Министерство образования и науки Украины
Национальный технический университет
«Харьковский политехнический институт»

Методические указания
к лабораторной работе № 10

«Идентификация личности по голосу»

по курсу «*Интерфейсы компьютерных систем управления*»

Харьков 2006

Министерство образования и науки Украины

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ЛАБОРАТОРНОЙ РАБОТЕ № 10

«ИДЕНТИФИКАЦИЯ ЛИЧНОСТИ ПО ГОЛОСУ»

ПО КУРСУ «ИНТЕРФЕЙСЫ КОМПЬЮТЕРНЫХ СИСТЕМ УПРАВЛЕНИЯ»

для студентов специальности 7.0910002

Утверждено
редакционно-издательским
советом университета,
протокол № 2 от 02.06.04

Харьков НТУ «ХПИ» 2006

Методические указания к лабораторной работе «Идентификация личности по голосу» по курсу «Интерфейсы компьютерных систем управления» для студентов специальности 7.0910002 / Сост. Д.А. Лунин, М.В. Гунбин. – Харьков: НТУ «ХПИ» 2006. – 13 с.

Составители: Д.А. Лунин, М.В. Гунбин

Рецензент А.В. Ивашко

Кафедра автоматики и управления в технических системах

Цель работы: написать программу контроля доступа на территории объектов или к источникам конфиденциальной информации и вычислительным ресурсам организации, предприятия и т.д.; основанную на идентификации личности по голосу.

Теоретическое обоснование работы:

В настоящий момент самый распространенный метод аутентификации пользователя является пароль. Идентификация личности по голосу хорошо вписывается в существующую технологию доступа с помощью пароля, как дополнительное средство защиты, так и самостоятельное.

Рассмотрим особенности идентификации личности, основанной на идентификации по парольной фразе, состоящей из задач идентификации и верификации голоса на некоторой ограниченной фонотеке.

Задача идентификации возникает, когда необходимо найти ближайший голос (или несколько голосов) из фонотеки к рассматриваемой фонограмме. Необходимость автоматизации этой задачи напрямую зависит от количества голосов в фонотеке, уровня эксперта и необходимой оперативности принятия решения. Обычно после задачи идентификации приходится решать вторую задачу, в которой подтверждается или опровергается принадлежность фонограммы конкретному голосу, т.е. задачу верификации. Решение задачи идентификации позволяет решать задачу верификации не на всей фонотеке, а только на группе ближайших голосов, что значительно сокращает время обработки фонограммы. Ниже приведены общие свойства процедур идентификации и верификации голоса.

Формальная запись процедуры идентификации группы ближайших голосов {f*} может быть дана в виде:

$$\{f^*\} = \underset{\{f\}\subset F}{\arg\min}[VID(x,f)] \qquad (1)$$

где:

x-идентифицируемый голос,

f - эталонные голоса из фонотеки F,

VID (x,f) - мера близости голосов.

Обозначим через P(f*) -вероятность совпадения предъявленного голоса с f*-ой эталонной, а пороговые уровни вероятности верификации голосов для классов <своих> и <чужих> как Pown и Pimp.

Процедуру верификации голоса VVER можно записать в виде:

VVER(P)= (2)

{IF P (f*) > Pown(P) THEN <свой - голос>}

{IF Pimp (f*) < P (f*) < Pown(f*) THEN <неопределенность>}

{IF P(f*) < Pimp (f*) THEN <чужой - голос>}

Первая продукция отражает процесс автоматической верификации <своего - голоса>. Вторая продукция описывает процесс возникновения неопределенности, когда при недостаточной или искаженной входной речевой информации (из-за слабого уровня входного сигнала или при наличии внешнего шума) вероятность верификации голоса P (P) ниже уровня Pown, но выше уровня Pimp. Третья продукция отражает процесс автоматического отклонения (режекции) <чужого голоса>. Метод идентификации и

верификации основывался на выборе совокупности характеристических функций речи и голоса Hi(x,f) и их свертки.

$$VID(x,f) = \sum_i [A_i * H_i(x,f)] \Big/ \sum_i A_i \qquad (3)$$

где:

Ai - суть веса характеристических функций речи Hi(x,f).

При выборе характеристических функций (метрик) речи Hi(x,s) следует иметь в виду, что они должны отражать сходство или различия анатомических констант голосов и динамических стереотипов порождения звуков. При определении весов характеристических функций речи Ai могут использоваться отношения выборочных средних для чужих и своих голосов: Hi(x,x*)/Hi(f,f*).

Опорные сегменты речи индивида, автоматически выделяемые в потоке слитной речи, описывались в пространстве спектральных, частотных и формантные характеристик голоса, а также его интегральных характеристик.

В соответствии с избранной выше структурой принятия решений на первом этапе вычисляется мера интегрального сходства входного образа голоса со всеми эталонными образами и выделяются К ближайших голосов, где К = N / 10, где N- количество голосов в фонотеке.

На втором этапе сравниваются входная последовательность опорных сегментов с последовательностями опорных сегментов эталонных голосов. При этом вычисляются меры фонетического и индивидуального сходства

образцов, а также уровни статистической вероятности (P(f*)} совпадения предъявленного голоса с К эталонными голосами.

На третьем этапе определяются три ближайших голоса и уровень статистической вероятности P(f*) наиболее правдоподобного кандидата. Далее осуществляется процедура верификации голоса наиболее правдоподобного кандидата.

Аппаратное обеспечение системы включает в себя интерфейс ввода/вывода звукового сигнала на базе звуковых карт семейства Sound Blaster 16 (формат ввода сигнала - 16(8)bit_PCM MONO, частота ввода 8000 Гц) и выше.

Программное обеспечение позволяет без применения спецпроцессора на основе последовательного алгоритма обнаружения звуковых сегментов речевого сигнала, вычислять спектры последовательности опорных сегментов и их моменты, формантные характеристики голосов, их интегральные портреты. Запоминаемые в процессе обучения несколько образцов представляют его локальные эталонные образцы голоса. Структура программной части системы представлена на рисунке 1.

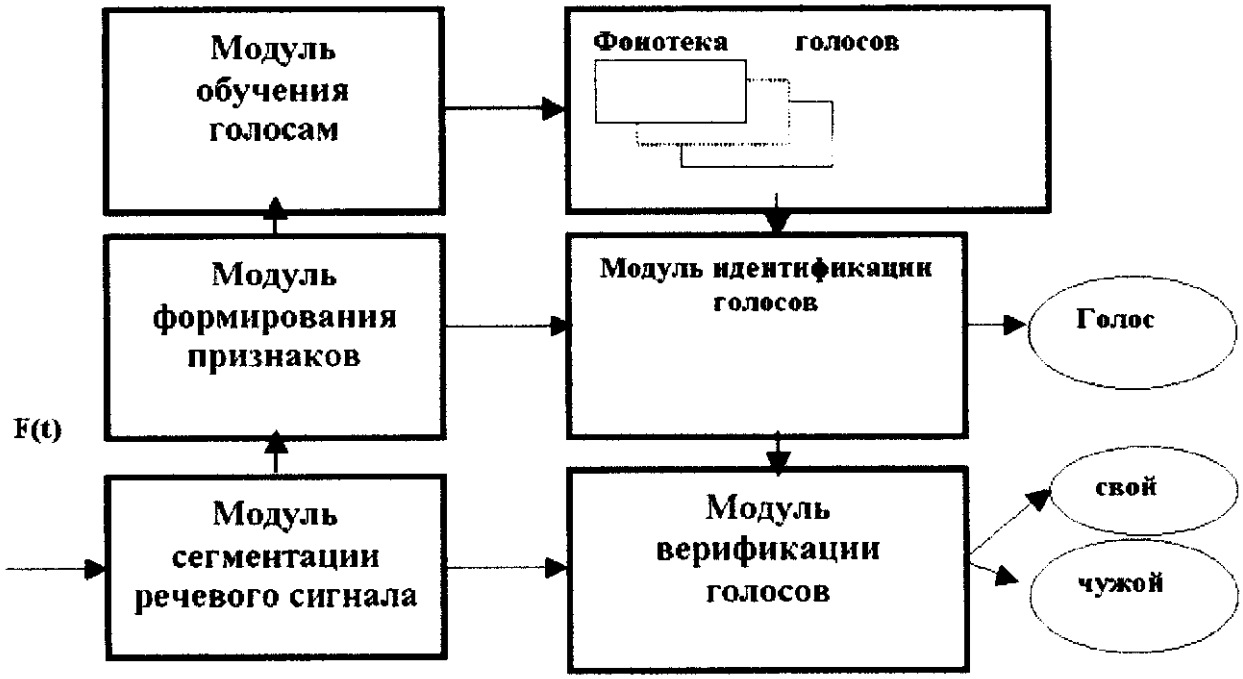

Рисунок 1 - Система оперативной идентификации и верификации голосов

Описание программы:

Требования к программному и аппаратному обеспечению:

- - Операционная среда Win95, Win98, WinNT, Win2000, WinXP;
- - Процессор 80486 и выше;
- - Звуковая карта 8 bit и выше;
- - Микрофон, аудиосистема;

Внешний вид программы изображён на рис. 2.

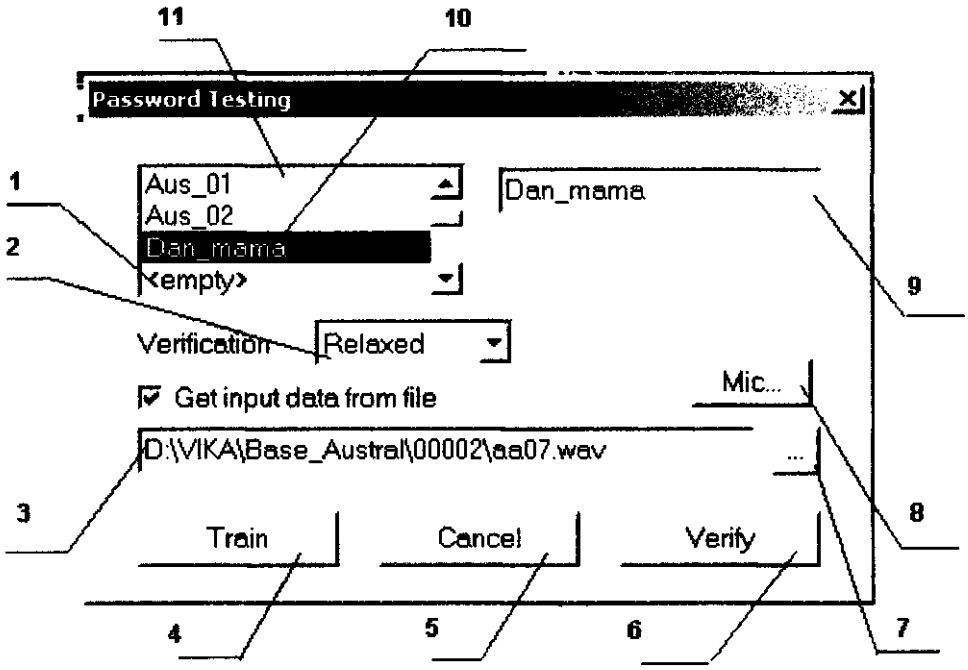

Рисунок 2 - Внешний вид программы

Инициализация.

При первом включении окна **9** и **11** пустые, файл с эталонами **(mdpwd.dat)** и файл истории **(mdpwd.log)** отсутствуют.

Обучение.

В окне **11** щелкнуть мышью на любое свободное место **1,** в окне **9** ввести свое имя, нажать **4.** После этого появится сообщение "Say you password (1)". Произнести пароль и нажать "Stop". Появится сообщение "Select Train to continue". И так продолжать 3 или 4 раза, пока не появится сообщение об успешной записи нового пользователя или о неуспешной попытке записи

"NOT train". В последнем случае новый эталон не создается и диктор в базу не записывается. При успешном завершении обучения в окне **11** появляется введенное имя диктора, эталон заносится в файл **mdpwd.dat,** в файле **mdpwd.log** появляется соответствующая запись.

Распознавание.

Для распознавания необходимо сначала в окне **11** выбрать имя диктора, под которым вы хотите войти в систему, дважды щелкнув мышкой (например, **10).** В окне **9** появится это имя. После этого нажать **6.** Появится сообщение "Listening...". Произнесите свой пароль и нажмите "Stop". Если необходимо повторное произнесение, то система выдаст сообщение "Select Verify to continue". Повторов может быть не больше трех. Если система не узнает диктора, то выдается сообщение "Alarm!!!". Если система узнала записанный в базу пароль, то выдается сообщение "Ok<User name>". <User name> - имя пользователя, которому принадлежит этот пароль.

Задание на лабораторную работу.

1. Начало и окончание ввода пароля осуществляется по нажатию соответствующих клавиш **4,5,6.**
2. После нажатия клавиши на начало ввода, говорить следует с небольшой паузой (< 0.5 сек). В этот момент система измеряет уровень шума.
3. Окно **2** используется для выбора порогов распознавания (Relaxed, Normal, Strict). Пороги влияют только на режим распознавания.
4. Окно **3** используется для выбора звукового файла **(WAV, PCM),** который можно использовать в качестве ввода вместо микрофона в режимах Обучения (нажать **4)** и Распознавания (нажать **6).** Для поиска файла можно воспользоваться кнопкой **7.**
5. Кнопка **8** вызывает программу **"AudiMic.exe",** которая позволяет следить за уровнем вводимого сигнала.
6. Программа реализует решение задачи открытой идентификации пользователей. Программа всегда производит сравнение введенного сигнала со всеми эталонами из базы, выбор ближайшего и окончательное принятие решения по порогу.
7. Файл **mdpwd.log** может содержать следующие записи.
   - <Имя диктора> accepted - диктор и пароль совпали (пароль меньше порога).
   - <Имя диктора> rejected - диктор и пароль не совпали (пароль больше порога).
   - <Имя диктора> rejected incorrectly! - диктор и пароль совпали, но пароль больше порога,
   - <Имя диктора> accepted incorrectly! - диктор не совпал но пароль в базе есть и он меньше порога.

Контрольные вопросы.

1. Особенности идентификации личности
2. Задача идентификации возникает
3. Аппаратное обеспечение системы распознавания.
4. Этапы распознавания.
5. Требования к программному и аппаратному обеспечению.