

ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

Խաչատուրյան Վահե Սամսոնի

Ձեռագիր ՍՏՈՐԱԳՐՈՒԹՅԱՆ ՃԱՆԱՉՄԱՆ ՄԵԹՈԴՆԵՐԻ ՄՇԱԿՈՒՄ ԵՎ
ԻՐԱԿԱՆԱՑՈՒՄ

Ե.13.04 «Հաշվողական մեքենաների, համալիրների, համակարգերի և ցանցերի մաթեմատիկական և ծրագրային ապահովում» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի զիտական աստիճանի հայցման ատենախոսության

ՍԵՂՄԱԳԻՐ

Երևան 2013

ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ НАН РА

Хачатурян Ваге Самсонович

РАЗРАБОТКА И РЕАЛИЗАЦИЯ МЕТОДОВ РАСПОЗНАВАНИЯ РУКОПИСНОЙ
ПОДПИСИ

АВТОРЕФЕРАТ

Диссертации на соискание ученой степени кандидата технических наук по специальности
05.13.04 – “ Математическое и программное обеспечение математических машин, комплексов,
систем и сетей ”

Ереван 2013

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում

Գիտական ղեկավար՝ Պաշտոնական ընդդիմախոսներ՝	Ֆիզ.մաթ.գիտ.դոկտոր տեխ.գիտ.դոկտոր տեխ.գիտ.թեկնածու	Հ. Գ. Սարուխանյան Գ. Հ. Խաչատրյան Ռ. Գ. Հակոբյան
---	--	--

Առաջատար կազմակերպություն՝	Երևանի կապի միջոցների գիտահետազոտական ինստիտուտ
----------------------------	--

Պաշտպանությունը կայանալու է 2013թ. հունիսի 14-ին, ժ 16:00-ին ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում գործող 037 «Ինֆորմատիկա և հաշվողական համակարգեր» մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակի 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:

Սեղմագիրը առաքված է 2013թ. մայիսի 14-ին:

Մասնագիտական խորհրդի
գիտական քարտուղար, ֆ.մ.գ.դ.



Հ. Գ. Սարուխանյան

Тема диссертации утверждена в Институте проблем информатики и автоматизации НАН РА

Научный руководитель:	доктор физ.мат.наук	А. Г. Саруханян
Официальные оппоненты:	доктор тех.наук кандидат тех.наук.	Г. Г. Хачатрян Р. Г. Акопян

Ведущая организация:	Ереванский научно-исследовательский институт средств связи
----------------------	---

Защита состоится 14 июня 2013г. в 16:00 на заседании специализированного совета 037 «Информатика и вычислительные системы» Института проблем информатики и автоматизации НАН РА по адресу: 0014, г. Ереван, ул. П. Севака 1.

С диссертацией можно ознакомиться в библиотеке ИПИА НАН РА.

Автореферат разослан 14-го мая 2013г.

Ученый секретарь специализированного
совета, д.ф.м.н.



А. Г. Саруханян

Աշխատանքի ընդհանուր բնութագիրը

Թեմայի արդիականությունը: Վերջին տարիներին նկատվում է կենսաչափական նույնականացման մեթոդների նկատմամբ հետաքրքրության և կիրառության կայուն աճ: Կենսաչափական տեխնոլոգիաները կիրառվում են անձի նույնականացման համար՝ օգտագործելով ֆիզիոլոգիական կամ վարքագծային առանձնահատկություններ, որոնք յուրահատուկ են յուրաքանչյուր անձի համար, և դրանք հնարավոր չէ մոռանալ, կորցնել կամ գողանալ: Կենսաչափական նույնականացման համակարգերում առավել հաճախ օգտագործվում են ծիածանաթաղանթի կամ ցանցաթաղանթի սկանավորումը, ձեռագիր ստորագրության ստուգումը, ձայնի կամ մատնահետքերի ճանաչումը և այլն:

Այս մեթոդների շարքում է ձեռագիր ստորագրության ստուգումը: Ստորագրության բնութագրիչները յուրահատուկ են և պրակտիկորեն չեն կարող կրկնօրինակվել: Այդ պատճառով ձեռագիրը մինչ օրս մնում է առավել շատ օգտագործվող նույնականացման մեթոդներից մեկը: Ձեռագիր ստորագրության նույնականացման ժամանակ իսկությունը հաստատելու համար վերցնում են փորձարկվող ստորագրության բազմաթիվ բնութագրիչներ և համեմատվում են ստուգողական ստորագրության հետ, որը պահվում է տվյալների բազայում: Եթե հասանելի են մի քանի ստուգողական ստորագրություններ, ապա մեծանում է առանձին հատկանիշի շեղման հաստատունության չափը, որն օգտագործվում է փորձարկվող ստորագրության մեջ ստացած շեղման գնահատման համար:

Ոչ բոլոր կենսաչափական մեթոդներն են հավասարապես կիրառելի տնտեսության մեջ և ծրագրերում: Օրինակ մատնահետքերի ճանաչումը և ծիածանաթաղանթի սկանավորումը կիրառելի չեն պրակտիկ շատ դեպքերում, օրինակ՝ բանկային և ֆինանսական ծառայություններում: Ձեռագիր ստորագրության ճանաչումը նույնականացման առավել հեշտ ընդունելի կենսաչափական մեթոդ է: Ստորագրման գործողությունը համընդհանուր ընդունված և սովորական գործողություն է ամենօրյա կյանքում: Ըստ էության, քիչ հավանական է, որ մարդիկ դեմ լինեն իրենց ստորագրության նույնականացմանը՝ համեմատած ուրիշ հնարավոր կենսաչափական մեթոդների: Սա թույլ է տալիս ձեռագիր ստորագրության համակարգչային ստուգումը ներմուծել մի շարք գոյություն ունեցող աշխատանքային գործընթացներում:

Ստորագրության նույնականացման համակարգը կարող է օգտագործվել բազմաթիվ բնագավառներում՝ արտադրականից մինչև բանկային՝ ապահովելով կազմակերպությանը և անհատներին ուժեղացված պաշտպանություն սեփական

միջոցների անօրինական օգտագործումից և փաստաթղթերի ու գործարքների վերահսկում, որոնք առկա են բիզնեսի գոյություն ունեցող պայմաններում:

Տեխնոլոգիական առաջընթացը մեծացրել է կենսաչափական համակարգերի ճշգրտությունը՝ դարձնելով դրանք առավել մատչելի որպես նույնականացման կենսունակ մեթոդ: Շնորհիվ ապահովության աստիճանի, հարմարավետության և բարձր անվտանգության, որը ապահովում է կենսաչափական մեթոդներն, այն արդեն ինտենսիվ օգտագործվում է որոշ ծրագրերում այլընտրանքային տեխնոլոգիաների հետ մրցակցային առավելության ապահովման համար:

Այդ պատճառով ինֆորմացիոն անվտանգության ապահովման նպատակով ձեռագիր ստորագրությամբ անձի նույնականացման հուսալի մեթոդների և ալգորիթմների մշակումը համարվում է ակտուալ խնդիր:

Ատենախոսության նպատակն է ուսումնասիրել և վերլուծել ձեռագիր ստորագրությամբ անձի նույնականացման գոյություն ունեցող մեթոդներն ու համակարգերը, բացահայտել նրանց առավելություններն ու թերությունները: Մշակել ձեռագիր ստորագրության հատկանիշների առանձնացման, ինչպես նաև նրանց համեմատման նոր և արդյունավետ ալգորիթմներ: Ստեղծել ձեռագիր ստորագրությունների ճանաչման ծրագրային համակարգ, կատարել նմանատիպ համակարգերի արդյունքների համեմատական վերլուծություն:

Հետազոտման մեթոդները: Տեսական ուսումնասիրությունները հիմնվել են թվային պատկերների մշակման, վերլուծման և կերպարների ճանաչման մեթոդների վրա: Ծրագրային համակարգի իրականացման համար օգտագործվել են բարձր մակարդակի ծրագրավորման լեզուներ (C#) և ծրագրային համակարգերի ստեղծման ժամանակակից մեթոդներ (Microsoft Visual Studio 2012, Microsoft SQL Server 2012):

Գիտական նորայթը: Առաջարկված է գլոբալ, դինամիկ տվյալներից և ստորագրության պատկերից հատկանիշների առանձնացման մեթոդներ: Մշակված է ստորագրությունների համեմատման բնութագրական վեկտորների համապատասխանեցման արդյունավետ ալգորիթմ: Մշակված են ստորագրությունների նույնականացման ալգորիթմներ՝ հիմնային վեկտորների մեթոդի SVM (Support vector machines), սահող շեմի (Sliding Threshold) և հեղինակի ընտրած շեմի (Global Threshold) օգտագործումով:

Գիրառական նշանակությունը: Առաջարկված մեթոդներն ու ալգորիթմները թույլ են տալիս բարձրացնել ձեռագիր ստորագրության նույնականացման համակարգի արտադրողականությունը՝ համեմատած

գրականությունում և պրակտիկայում ներկայացված համակարգերի հետ: Ստեղծված համակարգը կարող է օգտագործվել որպես հիմք՝ այս ոլորտի հետագա զարգացման համար: Ձեռագիր ստորագրությունների համեմատման մեթոդը թույլ է տալիս նշանակալիորեն մեծացնել նույնականացման արագությունը, թույլ է տալիս լուծել նույնականացման խնդիրը ռեալ ժամանակում:

Ներդրումներ: Ատենախոսության շրջանակներում մշակված ծրագրային փաթեթը ներդրվել է <<Սեվեն Սմարթս>> ընկերության կողմից մշակված “Focus Ortho” համակարգում:

Պաշտպանությանը ներկայացված են հետևյալ դրույթները.

- Հատկանիշների առանձնացման երեք մեթոդներ՝ լոկալ, գլոբալ և դինամիկ:
- Ստորագրությունների համեմատման բնութագրական վեկտորների համապատասխանեցման արդյունավետ ալգորիթմ:
- Ստորագրությունների նույնականացման ալգորիթմներ՝ հիմնային վեկտորների մեթոդի (SVM), սահող շեմի և հեղինակի ընտրած շեմի օգտագործումով:
- Առաջարկված հատկանիշների առանձնացման մեթոդների և բնութագրական վեկտորների համապատասխանեցման ալգորիթմի հիման վրա կառուցված ձեռագիր ստորագրությամբ անձի նույնականացման համակարգը:

Աշխատանքի ապրոքացիան

Ատենախոսության հիմնական արդյունքները գեկուցվել են կոմայտերագիտության և ինֆորմացիոն տեխնոլոգիաների միջազգային կոնֆերանսում (IPCV'12, Las Vegas, USA) և ՀՀ ԳԱԱ ԻԱՊԻ մասնագիտացված և ընդհանուր սեմինարներում:

Հրատարակումները: Ատենախոսության թեմայով հրատարակվել է 4 գիտական աշխատություն, որոնց ցանկը բերված է սեղմագրի վերջում:

Ատենախոսության կառուցվածքը և ծավալը: Ատենախոսությունը բաղկացած է ներածությունից, երեք գլուխներից, ամփոփիչ եզրակացությունից և 78 անվանում ընդգրկող գրականության ցանկից: Աշխատանքի ընդհանուր ծավալը կազմում է 109 էջ:

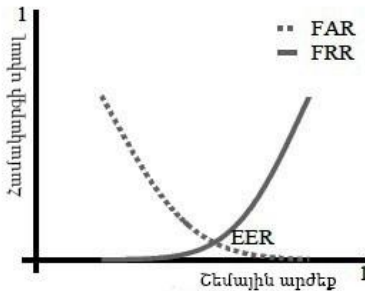
ԱՇԽԱՏԱՆՔԻ ՀԻՄՆԱԿԱՆ ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆԸ

Ատենախոսությունը կազմված է ներածությունից և երեք գլուխներից:

Ներածության մեջ հիմնավորված է ատենախոսության թեմայի արդիականությունը, ձևակերպված են ատենախոսության հիմնական նպատակները և ուսումնասիրության օբյեկտները:

Առաջին գլխում կատարված է կենսաչափական համակարգերի վերլուծություն, թվարկված են կենսաչափական հատկանիշներին ներկայացվող պահանջները: Բերված է կենսաչափական տեխնոլոգիաների համեմատություն կենսաչափական հատկանիշներին ներկայացվող պահանջներով:

Տրված են կենսաչափական համակարգի արտադրողականության գնահատման մեթոդները: Կենսաչափական համակարգի արտադրողականության գնահատման համար օգտագործվում է սխալների երկու տեսակ: Դրանք են՝ 1-ին տիպի սխալ (կեղծ ժխտում) և 2-րդ տիպի սխալ (կեղծ ընդունում): 1-ին տիպի սխալը տեղի է ունենում, երբ օրիգինալ ստորագրությունը սխալմամբ ժխտվում է, ինչը համապատասխանում է կեղծ ժխտման գործակցին (FRR-false rejection rate), իսկ 2-րդ տիպի սխալը տեղի է ունենում, երբ կեղծ ստորագրությունը ճանաչվում է որպես օրիգինալ, ինչը համապատասխանում է կեղծ ընդունման գործակցին (FAR-false acceptance rate): Նկ. 1-ում ներկայացված են FRR-ի և FAR-ի փոփոխման կորերը:



Նկար 1: FRR-ի և FAR-ի փոփոխման կորերը

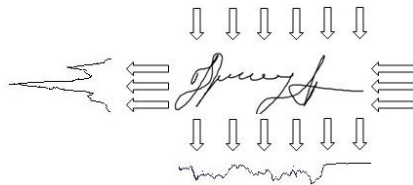
Նկարագրված են ձեռագիր ստորագրությունների նույնականացման տեխնոլոգիայի առավելությունները կենսաչափական տեխնոլոգիաների մյուս տեսակների նկատմամբ: Անցկացվում է ձեռագիր ստորագրության հիման վրա նույնականացման մեթոդների վերաբերյալ գրականության ուսումնասիրություն:

Երկրորդ գլխում մշակված են ձեռագիր ստորագրության ճանաչման մեթոդներ՝ ձեռագիր ստորագրությամբ անձի նույնականացման համակարգի մշակման համար:

§2.1-ում մշակվել են ձեռագիր ստորագրությունից հատկանիշների առանձնացման նոր մեթոդներ՝ հիմնված ստորագրության պատկերի պրոյեկցիաների, ձեռագիր ստորագրության դինամիկ և գլոբալ ինֆորմացիայի վրա: Ստորագրության պատկերի պրոյեկցիաները (ՄՊՊ) ստացվում է պատկերի պրոյեկցիաների հաշվարկման միջոցով՝ 0° -ից 180° -ի միջև հավասարաչափ բաշխված անկյունների միջոցով: Ենթադրենք, թե պատկերը ամբողջությամբ բաղկացած է N_ρ պիքսելներից: ՄՊՊ-ն ստացվում է N_ρ (ճառագայթների քանակ տվյալ անկյան համար) գուգահեռ ճառագայթների և N_θ (անկյունների քանակ) անկյունների միջոցով: Պիքսելների ընդհանուր քանակն, որոնք գտնվում են j -րդ ճառագայթի վրա նշանակենք R_j , $j=1, \dots, N_\rho N_\theta$: ՄՊՊ-ն կարող է ներկայացվել հետևյալ կերպ՝

$$R_j = \sum_{i=1}^{N_\rho} \omega_{ij}, \quad j = 1, 2, \dots, N_\rho N_\theta$$

որտեղ ω_{ij} ցույց է տալիս i -րդ պիքսելի մասնակցությունը j -րդ R_j -ում: Այդ պատճառով ամեն մի պրոյեկցիա պարունակում է այն R_j -երն, որոնք հաշվարկվում են տրված անկյան համար: ՄՊՊ-ի ճշգրտությունը կախված է N_θ -ից և N_ρ -ից:



Նկար 2: Ստորագրությունը և պրոյեկցիան, հաշվարկվում են 0° և 90° աստիճանի անկյունների համար:

Նախնական մշակման շատ պարզ և օգտակար գործընթաց է դիրքի կարգավորման իրականացումը: Կիրառվում է երկու ստորագրությունների կենտրոնների հավասարեցումը: Այս մեթոդն իրականացվում է ստորագրության ամեն առանձին կետից x և y կոորդինատների միջին արժեքների հաշվարկման միջոցով:

$$x' = x - \frac{1}{n} \sum_{i=1}^n x_i, \quad y' = y - \frac{1}{n} \sum_{i=1}^n y_i.$$

Այս բանաձևերում x -ը և y -ը սկզբնական կոորդինատներ են, x' -ը և y' -ը վերափոխված կոորդինատներն են, իսկ n -ը՝ պիքսելների քանակն է: Հատկանիշների առանձնացման ժամանակ շատ կարևոր գործընթաց է բնութագրական արժեքների կարգավորումը մինչև բնութագրական վեկտորների համեմատման փորձ կատարելը: Արժեքների կարգավորման համար օգտագործվել է z ցուցանիշը (z -score): Այն հաշվարկվում է տվյալների հավանական շեղման (σ) և միջին արժեքների (μ) օգտագործման միջոցով:

$$z = \frac{x - \mu}{\sigma}, \mu = \frac{1}{n} \sum_{i=1}^n x_i, \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

Կարգավորման այս մեթոդը ապահովում է այն փաստը, որ յուրաքանչյուր բնութագրիչի բաշխումը կունենա միջին զրո և միավորի ստանդարտ շեղում:

Գլոբալ տվյալները նպաստում են արտադրողականության արագ աճին՝ ամբողջ ստորագրությունում հասարակ բնութագրիչների հաշվարկման միջոցով: Մեր համակարգը հաշվում է գլոբալ բնութագրիչների հինգ հատկանիշ:

Աղյուսակ 1: Ստորագրության գլոբալ հատկանիշներ

Հատկանիշի նկարագրություն	Մեկնաբանություններ
Ts Ստորագրության ընդհանուր երկարություն	T-ժամանակահատված
N Ստորագրության հատվածներ	N-հատվածների քանակ
T _N Հատվածների երկարություն	T-ժամանակահատված
V Ստորագրության միջին արագություն	V-արագություն
V _N Հատվածներում միջին արագություն	

§2.2-ում մշակված է ձեռագիր ստորագրությունների համեմատման բնութագրական վեկտորների համապատասխանեցման ալգորիթմը, որն օգտագործվում է բնութագրական վեկտորների համապատասխանեցման համար, որոնք ստացվել են ՄՊՊ-ի և դինամիկ հատկանիշների առանձնացման նոր մեթոդներից: Նկ. 3-ում ներկայացված է վեկտորների համապատասխանեցման ալգորիթմի արդյունքների պատկերավոր ներկայացումը:

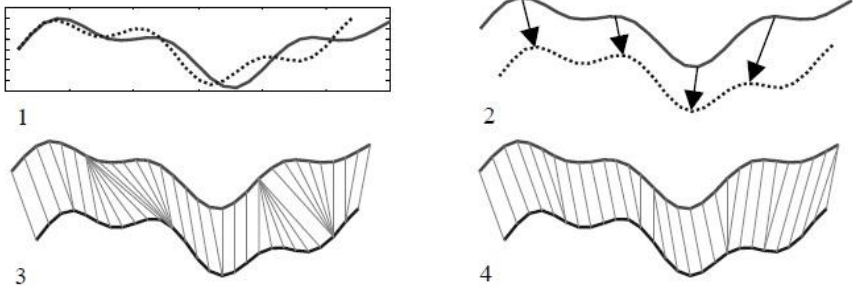
Դիցուկ յուրաքանչյուր հեղինակի շաբլոնում (տվյալների բազայում), գոյություն ունեն N թվով ստորագրություններ՝ $\{X_1, \dots, X_N\}$, որտեղ X_i -ն հեղինակի i -րդ ստորագրությունն է: Ստորագրության մոդելի ստեղծման համար անհրաժեշտ է առանձնացնել $N(N-1)/2$ քանակի գույգ հեռավորությունները N թվով ստուգողական ստորագրությունների հավաքածուի մեջ: Զույգ

հեռավորությունների հաշվարկումից հետո կսահմանենք հետևյալ ստատիստիկ մեծությունները w հեղինակի ստորագրության համար՝

I_{\min} –ն մինչև մոտակա ստորագրությունը միջին հեռավորությունը,

I_{\max} –ն մինչև ամենահեռու ստորագրությունը միջին հեռավորությունը:

Մշակված մոդելը w հեղինակի ստորագրության համար բաղկացած է N թվով ստորագրություններից, համակցված I_{\min} –ի և I_{\max} –ի հետ:



Նկար 3: Վեկտորների համապատասխանեցման ալգորիթմի արդյունքների պատկերավոր ներկայացում. 1) վեկտորները համադրված են մինևսյն սկզբնակետից, 2) ինտուիտիվ համադրում, 3) այլ համդրման ալգորիթմի արդյունք, 4) մշակված ալգորիթմի համադրման արդյունք

§2.3-ում մշակված են ստորագրությունների նույնականացման ալգորիթմներ՝ հիմնային վեկտորների մեթոդի (SVM), սահող շեմի և հեղինակի ընտրած շեմի օգտագործումով: Փորձարկվող ստորագրության X_{Test}^w մուտքի ժամանակ այն համեմատվում է w հեղինակի X_T^w ստուգողական ստորագրությունների հավաքածուի հետ:

$$X_T^w \in \{X_1, \dots, X_N\}$$

Այսպիսով, ստացվում է N թվով հեռավորություն, երբ փորձարկվող ստորագրությունը X_{Test}^w համեմատվում է w հեղինակի շարքնի N թվով ստուգողական ստորագրությունների: Հեռավորությունների հաշվարկից հետո սահմանում ենք հետևյալ ստատիստիկ մեծությունները՝

P_{\min}^w մինիմալ հեռավորություն,

P_{\max}^w մաքսիմալ հեռավորություն:

Ստորագրությունների նույնականացման հիմնային վեկտորների մեթոդի օգտագործման դեպքում՝ սկզբում ուսուցանվում է SVM դասակարգիչը: SVM դասակարգիչի ուսուցման համար առկա են M թվով իրական ստորագրություններ:

Յուրաքանչյուր ստորագրություն համեմատվում է հեղինակի շարժումում յուրաքանչյուր ստորագրության հետ: Դա բերում է NM քանակի հեռավորության որոշման, որոնցից ընտրվում են մինիմալ d_{\min} և մաքսիմալ d_{\max} հեռավորությունները: Ստացված հեռավորությունները կարգավորվում են I_{\min} և I_{\max} արժեքների օգնությամբ՝

$$N_{\min} = d_{\min} / I_{\min}, N_{\max} = d_{\max} / I_{\max} :$$

SVM-դասակարգիչի ուսուցանման համար օգտագործվում են N_{\min} և N_{\max} մեծությունները: P_{\min}^w և P_{\max}^w մեծությունները կարգավորվում են I_{\min} , I_{\max} արժեքների օգնությամբ՝

$$M_{\min} = P_{\min}^w / I_{\min}, M_{\max} = P_{\max}^w / I_{\max} :$$

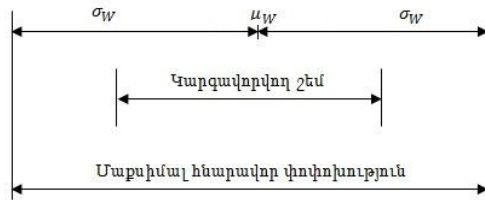
M_{\min} և M_{\max} մեծությունները օգտագործվում են փորձարկվող ստորագրության ընդունման կամ մերժման համար՝ ուսուցանված SVM-դասակարգիչի օգնությամբ:

Ստորագրությունների նույնականացման հեղինակի ընտրած շեմի մեթոդի օգտագործման դեպքում սահմանենք հետևյալ ստատիստիկ մեծությունները w հեղինակի ստորագրության համար՝

$$\mu_w = \frac{1}{M} \sum (\min D(X_A^w, X_T^w)), (A=1, \dots, M), (T=1, \dots, N)$$

$$\sigma_w^2 = \frac{1}{M} \sum (\min D(X_A^w, X_T^w) - \mu_w)^2, (A=1, \dots, M), (T=1, \dots, N)$$

որտեղ M – իրական ստորագրությունների քանակն է:



Նկար 4: Շեմի կարգավորվող արժեքների տիրույթ

Ստատիստիկ տվյալները օգտագործվում են շեմային հեռավորության ստացման համար, որը հետագայում օգտագործվում է մուտքային ստորագրության նույնականացման համար: σ_w ցույց է տալիս հեղինակի ստորագրության փոփոխականությունը, μ_w -ն միջին արժեքն է: Օրինակ, եթե շեմը տրված է $\sigma_w / 2$

արժեքով, ապա շեմային միջակայքերը հավասար կլինեն $[\mu_w - \sigma_w / 2, \mu_w + \sigma_w / 2]$ այսպիսով՝

$$[\mu_w - \sigma_w / 2] \leq P_{\min}^w \leq [\mu_w + \sigma_w / 2]$$

Եթե փորձարկվող ստորագրությունը ընկնում է տրված շեմային միջակայքերի միջև, ապա ստորագրությունը ընդունվում է կամ հակառակ դեպքում՝ մերժվում է:

Ստորագրությունների նույնականացման սահող շեմի մեթոդի օգտագործման դեպքում՝ հեղինակի համար սահող շեմի ստացման համար կարգավորվում է P_{\min}^w , որի կարգավորման համար օգտագործվում է μ_w մեծությունը հետևյալ կերպ՝

$$D_H(X_{Test}^w, X_{Ctrl}^w) = \frac{P_{\min}^w - \mu_w}{\mu_w}$$

որտեղ D_H -կարգավորված հեռավորությունն է:

Սահող շեմը τ , $\tau \in [0, \infty)$ -ի դեպքում օգտագործվում է փորձարկվող ստորագրությունների համար սխալների գործակցի որոշման համար: Երբ $D_H < \tau$ ունենում ենք՝

$$P_{\min}^w < \mu_w (1 + \tau)$$

փորձնական ստորագրությունն ընդունվում է, իսկ հակառակ դեպքում մերժվում է:

§2.4-ում մշակված է տարբեր մեթոդների համակցման եղանակները՝ անձի արդյունավետ նույնականացման համար: Արժեքների կարգավորումը լուծում է արժեքների տիրույթի և արժեքների բաշխման տարբերությունների խնդիրները: Այն փոխում է տարբեր համակարգերի համեմատման բաշխման արժեքների դիրքը և մաշտաբի պարամետրերը այնպես, որ բոլոր արժեքները ունենում են սահմանման միևնույն շրջան: Օգտագործվում է համակցման մեթոդ, որը հիմնված է կշռված գումարի կանոնի վրա, որը կարելի է նկարագրել հետևյալ կերպ՝

$$S_c = W_1 S_1 + W_2 S_2 + W_3 S_3$$

որտեղ S_1, S_2, S_3 -ը երեք համակարգերի համեմատման մինիմալ արժեքներն են, S_c -ը՝ համակցված արժեքը, W_1, W_2, W_3 ՝ կշիռը:

Այն նույն գործընթացը, որը նկարագրված է սահող կամ հեղինակի ընտրած շեմերի հաշվարկման համար կօգտագործվի արժեքների համակցման

համար: Ստատիստիկ արժեքները μ_c^w և σ_c^w հաշվարկվում են համակցված արժեքներից այն նույն գործընթացի միջոցով, ինչ որ μ_w և σ_w համար: Հեղինակի ընտրած շեմի համար՝ $[\mu_c^w - \sigma_c^w / 2] \leq S_c^w \leq [\mu_c^w + \sigma_c^w / 2]$, սահող շեմի համար՝ $S_c^w < \mu_c^w (1 + \tau)$:

Երրորդ գլխում ներկայացված են ձեռագիր ստորագրությամբ անձի նույնականացման համակարգի նկարագրությունը և փորձարկումների արդյունքները: Փորձարկումները անցկացվում են տվյալների հավաքածուների վրա, որը կոչվում է SVC2004 և բաղկացած է դինամիկ ստորագրություններից, որոնք հաշվարկվել էին ստորագրության ստուգման առաջին Միջազգային մրցման համար: Գրիչի ծայրի կոորդինատներն օգտագործվում են դինամիկ ստորագրությունը “իդեալական” ստորագրության պատկերի վերածելու համար, այսինքն այնպիսի ստորագրության, որը չի պարունակում ֆոնային շեղումներ:

SVC2004 տվյալների բազայի հիման վրա նույնականացման արդյունքները կարող են ստացվել չորս պայմաններից՝ (որակավորված/պատահական կեղծումներ տասը ստուգողական ստորագրությունների համար և որակավորված/պատահական կեղծումներ հինգ ստուգողական ստորագրությունների համար) դինամիկ տվյալների, գլոբալ տվյալների, ՄՊՊ մեթոդի և դրանց համակցման համար: Երեք մեթոդների մեծությունները կարգավորվում են մինչև ընդհանուր շրջան z-ցուցանիշի մեթոդի օգնությամբ, այնուհետև դրանք համակցվում են: Առանձին և համակցված համակարգի նույնականացումը իրականացվում է սահող, հեղինակի ընտրած շեմի և SVM-դասակարգիչի համար:

Աղյուսակ 2-ում ներկայացված է սահող շեմի, հեղինակի ընտրած շեմի և SVM-դասակարգիչի համար փորձարկվող տվյալների հավաքածուի նույնականացումը: Ցույց է տրված, որ բոլոր մեթոդների արտադրողականությունը մեծանում է մեծ քանակությամբ ստորագրությունների օգտագործման դեպքում՝ դրանք համակարգում գրանցելու ժամանակ, քանի որ ստուգողական տվյալների մեծ քանակության դեպքում կարգավորման ստատիստիկական որոշվում է առավել հստակ:

Սահող շեմի օգտագործմամբ փորձարկման արդյունքները ներկայացված են Աղյուսակ 2-ի երեք հատվածներից առաջինում: Օրինակ, որակավորված կեղծումների և 5 ստուգողական ստորագրությունների դիտարկման դեպքում համակցված համակարգը հանգեցնում է համարժեք սխալի գործակցին, որը

հավասար է 16%, այն դեպքում, երբ 14%-ին հավասար համարժեք սխալի գործակից համակցված համակարգը տալիս է 10 ստուգողական ստորագրությունների օգտագործման դեպքում: Այսպիսով բարելավում տեղի է ունենում ստուգողական ստորագրությունների քանակի ավելացման դեպքում:

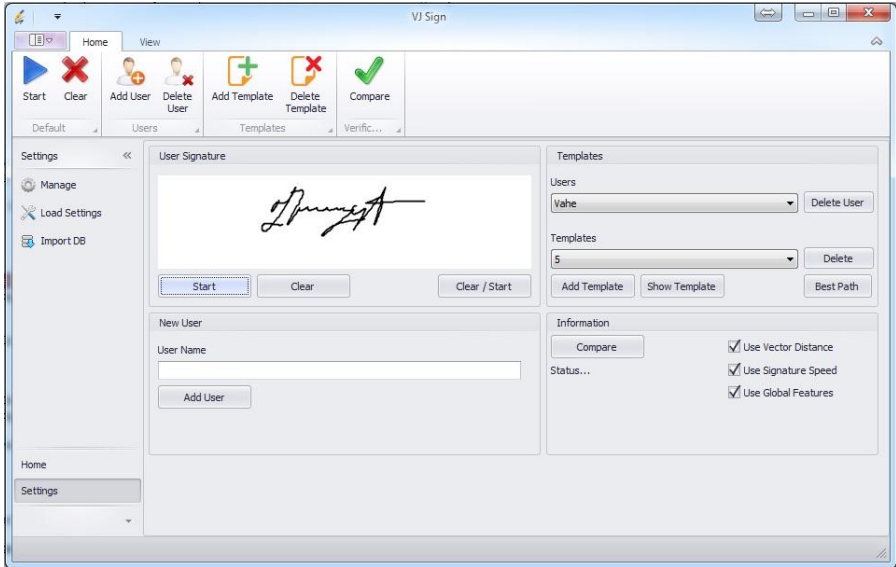
Աղյուսակ 2: Ստորագրության նույնականացում որակավորված/պատահական կեղծումների համար, SF-skilled forgeries (որակավորված կեղծումներ), RF-random forgeries (պատահական կեղծումներ)

	Համարժեք սխալի գործակից (EER) (%)											
	Սահող շեմ				Հեղինակի ընտրած շեմ				SVM-դասակարգիչ			
	5 ստ-ներ		10 ստ-ներ		5 ստ-ներ		10 ստ-ներ		5 ստ-ներ		10 ստ-ներ	
	SF	RF	SF	RF	SF	RF	SF	RF	SF	RF	SF	RF
ՄՊՊ մեթոդ	18	10	16	8	15	9	12	7	11	7	11	7
Դինամիկ տվյալների մեթոդ	20	12	18	11	18	9	17	7	16	7	15	7
Գլոբալ տվյալների մեթոդ	35	16	30	13	27	11	21	10	22	9	18	8
Մեթոդների համակցման դեպքում	16	9	14	7	13	8	13	6	12	7	10	6

Իրականացվել են միևնույն փորձարկումները, սակայն սահող շեմի փոխարեն օգտագործվել է հեղինակից կախված շեմը, ինչպես նկարագրված է աղյուսակ 2-ի երկրորդ մասում: Երևում է, որ արտադրողականությունը բարձրանում է երբ ստուգողական ստորագրությունների թիվը ավելանում է, և երբ դիտարկվում է երեք մեթոդների համակցումը: Հեղինակից կախված շեմի օգտագործումը ավելի լավ արդյունքի է հանգեցնում, քան սահող շեմի օգտագործումը: Իրականացվում են միևնույն փորձարկումները, սակայն շեմային արժեքի փոխարեն օգտագործվում է SVM-դասակարգիչը փորձնական ստորագրությունների դասակարգման համար: Աղյուսակ 2-ի երրորդ մասում ցուցադրված են ուսուցանված SVM-դասակարգիչի օգտագործմամբ փորձնական տվյալների հավաքածուի նույնականացման արդյունքները: Արտադրողականությունը բարելավվում է ստուգողական ստորագրությունների քանակի ավելացման և երեք մեթոդների համակցման դեպքում: SVM-

դասակարգիչի օգտագործումը ավելի լավ արդյունք է տալիս, քան հեղինակից կախված շեմի օգտագործումը:

Համակարգը (VJ Sign) մշակվել է C# լեզվով Visual Studio 2012 միջավայրում, որպես տվյալների բազա օգտագործվել է SQL Server 2012-ը, բազայի հետ աշխատանքն իրականացվել է LINQ հարցումների միջոցով:



Նկար 5: Համակարգի հիմնական ինտերֆեյսը

Հիմնական ինտերֆեյսը ցուցադրված է Նկ. 5-ում: Համակարգը հնարավորություն է տալիս ավելացնել նոր օգտագործողներ (Add User) և ջնջել գոյություն ունեցողներին (Delete User): Կարելի է խմբագրել օգտագործողների ստուգողական ստորագրությունները (Add Template, Delete Template), ստուգողական ստորագրությունների քանակի վրա սահմանափակումներ չկան: Հիմնական կարգավորումները դեկավարվում են Manage Settings պատուհանի միջոցով: Կարելի է օգտագործել ստորագրությունից հատկանիշների առանձնացման մեթոդները առանձին-առանձին կամ համակցված եղանակներով (Use Vector Distance, Use Signature Speed, Use Global Features):

Մյապիսով աշխատանքի հիմնական արդյունքները հետևյալն են՝

1. Մշակվել են ձեռագիր ստորագրության հատկանիշների առանձնացման երեք մեթոդներ՝ լոկալ, գլոբալ և դինամիկ [1,4]:
2. Առաջարկվել է ստորագրությունների համեմատման բնութագրական վեկտորների համապատասխանեցման արդյունավետ ալգորիթմ [3]:
3. Մշակվել են ստորագրությունների նույնականացման ալգորիթմներ՝ հիմնային վեկտորների մեթոդի (SVM), սահող շեմի և հեղինակի ընտրած շեմի օգտագործումով [2,3]:
4. Մշակվել է ձեռագիր ստորագրությամբ անձի նույնականացման ծրագրային համակարգ՝ կիրառելով առաջարկված հատկանիշների առանձնացման մեթոդները և բնութագրական վեկտորների համապատասխանեցման ալգորիթմը [2,4]:

Ատենախոսության թեմայով հրապարակված աշխատությունները

1. V. Khachaturyan, “Off-Line Signature Verification and Recognition” Mathematical Problems of computer science, vol 36, Yerevan, Armenia, 2012, pp. 115-120.
2. V. Khachaturyan, “Handwritten Signature Verification Using Hidden Markov Models” In Proceedings of The 2012 International Conference on Image Processing, Computer Vision, & Pattern Recognition, Volume I, WORLDCOMP’12, Las Vegas, Nevada, USA, July 16 - 19, 2012, pp. 347-351.
3. V. Khachaturyan, “Handwritten Signature Verification Using the DRT” Mathematical Problems of computer science, vol 39, Yerevan, Armenia, 2013, pp. 31-39.
4. V. Khachaturyan, “An Off-Line Signature Verification” Science and Education Publishing, Journal of Computer Sciences and Applications, 2013, Vol. 1, No. 2, Newark, De, USA, pp. 23-26, <http://pubs.sciepub.com/jcsa/1/2/2/index.html>

РАЗРАБОТКА И РЕАЛИЗАЦИЯ МЕТОДОВ РАСПОЗНАВАНИЯ РУКОПИСНОЙ ПОДПИСИ

РЕЗЮМЕ

В последние годы наблюдается устойчивая повышенность интереса к биометрическим методам идентификации и к их применению. Биометрические технологии используются для идентификации личности с помощью физиологических или поведенческих черт, которые являются уникальными для каждого индивидуума, и их невозможно забыть, потерять или украсть. В биометрических идентификационных системах наиболее часто используются сканирование радужной оболочки или сетчатки, рукописная верификация подписи, лица или распознавание голоса, отпечатков пальцев и т.д.

Среди этих методов есть биометрия рукописной подписи. Характеристики подписи неповторимы и практически не могут быть дублированы. Поэтому рукописный почерк до сих пор остается одним из наиболее сильных идентификаторов на сегодняшний день. В верификации рукописной подписи, для заключения о подлинности, извлекаются многочисленные характеристики тестируемой подписи и сравниваются с эталонной подписью, которая хранится в базе данных. Если доступно несколько подлинных эталонных подписей, увеличивается мера стабильности отдельного признака, что используется для оценивания вероятности отклонения, полученного в тестируемой подписи.

Не все биометрические методы равносильно приемлемы во всей индустрии и всех приложениях. Например, снятие отпечатки пальцев и сканирование радужной оболочки не приемлемо в промышленной практике, банковских и финансовых службах. Верификация рукописная подписи - это наиболее легко освоенный биометрический метод идентификации. Действие подписывания своим именем является общественно принятым и обычным действием в повседневной жизни. По существу, маловероятно, что индивидуумы будут протестовать против проверки своей подписи, по сравнению с другими возможными биометрическими анализами. Это позволяет компьютерную верификацию рукописной подписи внедрить во многие существующие рабочие процессы.

Система идентификации подписи может быть использована во многих областях: от производства до банковских, обеспечивая организаций и индивидуумов усиленной защитой от несанкционированного доступа и контроль документов и транзакций, которые имеются в существующих условиях бизнеса.

Технологический прогресс увеличил точность биометрических систем, делая их более доступными в качестве жизнеспособного метода верификации. Благодаря уровню надежности, удобству и высокой безопасности, которые обеспечивает

биометрия, она уже используется интенсивно в некоторых приложениях для обеспечения конкурентного преимущества с альтернативными технологиями.

Поэтому разработка надежных методов и алгоритмов автоматического идентификация личности по рукописной подписи в режиме реального времени, целью обеспечения информационной безопасности является актуальной задачей.

Цель диссертации исследовать и анализировать существующие системы и методы идентификации личности по рукописной подписи, выявить их преимущества и недостатки. Разработать новые и эффективные алгоритмы выделения и сопоставления признаков рукописной подписи. Разработать программную систему распознавания рукописных подписей, провести сравнительный анализ результатов похожих систем.

Научная новизна работы. Предложены методы выделения признаков из глобальных, динамических данных и с изображения подписи. Разработан эффективный алгоритм согласования характеристических векторов сравнения рукописных подписей. Разработаны алгоритмы верификации подписей с использованием SVM-классификатора (Support vector machines), с использованием скользящего порога (Sliding Threshold) и с использованием порога, зависящего от автора(Global Threshold).

Основные результаты работы следующие:

1. Разработаны три метода выделения признаков рукописной подписи: локальный, глобальный и динамический [1,4].
2. Предложен эффективный алгоритм согласования характеристических векторов сравнения рукописных подписей [3].
3. Разработаны алгоритмы верификации подписей с использованием SVM-классификатора, скользящего порога и с использованием порога, зависящего от автора [2,3].
4. Разработана програмная система идентификации личности по рукописной подписи с использованием предложенных методов выделения признаков и алгоритма согласования характеристических векторов [2,4].

DEVELOPMENT AND IMPLEMENTATION OF METHODS OF RECOGNITION OF A
HANDWRITTEN SIGNATURE

ABSTRACT

In the recent years the interest in biometric methods of identification and their use has steadily increased. Biometric technologies are used to establish a person's identity with the help of physiological or behavioral traits that are unique for each individual and cannot be forgotten, lost or stolen. In biometric identification systems the most commonly used methods are the iris or retina scanning, handwritten signature verification, face or voice recognition, fingerprints, etc.

Among these methods the biometrics of handwritten signature is also known. Signature characteristics are unique and cannot be duplicated in practice. Therefore, handwritten signature is known to be one of the strongest identifiers as of today. For the verification of handwritten signatures numerous characteristics of the test signature are extracted and compared with a genuine signature stored in the database to establish its authenticity. If more than one genuine signatures are available, the measure of the stability of a specific descriptor is increased that is used to estimate the probability of deviations derived from the test signature.

Not all the biometric methods are equally acceptable in the industry and the corresponding applications. For example, the fingerprinting and the iris scanning are not acceptable in the industrial practice, banking and financial services. Handwritten signature verification is the most easily assimilated biometric identification method. Signing one's name is socially accepted and is considered to be a usual activity in everyday life. People are unlikely to protest against the signature verification in comparison with other possible biometric analyzes. This allows the implementation of handwritten signature verification into many existing workflows.

A handwritten signature identification software system can be used in many spheres: from industrial to banking organizations ensuring enhanced protection for both organizations and individuals from unauthorized access and providing supervision of documents and transactions available in the existing business environment.

Technological progress has increased the accuracy of biometric systems making them more accessible as a viable method of verification. Due to the level of reliability, convenience and high security which biometrics provides, it is already intensively used in some applications to gain a competitive advantage against alternative technologies.

Therefore, the development of reliable methods and algorithms for automatic identification of an individual based on the handwritten signature in real time mode is an urgent task for ensuring information security.

The purpose of the dissertation is to research and analyze the existing systems and methods of identification based on handwritten signatures, to identify their strengths and weaknesses. Develop new and efficient algorithms for detecting and comparing the characteristics of a written signature. Develop a software system for the recognition of handwritten signatures; make a comparative analysis of similar systems.

Scientific novelty of the work. Methods have been suggested for sign extraction from the global and dynamic data and the image caption. An efficient algorithm has been implemented for matching the characteristic vectors to compare handwritten signatures. The algorithms of signature verification have been implemented using SVM-classifier (Support vector machines), Sliding Threshold and the Global Threshold.

The main results of the works are following:

1. Three methods have been developed to extract the characteristics of a written signature: local, global, and dynamic [1,4].
2. An effective algorithm has been suggested for matching the characteristic vectors and comparing handwritten signatures [3].
3. The algorithms of signature verification have been implemented using SVM-classifier, Sliding Threshold and the Global Threshold [2,3].
4. A handwritten signature identification software system has been developed using the suggested methods of sign extraction and characteristic vectors matching algorithm [2,4].

A handwritten signature in black ink, appearing to be 'D. Praveen', written in a cursive style with a long horizontal stroke at the end.

Ծավալը՝ 20 էջ: Տպաքանակ՝ 70:
ՀՀ ԳԱԱ ԻԱՊԻ կոմպյուտերային տպագրության լաբորատորիա
Հայաստան, Երևան, Պ. Մնակի փ., 1