VYSOKÁ ŠKOLA BÁŇSKÁ–TECHNICKÁ UNIVERZITA OSTRAVA
VŠB–TECHNICAL UNIVERSITY OF OSTRAVA

FAKULTA ELEKTROTECHNIKY A INFORMATIKY
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

KATEDRA TELEKOMUNIKAČNÍ TECHNIKY
DEPARTMENT OF TELECOMMUNICATIONS

Využití umělé inteligence pro vícenásobnou bezkontaktní biometrickou autentizaci

Multimodal biometric contactless authentication using the artificial intelligence

REFERÁT K DIZERTAČNÍ PRÁCE

AUTOR PRÁCE
AUTHOR — Ing. Jaromír Továrek

VEDOUCÍ PRÁCE
SUPERVISOR — doc. Ing. Miroslav Vozňák, Ph.D.

OSTRAVA, 2017

Dizertační práce byla napsána během prezenčního doktorského studia na Katedře telekomunikační techniky, Fakultě elektrotechniky a informatiky, Vysoké školy báňské – Technické univerzitě Ostrava.

Student: Ing. Jaromír Továrek
Katedra telekomunikační techniky
Fakulta elektrotechniky a informatiky
Vysoká škola báňská - Technická univerzita Ostrava
17. listopadu 15/2172, 708 00 Ostrava - Poruba
jaromir.tovarek@vsb.cz

Školitel: doc. Ing. Miroslav Vozňák, Ph.D.
Katedra telekomunikační techniky
Fakulta elektrotechniky a informatiky
Vysoká škola báňská - Technická univerzita Ostrava
17. listopadu 15/2172, 708 00 Ostrava - Poruba
miroslav.voznak@vsb.cz

Oponenti: _____

_____

_____

OBSAH

1 ÚVOD A MOTIVACE

Problematiku ověření identity osob můžeme rozdělit do tří skupin podle použité metody autentizace. První skupinu tvoří metody založené na prokázání identity pomocí vlastnictví (identifikační doklady, karty, čipy), druhá skupina využívá k prokázání identity znalostí (hesla, identifikační čísla), třetí skupinu představují metody využívající k ověření identity měřitelné biometrické charakteristiky (otisk prstu, geometrie obličeje, duhovka oka, sítnice oka, geometrie ruky, hlas atd.) Z principu jednotlivých metod můžeme odvodit největší nevýhody. U prokazování identity vlastnictvím je největší bezpečnostní riziko dáno možností odcizení nebo napodobení tokenu. Ověření identity pomocí znalostí nese riziko v odpozorování, uhodnutí případně odvození hesla nebo identifikačního čísla. V praxi se ke snížení těchto bezpečnostních rizik používá kombinace obou přístupů. Přesto se jako nejbezpečnější způsob jeví ověření identity pomocí specifické biometrické charakteristiky člověka.

1.1 Motivace

Přes vysokou bezpečnost ověřování identity pomocí biometrických charakteristik se i u této identifikace začínají objevovat první pokusy pachatelů o změnu nebo napodobení biometrické charakteristiky. Příkladem mohou být pokusy o změnu otisků prstů, plastické operace v oblasti tváře, napodobování hlasu apod. Pokud by se pachateli podařilo vytvořit napodobeninu některé z biometrických charakteristik, tak by to mohlo znamenat výraznou hrozbu pro všechny systémy založené na biometrické identifikaci. Z tohoto důvodu je snaha stálého zvyšování bezpečnosti biometrických systémů, tak aby k podobným pokusům pokud možno vůbec nedocházelo nebo se jim alespoň dalo v dostatečné míře čelit.

Jedním ze způsobů jak dosáhnout vyšší bezpečnosti biometrických aplikací je použití vícenásobné biometrie (Multiple Biometrics). Jedná se o kombinaci více biometrických charakteristik pro verifikaci v jednom systému. V současné době jsou nejpoužívanější kombinace otisk prstů - geometrie obličeje a geometrie oční duhovky - hlas. Lze očekávat, že v brzké době přibudou i další kombinace biometrických charakteristik.

Z tohoto důvodu je tato práce zaměřena na návrh vícenásobného biometrického autentizačního systému, který je založen na autentizace hlasem a autentizaci geometrií obličeje. Tyto biometrické metody byly vybrány s ohledem na jejich vlastnosti. Mezi tyto vlastnosti patří: přijatelnost pro uživatele (snímání biometrických dat je přirozené - promluva, fotka), dostatečná přesnost a jednoduchost snímání (stačí pouze fotoaparát a mikrofon). Jednoduchost snímání umožňuje využití systému v podstatě na jakémkoliv zařízení a v různých oblastech. Další výhodou této kombinace metod je to, že pohyb úst při verifikaci hlasem může být využit jako "Test živosti"(data pochází od skutečného uživatele) autentizující se osoby. Takto vzniklý systém může být nasazen jako vysoce bezpečný autentizační systém pro vstup do objektu nebo pro přihlášení k různým zařízením.

1.2 Cíle dizertační práce

Pro vypracování dizertační práce byly stanoveny dílčí cíle, které byly schváleny komisí při státní doktorské zkoušce. Tyto cíle jsou z důvodu větší přehlednosti zformulovány do 4 hlavních bodů:

- Návrh biometrického autentizačního systému založeného na verifikaci hlasem s využitím umělé inteligence.
- Návrh biometrické autentizace založené na rozpoznávání pomocí geometrie obličeje s využitím pokročilých metod vyhodnocování extrahovaných příznaků.
- Vytvoření komplexního vícenásobného biometrického autentizačního systému využívajícího verifikaci hlasem a verifikaci geometrií obličeje.
- Experimentální ověření funkčnosti navrženého vícenásobného autentizačního systému a porovnání přesnosti s aktuálně používanými systémy biometrické autentizace.

Zpracování jednotlivých bodů vede k naplnění hlavního dizertabilního cíle, kterým je vytvoření komplexního vícenásobného biometrického autentizačního systému, který je určen pro bezkontaktní autentizaci a může být využit jako přístupový systém v budovách či pro ověřování přístupu osob k různým zařízením.

2 NÁVRH HLASOVÉHO AUTENTIZAČNÍHO SYSTÉMU

Tato kapitola popisuje návrh biometrického autentizačního systému, který využívá k ověření totožnosti uživatele jeho hlas.

2.1 Princip návrhu hlasového autentizačního systému

Návrh systému probíhal jak z pohledu vhodných parametrů, tak z pohledu vhodného klasifikátoru. V rámci návrhu byly porovnány následující parametry: MFCC, delta MFCC, delta-delta MFCC, LPC a jejich kombinace. Bylo použito 13 koeficientů MFCC, ze kterých byly odvozeny jejich deriváty (delta MFCC a delta-delta MFCC) a 13 LPC koeficientů. Tyto parametry byly následně klasifikovány pomocí dvou klasifikátorů (MLNN a SVM). Struktura MLNN se skládala ze tří vrstev. Počet neuronů vstupní vrstvy odpovídal délce příznakového vektoru, počet neuronů ve skryté vrstvě byl nastaven na 10 a výstupní vrstva byla tvořena dvěma neurony. Jako aktivační funkce neuronů byla použita funkce sigmoid se strmostí 0.5. Struktura a nastavení parametrů neuronové sítě bylo provedeno na základě poznatků z předchozího výzkumu [Tov01]. Pro klasifikaci pomocí SVM byla použita polynomiální funkce konkrétně kubická. Volba signifikantních parametrů a klasifikátorů byla provedena na základě provedené rešerše s přihlédnutím ke konkrétním požadavkům na systém. Tyto požadavky vychází ze skutečnosti, že systém má být součástí komplexního vícenásobného biometrického systému.

Prvním krokem hlasové autentizace je předzpracování řečového signálu, které se obvykle skládá ze čtyř fází: odstranění stejnosměrné složky, preemfáze, segmentace (20 ms), váhování oknem (Hammingovo okno). V dizertační práce byl tento krok rozšířen o jednu fázi, která se nazývá odstranění nízkoenergetických segmentů. Druhým krokem je extrakce signifikantních parametrů po které je aplikována min-max normalizace. Po klasifikaci jednotlivých segmentů je provedena evaluace segmentů v rámci jedné nahrávky pomocí fúzní metody Majority voting [2]. Výsledky evaluace jsou porovnány s verifikačním prahem a následně je provedeno finální rozhodnutí o autentizaci. Jednotlivé přístupy jsou mezi sebou porovnávány pomocí FAR, FRR, EER, ROC a DET. Výsledky byly získány pro databázi Comtech, která vznikla v rámci dizertační práce. Implementace použitých metod pro návrh systému probíhala v prostředí Matlab.

2.2 Databáze Comtech

V rámci dizertační práce byla vytvořena na Katedře telekomunikační techniky, VŠB-TU Ostrava databáze řečových vzorků s názvem Comtech, která je určena pro návrh a evaluaci algoritmů použitých v systémech rozpoznávání řečníka (textově závislých i textově nezávislých). Jedná se o databázi českých řečových vzorků a tudíž je určena především pro

tuzemské použití. Databáze byla rozdělena do dvou částí. První část je tvořena nahrávkami 18 referenčních řečníků (13 mužů, 5 žen) a skládá se z 1080 foneticky vyvážených vět a 1080 přístupových frází (hesel). Druhá část obsahuje nahrávky neoprávněných uživatelů, takzvaných podvodníků (uživatelé, kteří se snaží o neoprávněný přístup). Tato část se skládá ze 190 foneticky vyvážených vět a 190 přístupových frází. Věk řečníků se pohyboval v rozmezí 25-50 let.

Řečové vzorky byly pořizovány během jednoho měsíce (03/2017). Nahrávání databáze (část referenční řečníci) bylo rozděleno do 6 sezení. Každý z řečníků byl posazen do kanceláře, kde byl umístěn VoIP telefon (Grandstream GXP2140). Uživatel zadal volbu pro vytočení IVR, které bylo nakonfigurováno pro provedení uživatele nahrávacím procesem. V první části je uživatel vyzván pro zadání svého identifikačního čísla, tímto číslem jsou označeny všechny nahrávky daného uživatele. Následně uživatel pronesl 10 vět a 10 přístupových frází, které byly zaznamenány. Tento proces se opakoval pro všechny referenční řečníky a všechna sezení. Část databáze tvořena neoprávněnými uživateli/podvodníky byla nahrána stejným způsoben, jen s tím rozdílem, že nahrávání probíhalo pouze v jednom sezení. Řečový signál byl zaznamenán ve WAV formátu se vzorkovací frekvencí 8kHz s 32 bitovým rozlišením.

Jako promluvy jednotlivých řečníků byly zvoleny foneticky vyvážené věty a krátké dvojslovné přístupové fráze. Každé sezení obsahovalo 10 různých vět. Doba trvání každé z vět byla přibližně 10 sekund. Jednotlivé věty byly vybírány z elektronických novin. Přístupová fráze byla stejná během všech sezení a pro všechny řečníky. Doba trvání přístupové fráze byla přibližně 2 sekundy a její znění bylo "Jaromír Továrek".

V rámci dizertační práce byly využity vzorky přístupových frází, které jsou určeny především pro textově závislé rozpoznávání řečníka (textově závislou autentizaci hlasem).

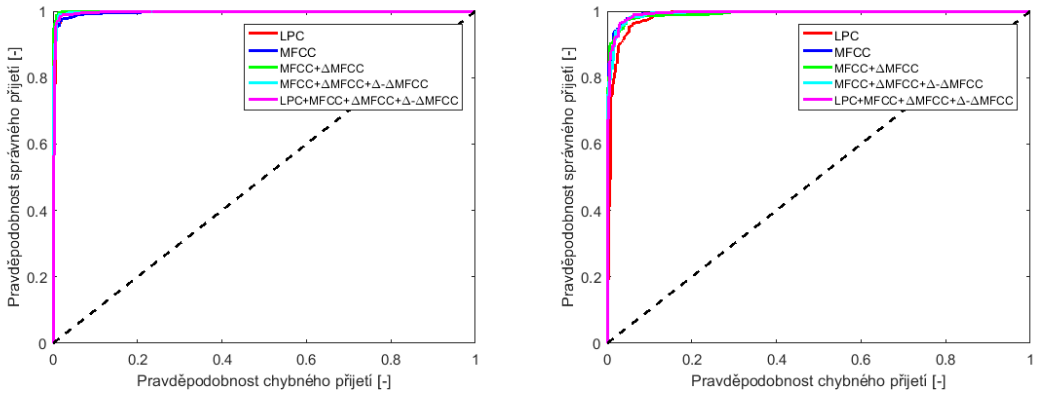
2.3 Experimentální výsledky

Pro každého z 18 referenčních řečníků byly natrénovány dva klasifikátory (SVM model a MLNN klasifikátor). Oba klasifikátory byly natrénovány k rozpoznávání mezi dvěma třídami (referenční uživatel, podvodník). Třída referenční uživatel byla trénována pomocí 40 přístupových frází (4 sezení databáze). Zbývajících 20 nahrávek bylo použito pro testování referenčního uživatele. Třída podvodníků byla trénována jako univerzální model pozadí. Tato třída byla trénována 40 nahrávkami od zbývajících 17 referenčních řečníků. Poměr trénovacích dat pro obě třídy byl přibližně stejný. Jako testovací data pro třídu podvodníků byly použity přístupové fráze z druhé části databáze (19 nahrávek od 19 podvodníků). Tyto nahrávky nepochází od referenčních řečníků, to znamená, že model pozadí není trénován nahrávkami od těchto uživatelů. Tento proces trénování klasifikátorů se opakoval pro různé řečové parametry a jejich kombinace. Výsledky pro oba klasifikátory a všechny typy parametrů jsou uvedeny v tabulce 2.1. Tabulka obsahuje hodnoty FAR,

FRR pro nastavený rozhodovací práh 0.5 (50%) a hodnotu EER. ROC křivky pro oba klasifikátory jsou zobrazeny na obrázku 2.1.

Tab. 2.1: Výsledky dosažené pomocí SVM a MLNN klasifikátorů pro použité řečové parametry (rozhodovací práh 50%).

Parametry	SVM			MLNN		
	FAR [%]	FRR [%]	EER [%]	FAR [%]	FRR [%]	EER [%]
LPC	5	0.3	1.7	13	1.4	5.3
MFCC	5	1.3	2.3	11	0.8	3.8
MFCC + ΔMFCC	2.3	0.27	**0.85**	7.8	1.6	4.3
MFCC + ΔMFCC + Δ-ΔMFCC	3.5	0.55	1.14	8.5	1.6	3.9
LPC + MFCC + ΔMFCC + Δ-ΔMFCC	2.9	1	1.7	6	1	**3.3**



(a) ROC křivky - SVM klasifikátor

(b) ROC křivky - MLNN klasifikátor

Obr. 2.1: ROC křivky - hlasová autentizace

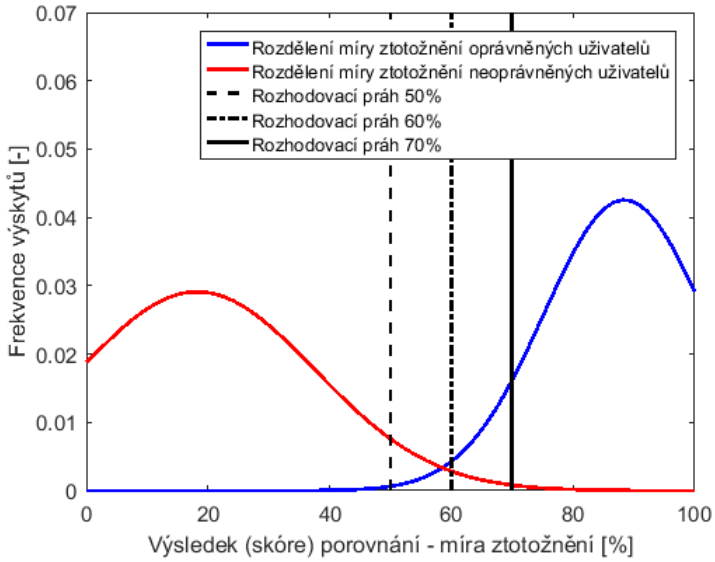
Nejlepších výsledků bylo dosaženo pro příznakový vektor tvořený parametry MFCC a delta MFCC při použití SVM klasifikátoru. Hodnoty FAR a FRR byly v tomto případě 2.3% a 0.27% pro rozhodovací práh nastavený na 50%. Tyto výsledky odpovídají přesnosti systému 98.7%. Pro tento příznakový vektor byla hodnota EER 0.85%, což odpovídalo nastaveného rozhodovacímu prahu 55%. Pro klasifikátor MLNN bylo dosaženo nejnižší hodnoty EER (3.3%) s použitím příznakového vektoru složeného ze všech použitých parametrů (LPC + MFCC + delta MFCC + delta-delta MFCC). Z pohledu přesnosti klasifikátorů dosáhl lepších výsledků klasifikátor SVM a to pro všechny varianty příznakových vektorů. Z tohoto důvodu jsou níže uvedeny podrobné výsledky pouze pro SVM klasifikátor s využitím příznakového vektoru složeného z MFCC a delta MFCC.

Na obrázku 2.2 je zobrazen histogram rozdělení skóre oprávněných a neoprávněných uživatelů s vyznačenými hodnotami rozhodovacího prahu. Z pohledu verifikace (autentizace) je snahou dosáhnout co nejnižší hodnoty FAR v ideálním případě hodnoty nulové. Aby bylo dosaženo snížení hodnoty FAR je potřeba zvýšit hodnotu rozhodovacího prahu.

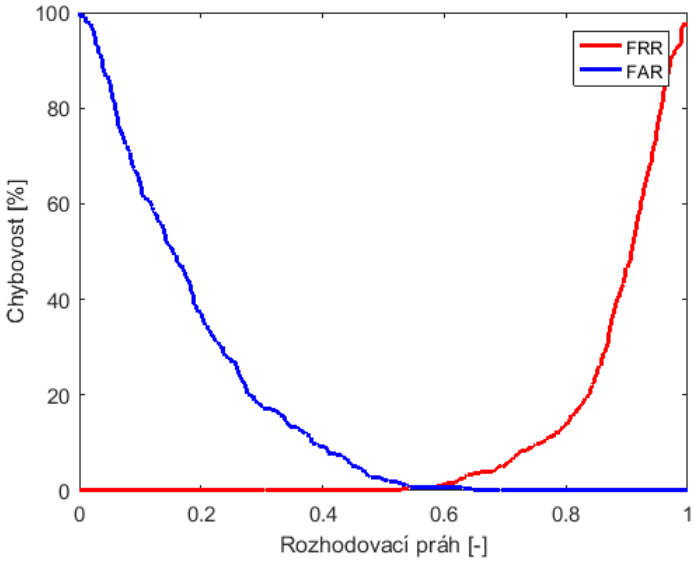
Obrázek 2.3 ukazuje závislost hodnot FAR a FRR na velikosti rozhodovacího prahu. Nulové hodnoty FAR bylo dosaženo po nastavení rozhodovacího prahu na hodnotu 65%. Tabulka 2.2 představuje kontingenční tabulku pro rozhodovací práh 50%. Jak je vidět na obrázku 2.3, hodnota FAR klesá se zvyšujícím se rozhodovacím prahem, naopak hodnota FRR v tomto případě roste. Obrázek 2.4 zobrazuje DET křivku s vyznačeným bodem EER.

Tab. 2.2: Kontingenční tabulka - SVM klasifikátor (MFCC + delta MFCC), rozhodovací práh 50%.

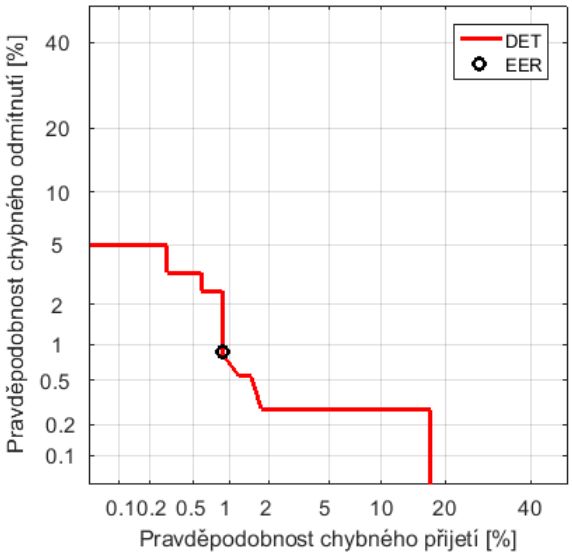
Skutečný výstup	*Oprávněný uživatel*	359 51.1%	8 1.2%	97.8%
	Neoprávněný uživatel	1 0.1%	334 47.6%	99.7%
		99.7%	97.7%	**98.7%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		



Obr. 2.2: Histogram rozdělení skóre oprávněných a neoprávněných uživatelů s vyznačenými hodnotami rozhodovacího prahu - hlasová autentizace.



Obr. 2.3: Závislost hodnot FAR a FRR na velikosti rozhodovacího prahu - hlasová autentizace.



Obr. 2.4: DET křivka - hlasová autentizace.

2.4 Zhodnocení dosažených výsledků

Z výše uvedených výsledků vyplývá, že z pohledu klasifikátorů je vhodnější pro hlasovou autentizaci využít klasifikátor SVM, který potřebuje pro dosažení nejlepších výsledků pouze minimální množství parametrů, na rozdíl od klasifikátoru MLNN, který pro dosažení minimální chybovosti vyžaduje všechny testované parametry. Tato skutečnost je dána vlastnostmi jednotlivých klasifikátorů. Při trénování klasifikátoru MLNN dochází

k promíchání trénovacích dat tak, aby nedocházelo k přetrénování klasifikátoru. Promíchání trénovacích dat způsobí porušení časové sekvence v rámci jedné nahrávky a tudíž klasifikátor vyžaduje i parametry popisující časové změny mezi jednotlivými segmenty (delta-delta MFCC). Při trénování SVM klasifikátoru nedochází k promíchávání trénovacích dat a tudíž není potřeba při trénování využít parametry popisující časové změny mezi segmenty.

Výstupem této kapitoly je tedy textově závislý hlasový autentizační systém postavený na klasifikátoru SVM s využitím příznakového vektoru složeného z MFCC a delta MFCC parametrů. Takto navržený systém produkuje nejnižší počet chyb pro databázi Comtech. Systém je následně využit jako část vícenásobného biometrického autentizačního systému popsaného v kapitole 4.

3 NÁVRH AUTENTIZAČNÍHO SYSTÉMU ZALOŽENÉHO NA OVĚŘENÍ IDENTITY POMOCÍ GEOMETRIE OBLIČEJE

V této kapitole je popsán návrh unimodálního biometrického autentizačního systému založeného na ověření totožnosti uživatele pomocí geometrie tváře.

3.1 Princip návrhu biometrického systému pro autentizaci geometrií obličeje

Princip návrhu systému byl obdobný jako v případě systému navrženého v kapitole 2. Stejně jako v předchozím případě byl návrh proveden jak z pohledu nejvhodnějších parametrů, tak z pohledu vhodného klasifikátoru. Z pohledu parametrů byly porovnávány tyto parametry: LBP, HOG a jejich kombinace. Pro extrakci HOG parametrů byly stanoveny následující podmínky: velikost buňky 8x8 pixelů, počet buněk v bloku 4, překryv buněk mezi sousedními bloky 1, počet binů orientovaného histogramu 9. Velikost okolí v případě extrakce LBP parametrů byla nastavena na 3x3 pixelů. Tyto podmínky byly stanoveny na základě předchozího výzkumu [Tov09]. Parametry byly vyhodnocovány pomocí klasifikátorů MLNN a SVM. Nastavení klasifikátorů MLNN a SVM bylo obdobné jako v kapitole 2. Volba stejných klasifikátorů jako v případě návrhu systému hlasové autentizace byla provedena na základě jejich vhodnosti pro řešení problému binární klasifikace (klasifikace do dvou tříd) a zároveň na snaze zjednodušit finální vícenásobný biometrický autentizační systém.

Prvním krokem autentizace pomocí geometrie obličeje je detekce tváře. Při návrhu systému byla využita detekční metoda Viola-Jones. Druhým krokem je předzpracování obrazu detekované tváře. Tento krok se skládá ze dvou částí: změna velikosti detekované tváře (nastavení fixní velikosti pro všechny obrazy - 120x120 pixelů) a převod obrazu do odstínů šedi (vyžadováno pro správnou extrakci parametrů). Další krokem je samotná extrakce parametrů, která je doplněna o min-max normalizaci. Extrahované parametry jsou následně klasifikovány příslušným klasifikátorem a na základě výsledků klasifikace je provedeno finální rozhodnutí a autentizaci. Jednotlivé přístupy jsou stejně jako v předchozí kapitole porovnávány pomocí FAR, FRR, ROC a DET. Výsledky byly získány pro databázi AR Face Database [4]. Návrh a implementace jednotlivých metod probíhala v prostředí Matlab.

3.2 AR Face Database

Databáze slouží pro návrh a evaluaci algoritmů používaných v systémech rozpoznávání tváří. Databáze obsahuje přes 4000 barevných čelních pohledů na tváře od 126 respon-

dentů (70 mužů, 56 žen). Nasnímané obrázky byly pořízeny během dvou sezení. Mezi jednotlivými sezeními byl časový rozestup 14 dní. Během každého sezení bylo pořízeno 13 čelních pohledů na tvář jednotlivých respondentů. Na respondenty nebyly kladeny žádné požadavky ohledně oblečení, účesů, brýlí atd. V rámci snímání byly pořízeny obrazy v za předem definovaných podmínek:

- výraz tváře - neutrální výraz, úsměv, zlost, křik,
- změna osvětlení - osvětlení zleva, osvětlení zprava, osvětlení z obou stran,
- částečné zakrytí tváře - sluneční brýle, šátek přes obličej.

V rámci návrhu systému bylo vybráno 18 referenčních uživatelů (13 mužů, 5 žen) a dalších 10 uživatelů, kteří představují podvodníky. Tento počet byl zvolen s ohledem na počet uživatelů řečové databáze Comtech a následný návrh vícenásobného biometrického systému.

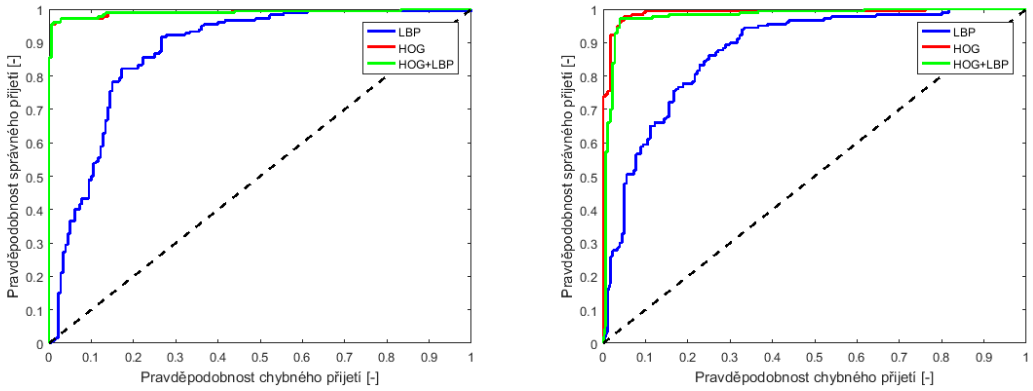
3.3 Experimentální výsledky

Stejně jako v případě autentizace hlasem byly natrénovány klasifikátory SVM a MLNN pro každého z 18 referenčních uživatelů. Klasifikátory byly natrénovány tak, aby rozpoznávaly mezi dvěma třídami (třída referenční uživatel, třída podvodník). Každý z klasifikátorů byl trénován pomocí 32 digitálních obrazů detekovaných tváří. Pro trénování třídy referenčního uživatele bylo použito 16 obrazů tváře daného uživatele (mix obrazů z obou sezení). Jako testovací data pro třídu referenčního uživatele bylo použito 10 zbývajících obrazů. Třída podvodníků byla trénována jako univerzální model pozadí. Pro model pozadí každého referenčního uživatele bylo použito 16 obrazů tváří, které pocházejí od zbývajících referenčních uživatelů. Jako testovací data v tomto případě byla použita od 10 vybraných uživatelů z databáze (tito uživatelé netvoří univerzální model pozadí). Tento proces trénování klasifikátorů se opakoval pro různé obrazové parametry a jejich kombinaci. Validační výsledky obou klasifikátorů pro všechny použité parametry jsou uvedeny v tabulce 3.1. Tabulka obsahuje hodnoty FAR, FRR pro rozhodovací práh 0.5 (50%) a hodnotu EER. ROC křivky pro jednotlivé klasifikátory jsou zobrazeny na obrázku 3.1.

Tab. 3.1: Výsledky dosažené pomocí SVM a MLNN klasifikátorů pro použité obrazové parametry (rozhodovací práh 50%)

Parametry	SVM			MLNN		
	FAR [%]	FRR [%]	EER [%]	FAR [%]	FRR [%]	EER [%]
LBP	35.0	6.6	17.8	26.6	13.9	21.7
HOG	2.7	3.3	**2.8**	6.0	2.2	**3.9**
LBP + HOG	2.7	3.3	**2.8**	7.2	2.7	3.9

Jak je patrné z tabulky 3.1 nejlepších výsledků bylo dosaženo pro obrazové parametry HOG při použití SVM klasifikátoru. Hodnoty FAR a FRR byly pro tuto kombinaci klasifikátoru a parametrů 2.7% a 3.3% pro rozhodovací práh 50%. Systém tedy dosáhl



(a) ROC křivka - SVM klasifikátor

(b) ROC křivka - MLNN klasifikátor

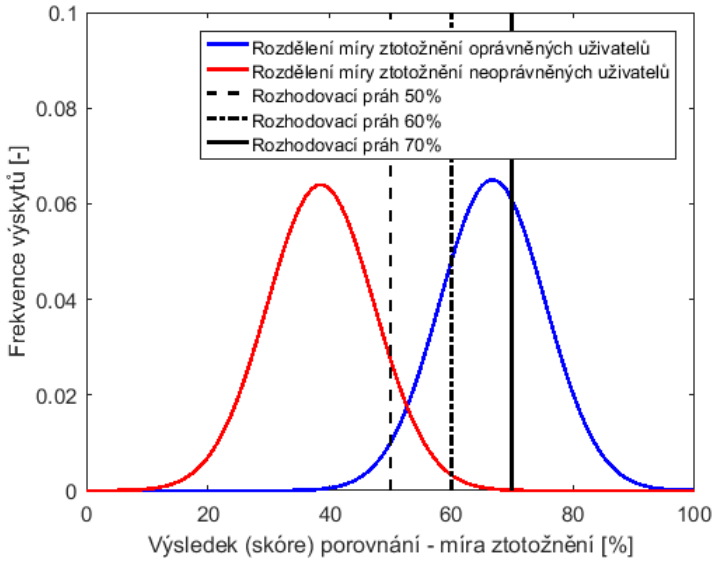
Obr. 3.1: ROC křivky - autentizace tváří

přesnosti 96.9%. Hodnota EER dosáhla v tomto případě minimální hodnoty 2.8%. Stejných výsledků bylo dosaženo také pro kombinaci parametrů HOG a LBP. Vzhledem k tomu, že kombinací parametrů nebylo dosaženo zvýšení přesnosti a snížení chyb, lze tvrdit, že parametry LBP nepřináší žádnou novou informaci popisující rozdíly mezi tvářemi. Pro klasifikátor MLNN bylo dosaženo nejlepších výsledků pro parametry HOG, hodnota EER byla v tomto případě 3.9%. V následující části jsou uvedeny podrobné výsledky pro SVM klasifikátor využívající parametry HOG.

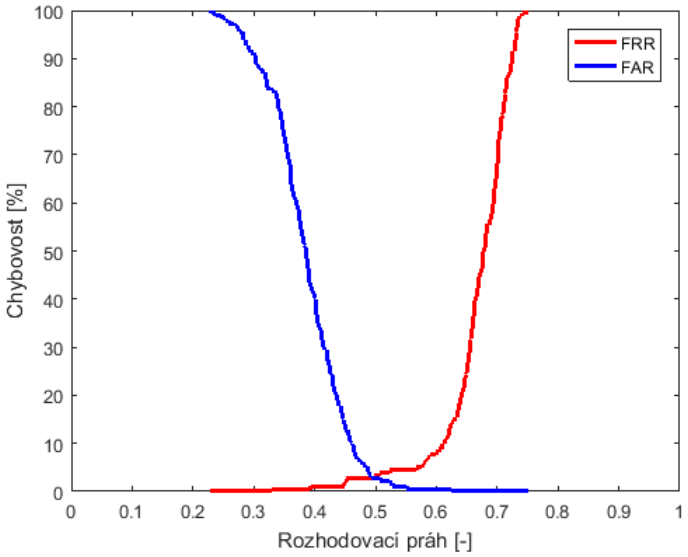
Obrázek 3.2 zobrazuje histogram rozdělení skóre s vyznačenými hodnotami prahu. Výsledky klasifikace a celková přesnost systému pro rozhodovací práh 50% je uvedena v kontingenční tabulce 3.2. Na obrázku 3.3 je zobrazen průběh hodnot FAR a FRR v závislosti na verifikačním prahu. Z grafu je patrné, že zvyšováním verifikačního prahu je sice dosaženo snížení hodnoty FAR, ale na druhou stranu dojde ke zvýšení hodnoty FRR a v důsledku toho klesne i přesnost celého systému. Pro verifikační práh 70% dosáhne systém hodnoty FRR 67.8%, což je v praxi nepoužitelné, protože by byli dva ze tří oprávněných uživatelů odmítnuti. Z tohoto důvodu je nutné volit kompromis při nastavení verifikačního prahu. Obrázek 3.4 zobrazuje DET křivku se znázorněným bodem EER.

Tab. 3.2: Kontingenční tabulka - SVM klasifikátor (parametry HOG), rozhodovací práh 50%.

Skutečný výstup	*Oprávněný uživatel*	174 48.3%	5 1.4%	97.2%
	Neoprávněný uživatel	6 1.7%	175 48.6%	96.6%
		96.6%	97.2%	**96.9%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		



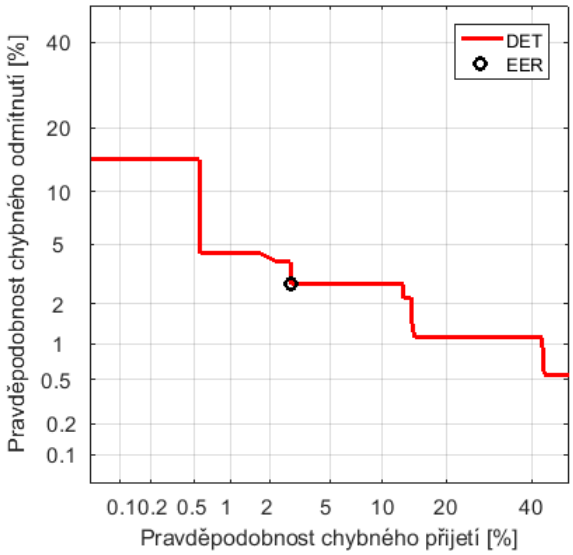
Obr. 3.2: Histogram rozdělení skóre oprávněných a neoprávněných uživatelů s vyznačenými hodnotami rozhodovacího prahu - autentizace tváří.



Obr. 3.3: Závislost hodnot FAR a FRR na velikosti rozhodovacího prahu - autentizace tváří.

3.4 Zhodnocení dosažených výsledků

Na základě uvedených výsledků se jeví jako nejvhodnější přístup pro návrh systému klasifikátor SVM s využití HOG parametrů. Při využití těchto přístupů bylo dosaženo nejnižších hodnot FAR, FRR a EER. Tento přístup je také následně využit pro návrh vícenásobného biometrického systému v kapitole 4. Dosažení lepších výsledků pro SVM klasifikátor je



Obr. 3.4: DET křivka - autentizace tváří.

dáno vlastnostmi klasifikátorů. V případě SVM jsou použity pro klasifikaci pouze podpůrné vektory. To znamená, že SVM klasifikátor nepotřebuje velkou trénovací sadu, pokud trénovací sada obsahuje vhodné podpůrné vektory. V případě MLNN klasifikátoru je důležitá velká trénovací sada, pomocí které jsou nastaveny váhy jednotlivých neuronů sítě. Z pohledu parametrů dosáhly nejlepších výsledků parametry HOG a to jak pro klasifikátor SVM, tak i pro klasifikátor MLNN. Dosažení lepších výsledků pro HOG parametry je dáno tím, že tyto parametry popisují soubor (histogram) změn (gradientů) v celém obraze, kdežto LPB parametry popisují pouze lokální vzory. Při detailní analýze výsledků bylo zjištěno, že nejčastější chyba klasifikace (SVM i MLNN klasifikátor) vzniká při klasifikaci obrázků, kde je osoba částečně zahalena do šátku. Dá se tedy předpokládat, že při odstranění těchto snímků z trénovací i testovací množiny bude dosaženo nižších hodnot FAR, FRR, EER a vyšší přesnosti. Odstranění těchto snímků je i v souladu z principem autentizace, kde uživatel chce být rozpoznán a tudíž nemá důvod zahalovat se šátkem.

4 NÁVRH KOMPLEXNÍHO VÍCENÁSOBNÉHO BIOMETRICKÉHO AUTENTIZAČNÍHO SYSTÉMU

Tato kapitola je věnována hlavní části dizertační práce a to návrhu vícenásobného biometrického autentizačního systému. Tento systém je založen na vhodné kombinaci hlasového autentizačního systému a systému využívajícího geometrii obličeje pro ověření totožnosti daného uživatele. Tyto systémy jsou postaveny na výsledcích a poznatcích z kapitol 2 a 3.

4.1 Princip návrhu vícenásobného biometrického autentizačního systému

Při návrhu vícenásobného autentizačního systému bylo vycházeno ze strategie vícenásobné biometrie (v jednom systému je využito více biometrických charakteristik). V rámci práce byly porovnány dva typy propojení (fúze) biometrických charakteristik: propojení na úrovni rozhodnutí o verifikaci a propojení na úrovni verifikační míry. V případě fúze na úrovni rozhodnutí o verifikaci byly ověřeny dva přístupy (AND a OR pravidla). Pro fúzi na úrovni verifikační míry bylo použito pravidlo o maximální pravděpodobnosti (Max rule) a pravidlo o sčítání pravděpodobností (Sum rule). V práci není uvažováno s klasifikačním přístupem pro fúzi na úrovni verifikační míry a to z důvodu nízkého počtu testovacích dat. Implementace jednotlivých typů fúzí včetně jejich vyhodnocení probíhala v prostředí Matlab.

V experimentu byla použita data pocházející z databází Comtech a AR Face Database. Referenčním řečníkům (18) z databáze Comtech byly přiřazeny referenční obličeje pocházející z AR Face Database. Totéž platilo pro 10 uživatelů vydávajících se za podvodníky. K testování systému bylo použito 10 vzorků od každého referenčního uživatele a 10 vzorků od 10 různých podvodníků. Pod pojmem vzorek je v tomto případě myšleno spojení 1 nahrávka + 1 obraz tváře. Pro část systému založeného na autentizaci hlasem byl využit přístup poskytující nejlepší výsledky uvedené v kapitole 2. Jedná se o přístup využívající parametry MFCC + delta MFCC a klasifikátor SVM. Část systému sloužící pro ověření identity na základě obrazu tváře byla postavena na výsledcích z kapitoly 3. Tento přístup zahrnoval parametry HOG a využitím SVM klasifikátoru.

4.2 Propojení na úrovni rozhodnutí o verifikaci

V rámci propojeni na úrovni rozhodnutí o verifikaci byly ověřeny dvě strategie: AND a OR pravidlo. V případě použití AND pravidla je totožnost referenčního uživatele potvrzena pouze v případě, že obě části systému (autentizace hlasem, autentizace geometrií obličeje) rozhodly, že se jedná o referenčního uživatele. V ostatních případech je uživatel označen jako podvodník. Toto pravidlo je vhodné v případě autentizace, kde chceme dosáhnout

nízké hodnoty FAR (přijetí podvodníků). Naopak v případě OR pravidla stačí, aby alespoň jedna část systému označila uživatele jako referenčního a ten je následně opravdu označen jako referenční. To znamená, že z pohledu autentizace bude dosaženo nízké hodnoty FRR (odmítnutí referenčních uživatelů). Vzhledem k tomu, že k finálnímu rozhodnutí dochází až po rozhodnutí dílčích podsystémů, kde je již nastaven verifikační práh jsou výsledky těchto fúzí popsány pouze pomocí kontingenční tabulky.

4.2.1 Experimentální výsledky pro fúzi pomocí AND pravidla

Aplikací AND pravidla na finální rozhodnutí dílčích podsystémů bylo dosaženo snížení počtu přijetí neoprávněných uživatelů (podvodníků) na nula. Přístup je povolen uživateli pouze v případě, že oba podsystémy rozhodly, že se jedná o oprávněného uživatele. Tento typ fúze je možné využít v případě, kdy chceme opravdu striktně odmítat přístup neoprávněných uživatelů (hodnota FAR je rovna nule). Výsledky pro AND pravidlo jsou uvedeny v kontingenční tabulce 4.1. Použitím této fúze je dosaženo celkové přesnosti systému 98.3%. Porovnání jednotlivých typů fúzí z pohledu hodnot FAR a FRR je uvedeno v tabulce 4.5.

Tab. 4.1: Kontingenční tabulka - fúze pomocí AND pravidla, rozhodovací práh podsystémů 50%.

Skutečný výstup				
	Oprávněný uživatel	174 48.3%	0 0.0%	100.0%
	Neoprávněný uživatel	6 1.7%	180 50.0%	96.7%
		96.6%	100.0%	**98.3%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		

4.2.2 Experimentální výsledky pro fúzi pomocí OR pravidla

V případě použití OR pravidla je kladen důraz na co pokud možno největší komfort oprávněných uživatelů (stačí aby alespoň jeden z podsystémů označil uživatele jako oprávněného a uživatel je následně označen za oprávněného i vícenásobným systémem). Na druhé straně pro podvodníka je mnohem jednodušší systém prolomit. Aplikováním OR pravidla je sice dosaženo nízké hodnoty FRR ale naproti tomu vzroste hodnota FAR. Výsledky jsou uvedeny v kontingenční tabulce 4.2. Celková přesnost systému v tomto případě je 96.7%.

Tab. 4.2: Kontingenční tabulka - fúze pomocí OR pravidla, rozhodovací práh podsystémů 50%.

Skutečný výstup	*Oprávněný uživatel*	180 50.0%	11 3.0%	94.2%
	Neoprávněný uživatel	0 0.0%	169 47.0%	100.0%
		100.0%	93.8%	**96.7%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		

4.3 Propojení na úrovni verifikační míry

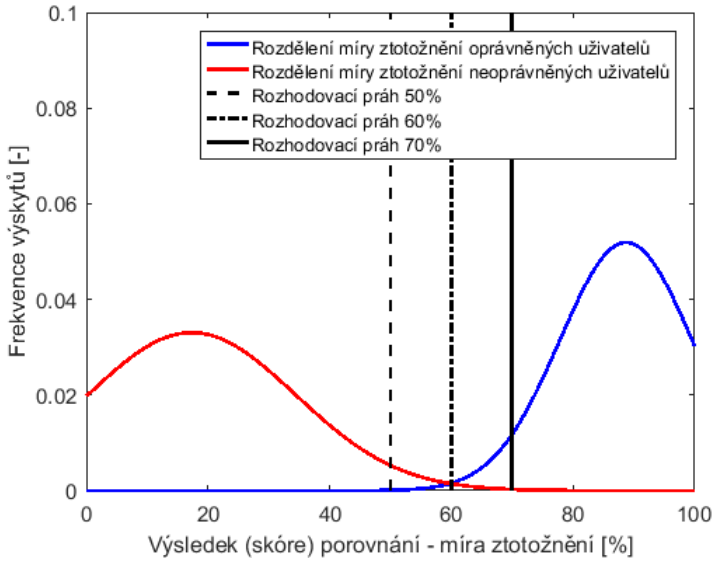
V případě fúze na úrovni verifikační míry, byl výzkum zaměřen na dvě kombinační metody: pravidlo o maximální pravděpodobnosti a pravidlo o sčítání pravděpodobností. V případě využití pravidla o maximální pravděpodobnosti je fúze dána pouze výběrem největší verifikační míry (skóre) z jednotlivých klasifikátorů pro danou třídu. Při použití pravidla o sčítání pravděpodobností je proveden součet pravděpodobností z jednotlivých klasifikátorů pro danou třídu. Zároveň může být každému klasifikátoru přiřazena určitá váha.

4.3.1 Experimentální výsledky pro fúzi pomocí pravidla o maximální pravděpodobnosti

Při použití pravidla o maximální pravděpodobnosti je rozhodnutí provedeno na základě porovnání maximální pravděpodobnosti jednoho z podsystémů a nastaveného rozhodovacího prahu. Výběrem maximální pravděpodobnosti je upřednostněn podsystém, který si je více "jistý"o přiřazení do dané třídy. Histogram rozdělení skóre po aplikaci fúze je zobrazen na obrázku 4.1. Z tabulka 4.3 je patrné, že se pomocí fúze podařilo snížit počet chybných rozhodnutí. S tím koresponduje i zvýšení přesnosti systému na 99.7%. ROC křivka systému je zobrazena na obrázku 4.2. Průběh hodnot FAR a FRR v závislosti na rozhodovacím prahu je uveden na obrázku 4.3. Hodnota EER byla v tomto případě 0.55% a je vyznačena v DET křivce na obrázku 4.4.

4.3.2 Experimentální výsledky pro fúzi pomocí pravidla o sčítání pravděpodobností

V případě fúze pomocí sčítání pravděpodobností je proveden součet posteriorních pravděpodobností jednotlivých podsystémů pro odpovídající třídy. Tento součet je následně porovnán s verifikačním prahem a na základě porovnání je provedeno finální rozhodnutí o autentizaci. Histogram rozdělení skóre pro tento typ fúze je zobrazen na obrázku 4.5. Z kontingenční tabulky 4.4 je patrné, že aplikací tohoto typu fúze nedosáhneme zvýšení



Obr. 4.1: Histogram rozdělení skóre oprávněných a neoprávněných uživatelů s vyznačenými hodnotami rozhodovacího prahu - Max rule.

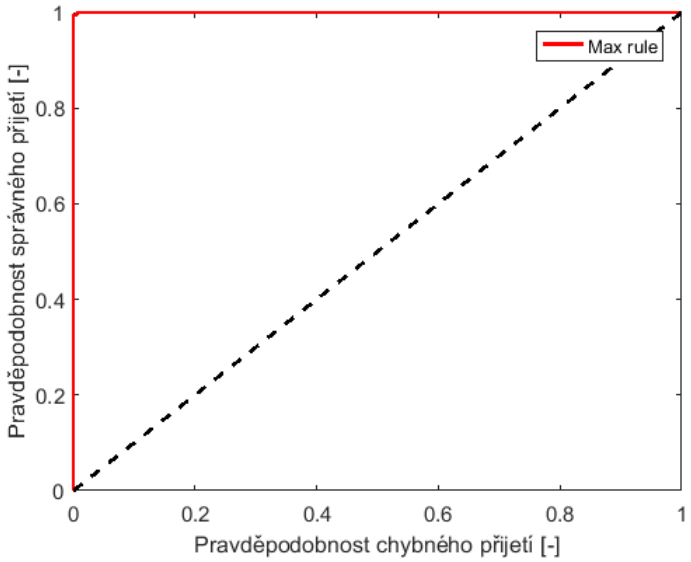
Tab. 4.3: Kontingenční tabulka - fúze pomocí Max rule, rozhodovací práh 50%.

Skutečný výstup	*Oprávněný uživatel*	180 50.0%	1 0.3%	99.4%
	Neoprávněný uživatel	0 0.0%	179 49.7%	100.0%
		100.0%	99.4%	**99.7%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		

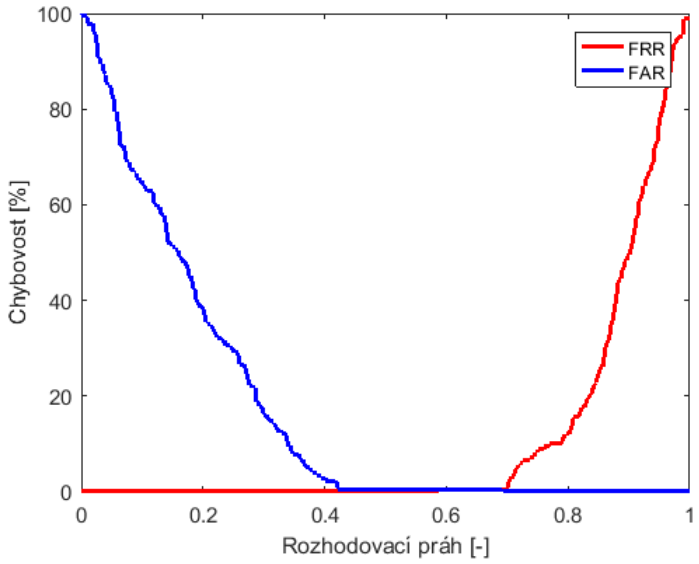
přesnosti oproti fúzi Max rule (systém stále produkuje 1 chybu). ROC křivka je zobrazena na obrázku 4.6. Zlepšení sytému v případě Sum rule je dosaženo z pohledu velikosti EER, která je v tomto případě 0.0%. To znamená, že pro vhodně nastavený rozhodovací práh (0.55 nebo 55%) je možné dosáhnout nulových hodnot FAR a FRR na použité testovací sadě. Tento fakt je patrný z obrázku 4.7. DET křivka není uvedena z důvodu nulové hodnoty EER.

4.4 Zhodnocení dosažených výsledků

Na základě výše uvedených výsledků lze konstatovat, že každý typ fúze přináší určité výhody z pohledu oblasti uplatnění systému. Porovnání jednotlivých typů fúzí je uvedeno v tabulce 4.5.

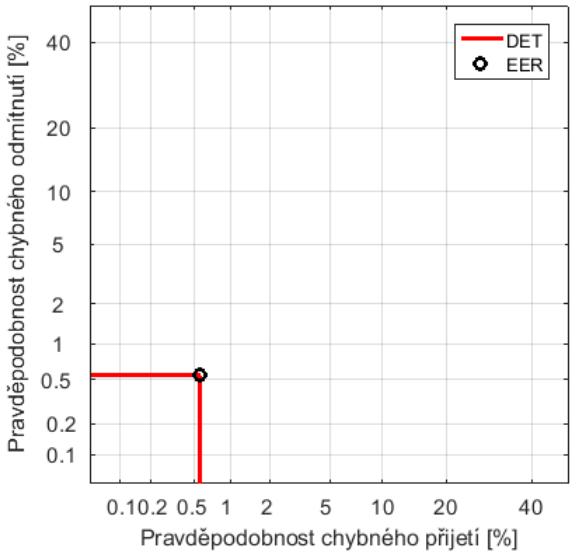


Obr. 4.2: ROC křivka - Max rule.

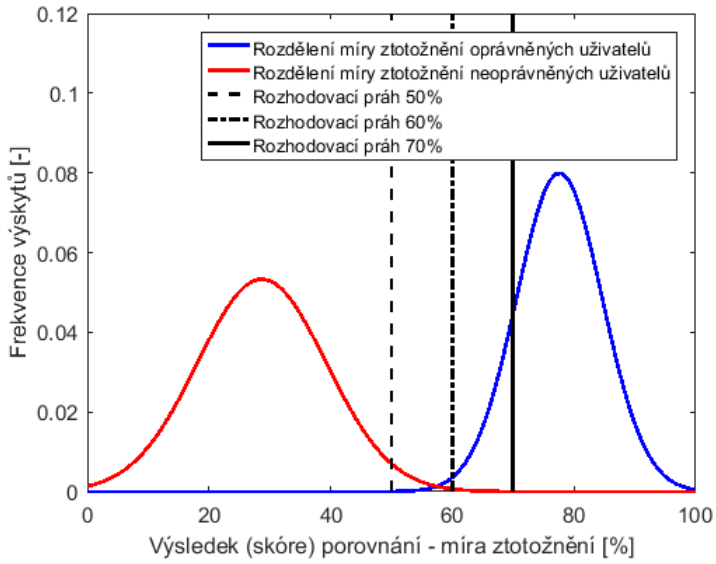


Obr. 4.3: Závislost hodnot FAR a FRR na velikosti rozhodovacího prahu - Max rule.

Použitím AND pravidla je dosaženo nejnižší hodnota FAR (0.0%). Z praktického hlediska to znamená, že v rámci testovací sady není označen žádný podvodník jako referenční uživatel. Na druhé straně je v tomto případě dosaženo nejvyšší hodnoty FRR (3.3%) což znamená, že určitý počet oprávněných uživatelů je označen za podvodníky. Vícenásobný systém s využitím tohoto typu fúze je vhodné využít v případě, kdy je kladen obrovský důraz na bezpečnost (pokud by byl přístup povolen podvodníkovi, mohlo by to mít fatální



Obr. 4.4: DET křivka - Max rule.

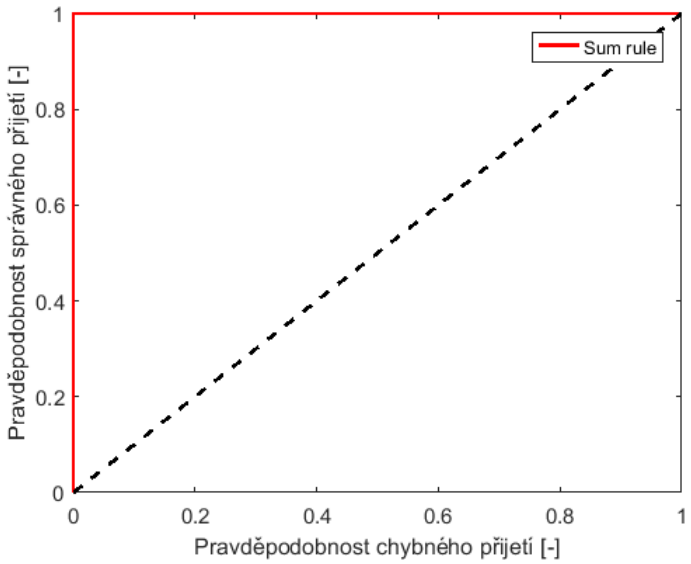


Obr. 4.5: Histogram rozdělení skóre oprávněných a neoprávněných uživatelů s vyznačenými hodnotami rozhodovacího prahu - Sum rule.

následky) například přístup k bankovnímu účtu. V případě použití OR pravidla je naopak dosaženo nejvyšší hodnoty FAR (6.2%). Hodnota FRR je v tomto případě 0.0%. Tento typ fúze je vhodné využít v případě, kdy je kladen důraz na komfort referenčních uživatelů. Referenční uživatelé nejsou obtěžování opětovnými pokusy o autentizaci, na druhou stranu se může stát, že některý z podvodníků bude označen jako referenční uživatel.

Tab. 4.4: Kontingenční tabulka - fúze pomocí Sum rule, rozhodovací práh 50%.

Skutečný výstup	*Oprávněný uživatel*	180 50.0%	1 0.3%	99.4%
	Neoprávněný uživatel	0 0.0%	179 49.7%	100.0%
		100.0%	99.4%	**99.7%**
		Oprávněný uživatel	*Neoprávněný uživatel*	
		Požadovaný výstup		

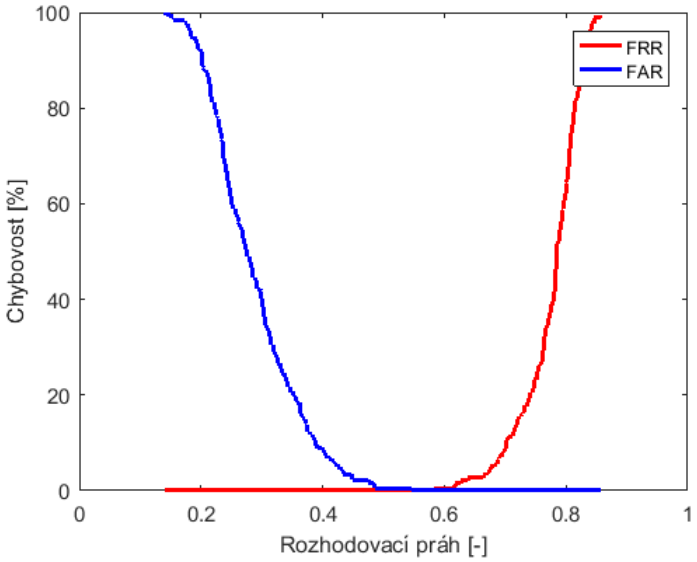


Obr. 4.6: ROC křivka - Sum rule.

Tab. 4.5: Tabulka hodnot FAR, FRR a EER pro jednotlivé typy fúzí pro rozhodovací práh 50%

Typ fúze	FAR [%]	FRR [%]	EER [%]
AND rule	0.0	3.3	-
OR rule	6.2	0.0	-
Max rule	0.6	0.0	0.55
Sum rule	0.6	0.0	0.00

Zatímco dvě předchozí metody přináší zlepšení systému pouze v závislosti na oblasti jeho použití (snížení hodnoty FAR nebo FRR), tak fúze typu pravidlo o maximální pravděpodobnosti a pravidlo o sčítání pravděpodobností přináší celkové zvýšení přesnosti bez ohledu na oblast použití. V obou případech propojení na úrovni rozhodnutí o verifikaci bylo dosaženo maximální přesnosti systému a to 99.7%. Tato přesnost odpovídá skuteč-



Obr. 4.7: Závislost hodnot FAR a FRR na velikosti rozhodovacího prahu - Sum rule.

nosti, že systém udělal jednu chybu při evaluaci testovací sady (360 vzorků). Přestože oba systémy dosáhly stejné přesnosti, tak z pohledu EER se jejich výsledky liší. Odlišnost hodnot EER je způsobena principem jednotlivých fúzí. V případě použití fúze Max rule systém dosáhl hodnoty EER 0.55%, kdežto pomocí Sum rule bylo dosaženo nulové hodnoty EER. To v praxi znamená, že při vhodně nastaveném rozhodovacím prahu systém založený na pravidle o sčítání pravděpodobností neprodukuje žádný typ chyby na příslušné testovací sadě. Hodnota verifikačního prahu pro dosažení nulové hodnoty EER je 0.55 (55%). Z tohoto pohledu lze označit vícenásobný systém založený na pravidle o sčítání pravděpodobností za nejlepší variantu.

5 EXPERIMENTÁLNÍ OVĚŘENÍ FUNKČNOSTI NAVRŽENÉHO VÍCENÁSOBNÉHO AUTENTIZAČNÍHO SYSTÉMU A POROVNÁNÍ PŘESNOSTI S AKTUÁLNĚ POUŽÍVANÝMI SYSTÉMY BIOMETRICKÉ AUTENTIZACE

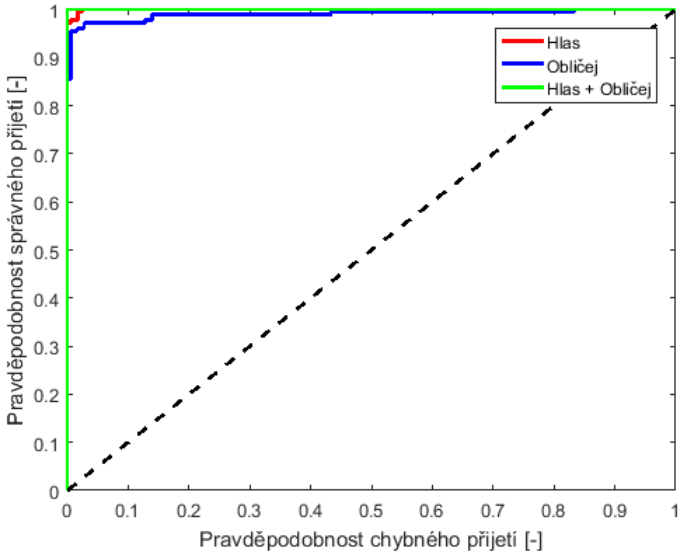
V první části této kapitoly je provedeno porovnání unimodálních biometrických autentizačních systému navržených v kapitolách 2 a 3 s vícenásobným biometrickým autentizačním systémem navrženým v kapitole 4. Druhá část kapitoly je zaměřena na porovnání navrženého vícenásobného systému s již existujícími obdobnými systémy.

5.1 Porovnání navržených biometrických systémů

Porovnání navržených unimodálních biometrických systémů s vícenásobným biometrickým systémem je provedeno jak z pohledu chybovosti a přesnosti, tak z pohledu bezpečnosti. V tabulce 5.1 jsou uvedeny výsledky hodnot FAR, FRR, EER a přesnost pro tři typy biometrických systému. První je biometrický systém založený na ověření totožnosti pomocí hlasu. Tento systém využívá příznakový vektor skládající se z parametrů MFCC a delta MFCC, pro klasifikaci je použit SVM klasifikátor. Druhý biometrický systém využívá pro verifikaci informací o geometrii obličeje. Systém opět používá pro klasifikaci SVM klasifikátor a jako příznakový vektor parametry HOG. Vícenásobný biometrický systém je založen na fúzi předchozích dvou systémů. Fúze je provedena pomocí Sum rule. Z tabulky je patrné, že fúzí unimodálních systémů je dosaženo snížení FAR, FRR, EER a zvýšení celkové přesnosti systému. Navíc v případě vícenásobného systému jsme schopni nastavením vhodného rozhodovacího prahu dosáhnout nulové hodnoty EER na testovací sadě, což v oblasti biometrických systémů znamená ideální případ (systém správně rozpozná všechny referenční uživatele a správně odmítne všechny podvodníky). Použitím vícenásobného biometrického je také dosaženo vyšší bezpečnosti. U vícenásobných systémů je mnohem obtížnější provést podvržení identity referenčního uživatele. Na obrázku 5.1 jsou zobrazeny ROC křivky jednotlivých biometrických systémů.

Tab. 5.1: Tabulka hodnot FAR, FRR, EER a přesnosti pro jednotlivé typy biometrických systémů - rozhodovací práh 50%

Systém	FAR [%]	FRR [%]	EER [%]	Přesnost [%]
Hlas	3.3	0.0	1.7	98.3
Obličej	2.7	3.3	2.8	96.9
Hlas + Obličej	**0.6**	**0.0**	**0.0**	**99.7**



Obr. 5.1: ROC křivky - biometrické systémy.

5.2 Porovnání navrženého vícenásobného biometrického systému s obdobnými existujícími systémy

V této podkapitole je uveden přehled nejvýznamnějších publikací z oblasti audio-vizuální verifikace/identifikace. Vzhledem k tomu, že téměř v každé práci je použita jiná databáze nebo jiné podmínky experimentu, tak není uvedeno přímé srovnání číselných hodnot, které by bylo v tomto případě bezvýznamné.

První z významných publikací této problematiky je [5] z roku 1993, kde autoři využívají pro kombinaci podsystému textově závislé identifikace hlasu a podsystému identifikace obličeje pravidlo o sčítání pravděpodobností. V experimentu byly použity parametry MFCC a jako obrazové parametry byly použity vzdálenosti mezi markantními body na tváři. Pro klasifikaci byla v obou podsystémech použita ANN. Rozhodnutí o identifikaci bylo provedeno na základě porovnání nastaveného prahu a výsledného skóre po fúzi. Nastavením vhodného rozhodovacího prahu autoři dosáhli hodnoty EER 1.5%. Těchto výsledků bylo dosaženo na vlastní databázi tvořené 10 uživateli.

V literatuře [6] autoři využívají pro kombinaci biometrických charakteristik pravidlo o násobení pravděpodobností. V experimentu použili vlastní databázi, která obsahovala vzorky od 33 osob. Podsystém založený na identifikaci hlasem využíval parametry MFCC + delta MFCC, které byly klasifikovány pomocí vektorové kvantizace. Obličejový podsystém pracuje s geometrickými informacemi o markatních bodech (pozice a šířka nosu, očí atd.). Tyto parametry byly klasifikovány pomocí bayesova klasifikátoru. Celková přesnost navrženého systému byla 95%.

Publikace [7] je rozšířením předchozí práce [6] se snahou zvýšení identifikační přes-

nosti. Autoři v tomto experimentu provedli rozdělení parametrů do více kategorií, kde každá kategorie byla klasifikována zvlášť (MFCC, delta MFCC, geometrické vlastnosti nosu, geometrické vlastnosti očí atd.) a poté provedli fúzi jednotlivých pravděpodobností pomocí pravidla o násobení pravděpodobností. Bylo tak dosaženo zvýšení přesnosti o 3% na celkových 98%.

Autoři [8] navrhli biometrický autentizační systém, který využívá tři biometrické charakteristiky (hlas, obličej, pohyb rtů). Rozhodnutí o verifikaci je provedeno na základě kombinace dvou ze tří podsystémů. V první fázi jsou vybrány dva ze tří podsystémů (klasifikátorů), které poskytují nejlepší výsledky. Následně je provedena kombinace skóre vybraných podsystémů a porovnání tohoto skóre s přednastaveným prahem. Využitím takového typu fúze dojde ke zvýšení robustnosti systému proti rušení.

V literatuře [1] je představen vícenásobný biometrický autentizační systém založený na verifikaci hlasem a verifikaci pomocí obličeje. Pro experiment byla použita databáze XM2VTS. Autoři pro podsystém verifikace hlasem využili parametry LPC, které klasifikovali pomocí skrytých markovových modelů. Pro reprezentaci obličeje byly použity deformační modely EGM (Elastic Graph Matching). Fúze je v tomto případě prováděna na úrovni verifikační míry a je na ní pohlíženo jako na klasifikační problém. Z toho důvodu byly pro kombinaci jednotlivých podsystému použity klasifikátory SVM a bayesův. Požitím SVM klasifikátoru bylo dosaženo hodnoty EER 1.2%, v případě druhého klasifikátoru 0.6%.

Publikace [9] popisuje systém, který využívá pro kombinaci podsystémů AND pravidlo. Jedná se tedy o typ fúze na úrovní rozhodnutí o verifikaci. V rámci experimentu byla použita vlastní databáze, která byla tvořena 30 účastníky. Autoři použitím AND pravidla dokázali minimalizovat hodnotu FAR. Jako parametry popisující hlas autoři použili koeficienty vycházející z vlnkové tranformace. Obrazové parametry představovaly vzhled očí. Oba typy parametrů byly klasifikovány pomocí vícevrstvé neuronové sítě.

Autoři v literatuře [3] využívají stejných metod fúze (klasifikátor SVM a bayesův klasifikátor) jako autoři v literatuře [1] s tím rozdílem, že v rámci jednotlivých podsystémů využívají jiné parametry a klasifikátory. Pro reprezentaci hlasu jsou použity parametry MFCC a jejich první a druhá derivace. Obličej je reprezentován parametry získanými pomocí metody hlavních komponent. V obou případě je pro klasifikaci využit klasifikátor GMM. Vzorky pro experiment byly získány z databáze VidTIMIT.

Literatura [10] popisuje biometrický identifikační systém, který je založen na fúzi pomocí pravidla o sčítání pravděpodobností. Hlasový podsystém využívá pro reprezentaci hlasu prozodických parametrů (základní frekvence, počet průchodů nulou, formanty atd.). Podsystém identifikace obličeje využívá parametrů popisujících geometrické vlastnosti markantních bodů. Klasifikace jednotlivých parametrů je provedena pomocí umělé neuronové sítě. Autoři použili v rámci experimentu vlastní databázi, která byla tvořena 20 respondenty. Na této databázi se podařilo dosáhnout identifikační přesnosti 97.5%.

V rámci publikace [13] byl popsán vícenásobný verifikační systém využívající pravi-

dlo o sčítání pravděpodobností. Autoři v tomto experimentu použili databázi VidTIMIT. Podsystém verifikace hlasem byl založen na použití LPC koeficientů, které byly klasifikovány GMM klasifikátorem. Podsystém verifikace obličeje využíval koeficienty získané pomocí 2DLDA (Two-Dimensional Linear Discriminative Analysis), tyto koeficienty byly následně klasifikovány pomocí K-NN klasifikátoru.

Výzkum v práci [11] je zaměřen na porovnání různých typů fúzí a s tím spojených normalizačních technik. Data pro experiment pochází ze dvou databází XM2VTS (obrazová data) a TIMIT (hlasová data). Autoři dosáhli nejlepších výsledků s využitím pravidla o sčítání pravděpodobností a pravidla o násobení pravděpodobností.

Další varianty vícenásobných biometrických systémů jsou uvedeny v [12], [14] a [15].

Na základě rešerše existujících systémů lze tvrdit, že navržený vícenásobný biometrický systém v rámci dizertační práce je jedinečný z pohledu použitých metod (žádný z výše uvedených systémů není založen na stejných technikách a metodách jako navržený systém). Zároveň navržený systém dosahuje vysoké přesnosti a nízké chybovosti v porovnání s existujícími systémy, což ho předurčuje k možnosti reálného nasazení.

6 ZÁVĚR A PŘÍNOS PRÁCE

Biometrické autentizační systémy se v současné době stávají synonymem samotné autentizace. S rostoucím využívání biometrických metod rostou i požadavky na systémy, které tyto metody využívají. Pro většinu biometrických metod již není limitující pouze jejich přesnost, ale především bezpečnost. Existuje již celá řada přístupů, jak je možné napodobit různé biometrické charakteristiky. Z tohoto důvodu se hledají nové cesty, jak vhodně zamezit podvržení identity. Jednou z těchto cest je využití vícenásobné biometrie. Díky vícenásobné biometrii je mnohem těžší nebo dokonce nemožné provést podvržení identity a zároveň je dosaženo zvýšení přesnosti celého systému. Tato skutečnost byla hlavním podnětem pro vypracování této dizertační práce. Dizertační práce byla zaměřena na návrh vícenásobného biometrického autentizačního systému, který vhodně kombinuje dvě biometrické charakteristiky a to hlas a tvář. Tyto dvě biometrické charakteristiky byly zvoleny s ohledem na jejich vlastnosti (přijatelnost pro uživatele, jednoduché bezkontaktní snímání, přesnost, rychlost, spolehlivost). Samotný návrh vícenásobného systému byl rozdělen do tří částí. V první části byl proveden návrh podsystému zajišťujícího autentizaci pomocí řečového signálu. V druhé části byl navržen druhý autentizační podsystém založený na ověření identity pomocí geometrie obličeje. V poslední části byly ověřovány různé fúzní strategie navržených podsystémů za účelem nalezení nejvhodnějšího architektury vícenásobného biometrického autentizačního systému.

Návrh hlasového autentizačního systému probíhal jak z pohledu hledání vhodných parametrů, tak z pohledu nalezení dostatečně přesného klasifikátoru. Při návrhu byly porovnávány parametry MFCC, delta MFCC, delta-delta MFCC, LPC a jejich kombinace. Klasifikace parametrů probíhala pomocí SVM klasifikátoru a MLNN klasifikátoru. Volba parametrů a klasifikátorů byla provedena na základě rešerše a na základě konkrétních požadavků kladených na systém. Jednotlivé kombinace použitých parametrů a příslušného klasifikátoru byly hodnoceny pomocí hodnot FAR, FRR, EER a přesnosti. Ověřování výsledků bylo prováděno vzhledem k databázi Comtech, která byla v rámci dizertační práce navržena. Nejlepších výsledků bylo dosaženo pro příznakový vektor složený z parametrů MFCC a delta MFCC při použití SVM klasifikátoru. Pro nastavený výchozí rozhodovací práh 50% dosáhly hodnoty FAR 2.3% a FRR 0.27%. Tyto hodnoty odpovídají přesnosti systému 98.7%. Hodnota EER pro tento příznakový vektor byla 0.85%, což odpovídá nastavenému rozhodovacímu prahu na 55%. Zvýšením rozhodovacího prahu na 65% bylo dosaženo nulové hodnoty FAR, což je z pohledu autentizace ideální případ (žádný z podvodníků není označen jako referenční uživatel). Tento hlasový autentizační systém je reálně nasazen jako součást komplexního systému pro zabezpečenou komunikaci v rámci projektu TA ČR TF01000091, kde slouží jako doplňková ochrana. Princip nasazení systému do komplexního řešení je uveden v práci [Tov08]. Teoretickým přínosem této části dizertační práce je skutečnost, že klasifikátor, který dodržuje časovou sekvenci zpracovávaných segmentů vyžaduje pro dosažení maximální přesnosti pouze minimální

počet signifikantních parametrů, kdežto klasifikátor, který tuto sekvenci poruší potřebuje pro dosažení maximální přesnosti maximální počet parametrů. Dalším přínosem je vznik české databáze řečových vzorků, která může být využita, jak pro textově nezávislé rozpoznávání řečníka, tak pro textově závislé.

Návrh autentizačního systému založeného na ověření identity pomocí geometrie obličeje byl obdobný jako návrh hlasového autentizačního systému. Opět bylo úkolem nalézt vhodné parametry a klasifikátor produkující nejnižší počet chyb. Analyzovanými obrazovými příznaky byly LBP, HOG a jejich kombinace. Klasifikace těchto parametrů byla provedena pomocí SVM klasifikátoru a MLNN klasifikátoru. Volba stejných klasifikátorů jako v případě hlasové autentizace je založena na jejich vhodnosti pro řešení binární klasifikace a také na snaze o zjednodušení výsledného vícenásobného systému. Jednotlivé varianty realizace systému byly mezi sebou porovnávány opět pomocí hodnot FAR, FRR, EER a přesnosti. Obrazová data pro návrh systému pocházela z databáze AR Face Database. Výsledkem návrhu byl systém založený na HOG parametrech s využitím SVM klasifikátoru. Takto navržený systém dosáhl hodnot FAR 2.7% a FRR 3.3% pro rozhodovací práh 50%, což odpovídá celkové přesnosti systému 96.9%. Hodnota EER byla pro tento systém 2.8%. Stejně jako u předchozího systému lze dosáhnout nižší hodnoty FAR pomocí zvýšení rozhodovacího prahu. Experimentální výsledky návrhu autentizačního systému založeného na ověření identity pomocí geometrie obličeje byly publikovány v práci [Tov09]. Teoretickým přínosem je v tomto případě poznatek o volbě vhodného klasifikátoru s ohledem na velikost trénovací sady.

Výsledný vícenásobný biometrický autentizační systém je založen na vhodné kombinaci navrženého hlasového autentizačního systému a autentizačního systému využívajícího geometrii obličeje pro ověření identity. V dizertační práci byly porovnávány dva typy úrovní fúze (propojení na úrovni rozhodnutí o verifikaci a propojení na úrovni rozhodnutí o verifikaci). Pro každou z úrovní byly analyzovány dvě strategie, které reprezentující danou úroveň. V rámci propojení na úrovni rozhodnutí o verifikaci byly experimentálně ověřeny strategie AND a OR. V případě propojení na úrovni verifikační míry byly analyzovány dvě strategie pravidlo o maximální pravděpodobnosti a pravidlo o sčítání pravděpodobností. Z pohledu velikosti hodnoty FAR bylo dosaženo nejlepšího výsledku použitím fúze pomocí AND pravidla, kde hodnota FAR dosáhla nulové hodnoty. To v praxi znamená, že systém neoznačí za oprávněného uživatele ani jednoho podvodníka. Na druhou stranu použitím AND pravidla dosáhneme nejvyšší hodnoty FRR 3.3% ze všech testovaných strategií. Z tohoto pohledu se jeví jako ideální řešení využití fúze založené na pravidle o sčítání pravděpodobností, kde bylo dosaženo hodnot FAR 0.6%, FRR 0.0% a přesnosti 99.7% pro rozhodovací práh 50%. Současně lze dosáhnout nulové hodnoty EER při nastavení rozhodovacího prahu na 55%. Z pohledu činnosti systému to znamená, že pro takto nastavený rozhodovací práh systém neprodukuje žádné chyby na příslušné testovací sadě. Při porovnání navrženého vícenásobného autentizačního systému s unimodálními systémy je patrné snížení chybovosti a zvýšení přesnosti v případě vícenásobného systému. Současně

je dosaženo vysokého stupně bezpečnosti, díky kterému je prakticky nemožné provést podvržení identity (podvodník by musel věrohodně napodobit hlas pronášející tajné přístupové heslo, který náleží referenčnímu uživatel a musel by současně dokonale napodobit obličej téže osoby). Z porovnání s existujícími obdobnými systémy je patrné, že navržený systém v rámci dizertační práce je jedinečný jak z pohledu použitých metod, tak z pohledu přesnosti a bezpečnosti. Dosažením nulové hodnoty EER je možné navržený systém porovnávat i s jinými typy vícenásobných biometrických systémů dosahujících vyšší přesnosti. Z tohoto pohledu se následně projeví výhody použitých biometrických charakteristik jako jsou přijatelnost pro uživatele, jednoduché bezkontaktní snímání a rychlost.

Navržený vícenásobný biometrický autentizační systém představuje unikátní přístup pro dosažení vysoce přesné a bezpečné autentizace, který může být využit téměř ve všech oblastech ověřování identity.

LITERATURA

[1] BEN-YACOUB, Souheil, Yousri ABDELJAOUED a Eddy MAYORAZ. Fusion of face and speech data for person identity verification. *IEEE Transactions on Neural Networks* [online]. 1999, **10**(5), 1065-1074 [cit. 2016-03-07]. DOI: 10.1109/72.788647. ISSN 1045-9227. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=788647&isnumber=17091

[2] KITTLER, Josef. Combining classifiers: A theoretical framework. *Pattern Analysis and Applications*. 1998, **1**(1), 18-27. DOI: 10.1007/BF01238023. ISSN 1433-755X. Dostupné z: http://dx.doi.org/10.1007/BF01238023

[3] SANDERSON, Conrad a Kuldip K. PALIWAL. Identity verification using speech and face information. *Digital Signal Processing*. 2004, **14**(5), 449-480. DOI: 10.1016/j.dsp.2004.05.001. Dostupné z: http://dx.doi.org/10.1016/j.dsp.2004.05.001

[4] MARTINEZ, A. a R. BENAVENTE. The AR Face Database. *CVC Technical Report #24*. 1998.

[5] CHIBELUSHI, Claude C., Farzin DERAVI a John MASON. Voice and facial image integration for speaker recognition. *IEEE International Symposium and Multimedia Technologies and Future Applications*. 1993

[6] BRUNELLI, Roberto, Daniele FALAVIGNA, Tomaso POGGIO a Luigi STRINGA. Automatic person recognition by acoustic and geometric features. *Machine Vision and Applications*. 1995, **8**(5) 317-325. DOI: 10.1007/BF01211493. ISSN 1432-1769. Dostupné z: https://doi.org/10.1007/BF01211493

[7] BRUNELLI, Roberto a Daniele FALAVIGNA. Person Identification Using Multiple Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1995, **17**(10), 955-965. DOI: 10.1109/34.464560. ISSN 0162-8828. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=464560

[8] DIECKMANN, Ulf, Peter PLANKENSTEINER a Thomas WAGNER. SESAM: A biometric person identification system using sensor fusion. *Pattern Recognition Letters*. 1997, **18**(9), 827-833. DOI: 10.1007/BFb0016009. ISSN 0167-8655. Dostupné z: https://doi.org/10.1007/BFb0016009

[9] POH, Norman a Jerzy KORCZAK. Hybrid Biometric Person Authentication Using Face and Voice Features. *Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication*. 2001, 348-353. DOI: 10.1007/3-540-45344-X_51. ISBN 978-3-540-45344-4. Dostupné z: https://link.springer.com/content/pdf/10.1007/3-540-45344-X_51.pdf

[10] KALA, Rahul, Harsh VAZIRANI, Anupam SHUKLA, Ritu TIWARI. Fusion of Speech and Face by Enhanced Modular Neural Network. *Proceedings of the Fourth International Conference on Information Systems, Technology and Management*. 2010, 363-372. DOI: 10.1007/978-3-642-12035-0_37. ISBN 978-3-642-12035-0. Dostupné z: https://link.springer.com/chapter/10.1007/978-3-642-12035-0_37

[11] FOUDA, Yasser M. Fusion of Face and Voice: An improvement. *International Journal of Computer Science and Network Security*. 2012, **12**(4), 37-43. Dostupné z: http://paper.ijcsns.org/07_book/201204/20120406.pdf

[12] GHAYOUMI, Mehdi. A review of multimodal biometric systems: Fusion methods and their applications. *2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS)*. 2015, 131-136. DOI: 10.1109/ICIS.2015.7166582. Dostupné z: http://ieeexplore.ieee.org/document/7166582/

[13] RAGHAVENDRA, Ramachandra, Ashok RAO a Hemantha G. KUMAR. Multimodal person verification system using face and speech. *Proceedings of the International Conference and Exhibition on Biometrics Technology*. 2010, 181-187. DOI: 10.1016/j.procs.2010.11.023. ISSN 1877-0509. Dostupné z: http://www.sciencedirect.com/science/article/pii/S1877050910003534

[14] SOLTANE, Mohamed a Mimen BAKHTI. Multi-modal biometric authentications: concept issues and applications strategies. *International Journal of Advanced Science and Technology*. 2012, **48**. Dostupné z: https://pdfs.semanticscholar.org/8e60/8ebd80c59c4a5231aeba5bcb9d60c82e0b7e.pdf

[15] ALEKSIC, Petar S. a Aggelos K. KATSAGGELOS. Audio-Visual Biometrics. *Proceedings of the IEEE*. 2006, **94**(11), 2025-2044. DOI: 10.1109/JPROC.2006.886017. ISSN 0018-9219. Dostupné z: http://ieeexplore.ieee.org/abstract/document/4052464/

CITOVANÉ PŘÍSPĚVKY AUTORA V PRÁCI

[Tov01] TOVAREK, Jaromir, Pavol PARTILA, Jan ROZHON, Miroslav VOZNAK, Jan SKAPA, Dominik UHRIN a Zdenka CHMELIKOVA. Optimization of multilayer neural network parameters for speaker recognition. *Proceedings of SPIE - The International Society for Optical Engineering*. 2016. DOI: 10.1117/12.2223545. Dostupné z: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9850/1/Optimization-of-multilayer-neural-network-parameters-for-speaker-recognition/10.1117/12.2223545.full. Konferenční článek (SJR=0.228)

[Tov02] TOVAREK, Jaromir, Pavol PARTILA, Miroslav VOZNAK, Martin MIKULEC and Miralem MEHIC. Detection of cardiac activity changes from human speech. *Proceedings of SPIE - The International Society for Optical Engineering*. 2015. DOI: 10.1117/12.2177282. Dostupné z: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9496/94960V/Detection-of-cardiac-activity-changes-from-human-speech/10.1117/12.2177282.full. Konferenční článek (SJR=0.23)

[Tov03] PARTILA, Pavol, Jaromir TOVAREK a Miroslav VOZNAK. Self-organizing map classifier for stressed speech recognition. *Proceedings of SPIE - The International Society for Optical Engineering*. 2016. DOI: 10.1117/12.2224253. Dostupné z: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9850/1/Self-organizing-map-classifier-for-stressed-speech-recognition/10.1117/12.2224253.full. Konferenční článek (SJR=0.228)

[Tov04] PARTILA, Pavol, Miroslav VOZNAK a Jaromir TOVAREK. Pattern recognition methods and features slection for speech emotion recognition system. *The Scientific World Journal*. 2015. DOI: 10.1155/2015/573068. ISSN 2356-6140. Dostupné z: http://www.hindawi.com/journals/tswj/2015/573068/. Článek v časopise (SJR=0.32)

[Tov05] PARTILA, Pavol, Jaromir TOVAREK, Jaroslav FRNDA, Miroslav VOZNAK, Marek PENHAKER a Tomáš PETEREK. Emotional Impact on Neurological Characteristics and Human Speech. *Advances in Intelligent Systems and Computing*. DOI: 10.1007/978-3-319-07773-4_52. Dostupné z: http://link.springer.com/10.1007/978-3-319-07773-4_52. Konferenční článek (SJR=0.149)

[Tov06] PARTILA, Pavol, Miroslav VOZNAK, Tomas PETEREK, Marek PENHAKER, Vilem NOVAK, Jaromir TOVAREK, Miralem MEHIC a Lukas VOJTECH. Impact of human emotions on physiological characteristics. *Proceedings of SPIE - The International Society for Optical Engineering*. 2014. DOI: 10.1117/12.2050679. Dostupné z: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9118/1/Impact-of-human-emotions-on-physiological-characteristics/10.1117/12.2050679.full. Konferenční článek (SJR=0.237)

[Tov07] PARTILA, Pavol, Jaromir TOVAREK, Miroslav VOZNAK a Jakub SAFARIK. Classification methods accuracy for speech emotion recognition system. *Advances in Intelligent Systems and Computing*. 2014. DOI: 10.1007/978-3-319-07401-6_44. Dostupné z: http://link.springer.com/10.1007/978-3-319-07401-6_44. Konferenční článek (SJR=0.149)

[Tov08] TOVAREK, Jaromir a Pavol PARTILA. Speaker identification for the improvement of the security communication between law enforcement units. *Proceedings of SPIE - The International Society for Optical Engineering*. 2017. DOI: 10.1117/12.2261796. Dostupné z: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10200/1/Speaker-identification-for-the-improvement-of-the-security-communication-between/10.1117/12.2261796.full. Konferenční článek (SJR=0.228)

[Tov09] TOVAREK, Jaromir, Miroslav VOZNAK, Jan ROZHON, Filip REZAC, Jakub SAFARIK a Pavol PARTILA. Different approaches for face authentication as part of a multimodal biometrics system. *Advances in Electrical and Electronic Engineering*. 2017. DOI: 10.15598/aeee.v16i1.2547. Článek v časopise (přijat k vydání). (SJR=0.247)

PUBLIKAČNÍ ČINNOST AUTORA

K doložení svých vědecko–výzkumných aktivit přikládám i aktuální stav záznamů v relevantních vědeckých databázích ke dni odevzdání tohoto dokumentu.

- Publikace v bibliografické databázi ISI - Web of Knowledge: **8**
- Publikace v bibliografické databázi SCOPUS: **12**
- h-index podle ISI - Web of Knowledge: **1** (2 citace/2 bez autocitací)
- h-index podle SCOPUS: **2** (11 citací/7 bez autocitací)

Celkový počet indexovaných výstupů v Rejstříku výsledků vědy a výzkumu RIV, viz https://www.rvvi.cz : 14

- Publikační výstupy: **11**
- Aplikované výsledky : **3**

Participace na řešení projektů během studia

- **TA ČR DELTA TF01000091** – Aplikovaný výzkum: Bezpečnost mobilních zařízení a komunikace (2015-2017).
- **SP2017/174** – Specifický výzkum: Sítě a jejich bezpečnost, modelování, simulace, vytěžování znalostí a komunikační technologie pro chytrá města.
- **SP2016/170** – Specifický výzkum: Vytěžování informací z komunikačních sítí, jejich modelování a simulace II.
- **SP2015/82** – Specifický výzkum: Vytěžování informací z komunikačních sítí, jejich modelování a simulace I.