Софийски Университет „Св. Кл. Охридски”

Факултет по Математика и Информатика

ДИПЛОМНА РАБОТА

на тема

„Разпознаване на on-line подписи”

Дипломант: Дилян Лилков Лазаров

Специалност: Био и медицинска информатика
Факултетен номер: М21898

Ръководител: ст.н.с. II ст. д-р Георги Йорданов Глухчев
ИИТ на БАН,
ръководител на секция „Разпознаване на образи”

Консултант: доц. д-р Антоний Тодоров Попов

Юли, 2007

Съдържание

Глава III

Класификация (методи) за разпознаване

Глава IV

Система за разпознаване на on-line подписи

Въведение

През последните години започна да се обръща все по-голямо внимание на сигурността, което поражда необходимостта от идентификация на личността в реално време (on-line) по подпис. Поради много причини подписът се оказва най-сигурното средство за ауторизация. Подписът е уникална активна биометрична характеристика на всеки човек. Основно динамичните признаци (времето за изчертаване, натиск на писалката), отличават подписът от пасивните биометрични характеристики като отпечатък от пръстите на ръката, ириса на окото, които не се променят с течение на времето. Съвременните компютърни технологии дават възможност да бъде получена детайлна информация, при което и точността при разпознаването на подписа да бъде по-голяма. Това може да бъде постигнато с графичен таблет (фиг.1), който поддържа важни характеристики като координатите (X,Y) на всяка точка, определена от върха на писалката, степента на натиск, момента на регистриране на всяка точка. При някои съвременни таблети могат да бъдат отчитани и ъгъла на наклона на писалката спрямо равнината и азимута й. Поради противоречивите резултати, описани в литературата, последните два параметъра няма да бъдат обект на изследване в настоящата дипломна работа.

Фиг.1 – Таблет PC и графичен таблет

Глава I

Съвременно състояние – кратък преглед на разработките до момента

Успешния електронен бизнес се базира на доверие в бизнес партньорите и сигурността в обмена на информация между тях. Ауторизацията е задължителна при достъп до операционните системи, мрежови ресурси и уеб приложения, електронно банкиране както и до различни други индивидуални софтурени решения.

Лидер в разработката на софтуер за разпознаване на подписи е немската компания **SoftPro** (http://www.signplus.com/en/). За използването на подписа като средство за ауторизация Softpro представят следните аргументи:

Елиминиране на фалшификацията - напоследък зачестиха хакването на мрежовите ресурси, с цел извличане на информация за дебитни и кредитни карти с които е извършвано плащане по интернет. Използвайки електронен подпис за ауторизация намаляваме риска от хакване на пароли. Причината е че от подписа се извличат биометрични характеристики (ниво на натиск и времето за изчертаване на подписа), които са строго индивидуални за всеки човек и трудно могат да бъдат фалшифицирани.

Честото забравяне на пароли – статистиката за американските компании показва, че от 30% - 50% от всички обаждания за помощна поддръжка са за забравена парола или PIN(персонален идентификационен номер). Изчислено е от Gather group , че годишно в големите компании се плащат повече от $ 350 на служител за подобни обаждания. Разбира се че тези разходи можеха да бъдат спестени ако компаниите внедрят подпис като средство за ауторизация, тъй като подписа няма как да бъде забравен.

Ауторизация при електронна обработка на документи – подписът е най-често изполваното средство за ауторизация. Организациите разпечатват електронните документи за да могат да бъдат подписани. За да бъдат след това архивирани документите, отново се налага преконвертиране в електронен формат. С внедряването на електрония подпис за ауторизация може да бъде избегнато конвертирането на документите от електронен вид в хартиен и обратно.

1.Какво се прави ?

SoftPro е разработила няколко модула, които улесняват процеса на ауторизация.

SignSecure е модул които позволява първоначално влизане в операционната система на компютъра или отдалечен достъп до компютърни мрежи. Предимствата при използване на този модул са:

- уникална ауторизация на потребителите
- използване на общ ауторизиращ метод и интуитивен начин за достъп до компютърните ресурси
- намаляване на цената за системно администриране

SignDoc се използва за подписване на електронните документи като подписа може да бъде основен или потвърждаване на подписа със сертификат за юридически цели.

Предимствата при използване на този модул са:

- Намаляването на използване на хартия за сканиране, принтиране, прикрепяне и изпращане
- Намалява времето изгубено от грешките свързани с устни уговорки за документите
- Ускорява, автоматизира и увеличава сигурността при процеса на ауторизация на документите

Модулът **SignTeller** се използва за потвърждаване на on-line (в реално време) транзакции

Предимствата при използване на този модул са:

- При ПОС устройства, терминали или пред офисите за on-line ауторизация
- Комбинирано потвърждаване чрез статични и динамични биометрични характеристики

SignWare мощна библиотека (SDK – software development kit) за създаване на клиентски приложения за ауторизация.

Предимствата при използване на този модул са:

- Лесно интегриране в съществуващи приложения
- Поддръжка на интерфейси към C++, C, Active X, .Net, Java, Web services

SoftPro препоръчва да се използват тези модули с графични таблети разработени от Wacom и Interlink и от таблет PC произведени от HP,Fujitsu Siemens Computers, Moution Computing, Toshiba. (фиг.2)

Фиг. 2 Видове устройства за извличане на биометричните данни на подписа

2. Принципи на разпознаването на подписи (Как се прави ?)

При разпознаване на подписи SoftPro анализира и обработва два вида характеристики - статични и динамични

Статични характеристики на подписа (фиг.2)

- Пресичания на щрихите
- Площ на затворения контур
- Намиране на броят затворени щрихи

Фиг.2 Видове статични характеристики

Динамични характеристики на подписа (фиг.3)

- Начало и край на щриха
- Отчитане на големината на натиска и времето за изчертаване на всеки щрих от подписа
- Намиране на скорост и ускорение при изчертаване на подписа

Фиг.3 Видове динамични характеристики на подписа

3. Какви са резултатите?

Днес много от световно известните банки American Express, Deutsche Bank, CityGroup, Discover Financial, HypoVereinsBank, JPMorgan Chase, OCBC са внедрили софтуера разработен от SoftPro в резултат на което милиони проверки на транзакции и документи се проверяват автоматично всеки ден.

Фиг. 4 Подпис на документ и съхраняване на данните в електронен вид

Глава II

Измерване на графични(геометрични) динамични признаци

Разпознаването на подписа се извършва по някой важни (значими) признаци (характеристики) – геометричен център, масов център, разстояние между масов и геометричен център, ъгълът, който сключва отсечката определена от масов и геометричен център спрямо положителната част на оста X, дисперсията за параметрите X, Y и P- големината на натиска, броят щрихи, дължина на подписа като брой пиксели, дължина на подписа като разлика по X между най-дясната и най-лявата точка (дължина на подписа по X), дължина на подписа D като сума от разстоянията между пикселите, сложността на щриха.

Преди да представим формулите за изчисление на признаците на подписа ще въведем следните означения:

- n – брой точки в подписа
- N – брой признаци (характеристики)
- K – брой подписи, по който се изчислява еталона
- M – брой хора участващи в обучението на системата

Ъответните геометрични характеристики се пресмятат по следните формули

1. Геометричен център :

$$X_G = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

$$Y_G = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{2}$$

2 .Масов център :

$$X_M = \frac{1}{p}\sum_{i=1}^{n} p_i X_i \tag{3}$$

$$Y_M = \frac{1}{p}\sum_{i=1}^{n} p_i Y_i \tag{4}$$

където p е максималната стойност на натиска (зависи от конкретния модел таблет), а P_i е натиска във i - тата точка отчетена от таблета.

3. Разстояние между геометричен и масов център:

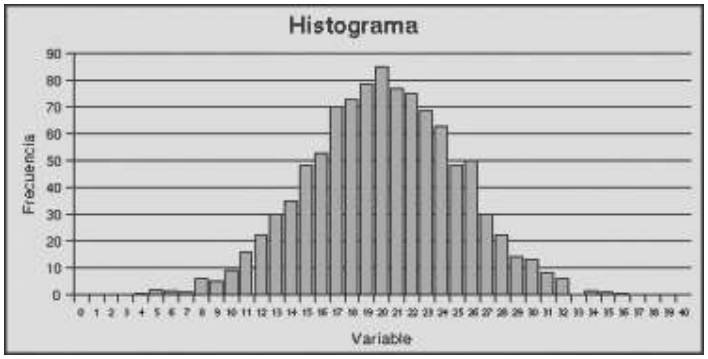
$$\left|\vec{R}\right| = \sqrt{\left(\left(X_M - X_G\right)^2 + \left(Y_M - Y_G\right)^2\right)} \tag{5}$$

4.Наклона спрямо оста X . Ъгълът се задава със следната формула:

$$arctg\ \frac{\left(Y_M - Y_G\right)}{\left(X_M - X_G\right)} \tag{6}$$

5. Дисперсия

За да бъде намерена дисперсията на един от параметрите X, Y, P-големината на натиска е необходимо да се изчисли хистограма $H(i)$ на съответния параметър. Хистограмата (Фиг. 5) показва по какъв начин са разпределени стойностите на параметъра, който се изследва.



Фиг. 5 Хистограма

Среден натиск:

$$P_h = \frac{1}{n}\sum_{i=1}^{n} iH(i) \tag{7}$$

Дисперсия на натиска:

$$\sigma_h^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} H(i)(i - P_h)^2 \quad (8)$$

6. Брой щрихи.

Щрихът представлява съвкупността от всички точки, които са записани без да се вдига писалката от таблета.

7. Дължина на подписа като брой пиксели

8. Дължина на подписа като разлика по X между най-дясната и най-лявата точка (дължина на подписа по X)

9. Дължина на подписа D като сума от разстоянията между пикселите

Разстоянието между k-1 пиксел и k пиксел

$$D_k = \sqrt{\left((X_k - X_{k-1})^2 + (Y_k - Y_{k-1})^2\right)} \quad (9)$$

$$D = \sum_{i=0}^{n} D_i \quad (10)$$

10. Сложността на щриха – отношението на параметрите изчислени в т. 9 и 8

Глава III

Класификация (методи) за разпознаване

1.Въведение

Биометричните техники за разпознаване направиха възможни значителни подобрения в обективното оценяване на количествени прилики между ръкописни образци, водещи към развиването на автоматични системи за on-line разпознаване на подписи [1, 2]. Те могат да бъдат разделени на [1]:

а) подходи, базирани на функцията, при които методологията на обработката на сигнали се прилага върху динамично получените времеви последователности от данни (например скорост, ускорение, сила или натиск);

б) подходи, базирани на признаците, където от получената информация се извличат статистически параметри. Могат също така да се определят различни нива на класификация, така че е възможно да се използват и да се комбинират базирани на формата глобални статични (например пропорции, център на тежестта, отношение на хоризонталните дължини), глобални динамични (например време за полагане на подписа, отношение на времето на допир към цялото време, средна скорост) или локални (посока на щриха, кривина или допирателна в наклона) параметри.

Най-новите действия, свързани с on-line разпознаване на подпис могат да се структурират в следните категории: динамично времево изкривяване, скрити марковски модели, невронни мрежи и йерархични подходи.

2 Видове сравнения

Процесът на разпознаване се базира или на измерване на сходство между даден подпис и представители на истински и фалшифицирани подписи, или на автоматично генерирани области в признаковото пространство, получени след обучение на невронна мрежа. Различни видове мерки за близост и критерии за

взимане на решение могат да бъдат приложени във връзка с избраните признаци, робастността на процедурата за изчисление, възможността за обобщение, изчислителното време и възможността да се пренебрегнат вътрешните разлики за всеки пишещ, като в същото време се запази разликата между различните хора.

2.1. Разстояния

Класифициращите правила най-често се базират на прости измервания на близостта като евклидово или махаланобисово разстояние. Те се използват когато стойностите на признаците могат да получат геометрична интерпретация като координати на точка в признаковото пространство.

В [3] като мяра за сходство се използва евклидово разстояние между глобални признаци – височина и ширина, ъгъл на наклона, вертикален център на тежестта, максимум на хоризонталната проекция, площ на подписа, изместване на основната линия. В [4] обратното разстояние на Махаланобис се използва за сравнение между вектора на изместване на профила на входното изображение и средния вектор на множеството от оригинални подписи.

2.2. Невронни мрежи

Възможността на невронните мрежи да бъдат обучени с примери привлича вниманието към тяхното прилагане като класификатор в разпознаването по подпис. В [5] се използва многослойна права невронна мрежа с 14 входни неврона, 18 скрити и 30 изходни. 7 инварианти, базирани на алгебричните централни моменти се изчисляват и нормализират, 7 глобални признаци се използват за захранване на невронната мрежа. За обучение се използва техниката на обратното разпространение на грешката. Подобна техника се прилага в [6], където функцията плътност на разпределението се

използва за описание на формите. Тя е инвариантна спрямо транслация и мащабиране и не изисква прекалено много изчислително време. Изчисленията показват, че класификаторът, обучен с обратно разпространение на грешката, е по-добър от класификатор с праг и може да се сравнява с класификатор по К най-близки съседи. Авторите считат, че този подход може успешно да бъде приложен като класификатор на първия етап на многонивова система за верификация. Подобно мнение е представено в [7], където се използват геометрични признаци и техника на мажоритарно гласуване. Изследването е насочено към първоетапна класификация на подписи, извършваща грубо сравнение на формата за филтриране на повечето неумели фалшификации. Линейна транслация, мащабиране и еластично съответствие се прилагат за сравнение на текущо и еталонно изображение. Леко изменени версии на оригинални подписи се използват като оригинални обучаващи примери. Верификаторът се състои от няколко прости 3-слойни перцептрони.

В [8] се използва алгоритъм на самоорганизиране на Кохонен. Извличат се 4 групи признаци: а) хоризонтални и вертикални признаци от бинарно и монохроматично изображение, б) глобална основна линия, положението на максималната стойност на хоризонталната и вертикалната проекция, в) натиск – 7 признака, д) наклон – 4 признака. Разпознаващата система е 5-слойна невронна мрежа. Третият слой представлява правила от размитата логика.

В [9] три различни типа глобални признаци се използват за класификацията. Това са моменти на проекциите и характеристики на горната и долната обвивка. Използвайки собствените стойности на ковариационната матрица се получава нормализирано представяне на подписа. Проектират се индивидуални класификатори с помощта на права невронна мрежа с два скрити слоя. За комбиниране на изходите от класификаторите се използва еднослойна невронна мрежа.

Ново представяне с бейсова мрежа се предлага за разпознаване по подпис в [10]. Идеята е да се моделират вариациите и зависимостите на компонентите в подписите. Проблемът се състои от две задачи: изучаване на структурата на мрежата и оценка на условните вероятности. Горен и долен профил на подписа са взети като обект на анализа. Те са разделени на

подсегменти на местата на промените в цвета. Тъй като е трудно да се оцени началната вероятност за едновременна поява на два сегмента използвайки честотата на последователни опити, се предлага евристична формула. Получените резултати показват, че новият метод се представя по-добре по отношение на грешка от I и II род в сравнение с корелация на регионите и обикновена бейсова мрежа.

Голямото предимство на класификатора с бейсова мрежа е, че позволява приносът на отделните признаци да бъде претеглян и ясно изразява натрупаното познание.

2.3. Размити множества

Методите с размити множества естествено отразяват вътрешните различия при признаците на подписите. Подходът, предложен в [11] е базиран на размито моделиране с използването на модела на Такаги-Сугено. За всеки признак се генерира размито множество. Признаците се размиват от експоненциална функция за принадлежност, която е модифицирана, за да включва структурни параметри. Експерименталните резултати показват, че всички видове фалшификации – груби, неумели и умели – са засечени с максимална точност (нулева грешка). Базата данни се състои от 1200 подписа.

2.4. Скрити марковски модели

В [12] е описана система, състояща се от две части, работещи съответно с глобални и локални признаци. За глобалните признаци се използва статистическо измерване на разстоянието, докато локалните признаци се обработват със скрити марковски модели. Посоката на наклона на щрихите и обвивките се използват като признаци. Локалните признаци са същите като глобалните, но се измерват в блокове с по-малък размер. Степента на сходство

се изчислява като нормализиран логаритъм на вероятността. Използва се фиксирана стратегия на сливане, базирана на сумиращо правило. Експериментите показват средна грешка от 10% за умелите фалшификации и около 2% за грубите фалшификации.

2.5. Многонивови класификатори

Многонивовите класификатори са доста подходящи в случаите, когато за описание на обекта се използват признаци от различно естество. Те са особено полезни за разпознаването по подпис, поради факта, че различни видове фалшификации се събират в един клас на отхвърляне.

В [13] се реализират 4 различни типа схеми за представяне на образците, а именно: геометрични признаци (ширина, отношение височина/ширина, наклон, брой вертикални линии, вертикален център на тежестта, максимум на хоризонталната проекция, площ на черните пиксели, изместване на основната линия, представяния чрез централни моменти, характеристики на обвивките и дървовидно структурирани Уейвлет-признаци. Изучени са класификатори, базирани на праг, включително традиционния класификатор по доверителни интервали (признаците са нормализирани, центрирани и мащабирани използвайки средните стойности и стандартните отклонения. Приемайки разпределението за нормално, е приет праг, отговарящ на вероятност 0,99), класификатор на съседство (модифициран класификатор по К най-близки съседи) и техните комбинации. Получена е верификационна точност от 90% за оригиналните подписи, над 98% за простите фалшификации и около 70-80% за умелите фалшификации. Базата данни съдържа около 650 подписа.

В [14] се въвеждат два нови вида сигнали за on-line верификация – ъгли на височина и азимут. Докладвано е подобрение на процента на двата вида грешка при равенство между тях от 14,2% на 1,8% от 480 оригинални и 480 подправени подписа на 24 субекта.

3. Изчисляване на еталона

В резултат на обучение на системата се получават K на брой записа на всички разглеждани параметри и е необходимо да се изчисли еталона, по който да се извършва разпознаването. За изчисляване на еталона се използват средно аритметично и средно квадратично за всеки параметър.

Средно аритметично:

$$S_A = \frac{1}{K}\sum_{i=1}^{K} P_i \tag{1}$$

Средно квадратично:

$$\sigma = \frac{1}{K-1}\sqrt{\sum_{i=1}^{K}\left(P_i - S_A\right)^2} \tag{2}$$

Еталонът се задава със стойностите $\left(S_{A1}, S_{A2}, S_{A3}, S_{A4}..S_{AN}\right)$ за всеки параметър, където N е броят характеристики (признаци) .

За да се разпознае даден подпис е необходимо да се намери най-малкото разстояние между вектора определен от характеристиките на подписа S до векторите, определени от всички изчислени еталони $\left(E_1, E_2, E_3, E_4..E_M\right)$, където M е броят участници в системата. Могат да бъдат използвани Евклидово или разстояние на Махаланобис, които се дефинират съответно с формулите:

4.Евклидово разстояние в двумерното пространство

$$R_E = \sqrt{\sum_{i=1}^{N}\left(S_i - P_i\right)^2} \tag{3}$$

N е брой характеристики (признаци), S_i - характеристика от наблюдавания подпис , P_i - характеристика от еталона

5. Разстояние на Махаланобис

$$R_M(x, P_i) = (x - P_i)^t \sum_i^{-1} (x - P_i) \quad (4)$$

Частен случай при статистически независими признаци, коварянтната матрица изглежда по следния начин:

$$\sum^{-1} = \begin{Vmatrix} 1/\sigma_{11}^2 & 0 & 0 \\ 0 & 1/\sigma_{22}^2 & 0 \\ 0 & 0 & 1/\sigma_{nn}^2 \end{Vmatrix} \quad (5)$$

След като заместим коварянтната матрица, разстоянието на Махаланобис придобива вида:

$$R(x, P_i) = \sum_{j=1}^{N} (x_j - P_{i,j})^2 \Big/ \sigma_{i,j}^2 \quad (6)$$

където векторът $x = (x_1, x_2 \ldots x_N)$ са съответните характеристики записани от графичния таблет, а P_i еталона.

Търсим минимума на разстоянията -

min($R_1, R_2, R_3, R_4 \ldots R_n$) и отнасяме подписа към еталона, за който този минимум е получен. По този начин смятаме, че подписа S е разпознат с най-голяма вероятност.

Глава IV

Система за разпознаване на on-line подписи

Въведение

За да заработи системата за “Разпознаване на online подписи” е необходимо да се изпълнят два основни етапа.

Етап 1:

Обучение на системата. Въвеждат се N на брой подписа за всеки, чийто подпис ще бъде разпознаван. Следствие на което системата изчислява еталон за всеки, по който ще бъде използван при разпознаване.

Етап 2:

Разпознаване на подписи. Въз основа на изградената база данни може да бъде направено сравнение с всички еталони и да се намери еталон, който е най-близо до подписа, който изследваме.

1. Реализация на програмния продукт

1.1 Проектиране

При проектиране на програмен продукт трябва много добре да бъде проучена предметната област. Поставения проблем да се раздели на отделни проблеми така че да могат да бъдат решени. Необходимо е да бъдат проучени няколко решения на проблемите и да се избере най-доброто решение. Под най-добро решение се има в предвид решение, за което е необходимо по-малко ресурси на компютърната система - оперативна памет и процесорно време. Основната цел при проектиране на програмен продукт е създаване на архитектура, която да бъде гъвкава. Благодарение на тази гъвкавост системата може да бъде разширяван с лекота. Не на последно място е необходимо да се избере най-подходящата технология за реализация на системата.

За софтуерната реализация са използвани следните технологии: **Visual Studio 2003 – C#, Microsoft SDK Tablet 1.5, Express edition MS-SQL 2005**

Архитектурата е трислойна клиент-сървър.

1.2 Клас диаграми

Първият слой е за връзка с базата данни. Това се осъществява чрез основния клас Root, в който са дефинирани няколко метода, които се наследяват от Signature и Person (Фиг. 6).

Find() – като параметър се задава условие за търсене, по което се конструират заявка Select за извличане на данни от таблиците на MS SQL server. Като резултат се конструират обекти и се връща масив от тези обекти

Create() – извиква се от наследниците на Root - Person или Signature и се конструира заявка Insert за създаване на нов запис в базата данни.

Update() – конструира заявка за обновяването на запис в базата данни

Delete() – в този метод се конструира заявка за изтриване на запис от базата

SetValue() – с помощта на този метод се инициализират данните на класа от входен параметър от тип Hashtable

GetValue() – методът връща HashTable обект, в който се записват данните на съответния клас

Фиг. 6 Класовете отговорни за комуникация с базата данни

Във втория слой е дефинирана бизнес логиката. Състои се от класовете SignTablet, SignCalculateData и SignRecognize (Фиг. 7).

Класът SignTablet се използва за извличане и интерпретиране на информацията от графичния таблет. Това е реализирано чрез метода GetDataFromStrokes(). За съхраняване на данните (серилизация) се използва метода SaveISF(), който серилизира данните от таблета във бинарен формат ISF. Използва се и променливата signCalculateData, която е от тип SignCalculateData.

В SignCalculateData се извършва калкулацията на всички параметри на подписа. За тази цел се използват променливите:

xPoints – масив в който се записва X координатите на всяка точка (пиксел)

yPoints – масив в който се записва Y координатите на всеки пиксел

normalPressure – масив, в който се записва натиска на писеца върху таблета за всеки пиксел. За всеки таблет максималната стойност е различна. Обикновено стойността варира от (0,255) или (0,511) при повечето таблети.

timeTick – масив, в който се записва времето за изчертаване на всяка точка. Стойностите са в милисекунди.

GetGeometricCenter() – в този метод се изчислява геометричният център на подписа

GetWeightCenter() – методът връща масовия център на подписа

NormalizeCordinate() – методът изчислява обхващащият правоъгълник на подписа, за да нормализира координатите спрямо локална координатна система

GetDistanece_GC_WC() – методът връща разстоянието между геометричен и масов център

GetAngle() – изчисляване на ъгъла между отсечката определена от геометричния и масовия и оста X

Histograma() – методът връща хистограмата на подадения параметър

GetCenterHistograma() – изчислява центъра на хистограмата

Dispersia() – намира дисперсията по X, Y или натиска P

SrednoAretmitichno() – методът изчислява средно аретмитичното на подадения параметър

SrednoKvadratichno() – методът връща средноквадратичното на входния параметър

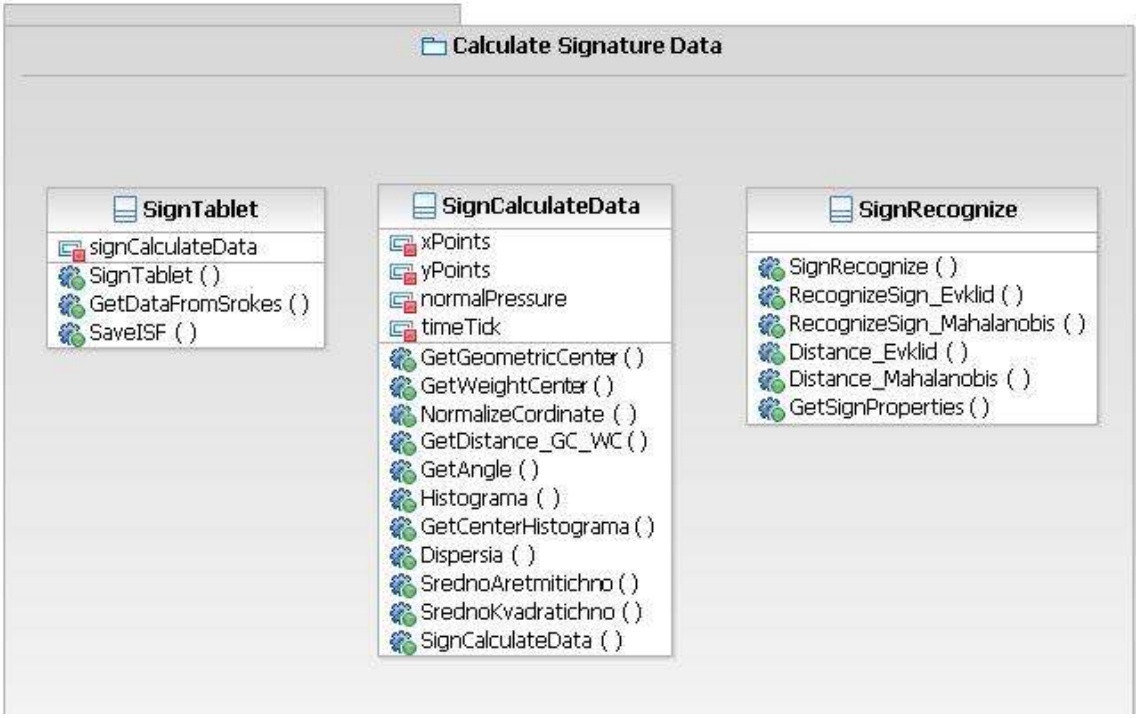
Класът SignRecognize е отговорен за разпознаване на подписа. Дефинирани са следните методи:

Distance_Evklid() – помощен метод за изчисляване на разстоянието между два вектора по метода на Евклид

Distance_Mahalanobis() – помощен метод за изчисляване на разстоянието между два вектора по метода на Махаланобис

RecognizeSign_Evklid() – методът намира най-малкото разстояние между изследвания вектор и векторите еталони използвайки функцията Distance_Evklid()

RecognizeSign_Mahalanobis() – методът намира най-малкото разстояние между изследвания вектор и векторите еталони използвайки функцията Distance_Mahalanobis()



Фиг. 7 Класовете описват бизнес логиката на системата

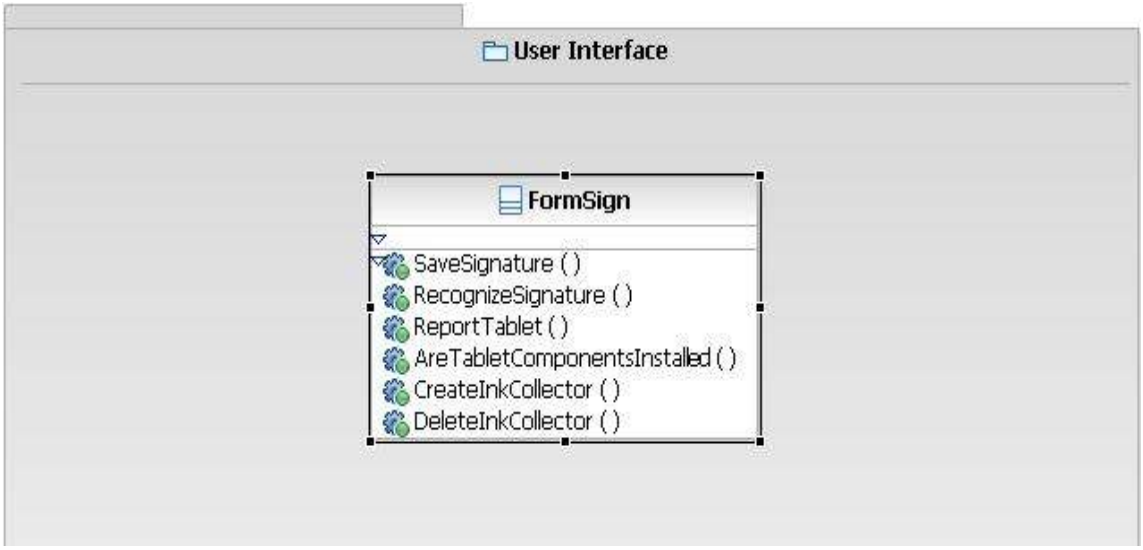
В третия слой е дефиниран потребителския интерфейс, като се обръща към функционалността на първите два слоя явно или неявно. В него се разграничават няколко по важни метода (Фиг. 8):

SaveSignature() – методът съхранява данните отчетени от таблета, когато програмата е в режим обучение

RecognizeSignature() – методът разпознава подписа на кой принадлежи по един от двата алгоритъма, когато програмата е в режим разпознаване

ReportTablet() – методът показва какви характеристики поддържа таблета

AreTabletComponentsInstalled() - проверява дали към персоналния компютър е инсталиран графичен таблет



Фиг. 8 Класът отговарящ за потребителски интерфейс

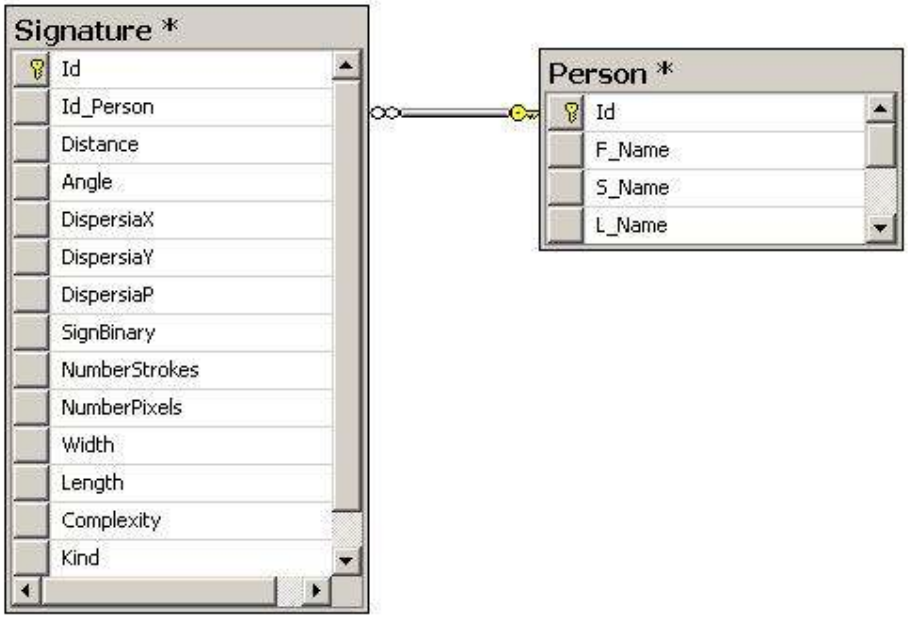
1.3 Структури от данни

Основния обект, който се поддържа от библиотеката на Microsoft Tablet SDK 1,5 е Strokes. Тази колекция се състои от обекти от тип Stroke, като всеки елемент съдържа множеството от точки, които са изчертани без да се вдига писалката. За всяка точка има записан пакет в който се намират данните – X,Y,Pressure и други характеристики поддържани от конкретния таблет. На диаграмата по долу (Фиг. 9) е представено схематично как са подредени данните.

Фиг. 9 Описание на структурата от данни за съхраняване на информацията от таблета

1.4 Съхраняване на данните

За съхраняване на информацията се използва MS-SQL 2005. За тази цел са проектирани две таблици – Person, Signature (Фиг. 10) . В таблицата Person се съхраняват трите имена, а в таблицата Signature изчислените характеристики на подписа и външен ключ Id_Person, с помощта на който се осъществява връзката между двете таблици.



Фиг. 10 Таблици в базата данни

1.5 Основни модули изграждащи системата

Системата се състои от два основни модула

- Обучение – чрез този модул системата записва данните извлечени от таблета в базата данни и изчислява еталона по който се извършва разпознаването.
- Разпознаване – системата разпознава подписа на кой принадлежи, по метода на Евклид или Махаланобис за намиране на разстояние между два вектора.

Заключение (изводи)

Реализирани са два метода за разпознаване на подписи

- Първият използва намиране на разстояние по формулата на Евклид
- Вторият използва намиране на разстояние по формулата на Махаланобис

И двата метода свеждат проблема до намиране на най-малкото разстояние на вектора от разглежданите параметри на изследвания подпис до векторите на изчислените еталони.

В реализираната програма се въвеждат подписите на краен брой участници, с помощта на което системата се обучава. След това се разпознава подписът изчертан с графичния таблет по един от по горе изброените методи. Резултатите от прилагането на системата върху извадка от 25 участника дава точност около 90%.

Архитектурата на приложението е клиент-сървър. Ясно се разграничават три слоя :

- Слой за съхраняване и извличане на данните (Data layer) – конструират се необходимите заявки по подадените параметри за извличане и съхраняване в системата за управление на база данни – MS SQL.
- Бизнес слой (Business layer) – калкулират се характеристиките на подписа по извлечените данни от графичния таблет.
- Потребителският интерфейс (Graphical User Interface) – този слой е реализиран с форма към която са добавени необходимите менюта и контроли за управление на програмата.

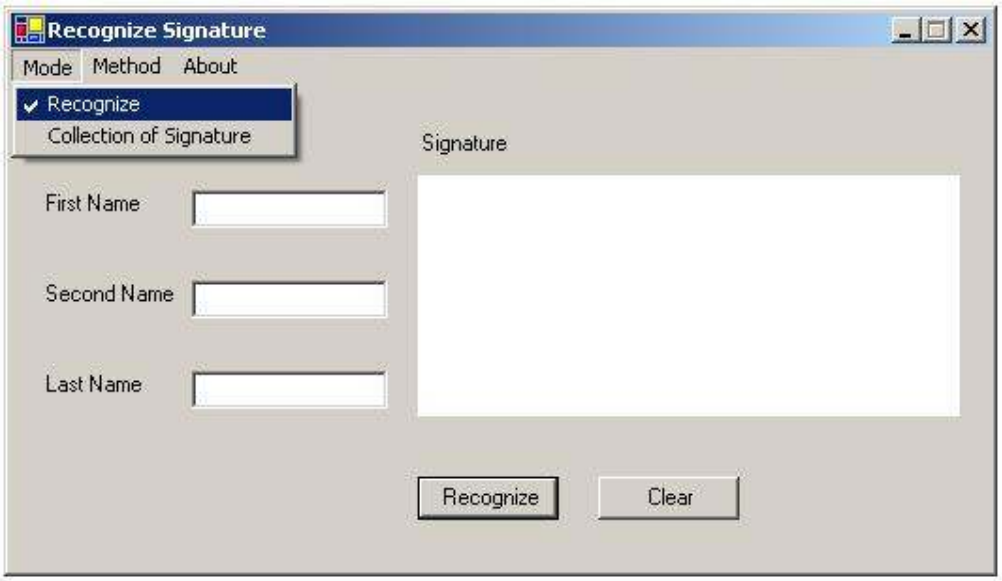
Тези резултати съответстват на първоначално поставената задача в дипломната работа.

Като насоки за бъдеща работа по разглежданата тема могат да се посочат следните:

- Добавяне на още значими характеристики на подписа
- Разработване на нови алгоритми за разпознаване
- Интегриране на системата с други системи, които използват ауторизация

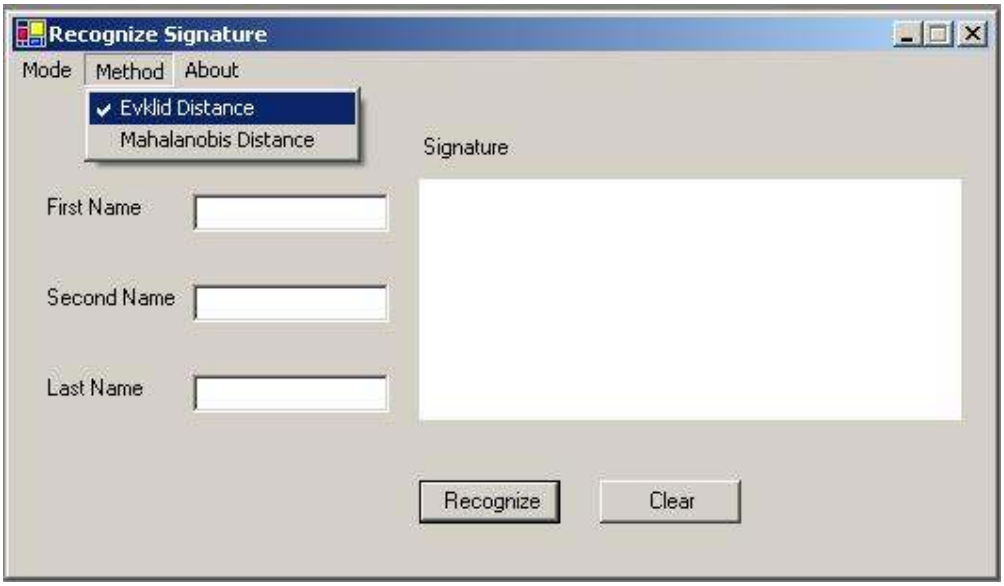
Приложение I

Системата работи в един от двата режима (Фиг. 11) – разпознаване по подразбиране или обучение чрез въвеждана на 10 подписа на нов участник.



Фиг. 11 Избиране на режим на работа

Избор на метод за разпознаване Евклидово или Махаланобисово разстояние (Фиг. 12).



Фиг. 12 Избор на метод за разпознаване

Разпознаване на подписи (Фиг. 13, 14)

Фиг. 13 Разпознаване на подпис

Фиг. 14 Разпознаване на подпис

Използвана литература

[1] Leclerc, F., R. Plamondon. “Automatic Verification and Writer Identification: The State of the Art 1989-1993”. Pattern Recognition and Artificial Intelligence, vol. 8, 1994, pp. 643–660.

[2] Plamondon, R., G. Lorette. “Automatic Signature Verification and Writer Identification - the State of the Art”. Pattern Recognition, vol. 22, no. 2, 1989, pp. 107–131.

[3] Qi, Y., B.R. Hurt. “Signature verification using global and grid features”. Pattern Recognition, vol. 27, no. 12, 1994, pp. 1621-1629.

[4] Fierrez-Aguilar, J., N. Alonso-Hermira, G. Moreno-Marquez, J. Ortega-Garcia. “An off-line signature verification system based on fusion of local and global information”. (D.Maltoni and A.K.Jain Eds) BioAW, 2004, pp. 295-306.

[5] Oz, C., F. Ercal, Z. Demir. “Siganture recognition and verification with ANN”. Proc. ELECO’2003, Bursa, Turkey, pp. 327-331.

[6] Drouhard, J-P., R. Sabourin, M. Godbout. “A neural network approach to the off-line signature verification using directional PDF”. Pattern Recognition, vol. 29. no. 3, 1996, pp. 415-424.

[7] Huang, K., H. Yan. “Off-line signature verification based on geometric feature extraction and neural network classification”. Pattern Recognition, vol. 30, no. 1, 1996, pp. 9-17.

[8] Quek, C., R.W. Zhou. “Anti-forgery: a novel pseudo-outer product based fuzzy neural network driven signature verification system”. Pattern Recognition Letters, vol. 23, 2002, pp.1795-1816

[9] Bajaj, R., S. Chaudhury. “Signature verification using multiple neural classifiers”. Pattern Recognition, vol. 30, no. 1, 1997, pp. 1-7.

[10] Xiao, X., G. Leedham. “Signature verification using a modified Bayesian network”. Pattern Recognition, vol. 35, 2002, pp. 983-995.

[11] Hanmandlu, M., M. Hafizuddin, M. Yusof, V. Krishna Madasu. “Off-line signature verification and forgery detection using fuzzy modeling”. Pattern Recognition, vol. 38, no. 3, 2005, pp. 341-356.

[12] Fierrez-Aguilar, J., N. Alonso-Hermira, G. Moreno-Marquez, J. Ortega-Garcia. “An off-line signature verification system based on fusion of local and global information”. (D.Maltoni and A.K.Jain Eds) BioAW, 2004, pp. 295-306.

[13] Ramesh, V.E., M. Narasimha Murty. “Off-line signature verification using genetically optimized weighted features”. Pattern Recognition, vol. 32, 1999, pp. 217-233.

[14] Hangai, S., S. Yamanaka, T. Hamamoto. “On-Line Signature Verification based on Altitude and Direction of Pen Movement”. Proc. IEEE International Conference on Multimedia and Expo (ICME), vol. 1, 2000, pp. 489–492.

[15] Ангелов, А., Д. Несторов, С. Бончев, Г. Глухчев, Д. Каменов, П. Велева. “Система за автоматизиран анализ на почерк”. Информационен бюлетин на НИКК-МВР, София, 1997, стр. 127-133.

[16] Атанасов, К., “Въведение в теорията на обобщените мрежи”. Понтика-принт, Бургас, 1992.

[17] Венков, П., “Анализ и разпознаване на изображения и сцени”. Технически университет - София, 1996.

[18] Гочев, Г., “Компютърно зрение”. Технически университет - София, 1993.

[19] Ланцман, Р.М., “Кибернетика и криминалистическая экспертиза почерка”. ‘Наука’, Москва, 1968.

[20] Минчев, Д., “Съдебно-графическа експертиза”. МВР-ДНМ, Научно-изследователски институт по криминалистика и криминология, София, 1990.

Уеб страници:

SoftPro лидер в разработката на софтуер за графични таблети и таблет PC
http://www.signplus.com/en/

Форум на Microsoft за програмиране на графични таблети
http://www.microsoft.com/communities/newsgroups/en-us/default.aspx

Статия на Microsoft за изграждане на приложения работещи с таблети
http://www.microsoft.com/mspress/books/sampchap/5958d.aspx

Статия за намиране на разстоянието по формулата на Евклид
http://en.wikipedia.org/wiki/Euclidean_distance

Статия за намиране на разстоянието по формулата на Махаланобис
http://en.wikipedia.org/wiki/Mahalanobis_distance

Д.Вындев, Приложна статистика 1 и 2, http://www.fmi.uni-sofia.bg/fmi/statist/

Уеб страница за статистика - http://cmamucmuka.hit.bg/