

VILNIAUS UNIVERSITETAS

Andrej Kisel

ASMENS IDENTIFIKAVIMAS PAGAL PIRŠTŲ ATSPAUDUS IR BALSĄ

Daktaro disertacijos santrauka
Fiziniai mokslai, informatika (09 P)

Vilnius, 2010

Disertacija rengta 2005 – 2010 metais Vilniaus universitete

Mokslinis vadovas:

doc. dr. Algirdas Bastys (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P)

Disertacija ginama Vilniaus universiteto Informatikos mokslo krypties taryboje:

Pirmininkas

prof. habil. dr. Feliksas Ivanauskas (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P)

Nariai:

prof. habil. dr. Henrikas Pranevičius (Kauno technologijos universitetas, fiziniai mokslai, Informatika – 09P)

prof. dr. Romas Baronas (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P)

prof. dr. Olegas Vasilecas (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija – 07T)

doc. dr. Rimantas Vaicekuskas (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P)

Oponentai:

prof. dr. Vytautas Kleiza (Kauno technologijos universitetas, fiziniai mokslai, informatika – 09P)

doc. dr.(HP) Artūras Štikonas (Vilniaus Universiteto Matematikos ir Informatikos institutas, fiziniai mokslai, matematika – 01 P)

Disertacija bus ginama viešame Informatikos mokslo krypties tarybos posėdyje 2010 m. gruodžio mėn. 20 d. 15 val. Vilniaus universiteto Informacinių technologijų centre.

Adresas: Šaltinių, g. 1A, LT – 03214, Vilnius

Disertacijos santrauka išsiuntinėta 2010 m. .mėn. d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje.

VILNIUS UNIVERSITY

Andrej Kisel

PERSON IDENTIFICATION BY FINGERPRINTS AND VOICE

Summary of Doctoral Dissertation
Physical Sciences, Informatics (09 P)

Vilnius, 2010

Doctoral dissertation was prepared at Vilnius University in 2005 – 2010

Scientific Supervisor:

Assoc. Prof. Dr. Algirdas Bastys (Vilnius University, Physical Sciences, Informatics – 09P)

The dissertation is being defended at the Council of Scientific Field of Informatics at Vilnius University:

Chairman

Prof. Dr. Habil. Feliksas Ivanauskas (Vilnius University, Physical Sciences, Informatics – 09P)

Members:

Prof. Dr. Habil. Henrikas Pranevičius (Kaunas University of Technology, Physical Sciences, Informatics – 09P)

Prof. Dr. Romas Baronas (Vilnius University, Physical Sciences, Informatics – 09P)

Prof. Dr. Olegas Vasilecas (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering – 07T)

Assoc. Prof. Dr. Rimantas Vaicekuskas (Vilnius University, Physical Sciences, Informatics – 09 P)

Oponents:

Prof. Dr. Vytautas Kleiza (Kaunas University of Technology, Physical Sciences, Informatics – 09P)

Assoc. Prof. Dr. Artūras Štikonas (Vilnius University, Physical Sciences, Mathematics – 01 P)

The dissertation will be defended at the public meeting of the Council of Scientific Field of Informatics at Vilnius University Information Technology Center, at 15 p.m. on December 20 2010

Address: Šaltiniu g. 1A, LT – 03214, Vilnius

Turinys

1	Įvadas	1
1.1	Tyrimų sritis.....	1
1.2	Problemos aktualumas.....	1
1.3	Tyrimo objektai	2
1.4	Tyrimo tikslai ir uždaviniai.....	2
1.5	Mokslinis naujumas.....	3
1.6	Praktinė darbo reikšmė.....	3
1.7	Darbo rezultatų aprobavimas	4
1.8	Ginamieji teiginiai	4
1.9	Publikacijos disertacijos tema	5
1.10	Disertacijos struktūra	5
2	Piršto atspaudų sintezavimas	6
2.1	Įvadas	6
2.2	SFINGE metodas.....	6
2.3	Modifikuotas SFINGE metodas	6
2.4	Tikrų ir sintetinių atspaudų panašumas.....	6
2.5	Piršto atspaudų požymių išskyrimo algoritmo kokybės įvertinimas	7
2.6	Eksperimentai.....	7
2.7	Išvados.....	7
3	Pirštų atspaudų palyginimas	8
3.1	Įvadas	8
3.2	Pirštų atspaudų palyginimas panaudojant lokalius požymius	8
3.3	Lokalus palyginimas	8
3.3.1	Vienas kitą atitinkančių požymių suradimas	10
3.4	Vienas kitą atitinkančių struktūrų patvirtinimas.....	10
3.5	Slenksčių reikšmių nustatymas	12
3.6	Kokybės įvertinimas	12
3.7	Rezultatai.....	13
3.8	Išvados.....	15
4	Asmens atpažinimas pagal balsą	15
4.1	Įvadas	15
4.2	Tiesinės prognozės modelio grupinės delsos požymiai	16

4.2.1	Tiesinė prognozė	16
4.2.2	TP modelio spektro fazė.....	16
4.2.3	TP modelio spektro fazės savybės.....	17
4.3	Kalbos pavyzdžių panašumo metrika.....	18
4.3.1	Požymių statistikos.....	18
4.3.2	Dviejų trumpų kalbos fragmentų panašumas.....	19
4.4	Rezultatai.....	20
4.4.1	Grupinės delsos požymių grafinė iliustracija	20
4.4.2	Duomenų bazės ir rezultatai	22
5	Kelių biometrikų sujungimas	23
5.1	Įvadas	23
5.2	Testavimo duomenys	23
5.3	Sujungimas	23
5.3.1	Pirštų sujungimas	23
5.3.2	Pirštų ir balso apjungimas	25
6	Išvados	27
	Apie autorių	28
	Person Identification by Fingerprints and Voice	30
	Research Area.....	30
	Problem Relevance.....	30
	Research Objects	31
	The Objectives and Tasks of the Research	32
	Scientific Novelty	32
	Practical Importance of the Work	32
	Approval of Research Results.....	33
	Defended propositions.....	33
	Publications	34
	Thesis Structure.....	34
	Summary and Conclusions	35
	About the Author	36

1 Įvadas

1.1 Tyrimų sritis

Biometrinės technologijos pastaraisiais metais tampa vis populiareesnės. Asmens identifikavimas panaudojant unikalius požymius (gautus iš pirštų atspaudų, veido, balso ar akies rainelės) padaro žmogaus identifikavimą greitu, paprastu ir patikimu. Daugelyje šalių biometriniai duomenys yra įtraukiami į pasus ir asmens tapatybės korteles. Biometriniai duomenys yra naudojami: įstaigose, tam, kad būtų tiksliai apskaičiuota darbo trukmė (asmuo identifikuojamas įeinant ir išeinant iš darbo); žmogaus tapatybė nustatoma per rinkimus, siekiant įsitikinti jog kiekvienas asmuo balsavo tik vieną kartą; bankuose ir laisvės atėmimo įstaigose, siekiant padidinti saugumą.

Pastovus biometrinių technologijų augimas daro biometriją labai patrauklia tyrimų sritimi. Šio darbo tyrimų sritis yra pirštų ir balso biometrija: piršto atspaudų sintezavimas, piršto atspaudų palyginimas, asmens atpažinimas pagal balsą ir kelių biometrinių duomenų sujungimas.

1.2 Problemos aktualumas

Lengva pastebėti, kaip auga biometrinių technologijų populiarumas: pradinio lygio nebrangūs nešiojamieji kompiuteriai ir netgi mobilieji telefonai dabar turi integruotą pirštų atspaudų skenerį, kuris užtikrina duomenų saugumą ir palengvina darbą su įrenginiu, nes nereikia atsiminti slaptažodžių. Integruotos kameros panaudojamos asmens atpažinimui pagal veidą, o mikrofonas – pagal balsą. Šios technologijos palengvina darbą ir užtikrina saugumą. Dėl šių priežasčių mokslo institucijos ir privačios kompanijos investuoja laiką ir pinigus į biometrinių technologijų tobulinimą.

Siekiant palyginti skirtingus biometrinius algoritmus ir stebėti pažangą biometrinių tyrimų srityje buvo surengti tarptautiniai konkursai: FVC (FVC 2000, FVC 2002, FVC 2004, FVC 2006 ir FVC ongoing), NIST (National Institute of Standards and Technology) MINEX (MINEX, MINEX II ir Ongoing MINEX) ir PIV pirštų atspaudams; NIST Face Recognition Vendor Tests (FRVT) veidams; NIST Speaker Recognition Evaluation (SRE)

balsui; NIST Iris Challenge Evaluation (ICE), Independent Testing of Iris Recognition Technology (ITIRT) akių rainelėms. Šie konkursai parodo, kad nors pažanga tokiuose aspektuose, kaip patikimumas, greitis ir sąveika yra įspūdinga, liko dar daug sudėtingų neišspręstų problemų.

Visos biometrinės technologijos priklauso nuo įvesties signalo kokybės: jeigu gautas piršto atspaudas yra nekokybiškas, blogo kontrasto arba deformuotas, įrašyta frazė yra nepakankamo garsumo, akies rainelė yra dalinai paslėpta dėl voko arba atspindžių, veido nuotrauka gauta nepakankamomis apšvietimo sąlygomis arba panaudojant nekokybišką kamerą, identifikavimo uždavinys tampa sunkesniu. Pagrindinis iššūkis šiuolaikiniams biometriniams algoritmams yra įveikti šias sudėtingas sąlygas ir išgauti kiek įmanoma daugiau patikimų požymių. Nauji metodai padeda algoritmų kūrėjams geriau suprasti ir spręsti algoritmų problemas.

Šis darbas yra apie asmens atpažinimą pagal pirštų atspaudus ir balsą. Žmogaus atpažinimas pagal pirštų atspaudus yra labai populiarus: pirštų atspaudų skaitytuvai yra pigūs, juos paprasta naudoti ir identifikavimo procesas yra greitas.

Žmogaus atpažinimas pagal balsą irgi yra lengvai prieinamas kiekvienam, nes dažniausiai nereikia papildomos aparatūros: dauguma kompiuterių turi garso sąsają su galimybe prijungti mikrofoną, mikrofonai yra integruoti į internetinės kameras, ausinės ir mobiliuosius telefonus.

1.3 Tyrimo objektai

Darbo tyrimų objektai yra: pirštų atspaudų požymių išskyrimo algoritmo kokybės įvertinimas, deformuotų pirštų atspaudų sulyginimas, asmens atpažinimas panaudojant balsą ir kelių biometrinių duomenų panaudojimas.

1.4 Tyrimo tikslai ir uždaviniai

Tyrimo tikslas – kompleksiškai išanalizuoti mokslinių tyrimų sritį ir išspręsti sudėtingas problemas. Išanalizavus piršto atspaudų požymių išskyrimo algoritmo kokybės įvertinimo problemas bei siekiant įvertinti jų tikslumą, buvo nuspręsta panaudoti sintetinius pirštų atspaudus. Išanalizavus pirštų atspaudų palyginimo problemas buvo nuspręsta sukurti naują metodą kuris galėtų sulyginti deformuotus pirštus. Išanalizavus

žmogaus atpažinimo pagal balsą problematiką buvo nuspręsta sukurti naują asmens identifikavimo pagal balsą algoritmą. Ištyrus bendras biometrinių sistemų problemas buvo nuspręsta sujungti atpažinimą pagal pirštų atspaudus ir balsą.

1.5 Mokslinis naujumas

Naujas pirštų atspaudų atvaizdo sintezės metodas skirtingai nuo žinomų sintezės metodų gali sukurti piršto atspaudų paveiksluką su iš anksto nustatytais požymiais. Tokie vaizdai su žinomomis charakteristikomis leidžia įvertinti pirštų atspaudų požymių gavimo algoritmus nepriklausomai nuo pirštų atspaudų požymių palyginimo algoritmų. Darbe siūlomas naujas sintetinių pirštų praktinis taikymas: jie gali būti panaudojami įvertinant vaizdų kokybę tam tikroje duomenų bazėje arba siekiant palyginti pirštų atspaudų skaitytuvų kokybę.

Naujas pirštų atspaudų palyginimo algoritmas, kuris yra aprašytas antrajame skyriuje, naudoja lokalius požymius, iš jų sudarytas struktūras ir sugeba atpažinti netgi deformuotus pirštų atspaudus.

Naujas žmogaus atpažinimo pagal balsą algoritmas susideda iš naujų požymių ir naujo šių požymių palyginimo būdo.

Šiame darbe įrodoma, kad nėra koreliacijos tarp pirštų atspaudų panašumų ir balso įrašų panašumų. Toks šių biometrinių duomenų nepriklausomumas reiškia, kad jie gali būti sėkmingai panaudoti kartu norint pagerinti atpažinimo kokybę.

1.6 Praktinė darbo reikšmė

Aprašyti šiame darbe metodai gali būti naudojami siekiant išspręsti daugelį sunkių užduočių.

Pirštų atspaudų sintezė (2-as skyrius) gali būti panaudota didelių pirštų atspaudų duomenų bazių sukūrimui bei pirštų atspaudų požymių išskyrimo algoritmo kokybės įvertinimui. O kadangi yra įmanoma sukurti pirštų atspaudus su iš anksto nustatytais savybėmis ir požymiais, sintetinius pirštus galima panaudoti įvertinant tokias pirštų atspaudų atpažinimo algoritmo savybes, kaip: stabilumas triukšmo atžvilgiu ir iš piršto gaunamų požymių parametrų tikslumas.

Naujas pirštų atspaudų palyginimo metodas (3-as skyrius) leidžia patikimai palyginti netgi deformuotus pirštų atspaudus.

Žmogaus atpažinimas panaudojant pirštų atspaudus ir balsą kartu (4-as skyrius) leidžia sukurti patikimesnes ir lankstesnes biometrines sistemas.

1.7 Darbo rezultatų aprobavimas

Tyrimo rezultatai buvo paskelbti tarptautiniame žurnale Informatica. Tai pat darbo rezultatai buvo pristatyti tarptautinėje informacinių technologijų konferencijoje (IT2010).

1.8 Ginamieji teiginiai

1. Siekiant įvertinti biometrinių sistemų išplečiamumą (100 tūkstančių ir daugiau biometrinių įrašų) siūloma naudoti sintezuotus pirštų atspaudus su iš anksto nustatytomis savybėmis ir požymiais. Darbe rasti sprendimai įgalino kokybiškai ir greitai (iki 5-ių kartų greičiau nei žinomas geriausias algoritmas) sintezuoti pirštų atspaudus.
2. Sukurta metodika, kuri įvertina pirštų atspaudų požymių išskyrimo algoritmų, pirštų atspaudų skenerių arba duomenų bazių kokybę.
3. Deformuotų pirštų palyginimui siūloma atsisakyti tradiciškai naudojamoms pirštų atspaudų vaizdų registracijos. Rasta invariantinių pirštų atspaudų požymių palyginimo metrika kuri nenaudoja registracijos ir pagerina deformuotų pirštų atspaudų palyginimo kokybę.
4. Rasti nauji balso grupinės delsos požymiai. Pasiūlyta ir pagrįsta bendros informacijos kiekiu paremta požymių palyginimo metrika ir įrodytas jos privalumas lyginant su tradicinėmis žmogaus atpažinimo pagal balsą metrikomis.
5. Sudaryta originali pirštų atspaudų ir balso įrašų duomenų bazė, kuria remiantis įrodyta, kad pirštų atspaudų ir balso įrašų panašumai statistiškai nekoreliuoja. Parodyta, kad pirštų atspaudų ir balso įrašų multibiometrika pagerina identifikavimo kokybę apie 10 kartų.

1.9 Publikacijos disertacijos tema

Publikacijos, esančios periodiniuose recenzuojamuose tarptautiniuose leidiniuose (ISI):

1. Andrej Kisel, Alexej Kochetkov, Justas Kranauskas: Fingerprint Minutiae Matching without Global Alignment Using Local Structures. INFORMATICA, 2008, Vol. 19, No. 1, pp. 31-44, ISSN 0868-4952.
2. Algirdas Bastys, Andrej Kisel, Bernardas Salna: The Use of Group Delay Features of Linear Prediction Model for Speaker Recognition. INFORMATICA, 2010, Vol. 21, No. 1, pp. 1-12, ISSN 0868-4952.

Publikacijos, esančios periodiniuose recenzuojamuose tarptautiniuose konferencijų leidiniuose (ISI):

1. Andrej Kisel: Fast Fingerprint Image Synthesis. Proceedings of 16th International Conference on Information and Software Technologies. April 21st - 23rd 2010, Kaunas University of Technology, Lithuania, ISSN 2029-0063 pp. 107-115.

Recenzuojami straipsniai periodiniuose recenzuojamuose tarptautiniuose leidiniuose (ISI):

1. Andrej Kisel: Multibiometrics using fingerprints and voice. Information technology and control, Kaunas University of Technology, Lithuania.

1.10 Disertacijos struktūra

Darbą sudaro 3 pagrindinės dalys: pirštų atspaudų atpažinimas (2-as ir 3-as skyriai), asmens atpažinimas pagal balsą (4-as skyrius) ir asmens atpažinimas pagal pirštus ir balsą kartu (5-as skyrius).

2-ame skyriuje yra aprašytas greitas piršto atspaudų sintezės metodas kuris gali būti panaudotas sukurti itin dideles pirštų atspaudų duomenų bazes ir įvertinti piršto atspaudų požymių išskyrimo algoritmo kokybę.

3-ame skyriuje piršto atspaudų palyginimo algoritmas, kuris gali patikimai sulgyinti deformuotus pirštų atspaudus.

4-as skyrius pristato tiesinės prognozės modelio grupinės delsos požymių panaudojimą atpažinti asmens balsą.

Asmens atpažinimas pagal pirštus ir balsą kartu yra pristatytas 5-ame skyriuje.

Paskutinis 6-as skyrius užbaigia darbą rezultatų santrauka ir išvadomis.

2 Piršto atspaudų sintezavimas

Šis skyrius pristato greitą piršto atspaudų sintezės metodą, kuris sugeba sukurti piršto atspaudų vaizdą su iš anksto nustatytais charakteristikomis ir požymiais.

2.1 Įvadas

Įvade aprašyti plačiai naudojami ir žinomi piršto atspaudų požymių išskyrimo algoritmų kokybės įvertinimo metodai bei pasiūlomas piršto atspaudų sintezės panaudojimas jų kokybei nustatyti.

2.2 SFINGE metodas

Šiame poskyryje yra aprašytas vienas iš pažangiausių sintezės metodų – SFINGE metodas.

2.3 Modifikuotas SFINGE metodas

Modifikacijos yra pasiūlytos tam, kad būtų galima sukurti piršto atspaudą su iš anksto nustatytais požymiais – pradinis paveikslukas yra ne atsitiktinis, o sukonstruotas taip, kad filtravimo metu piršto atspaudų požymiai atsirastų norimose vietose. Tai leidžia sintezę naudoti požymių išskyrimo algoritmo kokybės įvertinimui. Taip pat šiame poskyryje yra aprašyta iki penkių kartų sintezės procesą pagreitinanti optimizacija – adaptyvus filtravimas (filtravimo eiga yra apskaičiuojama prieš filtruojant panaudojant Euklido atstumų žemėlapij.

2.4 Tikrų ir sintetinių atspaudų panašumas

Siekiant įsitikinti, jog sintetiniai pirštai gali būti panaudoti vietoje tikrų pirštų, buvo atliktas papildomas eksperimentas: buvo pasirinkta duomenų bazė iš 800 pirštų (100 žmonių po 10 pirštų per 8-ias skenavimo sesijas); požymiai buvo išskirti iš pirštų; sintetiniai pirštai su požymiais iš tikrų pirštų požymių buvo sukurti; tikri ir sintetiniai pirštai buvo lyginami automatinio pirštų atspaudų atpažinimo algoritmu. Buvo

apskaičiuojama, kokia yra tikimybė, kad piršto atspaudas panašumas į kitus N to paties piršto atspaudus yra didesnis negu piršto panašumas į ne to paties piršto atspaudus.

Žemiau pateikta lentelė parodo, kad sintetiniai pirštai yra labai panašūs į tikrus pirštų atspaudus ir gali būti panaudoti vietoje jų.

N	1	2	3	4	5	6	7	8
Sintetiniai su Tikrais	100	100	100	99.625	99.625	99	97.375	89.125
Sintetiniai su Sintetiniais	100	100	100	100	99.875	99.125	97.125	89.625
Tikri su Tikrais	100	100	100	100	100	99.75	98.875	96.375

2.5 Piršto atspaudų požymių išskyrimo algoritmo kokybės įvertinimas

Šis poskyris aprašo du galimus sintetinių pirštų panaudojimo būdus ir pateikia panaudojimo pavyzdžius.

2.6 Eksperimentai

Šiame poskyryje aprašoma, kaip pirštų atspaudų atpažinimo kokybė buvo įvertinta abiem pasiūlytais sintezuotų pirštų panaudojimo atvejais.

2.7 Išvados

Skyriaus pabaigoje daromos tokios išvados: naujas sintezės metodas yra greitas (iki penkių kartų greitesnis už žinomą sintezės algoritmą), sugeba sintezuoti pirštų atspaudus su iš anksto nustatytais požymiais: pozicija, kampas, tipas; metodas gali būti panaudotas ne tik didelių pirštų atspaudų duomenų bazių kūrimui, bet ir pirštų palyginimo algoritmų kokybės įvertinimui ir silpnų pusių (stabilumas triukšmui, pjūviams, deformacijoms) išaiškinimui; eksperimentai rodo, jog gaunamo atvaizdo kokybė yra geresnė negu tikrų pirštų atspaudų kokybė ir kadangi informacija apie piršto požymius yra žinoma, tai leidžia patikimiau įvertinti algoritmo kokybę.

Pristatytos dvi požymių išskyrimo algoritmo arba duomenų bazės kokybės nustatymo matavimo priemonės pritaikytos skirtingoms duomenų bazėms. Kokybė įvertinama intervale nuo 0 iki 100%, kur 100% atitinka aukščiausią kokybę. Gauti įvertinimai atitinka žinomų duomenų bazių charakteristikas.

3 Pirštų atspaudų palyginimas

Šis skyrius pristato naują pirštų atspaudų palyginimo metodą, kuris sugeba palyginti pasuktus, pastumtus ir deformuotus pirštus.

3.1 Įvadas

Įvade aptariamos šiuolaikinių pirštų atspaudų palyginimo metodų problemos. Daroma išvada, kad nors posūkio ir pozicijos atstatymas gali ženkliai palengvinti pirštų palyginimą, tais atvejais, kai pirštas yra deformuotas, to padaryti neįmanoma.

3.2 Pirštų atspaudų palyginimas panaudojant lokalius požymius

Šis poskyris aprašo siūlomą palyginimo algoritmą, kuris naudoja lokalius požymius ir neatlieka piršto atspaudos posūkio ir poslinkio įvertinimo. Pirmame etape abiejuose pirštuose yra sudaromi lokalūs požymiai iš požymių ir jų kaimyninių požymių. Vieno piršto lokalūs požymiai yra sulyginami su kito piršto lokaliais požymiais ir nustatomi vienas kitą atitinkantys požymiai. Kitame etape yra sudaromi antros eilės požymiai ir yra įvertinamas lokalių požymių atitikimo tikslumas. Visi žingsniai yra invariantiški piršto atspaudos posūkiams, postūmiams ir deformacijoms. Eksperimentai su viešai prieinamomis ir vidinėmis duomenų bazėmis parodė kad naujo metodo kokybė yra geresnė negu žinomų metodų, kurie naudoja posūkio ir postūmio atstatymą.

3.3 Lokalus palyginimas

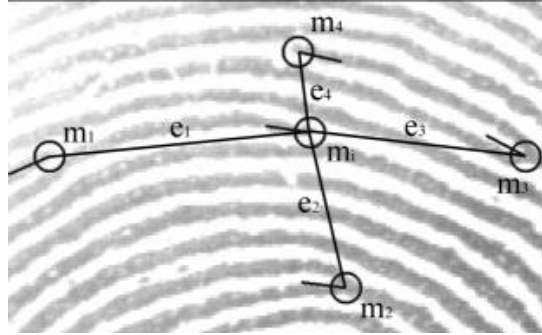
Šiame poskyryje yra aprašytas lokalių požymių sudarymas ir palyginimas.

Lokali struktūra (1-as pav.) asocijuota su požymiu m_i (aprašomu vektoriumi, kuris prasideda taške (x_i, y_i) su kryptimi φ_i), kai maksimalus atstumas yra d_{max} ir maksimalius artimiausių kaimyninių skaičius – n_{max} yra grafas $S_i = (V_i, E_i)$ sudarytas iš:

$$\begin{aligned} V_i &= \{m_j \mid \text{atstumas}(m_i, m_j) < d_{max}\}, |V_i| \leq n_{max}; \\ E_i &= \{e_{ij} \mid e_{ij} \text{ sujungia } m_i \text{ ir } m_j\}. \end{aligned} \quad (3.1)$$

Kur e_{ij} yra tokia duomenų struktūra: $(i, j, \text{atstumas}(m_i, m_j), \phi_{ij})$, ϕ_{ij} yra kampas tarp m_i ir m_j kryptių.

Tokia lokali struktūra yra invariantiška posūkio ir postūmio atžvilgiu ir išlieka stabili esant deformacijoms.



1 paveikslukas: Lokali struktūra asocijuota su požymiu m_i .

Kadangi požymių išskyrimo algoritmai daro klaidas, t.y. klaidingai išskiria arba atmeta požymius, lokaliuos struktūros negali būti palygintos tiesiogiai. Tam naudojama palyginimo funkcija $CS(S_i^T, S_j^S)$, kuri matuoja panašumą tarp dviejų lokalių struktūrų S_i^T ir S_j^S iš pirštų T (test) ir S (sample). Jeigu požymių išskyrimo metu klaidų nėra, struktūras galima palyginti tiesiog surūšiuojant kaimyninius požymius pagal laikrodžių rodyklę pradedant nuo centrinio požymio krypties ir pritaikant žemiau aprašytą funkcija:

$$CS(S_i^T, S_j^S) = \sum_{k=1}^{|E_i^T|} CE(e_{ik}^T, e_{jk}^S), \quad (3.2)$$

$$CE(S_j^T, S_j^S) = \begin{cases} 0, \text{ kai } \Delta d \geq \Delta d_{max} \text{ ar } \Delta \varphi \geq \Delta \varphi_{max}; \\ w_d \frac{\Delta d_{max} - \Delta d}{\Delta d_{max}} + w_\varphi \frac{\Delta \varphi_{max} - \Delta \varphi}{\Delta \varphi_{max}}, \text{ priešingu atveju.} \end{cases} \quad (3.3)$$

$$\Delta d = |\text{atstumas}(m_i^T, m_j^T) - \text{atstumas}(m_i^S, m_j^S)|,$$

$$\Delta \varphi = \min(|\varphi_i - \varphi_j|, 2\pi - |\varphi_i - \varphi_j|),$$

$\Delta d_{max}, \Delta \varphi_{max}$ – slenksčiai, w_d, w_φ – parametrai.

Prieš lyginant struktūras, jeigu požymių išskyrimo metu buvo padarytos klaidos, yra surandamas ilgiausias bendras kaimyninių požymių posekis (IBP)

$$\overline{CS}(S_i^T, S_j^S) = \sum_{e_{ik}^T, e_{jl}^S \in IBP(Ev_i^T, Ev_j^S)} CE(e_{ik}^T, e_{jl}^S). \quad (3.4)$$

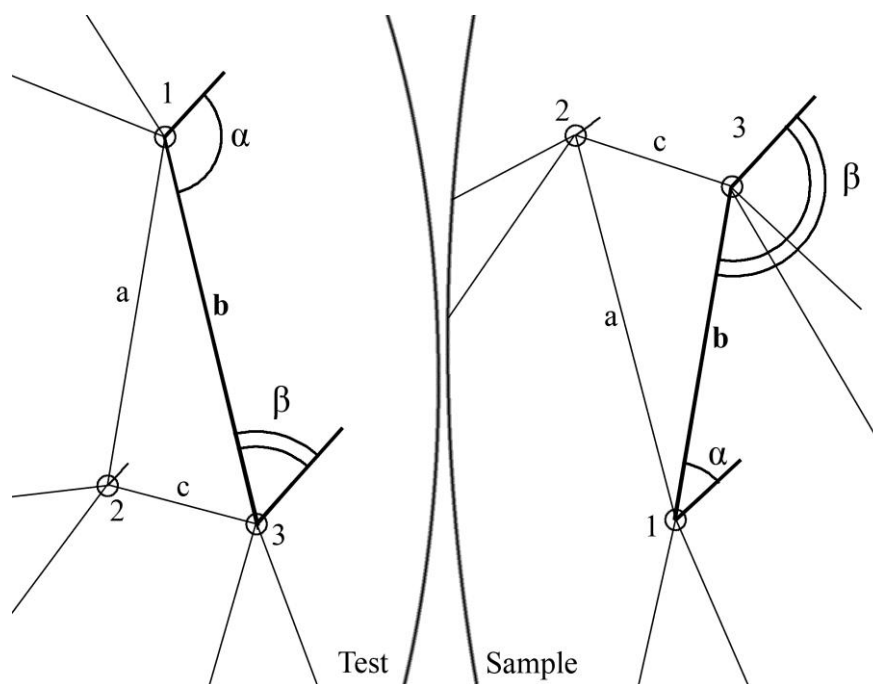
3.3.1 Vienas kitą atitinkančių požymių suradimas

Po to kai panašumas tarp kiekvienos lokalsios struktūros iš T piršto atspaudu ir lokalių struktūrų iš S piršto atspaudu yra paskaičiuotas, surandamas toks vienas kitą atitinkančių lokalių struktūrų porų rinkinys, kad kiekviena lokali struktūra iš vieno piršto atitiktų ne daugiau negu vieną lokalią struktūrą iš kito piršto:

$$C = \{(S_i^T, S_j^S) | S_i^T \in T, S_j^S \in S\}, |C| \leq \min(N, M). \quad (3.5)$$

3.4 Vienas kitą atitinkančių struktūrų patvirtinimas

Nors lokalsios struktūros gali turėti didelį panašumą, jos gali būti neteisingai lokalizuotos pirštų atspauduose t.y. transformacija, kuri galėtų jas sutapatinti neegzistuoja (2-as pav.).



2 paveikslukas: Atitinkančios lokalsios struktūros (sužymėtos 1, 2 ir 3).

Siekiant atmesti minėtus atvejus yra naudojamos aukštesnio lygio struktūros. Aukštesnio lygio struktūros HS_{ij}^T ir HS_{ij}^S yra apibrėžiamos kaip atitinkančių struktūrų poros iš atitinkančių struktūrų rinkinio, kuris buvo aprašytas anksčiau:

$$\begin{aligned} HS_{ij}^T &= (c_i^T, c_j^T), 1 \leq i < j \leq |C|, \\ HS_{ij}^S &= (c_i^S, c_j^S), 1 \leq i < j \leq |C|. \end{aligned} \quad (3.6)$$

Aukštesnio lygio struktūros yra pažymėtos lotyniškais raidėmis a , b ir c 2-ame pav.

Dviejų aukštesnio lygio struktūrų HS_{ij}^T ir HS_{ij}^S panašumas SS apibrėžiamas taip:

$$SS(HS_{ij}^T, HS_{ij}^S) = \begin{cases} 0, \text{ kai } \Delta l \geq \Delta l_{max} \text{ ar } \Delta \alpha \geq \Delta \alpha_{max} \text{ ar } \Delta \beta \geq \Delta \alpha_{max}; \\ w_l \frac{\Delta l_{max} - \Delta l}{\Delta l_{max}} + w_\alpha \frac{\Delta \alpha_{max} - \Delta \alpha}{\Delta \alpha_{max}} + w_\beta \frac{\Delta \alpha_{max} - \Delta \beta}{\Delta \alpha_{max}}, \text{ priešingu atveju.} \end{cases} \quad (3.7)$$

m_i^T, m_j^T – požymiai sudarantys struktūrą HS_{ij}^T ,

m_i^S, m_j^S – požymiai sudarantys struktūrą HS_{ij}^S ,

$$\Delta l = |\text{atstumas}(m_i^T, m_j^T) - \text{atstumas}(m_i^S, m_j^S)|,$$

$$\Delta \alpha = \min(|\alpha^T - \alpha^S|, 2\pi - |\alpha^T - \alpha^S|),$$

$$\Delta \beta = \min(|\beta^T - \beta^S|, 2\pi - |\beta^T - \beta^S|),$$

$\alpha^T, \alpha^S, \beta^T, \beta^S$ – kampai tarp atkarpų jungiančių lokalias struktūras ir lokalių struktūrų centrinių požymių kryptių (2-as pav.),

$\Delta l_{max}, \Delta \alpha_{max}$ – slenksčiai, w_l, w_α – parametrai.

Dviejų pirštų atspaudų panašumas apibrėžiamas kaip aukštesnio lygio struktūrų (praėjusių patvirtinimo žingsnį) panašumų (sukombinuotų su lokalių struktūrų sudarančių aukštesnio lygio struktūras panašumais) suma:

$$\text{panašumas} = \frac{\sum_{i,j} f(SS(HS_{ij}^T, HS_{ij}^S), CS(S_i^T, S_i^S), CS(S_j^T, S_j^S))}{g(N, M)}, \quad (3.8)$$

kur S_i^T, S_i^S – lokalias struktūros, kurios sudaro HS_{ij}^T ,

S_j^T, S_j^S – lokalias struktūros, kurios sudaro HS_{ij}^S ,

Panašumo funkcija $f(SS(HS_{ij}^T, HS_{ij}^S), CS(S_i^T, S_i^S), CS(S_j^T, S_j^S))$ gali būti viena iš žemiau pateikiamų:

$$SS(HS_{ij}^T, HS_{ij}^S) + \frac{1}{2} (CS(S_i^T, S_i^S) + CS(S_j^T, S_j^S)); \quad (3.9)$$

$$SS(HS_{ij}^T, HS_{ij}^S) \cdot CS(S_i^T, S_i^S) \cdot CS(S_j^T, S_j^S); \quad (3.10)$$

$$SS(HS_{ij}^T, HS_{ij}^S) \cdot \sqrt{CS(S_i^T, S_i^S) \cdot CS(S_j^T, S_j^S)}. \quad (3.11)$$

$g(N, M)$ funkcija yra naudojama panašumo reikšmių normalizavimui, kai požymių skaičiai pirštuose skiriasi. šiame tyrime buvo naudojama $g(N, M) = N * M$.

3.5 Slenksčių reikšmių nustatymas

Šis poskyris skirtas slenksčių bei parametų parinkimui. Siekiant pademonstruoti, kad slenksčių reikšmės gali būti keičiamos $\pm 50\%$ be žymių atpažinimo kokybės pakeitimų, buvo atliktas papildomas eksperimentas.

3.6 Kokybės įvertinimas

Panaudojus viešai prieinamą duomenų bazę NIST Special Database 29 (SD29) buvo įvertinta atpažinimo kokybė. Duomenų bazę sudaro skenuotų pirštų atspaudų kortelės po dešimt pirštų kiekvienoje (pirštai buvo gauti dviem būdais – tiesiog prispaudžiant ir prispaudžiant bei persukant tam, kad būtų gautas viso piršto galiuko atspaudas. Tiesioginiam palyginimui buvo pasirinktas viešai prieinamas NIST (JAV Standartų ir technologijų instituto) pirštų atspaudų palyginimo algoritmas (toliau vadinamas BOZORTH3). Požymiai buvo išskiriami komerciniu aukštos kokybės algoritmu (toliau vadinamu COMM). Pasiūlytas pirštų atspaudų požymių palyginimo metodas paprastumo dėlei toliau vadinamas LSM (Lokalių struktūrų metodas).

SD29 duomenų bazė yra sudaryta iš nuskenuotų pirštų atspaudų kortelių. Papildomi eksperimentai buvo atlikti tam, kad būtų palyginta atpažinimo kokybė, kai naudojami šiuolaikiniai skaitmeniniai pirštų atspaudų skaitytuvai.

Žemiau aprašytos duomenų bazės buvo panaudotos kokybės palyginimui:

1. DB1 iš FVC2002 pirštų atspaudų atpažinimo konkurso surinkta panaudojant optinį skenerį „Identix TouchView II“. Duomenų bazę sudaro 800 paveiksliukų;
2. DB2 iš FVC2002 konkurso surinkta „Biometrika FX2000“ skaitytuvo pagalba (irgi 800 paveiksliukų);
3. Vidinė kompanijos „Neurotechnology“ duomenų bazė surinkta „Identix DFR2090“ skaitytuvo pagalba (toliau vadinama INTERNAL1). (1400 paveiksliukų);
4. Vidinė kompanijos „Neurotechnology“ duomenų bazė surinkta „Cross Match Verifier 300“ skaitytuvo pagalba (toliau vadinama INTERNAL2) (1400 paveiksliukų).

3.7 Rezultatai

NIST VTB pirštų atspaudų palyginimo sistema su Bozorth98 pirštų atspaudų palyginimo algoritmu (ankstesnė BOZORTH3 algoritmo versija) dalyvavo JAV surengtame pirštų atspaudų atpažinimo konkurse (FpVTE 2003), pademonstravo prilygstančią komerciniams algoritmams kokybę ir netgi aplenkė dalį jų. Eksperimentai parodė, kad pasiūlytas LSM metodas veikia geriau negu BOZORTH3 pirštų atspaudų palyginimo algoritmas. Buvo sudaryta 18 ROC kreivių skirtingoms SD29 duomenų bazės dalims: P2P (pirštai gauti juos prispaudžiant lyginami su pirštais gautais prispaudžiant); P2R (pirštai gauti prispaudžiant lyginami su pirštais gautais prispaudžiant ir persukant); R2R (pirštai gauti prispaudžiant ir persukant lyginami su pirštais gautais prispaudžiant ir persukant). Pirštai žymimi taip: RT – dešinysis nykštys; LT – kairysis nykštys; RI – dešinysis rodomasis pirštas; LI – kairysis rodomasis pirštas; RM – dešinysis vidurinis pirštas; LM – kairysis vidurinis pirštas).

Rezultatai – klaidingų atmetimų skaičius, kai klaidingų atpažinimų skaičius yra 0.01%, pateikiami 1-oje ir 2-oje lentelėse.

1 lentelė: COMM+BOZORTH3.

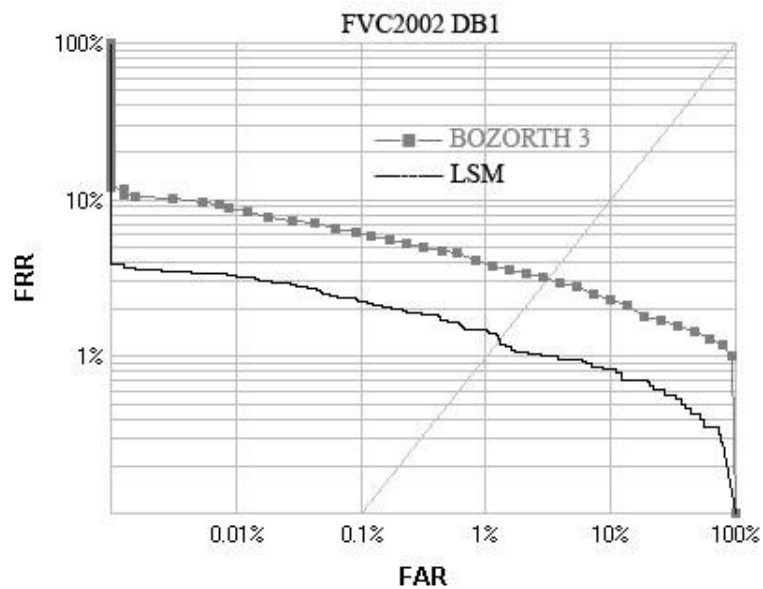
dalis	RT	LT	RI	LI	RM	LM	vidurkis
P2P	7.1	6.2	18.9	15.0	14.8	14.1	12.68
P2R	12.7	12.8	14.7	18.6	14.6	12.2	14.27
R2R	16.3	11.3	6.6	5.9	8.3	6.0	9.08

2 lentelė: COMM+ LSM.

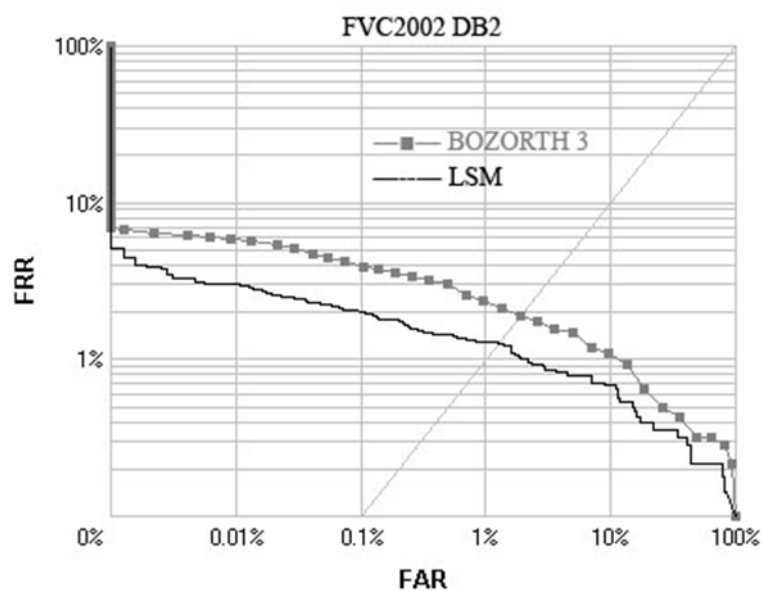
dalis	RT	LT	RI	LI	RM	LM	vidurkis
P2P	7.14	9.89	14.8	11.5	15.9	13.2	12.07
P2R	5.36	9.34	9.07	12.5	9.62	9.34	9.205
R2R	4.95	7.69	4.4	6.59	5.49	8.79	6.318

Pasiūlytas metodas sumažina klaidingų atmetimų procentą, kai klaidingų atpažinimų procentas yra 0.01 vidutiniškai nuo 12.01% iki 9.20%.

3 pav. ir 4 pav. demonstruoja atpažinimo kokybę (ROC kreives), kai pirštų atspaudai gauti naudojant skaitmeninius pirštų atspaudų skaitytuvus.



3 paveikslukas: COMM+BOZORTH3 palyginamas su COMM+LSM ant FVC2002 DB1 duomenų bazės.



4 paveikslukas: COMM+BOZORTH3 palyginamas su COMM+LSM ant FVC2002 DB2 duomenų bazės.

3-oje lentelėje yra pateikta rezultatų suvestinė, kai lyginimo kriterijus yra klaidingų atmetimų procentas, kai klaidingų atpažinimų procentas yra 0.01.

3 lentelė: klaidingų atmetimų procentas kai klaidingų atpažinimų procentas yra 0.01.

Duomenų bazė	BOZORTH3	LSM
SD29	12.01	9.20
FVC2002 DB1	8.73	3.29
FVC2002 DB2	5.82	3.00
INTERNAL1	5.65	2.3
INTERNAL2	1.78	0.44

3.8 Išvados

Šiame skyriuje yra pristatytas naujas pirštų atspaudų palyginimo metodas, kuris gali palyginti deformuotus pirštų atspaudus.

Atpažinimo kokybė buvo įvertinta naudojant dideles duomenų bases. Buvo pademonstruota, kad metodas yra tolerantiškas piršto atspaudų posūkiams, poslinkiams ir deformacijoms. Metodo kokybė buvo palyginta su BOZORTH3 pirštų palyginimo metodu ir buvo nustatytas žymus kokybės pagerėjimas.

4 Asmens atpažinimas pagal balsą

4.1 Įvadas

Šis skyrius aprašo naują asmens atpažinimo pagal balsą metodą.

Kalbos požymiams apibrėžti buvo pasiūlyta panaudoti tiesinės prognozės modelio perdavimo funkcijos spektro fazę. Požymiai išreiškiami skaičių poromis, kurios aprašo tiesinės prognozės modelio grupinės delsos ekstremumo taškus. Pirmasis poros komponentas yra grupinės delsos ekstremumo taško abscisė, o antrasis yra spektro ekstremumo taške juostinio pločio įvertis.

Pasiūlyta balsų panašumo įvertinimo metrika, kuri apibrėžiama panaudojant įvestus grupinės delsos požymius. Metrika adaptuota nepriklausomam nuo teksto balso atpažinimui laikantis nuostatos, kad pateikiamą kalbos signalą gali sudaryti kelių asmenų balsai.

4.2 Tiesinės prognozės modelio grupinės delsos požymiai

4.2.1 Tiesinė prognozė

Tiesinės prognozės (TP) modelyje kalbos signalo kadro imtys išreiškiamos lygybe

$$x_n = \sum_{i=1}^P a_i x_{n-i} + G e_n, \quad (4.1)$$

kur a_1, a_2, \dots, a_p yra tiesinės prognozės koeficientai (TPK), P yra modelio eilė, G ir e_n yra žadinimo stiprumas ir šaltinis. TPK randami adaptyviai kiekvienam 20-30 ms trukmės kalbos kadrai minimizuojant vidutinę kvadratinę energiją. Paprastumo dėlei laikysime, kad TP modelio eilė yra nelyginė, t.y. $p = 2M - 1$. TP modelio kompleksinis spektras arba perdavimo funkcija apibrėžiama lygybe:

$$H_z = \frac{G}{A(z)}, \quad (4.2)$$

kur

$$A(z) = 1 - \sum_{i=1}^{2M-1} a_i z^{-i} \quad (4.3)$$

yra atvirkštinis filtras. TP modelio spektras yra kalbos spektro gaubiamoji, kurios detalumas priklauso nuo modelio eilės $2M - 1$; didėjant M didėja gaubiamosios detalumas.

4.2.2 TP modelio spektro fazė

Apibrėžkime simetrinį $p(z)$ ir antisimetrinį daugianarius $q(z)$ tokiomis formulėmis:

$$p(z) = \frac{z^M A(z) + z^{-M} A(z^{-1})}{2}, \quad (4.4)$$

$$q(z) = \frac{z^M A(z) - z^{-M} A(z^{-1})}{2i}, \quad i = \sqrt{-1}. \quad (4.5)$$

$p(z)$ ir $q(z)$ daugianariai susieti su linijinio spektro (LS) daugianariais $P(z)$ ir $Q(z)$ tokiomis formulėmis:

$$P(z) = A(z) + z^{-2M} A(z^{-1}) = 2z^{-M} p(z), \quad (4.6)$$

$$Q(z) = A(z) - z^{-2M} A(z^{-1}) = 2iz^{-M} q(z). \quad (4.7)$$

Vienetiniame apskritime $p(z)$ ir $q(z)$ įgyja realias reikšmes,

$$|A(z)|^2 = p(z)^2 + q(z)^2, \quad (4.8)$$

ir

$$p(z) + q(z)i = z^M A(z). \quad (4.9)$$

Iš (4.8) ir (4.9) lygybių gauname, kad TP modelio perdavimo funkcijos spektro atsakas ir fazė tenkina tokias lygtis:

$$|H(z)| = \frac{G}{\sqrt{p(z)^2 + q(z)^2}}, \quad (4.10)$$

$$\text{ir} \quad (\arg H)(e^{i\omega}) = \Phi(\omega) = M\omega - \arctan\left(\frac{q(e^{i\omega})}{p(e^{i\omega})}\right), \quad (4.11)$$

$$\omega \in [0, 2\pi), \quad i = \sqrt{-1}.$$

4.2.3 TP modelio spektro fazės savybės

TP polinio modelio kompleksinį spektrą galima užrašyti tokia forma:

$$H(z) = \frac{G}{\prod_{m=1}^P (1 - r_m e^{i\alpha_m} z^{-1})}, \quad (4.12)$$

kur $r_m e^{i\alpha_m}$ yra m-ojo poliaus padėtis, o $\alpha_m \in [0, 2\pi)$ yra poliaus kampinis dažnis. Iš (4.12) lygybės išplaukia, kad m-ojo poliaus indėlis į TP modelio kompleksinio spektro fazę išreiškiamas adityviu nariu

$$\arctan\left(\frac{r_m \sin(\omega - \alpha_m)}{1 - r_m \cos(\omega - \alpha_m)}\right).$$

Todėl spektro fazės pirmosios ir antrosios eilės išvestinės tenkina tokias lygtis:

$$\frac{d\Phi(\omega)}{d\omega} = \sum_m \frac{r_m \cos(\omega - \alpha_m) - r_m}{1 - 2r_m \cos(\omega - \alpha_m) + r_m^2}, \quad (4.13)$$

$$\frac{d^2\Phi(\omega)}{d\omega^2} = - \sum_m \frac{r_m (1 - r_m^2) \sin(\omega - \alpha_m)}{(1 - 2r_m \cos(\omega - \alpha_m) + r_m^2)^2}. \quad (4.14)$$

TP modelio grupinė delsa (angl. group delay) vadinama fazės pirmosios eilės išvestinė, paimta su minuso ženklu. Polių reikšmės tiesiogiai neskaičiuojamos (nes jų apskaičiavimas yra brangi ir nestabili procedūra skaičiavimo kompiuteriu prasme) ir grupinės delsos reikšmės įvertinamos skaitmeniškai diferencijuojant (4.11) tapatybę. Iš (4.13) lygybės išplaukia, kad greta „stiprių“ polių, kurių r_m artimas 1 galima tikėtis grupinės delsos maksimumo taške ω_m , kurio reikšmė artima kampiniam dažniui α_m . Lokalaus maksimumo taškas ω_m gali būti surandamas prilyginus fazės antrosios eilės išvestinę nuliui ir imant artimiausią α_m reikšmei šaknį. Iš (4.13) gauname:

$$\Phi'(\omega_m) \approx \frac{r_m}{1 - r_m}, \quad (4.15)$$

$$\text{ir} \quad r_m \approx \frac{\Phi'(\omega_m)}{1 + \Phi'(\omega_m)}. \quad (4.16)$$

Remiantis gautomis išraiškomis, kalbos kadro grupinės delsos savybes apibrėžiame aibe porų

$$\left(\omega_m, \frac{1}{1 + \Phi'(\omega_m)}\right) = (\omega_m, \delta_m), \quad (4.17)$$

kur $\{\omega_m\}$ yra fazės antrosios eilės išvestinės nulių kirtimų (šaknų) priklausančių intervalui $(0, \pi)$ aibė ir

$$\delta_m = 1 - \frac{\Phi'(\omega_m)}{1 + \Phi'(\omega_m)} = \frac{1}{1 + \Phi'(\omega_m)} \quad (4.18)$$

apibrėžia kalbos kadro spektro formantės plotį.

4.3 Kalbos pavyzdžių panašumo metrika

Padarykime prielaidą, kad turime du skaitmeninius kalbos pavyzdžius $\{x_n\}$ ir $\{y_n\}$ kurių panašumą reikia įvertinti. Tarkime imtis $\{x_n\}$ yra žinomo diktoriaus X ir $\{y_n\} = Y$ imtys priklauso vienam, dviems ar keliems tiriamiesiems. Panašume metrika turi įvertinti tikimybę, kad diktorius X dalyvauja Y kalbos pavyzdžio pokalbyje. Toks kalbančiojo identifikavimo scenarijus būdingas kriminalistikai, kur naudojamos automatinės kalbančiojo identifikavimo sistemos ieškant galimą įtariamąjį iš anksto sukauptose padariusių nusikaltimus kalbos pavyzdžių duomenų bazėse.

4.3.1 Požymių statistikos

Mūsų apibrėžtos k -jo kalbos kadro grupinės delsos savybės sudarytos iš porų (f_m^k, δ_m^k) , kur f_m^k yra m -asis grupinės delsos maksimumo dažnis ir δ_m^k spektro formantės plotis ekstremumo taške. Turimas diktoriaus kalbos pavyzdys yra dalinamas į trumpus vienos sekundės trukmės intervalus ir yra įvertinamas grupinės delsos savybių skirstinys kiekviename 1 sek. trukmės laiko intervale. Kadangi atstumas tarp dviejų kalbos kadru yra 0.01. sek., tai 1 sek. trukmės intervale turime bendrai apie $100(M - 1)$ (f_m^k, δ_m^k) porų. $(f_m^k, \delta_m^k) \in (0, \frac{FS}{2}) \times (0, 1)$ Skirstinys įvertinamas padalinant $(0, \frac{FS}{2}) \times (0, 1)$ į $N \times L$ stačiakampių dalių ir apskaičiuojant porų (f_m^k, δ_m^k) patenkančių į kiekvieną stačiakampį kiekį.

4.3.2 Dviejų trumpų kalbos fragmentų panašumas

Dviejų vienos sekundės trukmės kalbos pavyzdžių panašumas apibrėžiamas bendrąja tų kalbos intervalų grupinės delsos požymių skirstinių informacija (angl. mutual information). Tarkime $I = N * L$ yra bendras skirstinio histogramos stačiakampių $\{B_i\}_{i=1}^I$ kiekis, o $C_X^x = \{c_i^x\}_{i=1}^I$ ir $C_Y^y = \{c_i^y\}_{i=1}^I$ požymių vektoriai, kurių komponentai yra grupinės delsos porų priklausančių B_i stačiakampiui kiekis. Pagal apibrėžimą visi c_i^x ir c_i^y atitinka X ir Y kalbėtojų $[x, x + 1)$ ir $[y, y + 1)$ kalbos sekundžių intervalus. Tegul H_X^x and H_Y^y yra C_X^x and C_Y^y skaitliukų Šenono entropijos, t.y.

$$H_X^x = - \sum_{i=1}^I c_i^x / |C_X^x| \log_2(c_i^x / |C_X^x|), \quad (4.19)$$

$$H_Y^y = - \sum_{i=1}^I c_i^y / |C_Y^y| \log_2(c_i^y / |C_Y^y|), \quad (4.20)$$

$$|C_X^x| = \sum_{i=1}^I c_i^x, \quad |C_Y^y| = \sum_{i=1}^I c_i^y. \quad (4.21)$$

Tegul $C_{X,Y}^{x,y} = \{c_i^x + c_i^y\}_{i=1}^I$ žymi jungtinius C_X^x ir C_Y^y skaitliukus

$$\text{ir} \quad H_{X,Y}^{x,y} = - \sum_{i=1}^I c_i^{x,y} / |C_{X,Y}^{x,y}| \log_2(c_i^{x,y} / |C_{X,Y}^{x,y}|) \quad (4.22)$$

yra jungtinė $C_{X,Y}^{x,y}$ skaitliukų entropija. Tuomet entropijoms H_X^x, H_Y^y teisingas toks teiginys:

Teorema 1. Bet kokiems skaitliukams C_X^x ir C_Y^y ir jų jungtiniams $C_{X,Y}^{x,y}$ teisingos tokios nelygybės:

$$pH_X^x + qH_Y^y \leq H_{X,Y}^{x,y} \leq pH_X^x + qH_Y^y + H_{p,q}, \quad (4.23)$$

$$\text{kur} \quad p = \frac{|C_X^x|}{|C_{X,Y}^{x,y}|}, \quad q = \frac{|C_Y^y|}{|C_{X,Y}^{x,y}|} = 1 - p \quad (4.24)$$

$$\text{ir} \quad H_{p,q} = -p \log_2 p - q \log_2 q. \quad (4.25)$$

Apibrėžimas 1. X kalbos $[x, x + 1)$ sekundės ir Y kalbos $[y, y + 1)$ sekundžių intervalų panašumu vadinsime skaičių

$$\rho(X_{[x,x+1)}, Y_{[y,y+1)}) = 1 + \frac{pH_X^x + qH_Y^y - H_{X,Y}^{x,y}}{H_{p,q}}. \quad (4.26)$$

Iš 1 teoremos išplaukia, kad $X_{[x,x+1)}$ ir $Y_{[y,y+1)}$ kalbos intervalų panašumas yra visuomet neneigiamas ir nedidesnis už 1. Apibrėškime kalbos intervalo $Y_{[y,y+1)}$ panašumą į bendrą X kalbos pavyzdį.

Apibrėžimas 2. Kalbos intervalo $Y_{[y,y+1)}$ panašumas į X kalbos pavyzdį yra skaičius

$$\rho(X, Y_{[y,y+1)}) = \frac{\sum_{x=0}^{T_x-1} \rho(X_{[x,x+1)}, Y_{[y,y+1)})}{T_x}. \quad (4.27)$$

Čia T_x yra X kalbos trukmė sekundėmis. Kitaip tariant $\rho(X, Y_{[y,y+1)})$ panašumas yra vidutinis $Y_{[y,y+1)}$ panašumas į visus $X_{[x,x+1)}$ kalbos intervalus.

Paskutinis apibrėžimas skirtas apibendrintam X ir Y kalbos pavyzdžių panašumui.

Apibrėžimas 3. X ir Y kalbos pavyzdžių panašumas yra skaičius

$$\rho(X, Y) = \text{vidutinė reikšmė pusės didžiausių } \rho(X, Y_{[y,y+1)}), \quad (4.28)$$

$$y = 0, 1, \dots, T_Y - 1.$$

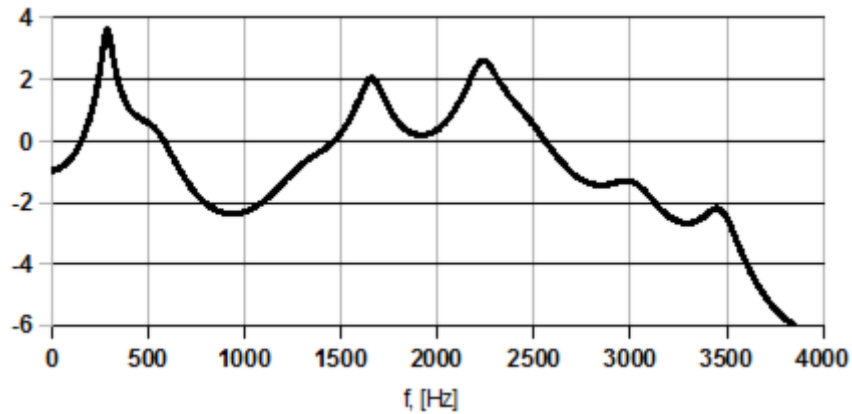
Pateikta panašumo metrika $\rho(X, Y)$ yra asimetrinė (bendru atveju $\rho(X, Y) \neq \rho(Y, X)$). Tai paaiškinama X ir Y duomenų asimetrija: X kalba priklauso vienam kalbėtojui (diktoriui), o Y kalbos pavyzdyje gali kalbėti vienas, du ar daugiau asmenų. Visi pateikti panašumų apibrėžimai remiasi bendrosios informacijos sąvoka, yra neneigiami ir neviršija 1. Jei X ir Y yra visiškai skirtingi, t.y. visi X ir Y grupinės delsos taškai priklauso nesikertantiems stačiakampiams B_i , tuomet visiems x ir y $H_{x,y}^{x,y} = pH_x^x + qH_y^y + H_{p,q}$ ir $\rho(X, Y) = 0$. Priešingu atveju, kai visi skaitikliai yra proporcingi, t.y. ($\forall x, y, i: c_i^x = \text{const } c_i^y$), gauname $H_{x,y}^{x,y} = pH_x^x + qH_y^y = H_x^x$ ir $\rho(X, Y) = 1$. Vadinasi, pasiūlyta panašumo metrika $\rho(X, Y)$ turi tikimybinę interpretaciją, t. y. panašumą $\rho(X, Y)$ galime interpretuoti tikimybe, kad X kalbėtojas dalyvauja Y kalbos dialoge.

4.4 Rezultatai

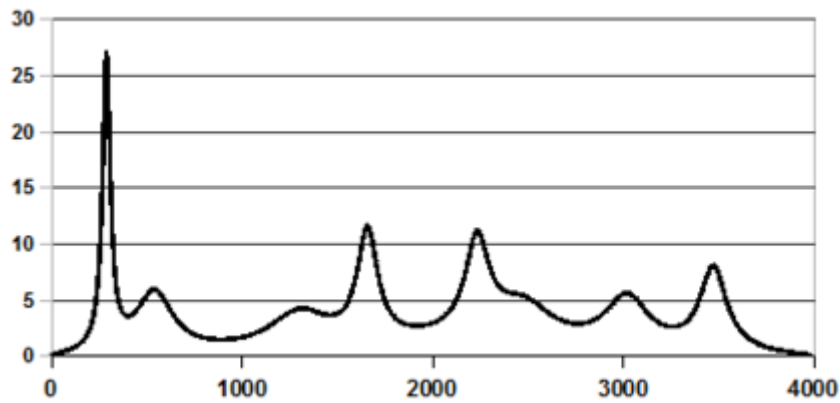
4.4.1 Grupinės delsos požymių grafinė iliustracija

5 pav. iliustruoja vieno kalbos kadro galingumo spektrą įvertintą TP modeliu. To paties kadro TP modelio spektro fazės išvestinė pateikta 6-ame pav. Lyginant TP modelio galingumo spektrą ir kompleksinio spektro fazę matyti, kad pastarasis išryškina dvi papildomas formantes (spektro maksimumus). Likusios penkios formantės dažnių ašyje

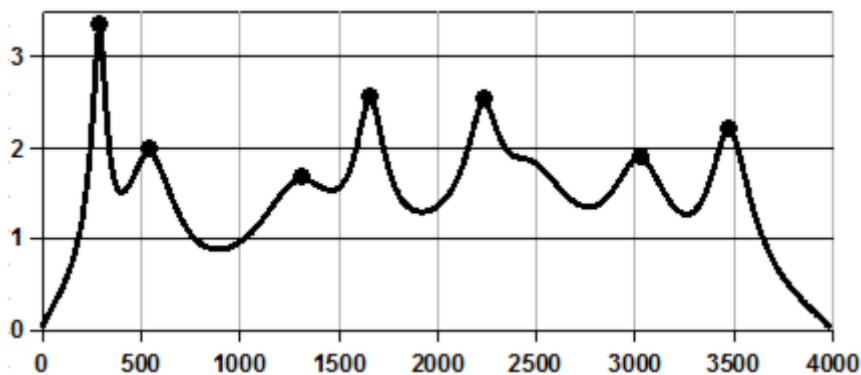
turi panašias pozicijas, tačiau grupinės delsos ekstremumai yra labiau lokalizuoti, kas didina jų skiriamąją gebą atskirti panašaus spektro kalbos fragmentus.



5 paveikslukas: Vieno kalbos kadro TP modelio galingumo spektro logaritmas.



6 paveikslukas: To paties kalbos kadro grupinė delsa.



7 paveikslukas: Kalbos kadro $-\log$ formančių pločio $-\log$ su pažymėtais grupinės delsos savybių taškais $(f_m, -\log(\delta_m))$.

(4.16) formulė apytiksliai įvertina poliaus taško atstumą pasirinktos formantės $f = f_m$ iki vienetinio apskritimo. 7 pav. pateiktas tų atstumų $-\log$ ir papildomai grafike

paženklinti taškai formantes atitinkantys taškai. Pažymėtų taškų koordinatės apibrėžia poras $(f_m, -\log(\delta_m))$, kurios sudaro kalbos kadro požymių vektorių.

4.4.2 Duomenų bazės ir rezultatai

Keletas skirtingų asmens atpažinimo pagal balsą metodikų buvo palyginti naudojant RUSBASE ir NFISRE duomenų bases. Panaudojant NFISRE duomenų bazę buvo pasiektas idealus atpažinimas.

4 lentelė: Asmens atpažinimo kokybė duomenų bazė RUSBASE, vyrų balsai, 1-a frazė, su skirtingais metodais ir požymiais.

Metodas	Požymiai	EER [%]
VQ	MFCC	8.8
GMM	MFCC	5.8
GMM	F&A	5.1
Foneminis	F&A	2.32

Naujas atpažinimo metodas buvo palygintas su Gauso Mišinių Modeliu (GMM), kuris naudoja Mel Kepstro Koeficientus (MFCC), formantes ir antiformantes (F&A). 4-a lentelė pateikia atpažinimo rezultatus (lygių klaidų procentą (EER)) gautus panaudojant RUSBASE bazės, vyrų balsai, 1-a frazė, naudojant MFCC, F&A, vektorių kvantavimo metodiką (VQ) ir Gauso mišinių modelio GMM atpažinimo metodus. Lygių klaidų procentas (EER) kinta nuo 2.32% iki 8.8 %. Naujas metodas su tais pačiais duomenimis padaro tik 0.042% klaidų.

5-a lentelė pateikia pilnus RUSBASE atpažinimo rezultatus. Zero FAR yra toks klaidingų atmetimų procentas, kai klaidingų atpažinimų procentas yra nulis. Zero FRR yra toks klaidingų atpažinimų lygis, kai klaidingų atmetimų lygis yra nulis.

5 lentelė: Kalbančiojo atpažinimas naudojant naują metodą (RUSBASE duomenų bazė, frazės 1-5).

Frazė	Balsas	Zero FAR[%]	EER[%]	Zero FRR[%]
1	vyras	1.8	0.042	0.12
1	moteris	1.96	0.042	0.07
2	vyras	0.8	0.084	0.12
2	moteris	2.17	0.2	1.37
3	vyras	3.19	0.058	0.09
3	moteris	1.96	0.033	0.06
4	vyras	0.6	0.01	0.02
4	moteris	4.6	0.112	0.15
5	vyras	2.79	0.199	0.59
5	moteris	0.44	0.007	0.01

Naujas kalbančiojo atpažinimo pagal balsą metodas demonstruoja iki 20 kartų geresnį atpažinimą, kai vertinimo kriterijus yra lygių klaidų procentas (EER).

5 Kelių biometrių sujungimas

5.1 Įvadas

Įvade yra išdėstytos biometrinių sistemų problemos, o joms spręsti yra pasiūlomas kelių biometrinių įrašų sujungimas.

5.2 Testavimo duomenys

Tam, kad būtų išsamiai įvertinta apjungto algoritmo kokybę, buvo sudaryta speciali duomenų bazė, susidedanti iš pirštų atspaudų ir balso įrašų. Šiame poskyryje aprašyti duomenų bazės struktūra ir sudarymo būdas.

5.3 Sujungimas

Šiame poskyryje yra palyginti du apjungimo scenarijai: dviejų pirštų panaudojimas bei piršto ir balso panaudojimas. Pirštų atspaudų požymiai buvo išskiriami naudojant komercinį požymių išskyrimo algoritmą ir lyginami panaudojant algoritmą, kuris buvo aprašytas trečiame skyriuje.

5.3.1 Pirštų sujungimas

Siekiant ištirti skirtingų pirštų panašumų nepriklausomumą, buvo suskaičiuotos koreliacijos tarp pirštų panašumų (6-ta lentelė). Pirštai yra sunumeruoti taip: 1, 2, 3, 4, 5

yra dešinėsios rankos pirštai pradedant nuo nykščio; 6, 7, 8, 9, 10 yra kairės rankos pirštai pradedant nuo nykščio.

6 lentelė: Koreliacijos tarp skirtingų pirštų panašumų.

Pirštas	1	2	3	4	5
1	1	.21 ± .04	.21 ± .04	.3 ± .04	.19 ± .04
2	.21 ± .04	1	.33 ± .04	.34 ± .03	.33 ± .04
3	.21 ± .04	.33 ± .04	1	.44 ± .04	.37 ± .04
4	.30 ± .04	.34 ± .03	.44 ± .04	1	.37 ± .04
5	.19 ± .04	.33 ± .04	.37 ± .04	.37 ± .037	1
6	.24 ± .04	.11 ± .04	.16 ± .04	.19 ± .05	.09 ± .05
7	.10 ± .04	.10 ± .05	.07 ± .05	.09 ± .04	.06 ± .05
8	.10 ± .04	.13 ± .04	.14 ± .05	.26 ± .04	.20 ± .05
9	.14 ± .04	.08 ± .04	.15 ± .04	.29 ± .04	.22 ± .04
10	.27 ± .04	.20 ± .04	.12 ± .04	.12 ± .04	.26 ± .04

6 lentelės pratęsimas.

Pirštas	6	7	8	9	10
1	.24 ± .04	.10 ± .04	0,05 ± .04	.14 ± .04	.27 ± .04
2	.11 ± .04	.10 ± .05	0,13 ± .04	.07 ± .04	.20 ± .04
3	.16 ± .04	.07 ± .05	0,14 ± .05	.15 ± .04	.12 ± .04
4	.19 ± .05	.09 ± .04	0,26 ± .04	.29 ± .04	.12 ± .04
5	.09 ± .05	.06 ± .04	0,20 ± .05	.22 ± .04	.26 ± .04
6	1	.17 ± .04	0,20 ± .04	.20 ± .04	.30 ± .04
7	.17 ± .04	1	0,25 ± .05	.12 ± .05	.21 ± .04
8	.20 ± .04	.25 ± .05	1	.30 ± .03	.21 ± .04
9	.20 ± .04	.12 ± .05	0,30 ± .04	1	.32 ± .04
10	.30 ± .04	.21 ± .04	0,21 ± .04	.37 ± .04	1

Taip pat buvo įvertinta pirštų atpažinimo kokybė naudojant vieną ir du pirštus (7-ta lentelė pateikia atpažinimo pagal vieną pirštą kokybę, o 8-a pateikia atpažinimo pagal du pirštus kokybę).

7 lentelė: Vieno piršto atpažinimo klaidingų atmetimų procentas kai klaidingų atpažinimų procentas yra nulis (Zero FAR) ir vienodų klaidų procentas (EER).

Piršto pozicija	Zero FAR[%]	EER[%]
1	1.67	0.64
2	0.56	0.31
3	4.17	1.80
4	6.76	2.51
5	5	1.72
6	2.78	1.94
7	1.39	0.86
8	4.35	1.63
9	7.96	3.15
10	5.93	2.50

8 lentelė: Dviejų pirštų atpažinimo klaidingų atmetimų procentas kai klaidingų atpažinimų procentas yra nulis (Zero FAR) ir vienodų klaidų procentas (EER).

Pirštai	Zero FAR[%]	EER[%]
1+6	0.09	0.08
2+7	0	0
3+8	0.46	0.17
4+9	0.83	0.18
5+10	0.28	0.11

Lentelės parodo, kad atpažinimo kokybė naudojant du pirštus tampa dešimt ir daugiau kartų geresnė.

5.3.2 Pirštų ir balsų apjungimas

Balsų atpažinimo algoritmo kokybė buvo įvertinta naudojant šiuos scenarijus:

- fiksuotos frazės (F) lyginamos su fiksuotomis frazėmis (F su F);
- ne fiksuotos frazės (NF) lyginamos su ne fiksuotomis frazėmis (NF su NF);
- ne fiksuotos frazės lyginamos su fiksuotomis frazėmis (NF su F);
- fiksuotos frazės (F) lyginamos su ne fiksuotomis frazėmis (F su NF).

9-ta lentelė parodo koks buvo klaidingų atmetimų procentas, kai klaidingų atpažinimų procentas yra 0 (vidurinis stulpelis) ir koks buvo lygių klaidų procentas (EER).

9 lentelė: Klaidingų atmetimų lygis, kai klaidingų atpažinimų lygis yra 0 (Zero FAR) ir vienodas klaidų lygių procentas (EER) asmens atpažinimo pagal balsą algoritmo.

Scenarijus	Zero FAR[%]	EER[%]
NF su NF	0	0
F su F	0	0
NF su F	78.26	10
F su NF	47.82	9.75

Iš lentelės galima padaryti išvadą, kad kai to paties tipo frazės yra naudojamos, atpažinimas yra idealus (algoritmas klaidų nedaro), bet kai frazių tipas yra skirtingas, kokybė yra ne tokia gera kaip pirštų atspaudų algoritmo.

Tam, kad būtų ištirtas pirštų atspaudų ir balso įrašų panašumų nepriklausomumas, buvo įvertinta koreliacija tarp pirštų atspaudų panašumų ir balso įrašų panašumų, remiantis kuria buvo padaryta išvada, kad pirštai ir balsas yra nepriklausomi ir jų apjungimas galėtų pagerinti atpažinimą.

Buvo įvertinta apjungto algoritmo kokybė – 10-ta ir 11-ta lentelės pateikia algoritmų kokybės palyginimą pagal skirtingus kriterijus.

10 lentelė Trijų scenarijų palyginimas (klaidingų atmetimų procentas, kai klaidingų atpažinimų procentas yra nulis).

Vienas Pirštas		Du Pirštai		Pirštas ir balsas	
1	1.67	1+6	0.09	1	0%
2	0.56	2+7	0	2	0%
3	4.17	3+8	0.46	3	0.22%
4	6.76	4+9	0.83	4	1.89%
5	5	5+10	0.28	5	0.67%
6	2.78	6+1	0.09	6	1.11%
7	1.39	7+2	0	7	0%
8	4.35	8+3	0.46	8	0.56%
9	7.96	9+4	0.83	9	1%
10	5.93	10+5	0.28	10	0.33%

11 lentelė: Trijų scenarijų palyginimas (lygių klaidų procentas).

Vienas Pirštas		Du Pirštai		Pirštas ir balsas	
1	0.64	1+6	0.08	1	0%
2	0.31	2+7	0	2	0%
3	1.80	3+8	0.17	3	0.11%
4	2.51	4+9	0.18	4	0.43%
5	1.72	5+10	0.11	5	0.25%
6	1.94	6+1	0.08	6	0.45%
7	0.86	7+2	0	7	0%
8	1.63	8+3	0.17	8	0.23%
9	3.15	9+4	0.18	9	0.22%
10	2.50	10+5	0.11	10	0.21%

Iš lentelių galima matyti, kad pirštų ir balso sujungimas veikia geriau negu vieno piršto atpažinimas nepriklausomai nuo to, koks įvertinimo kriterijus naudojamas, ir netgi kartais geriau negu dviejų pirštų atpažinimas.

6 Išvados

1. Parodyta, kad galima sintezuoti pirštų atspaudus su iš anksto nustatytais savybėmis ir požymiais. Įrodyta, kad gautų sintezuotų pirštų atspaudai nesiskiria nuo imituojamų tikrų pirštų (požymių prasme sintezuotų ir tikrų pirštų panašumas siekia 90%). Pritaikius adaptyvų filtravimą kelis kartus buvo pagreitintas pirštų sintezavimo metodas. Didelis sintezavimo greitis leido per trumpą laiką sukurti 100 tūkstančių sintezuotų pirštų atspaudų duomenų bazę, kuri buvo panaudota testuojant vieno stambaus biometrinio projekto sistemos patikimumą ir greitį.
Rasta nauja pirštų atspaudų skenerių ir požymių išskyrimo algoritmų kokybės įvertinimo metodika, naudojant kurią gaunami rezultatai dera su pirštų atspaudų ekspertų skenerių kokybės įvertinimais.
2. Rasti sprendimai kaip pagerinti deformuotų pirštų atspaudų palyginimo kokybę (iki 4-ių kartų su $p < 0.0005$). Atsisakius tradiciškai naudojamo pirštų atspaudų registracijos žingsnio, sumažėjo algoritmo parametru

kiekis ir nuo likusių parametru rasta pirštų atspaudų palyginimo metrika priklauso tolygiai. Algoritmo sudėtingumas neviršija $O(\max(N, M)^3)$, kur N ir M yra lyginamų pirštų požymių skaičius.

3. Rasti nauji balso požymiai gauti iš kalbos kadro kompleksinio spektro fazės. Pasiūlyta ir pagrįsta bendros informacijos kiekiu paremta asimetrinė požymių palyginimo metrika. Naudojant viešai prieinamą balso duomenų bazę buvo pasiektas 0.042 lygių klaidų procentas, kas ženkliai geriau nei 2.32% – 8.8% gaunamų tradiciniais Gauso mišinių, vektorinio kvantavimo ir formančių ir antiformančių metodais.
4. Surinkta originali pirštų atspaudų ir balso įrašų duomenų bazė. Ja naudojantis įrodyta, kad pirštų atspaudų ir balso įrašų panašumai statistiškai nekoreliuoja; gauta koreliacijos reikšmė yra $-0.1100 \pm .251$, kas leidžia teigti, kad nulinė hipotezė apie panašumo reikšmių nekoreliavimą yra neatmestina. Atlikus analogiškus tyrimus dviejų pirštų atspaudų panašumams, gautos koreliacijų reikšmės skirtingų pirštų poroms kinta nuo 0.06 ± 0.05 iki 0.44 ± 0.04 , kas įrodo statistiškai patikimą teigiamą dviejų pirštų panašumų koreliaciją visų galimų 45-ių pirštų porų atvejais. Parodyta, kad pirštų atspaudų ir balso įrašų multibiometrika pagerina identifikavimo kokybę apie 10 kartų, kas maždaug atitinka dviejų pirštų atspaudų multibiometrikos kokybę.

Apie autorių

Andrej Kisel gimė 1981 m. vasario 10 d. Vilniuje.

1999 baigė Vilniaus Naujamiesčio vidurinę mokyklą (8-oji vid.).

1999-2005 įgijo informatikos bakalauro ir magistro kvalifikacinius laipsnius Vilniaus Universiteto Matematikos ir Informatikos fakultete.

2005-2009 – Vilniaus Universiteto doktorantas.

Nuo 2002 metų dirba biometrinių algoritmų kūrėju UAB „Neurotechnology“.

Person Identification by Fingerprints and Voice

Research Area

Biometric technologies are becoming very common in everyday life. The use of distinctive and unique features that can identify a person (such as fingerprints, palm prints, face, iris or voice) makes it possible to determine an identity of a person in easy and convenient way. Many countries integrate biometric features into the passports and identity cards. Biometrics is used at companies to track working time; identity is checked during elections to prevent multiple voting, at banks and in prisons to enforce security.

The use of biometric technology grows every day and is forecasted to grow in coming years what makes biometrics a very attractive branch of science. The research area of this work is fingerprint and voice biometrics: fingerprint image synthesis for fingerprint extraction algorithm performance evaluation, distortion tolerant fingerprint matching, speaker recognition and multibiometrics.

Problem Relevance

It is easy to notice that biometric technologies are spreading across the world. Even low cost notebooks and mobile phones have integrated fingerprint scanners and users can log on with fingerprint instead of password. Integrated webcams are used to identify a person by face, and microphones are used to provide access to the system by the voice. All these technologies provide faster and reliable access to data, bank account or computer than password, because passwords can be stolen, forgotten, lost or unlocked by specific software. These are the reasons why many universities, companies and institutions invest time and money in research and development of biometric algorithms.

Several international competitions were arranged to compare different biometric algorithms and track progress in biometric research: FVC (FVC 2000, FVC 2002, FVC 2004, FVC 2006 and FVC ongoing), NIST (National Institute of Standards and Technology) MINEX (MINEX, MINEX II and Ongoing MINEX) and PIV for fingerprints;

NIST Face Recognition Vendor Tests (FRVT) for faces; NIST Speaker Recognition Evaluation (SRE) for voice biometrics; NIST Iris Challenge Evaluation (ICE), Independent Testing of Iris Recognition Technology (ITIRT) for irises are the largest and most known biometric competitions. These competitions show that in spite the progress in such aspects as reliability, speed and interoperability is impressive, there are many difficult problems left to overcome.

All biometric technologies are dependent on input quality: If obtained fingerprint image is noisy, low contrast or deformed; recorded voice phrase is of low volume or very short, iris image is obstructed by eyelids, reflections or glasses, face image is acquired in poor lightning conditions or using low quality camera, the task of verification becomes more difficult. The main challenge of modern biometric algorithms is to overcome these difficult conditions and extract as much data as possible. Innovative methods help algorithm developers better understand the weaknesses of their algorithms and address them. Algorithm developers have to take into consideration that when the popularity of biometric technology increase, requirements to algorithm accuracy also increase. Error rate of one percent may be suitable for a small company using time attendance system based on biometrics, but will make a lot of problems to a bank with millions of customers or during elections to prevent multiple voting.

This work is about fingerprint and voice biometrics. Fingerprint biometrics is the most popular biometrics: fingerprint scanners are cheap, easy to use and the process of verification is fast.

Voice biometrics is the most available biometrics, because no additional hardware is needed. Most computers have audio interface with possibility to plug microphone, microphones are integrated into webcams, headphones and mobile phones.

Research Objects

The thesis research objects are: performance evaluation of fingerprint extraction algorithm using fingerprint synthesis, fingerprint matching method that is able to match deformed fingerprints, person identification using voice and fusion of both biometrics.

The Objectives and Tasks of the Research

The aim of the research was to complexly analyze research area and address difficult problems. In the first part of the work fingerprint extraction algorithm development problems are analyzed and fingerprint image synthesis is suggested to overcome that problems. In the second part of the work fingerprint matching algorithm problems are analyzed and new matching algorithm is proposed to deal with them. New person identification by voice method is addressed in the third part of the work and multibiometrics using fingerprints and voice is proposed to increase identification accuracy.

Scientific Novelty

The new method of fingerprint image synthesis is introduced in first chapter. Differently from already existing synthesis methods it can generate fingerprint images with predefined features. Such images with known characteristics allow evaluating the performance of fingerprint extraction algorithm independently from fingerprint matching algorithm. A new practical application for synthetic fingerprints is suggested: they can be used to estimate the quality of images in a given database or the quality of a fingerprint scanner.

New fingerprint matching algorithm that is described in the second chapter does not perform fingerprint registration (evaluation of rotation and translation) and is capable to match fingerprints with elastic deformations.

Multibiometrics using new person identification by voice method and new fingerprint matching method is described in next chapters. The performance was analyzed using specially prepared multibiometric database.

This work is the first attempt to prove that there is no correlation observed between similarities based on fingerprints and similarities based on voice. Such independence of two biometrics means that they can be successfully combined into multibiometrics.

Practical Importance of the Work

Methods described in this work can be used to solve many difficult tasks.

Fingerprint image synthesis (chapter 2) can be used to generate large fingerprint databases, to evaluate the performance of fingerprint extraction algorithm. Since it is possible to generate a fingerprint image with predefined properties and features, it becomes easy to evaluate such properties of fingerprint extraction algorithm as stability to noise and accuracy of extracted features.

New fingerprint matching method (chapter 3) allows accurate matching of plain and rolled fingerprints with elastic deformations that are common in rolled fingerprints and sometimes occur in plain fingerprints.

Multibiometrics using fingerprints and voice (chapters 4 and 5) can provide more flexible and accurate way of person identification.

Approval of Research Results

Research results were published in valuable international journal Informatica. The conference papers were presented and an oral presentation in INFORMATION TECHNOLOGIES (IT2010) conference was done.

Defended propositions

1. New fingerprint image synthesis method can generate fingerprints with predefined features. Such fingerprints can be used to test and develop biometric systems.
2. Synthetic fingerprints appear very realistic and can be used as a valid substitute for real fingerprint for testing fingerprint recognition algorithms.
3. Fingerprint image synthesis uses iterative convolution with large kernel that is a very time consuming operation. An optimization that speeds up synthesis process several times was proposed.
4. A method to evaluate the performance of fingerprint extraction algorithm using synthetic fingerprints can be used evaluate extractor's performance.
5. Fingerprint matching method that does not perform fingerprint registration and is able to match deformed plain and rolled fingerprints with better accuracy.

6. New speaker identification method outperforms traditional speaker identification methods.
7. Since fingerprint and voice similarities do not correlate much, multibiometric using both fingerprints and voice can further increase identification accuracy.

Publications

International journals which are included into the International Master Journal List (ISI):

1. Andrej Kisel, Alexej Kochetkov, Justas Kranauskas: Fingerprint Minutiae Matching without Global Alignment Using Local Structures. *INFORMATICA*, 2008, Vol. 19, No. 1, pp.31-44 ISSN 0868-4952
2. Algirdas Bastys Andrej Kisel, Bernardas Salna: The Use of Group Delay Features of Linear Prediction Model for Speaker Recognition. *INFORMATICA*, 2010, Vol. 21, No. 1, pp.1-12 ISSN 0868-4952

International journals which are included in the Scientific Master Journal Proceeding List (ISI):

1. Andrej Kisel: Fast Fingerprint Image Synthesis. Proceedings of 16th International Conference on Information and Software Technologies. April 21st - 23rd 2010, Kaunas University of Technology, Lithuania, ISSN 2029-0063 pp. 107-115.

Journal submissions under review:

1. Andrej Kisel: Multibiometrics using fingerprints and voice. Information technology and control, Kaunas University of Technology.

Thesis Structure

The thesis consists of 3 main parts: fingerprint biometrics (chapters 2 and 3), voice biometrics (chapter4) and multibiometrics (chapter 5).

The 2nd chapter describes fast fingerprint image synthesis method that can be used to create large fingerprint databases and to evaluate the performance of fingerprint extraction methods.

The 3d chapter is devoted to a fingerprint matching method that is robust to deformations and does not perform fingerprint alignment.

The 4th chapter introduces the use of group delay features of linear prediction model for speaker recognition.

The 5th chapter presents multibiometrics using fingerprints and voice.

The 6th chapter completes thesis with brief summary and conclusions. At the end of the work a bibliography list is presented.

Summary and Conclusions

Fingerprint image synthesis method and two biometric algorithms (person identification by fingerprints and speaker identification by voice) are described in detail. The speed of an earlier known synthesis algorithm has increased more than three times. A new practical application of synthetic fingerprints to estimate the quality of a fingerprint image by comparing it to an ideal noise free synthetic fingerprint was found. Problems in fingerprint matching are analyzed, and necessity of registration (evaluation of rotation and translation) is discussed.

New Fingerprint matching algorithm is designed to match deformed fingerprints. It consists of simple and intuitive steps. The proposed implementation of the steps is straightforward and flexible, does not use registration, and therefore is capable of matching deformed fingerprints. The advantage of proposed fingerprint matching algorithm is validated on a set of popular publically available fingerprint databases. One-tailed t-test showed improvement in comparing with NIST matcher with $p < 0.0005$.

Speaker recognition algorithm uses phase of transfer function defined by linear prediction model for derivation of features of utterances. The features represent extremes of the group delay of the LP model. Similarity measure between two speech utterances was defined as a mutual information of the two group delay feature distributions.

The proposed voice recognition algorithm showed up to ten times better performance in comparison with commonly accepted on Gaussian Mixture Model based voice recognition algorithm.

Multibiometrics using fingerprints and voice was also presented. Experiments showed significant decrease in EER and $FRR@FAR = 0\%$ when two fingerprints or fingerprint and voice are used.

Voice and fingerprint fusion performs almost as well as two fingerprints fusion.

This study is the first to demonstrate that similarities between voice samples and similarities between fingerprints do not have any correlation what makes them ideal for multibiometric.

About the Author

Andrej Kisel received both the Bachelor and Master degrees in Informatics at Vilnius University, Faculty of Mathematics and Informatics in 2003 and 2005.

In 2005-2009 – Phd student at Vilnius University.

From 2002 works at “Neurotechnology“ as a biometric algorithms developer.