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd edition. Prentice-Hall, 1998. (Russ. ed.: Saimon Khaikin. *Neironnye seti : Polnyi kurs*. 2th ed. Moscow, 2008. 1103 p.).
- 6. Trofimov A.G., Skrugin V.I. Adaptivnyi klassifikator mnogomernykh nestatsionarnykh signalov na osnove analiza dinamicheskikh patternov [Adaptive classificator of multidimensional non stationary signals based on dynamical patterns analysis]. *Nauka i obrazovanie MGTU im. N.E. Baumana* [Science and Education of the Bauman MSTU], 2010, no. 8. Available at: http://technomag.edu.ru/doc/151934.html, accessed 25.11.2012.
- 7. Balakrishnan A.V. *Kalman filtering theory*. New York, Optimization Software, Publications Division, 1987. 253 p. (Russ. ed.: Balakrishnan A.V. *Teoriia fil'tratsii Kalmana*. Moscow, Mir, 1988. 168 p.).
- 8. Akulov L.G., Mukha Iu.P. Metody obrabotki elektroentsefalograficheskikh dannykh [Methods of processing of electroencephalographic data]. *Izvestiia VolgGTU. Ser. Elektronika, izmeritel'naia tekhnika, radiotekhnika i sviaz'* [Proc. of VolgGTU. Ser. Electronics, measuring equipment, radio engineering and communications], 2008, vol. 4, no. 2, pp. 66-69.
- 9. Greshilov A.A., Stakun V.A., Stakun A.A. *Matematicheskie metody postroeniia prognozov* [Mathematical methods of making forecast]. Moscow, Radio i sviaz', 1997. 112 p.

10.7463/1112.0500433