

Zpracování akustických signálů v bezpečnostních technologiích

Bc. Martin Mlčoch

Diplomová práce
2019



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
akademický rok: 2018/2019

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Martin Mlčoch**
Osobní číslo: **A17253**
Studijní program: **N3902 Inženýrská informatika**
Studijní obor: **Bezpečnostní technologie, systémy a management**
Forma studia: **prezenční**

Téma práce: **Zpracování akustických signálů v bezpečnostních technologiích**
Téma anglicky: **Acoustic Signal Processing in Security Technologies**

Zásady pro vypracování:

- 1. Zpracujte literární rešerši zaměřenou na systémy využívající akustické signály v bezpečnostních technologiích.**
- 2. Uvedte a popište základní vlastnosti zvuku využitelné pro analýzu audio signálu.**
- 3. Popište fyzikální princip šíření zvuku v prostředí.**
- 4. Vytvořte algoritmus pro analýzu audio signálu v časové a frekvenční doméně.**
- 5. Realizujte algoritmus vyhodnocující odlišnost různých audio signálů.**
- 6. Ověřte spolehlivost realizovaného algoritmu.**

Rozsah diplomové práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. HOWARD, Carl Q. a Benjamin S. CAZZOLATO. *Acoustic analyses using Matlab and Ansys*. Boca Raton: Taylor & Francis Group, 2015. ISBN 978-1-4822-2325-5.
2. LUKÁŠ, Luděk. *Bezpečnostní technologie, systémy a management I*. 1. vyd. Zlín: VeRBuM, 2011, 316 s. ISBN 978-80-87500-05-7.
3. UHLÁŘ, Jan. *Technická ochrana objektů*. vyd. Praha: Policejní akademie České republiky v Praze, 2009, 229 s. ISBN 978-80-7251-313-0.
4. *Akustika, vznik a šíření zvuku, frekvenční analýza a syntéza, sluchový vjem zvukového signálu*. Vysoká škola báňská Technická univerzita Ostrava[online]. 13 duben 2005. Dostupné z: http://homen.vsb.cz/ber30/texty/varhany/anatomie/pistaly_akustika.htm
5. PEETERS, Geoffroy. *A Large Set of Audio Features for Sound Description*[online]. 2004, , 1-25 [cit. 2018-11-22]. Dostupné z: http://recherche.ircam.fr/equipes/analyse-synthese/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf?fbclid=IwAR3iETsatY5t
6. ROBINSON, Tony. *Speech Analysis*[online]. 1998 [cit. 2018-11-22]. Dostupné z: https://www.researchgate.net/publication/2443731_Speech_Analysis
7. *Navrhujeme odolný detektor tříštění skla - 1. část*. cz[online]. 22 prosinec 2010. Dostupné z: <http://vyvoj.hw.cz/teorie-a-praxe/navrhujeme-odolny-detektor-tristeni-skla-1cast.html>

Vedoucí diplomové práce:

doc. Mgr. Milan Adámek, Ph.D.
Ústav bezpečnostního inženýrství

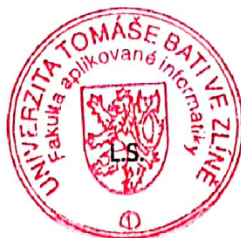
Datum zadání diplomové práce:

30. listopadu 2018

Termín odevzdání diplomové práce:

17. května 2019

Ve Zlíně dne 14. prosince 2018



doc. Mgr. Milan Adámek, Ph.D.
děkan

doc. RNDr. Vojtěch Křesálek, CSc.
ředitel ústavu

Jméno, příjmení: Martin Mlčoch

Název diplomové práce: Zpracování akustických signálů v bezpečnostních technologiích

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou práci nebo poskytnout licenci k jejímu využití jen připouští-li tak licenční smlouva uzavřená mezi mnou a Univerzitou Tomáše Bati ve Zlíně s tím, že vyrovnání případného přiměřeného příspěvku na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše) bude rovněž předmětem této licenční smlouvy;
- beru na vědomí, že pokud bylo k vypracování diplomové práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně, dne 5.5.2019

MARTIN MLČOCH, v. r.
.....
podpis diplomanta

ABSTRAKT

Diplomová práce je zaměřena na problematiku využití rozpoznávání zvuku v bezpečnostních technologiích. Hlavním úkolem práce je seznámit se se základními principy šíření, analýzou a využitím akustických signálů. Teoretická část je zaměřena na fyzikální princip šíření zvuku v prostředí a na možnosti využití rozpoznávání zvuku v bezpečnostních technologiích. Dále jsou popsány charakteristické vlastnosti zvuku využitelné pro analýzu audio signálu. V praktické části jsou vytvořeny algoritmy pro popis audio signálu v časové a frekvenční doméně a algoritmus schopný odlišit různé audio signály. Algoritmus je otestován a je stanovena jeho úspěšnost rozpoznání audio signálu.

Klíčová slova:

Zpracování signálu, audio signál, fyzikální princip, algoritmus, biometrický systém, digitalizace, parametry, analýza, klasifikace, matlab.

ABSTRACT

The diploma thesis is focused on problematics of using sound recognition systems in security technologies. The main task is to acquaint with basic principles of spreading, analysis and utilization of acoustic signals. The theoretical part is focused on the physical principle of sound spreading in the environment and the possibility of using sound recognition systems in security technologies. Further, there are described the characteristic features of sound usable for analysis of audio signal. In the practical part there are developed algorithms for description of audio signals in time and frequency domain, and also algorithm able to distinguish different audio signals. The algorithm is tested and there is determined its success of the audio signal recognition.

Keywords:

Signal processing, audio signal, physical principle, algorithm, biometric system, digitization, features, analysis, classifier, matlab.

Poděkování

Tímto chci poděkovat svému vedoucímu panu doc. Mgr. Milanu Adámkovi, Ph.D. za vedení, odborné rady a poskytnutý čas, při zpracování mé diplomové práce.

Prohlašuji, že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

ÚVOD.....	9
I TEORETICKÁ ČÁST.....	10
1 VYUŽITÍ SYSTÉMŮ ROZPOZNÁVÁNÍ ZVUKU V BEZPEČNOSTNÍCH TECHNOLOGIÍCH	11
1.1 DETEKTORY ROZBITÍ SKLA	11
1.1.1 Kontaktní detektory rozbití skla.....	11
1.1.2 Bezkontaktní detektory rozbití skla	12
1.1.2.1 Zpracování akustického signálu detektorem.....	13
1.1.2.2 Algoritmus vyhodnocení akustické události.....	16
1.2 BIOMETRICKÉ SYSTÉMY VYUŽÍVAJÍCÍ ROZPOZNÁVÁNÍ HLASU.....	17
1.2.1 Biometrické systémy	18
1.2.1.1 Identifikace	19
1.2.1.2 Verifikace.....	20
1.2.2 Hlasové rozpoznávání	20
1.2.2.1 Rozpoznávání řeči.....	20
1.2.2.2 Rozpoznávání mluvěcího	21
1.2.2.3 Speciální rozpoznávání	21
1.2.2.4 Postup zpracování řečového signálu a rozpoznání uživatele.....	21
1.2.2.5 Charakteristiky hlasového rozpoznávání	22
1.3 AUTENTIZACE NA ZÁKLADĚ ROZPOZNÁVÁNÍ HLASU V BANKOVNÍM SEKTORU	22
1.3.1 Průběh hlasového ověření	23
1.3.2 Bezpečnost hlasového ověření	23
1.4 ROZPOZNÁVÁNÍ ZVUKU VE FORENZNÍM VYŠETŘOVÁNÍ.....	24
1.4.1 Objekty kriminalistické audio analýzy.....	25
2 ZPRACOVÁNÍ SIGNÁLŮ.....	27
2.1 ZÁZNAM ZVUKU	27
2.1.1 Piezoelektrický mikrofon	27
2.1.2 Elektretový mikrofon	28
2.2 FILTRACE SIGNÁLU.....	29
2.2.1 Frekvenčně selektivní filtry.....	29
2.2.2 Analogové filtry	30
2.2.3 Číslicové filtry.....	30
2.2.4 Antialiasing filtr	32
2.3 DIGITALIZACE	32
2.3.1 Vzorkování signálu	32
2.3.2 Kvantování signálu.....	33
2.4 FOURIEROVA TRANSFORMACE.....	34
2.4.1 Diskrétní Fourierova transformace.....	34
3 ZÁKLADNÍ PARAMETRY ZVUKU PRO AUDIO ANALÝZU	36
3.1 PARAMETRIZACE.....	36
3.2 ODSTRANĚNÍ STEJNOSMĚRNÉ SLOŽKY.....	36
3.3 ROZDĚLENÍ NA RÁMCE	37
3.3.1 Okénkové funkce	37

3.4	SPEKTROGRAM.....	37
3.5	STŘEDNÍ KRÁTKODOBÁ ENERGIE.....	37
3.6	KORELACE	37
3.7	POČET PRŮCHODŮ NULOU	38
3.8	FORMANTOVÉ FREKVENCE	38
3.9	CEPTSTRUM.....	38
3.10	MEL-FREKVENČNÍ SPEKTRÁLNÍ KOEFICIENTY	39
3.11	LINEÁRNÍ PREDIKCE	40
3.12	DALŠÍ PARAMETRY VHODNÉ PRO AUDIO ANALÝZU	41
4	STROJOVÉ UČENÍ.....	45
4.1	SKRYTÉ MARKOVOVY MODELY	45
4.2	SUPPORT VECTOR MACHINE	46
4.2.1	Lineární rozdělení	47
4.2.2	Nelineární rozdělení	47
4.3	NEURONOVÉ SÍTĚ	47
5	FYZIKÁLNÍ PRINCIP PRO ZPRACOVÁNÍ AKUSTICKÝCH SIGNÁLŮ.....	49
5.1	AKUSTIKA	49
5.1.1	Šíření zvuku v prostoru	50
5.1.2	Šíření zvuku v pevných látkách a kapalinách	51
5.2	VZNIK MLUVENÉHO SLOVA	52
II	PRAKTICKÁ ČÁST	54
6	ALGORITMUS PRO ANALÝZU AUDIOSIGNÁLU	55
6.1	MATLAB.....	55
6.2	ANALÝZA AUDIOSIGNÁLU	55
7	ALGORITMUS PRO ROZPOZNÁVÁNÍ ZVUKŮ	59
7.1	EXTRAKCE A VÝBĚR REFERENČNÍCH PARAMETRŮ.....	59
7.2	KLASIFIKACE.....	63
7.2.1	Algoritmus pro rozpoznání zvuků typických pro vloupání do objektu.....	63
7.2.1.1	Použité parametry pro analýzu záznamů	64
7.2.1.2	Výstup z algoritmu pro rozpoznání zvuků typických pro vloupání do objektu	65
7.2.2	Simulace detektoru rozbití skla.....	65
7.2.2.1	Použité parametry pro simulaci detektoru rozbití skla	67
7.2.2.2	Výstup z algoritmu pro simulaci detektoru rozbití skla.....	67
8	ZHODNOCENÍ EFEKTIVITY VYTVOŘENÝCH ALGORITMŮ.....	68
8.1.1	Efektivita algoritmus pro rozpoznání zvuků typických pro vloupání do objektu.....	68
8.2	EFEKTIVITA ALGORITMU PRO SIMULACI DETEKTORU ROZBITÍ SKLA	71
8.2.1	Testování vzniku falešných poplachů při běžném provozu	71
8.2.2	Testování reakce na zvuky rozbití skla	72
	ZÁVĚR	73
	SEZNAM POUŽITÉ LITERATURY.....	75

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....	79
SEZNAM OBRÁZKŮ	80
SEZNAM TABULEK.....	82
SEZNAM PŘÍLOH.....	83

ÚVOD

S rozvojem společnosti se stále zvětšuje objem soukromého vlastnictví a s tím i potřeba ochrany majetku. Technologický vývoj přináší mnoho možností, jak zajistit ochranu majetku. Tyto možnosti jsou stále sofistikovanější a jejich vývoj a následný přenos technologií do komerčního prostředí je náročnější.

Systemy rozpoznávání zvuku jsou v průmyslu bezpečnostních technologií stále poměrně novinkou. Základní aplikaci představují elektroakustické detektory. Tyto detektory jsou důležitou součástí plášťové ochrany objektu. Stále větší části plášťů budov jsou tvořena skleněnými výplněmi. Detektory rozbití skla jsou určeny k detekci pokusu o vniknutí skrz tyto skleněné výplně. Další významnou pokročilejší aplikací je autentizace osob pomocí hlasu. Tato technologie bývá začleňována do systému kontroly vstupů pro použití v bezpečnostních aplikacích. Standartní přístupové systémy jsou postaveny na znalosti hesla nebo vlastnictví identifikační karty. Nevýhodou tohoto způsobu ověření je, že heslo může být prolomeno a identifikační karta ukradena. Proto se stále častěji v systémech kontroly vstupů používá biometrické identifikační prvky. Do této kategorie patří i rozpoznávání pomocí hlasu. Hlas každého člověka je unikátní a nemění se až do smrti. Hlasová identifikace se dále může využít při forenzním vyšetřování. Další uplatnění nacházejí tyto technologie v bankovním sektoru k autentizaci uživatele prostřednictvím hlasového asistenta nebo k autorizaci uživatele k určitým bankovním úkonům.

Ve své diplomové práci se zaměřím na problematiku aplikace systémů rozpoznávání zvuku v bezpečnostních technologiích. Nejprve se budu zabývat konkrétními aplikacemi rozpoznávání zvuku v bezpečnostních technologiích. Patří zde elektroakustické detektory, systémy kontroly vstupu využívající biometrické ověření a komerční aplikace zahrnující bankovní sektor. Déle se zaměřím na parametry zvuku, pomocí kterých lze rozlišit jeden zvuk od jiných zvuků. Tyto parametry jsou extrahovány z digitální podoby audio záznamu a charakterizují signál v časové a frekvenční doméně. Pro práci se zvukem je nutné pochopit fyzikální principy akustiky, kterým se budu také věnovat. Hlavním cílem diplomové práce je vytvořit algoritmus, který bude schopný odlišit podle vlastností získaných analýzou záznamů různé audio signály. Algoritmy jsou vytvořeny pomocí programovacího prostředí softwaru Matlab. Součástí řešení jsou výstupy zobrazující jednotlivé vlastnosti audio signálu ve frekvenční a časové doméně. Výsledkem audio analýzy je stanovení úspěšnosti algoritmu při rozpoznávání audio signálů.

I. TEORETICKÁ ČÁST

1 VYUŽITÍ SYSTÉMŮ ROZPOZNÁVÁNÍ ZVUKU V BEZPEČNOSTNÍCH TECHNOLOGIÍCH

Využití analýzy zvuku má poměrně široké možnosti uplatnění. Rozpoznávání zvuku se stále častěji rozšiřuje nejen do oblasti bezpečnostních technologií, ale i do oblastí běžného života. Typickým příkladem využití je ovládání různých zařízení prostřednictvím hlasových povelů. V bezpečnostních technologiích se za nejjednodušší aplikaci dají považovat elektroakustické detektory. Mezi sofistikovanější aplikace patří biometrické systémy rozpoznávání hlasu, které se typicky implementují do systémů kontroly vstupů. Tyto systémy nacházejí své uplatnění i v bankovním sektoru k autentizaci uživatele k různým úkonům.

1.1 Detektory rozbití skla

Detektory rozbití skla se používají k detekci vniknutí narušitele skrz skleněné výplně pláštěů budov. Detektory v závislosti na typu reagují na mechanické poškození skleněných ploch. Typicky se tento typ detektorů aplikuje na skleněné okenní, dveřní výplně a výlohy obchodů. Detektor reaguje na akustické směny v prostředí. V případě kontaktní varianty detektoru je to detekce vibrací způsobených například řezáním skla diamantovým nožem nebo samotným rozbitím skleněné výplně. Bezkontaktní varianta detektoru reaguje na akustické frekvence typické pro náraz předmětu na skleněnou plochu a následné rozbití skleněné výplně. V tomto případě se zvuk šíří vnitřním prostorem střeženého objektu. Detektory rozbití skla se dělí podle toho, jak působí na střežený prostor na aktivní a pasivní. Podle principu funkce jsou pak detektory k dostání v kontaktní a bezkontaktní variantě.

1.1.1 Kontaktní detektory rozbití skla

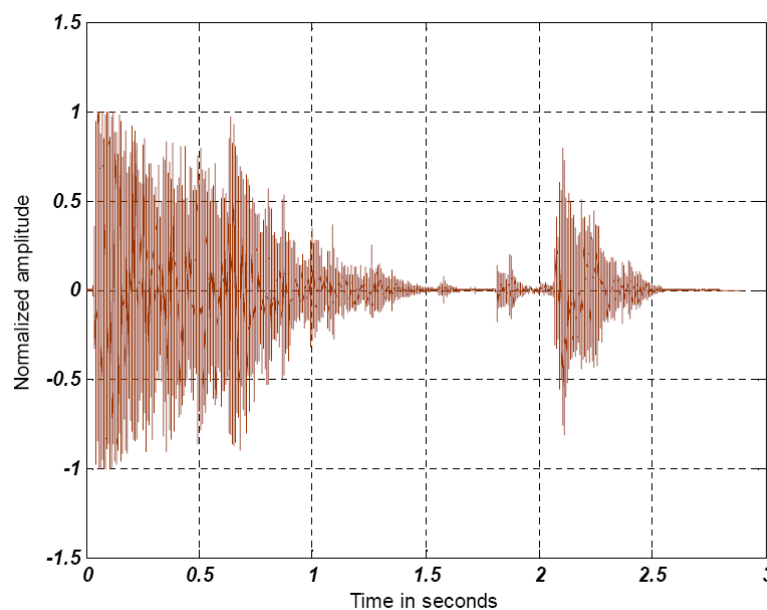
Tento typ detektorů je založen na detekci vibrací nebo rozbití skleněné výplně na základě piezoelektrického jevu. Detektory jsou k dispozici jako pasivní i aktivní. V případě pasivního typu detektoru rozbití skla je piezo snímač umístěn na povrchu skleněné výplně a je nastaven na rezonanční kmitočet v pásmu 40–120 kHz. Tyto frekvence jsou typické pro mechanické poškození skleněné plochy (řezání skla diamantem, praskání skla, rozbití skleněné výplně), které se šíří pevnými látkami jako podélné a příčné vlnění. V případě narušení skleněné plochy je vlnění vyhodnoceno a následně dojde k vyhlášení poplachu. Detektory se při instalaci umísťují ve vzdálenosti 50 mm od rámu okna a jejich detekční schopnost je omezena do 1,5–3 m. Při potřebě zabezpečit větší skleněnou plochu je tedy nutné využít většího počtu detektorů. [1]

Kontaktní detektory rozbití skla jsou k dostání i v aktivním provedení. Detektor se v tomto případě skládá z vysílací a přijímací části. Vysílač generuje ultrazvukové vlnění, které je na druhé straně skleněné plochy přijímáno přijímačem. V přijímači je vyhodnocována změna ultrazvukového signálu. V případě změn v povrchové struktuře skleněné plochy se mění i ultrazvukové vlnění. Tento typ detektoru je schopný spolehlivě pokrýt skleněnou výplň o ploše až 25 m². [1]

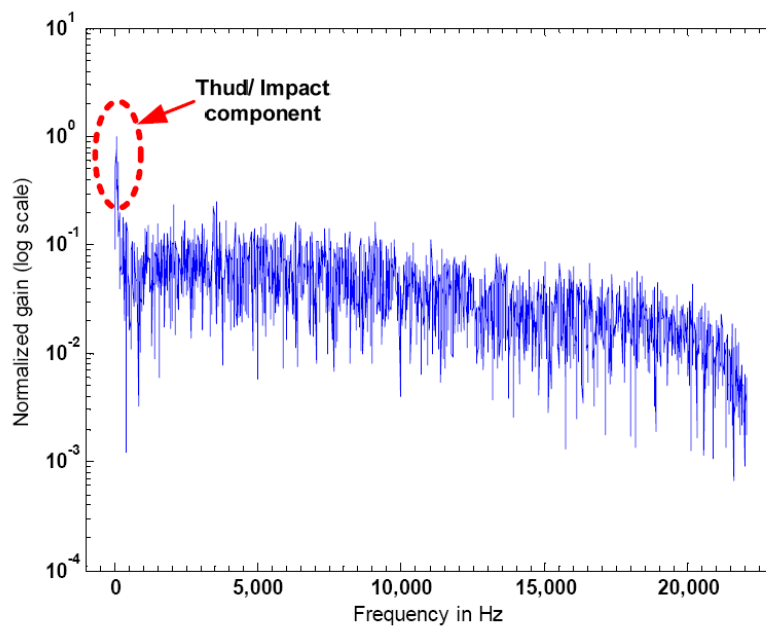
1.1.2 Bezkontaktní detektory rozbití skla

Bezkontaktní detektory jsou vyráběny v pasivním a aktivním provedení. Funkce pasivní varianty detektoru je založena na snímání vlnění, které se šíří prostorem od střežené skleněné plochy až k samotnému detektoru. Zachycení vlnění šířící se prostorem je realizováno prostřednictvím mikrofonu. V praxi se používá piezoelektrický nebo elektretový mikrofon. Kvalitnější příjem zvuku poskytuje elektretový mikrofon, který je schopný zaznamenat zvuk o frekvenci 20 Hz–20 kHz. Po přijetí akustického vlnění pomocí mikrofonu následuje soustava filtrů zajišťující průchod jen té části spektra, která je typická pro rozbití skleněné plochy. [1]

Poplach je pak spuštěn na základě detekce nízkofrekvenčního krátkého intenzivního úderu předmětu na skleněnou plochu a následného rozbití skleněné tabule. Pro úder na skleněnou plochu jsou typické frekvence 100–300 Hz. [2]



Obr. 1 Časový průběh akustického signálu rozbití skleněné výplně. [3]



Obr. 2 Frekvenční spektrum akustického signálu rozbití skleněné výplně. [3]

V druhé fázi následuje vlivem úderu na skleněnou plochu její rozbití. Dochází k tříštění a lámání skleněné tabule a střepy dopadají na podlahu. Pro tuto fázi je charakteristický delší časový průběh, akustické vlnění má nižší intenzitu a frekvence se pohybuje mezi 12–15 kHz. [2]

Druhým provedením bezkontaktního detektoru rozbití skla je aktivní. Aktivní detektor rozbití skla se skládá z vysílací a přijímací části. Detektor pracuje v infračerveném nebo ultrazvukovém spektru. Fyzikální princip detekce je založen na Dopplerově jevu, kdy se vyhodnocuje změna frekvence vlnění odraženého od skleněné plochy. Detektor má v paměti uložené vzorky frekvencí odpovídající řezání nebo rozbití skleněné výplně. Tyto vzorky jsou porovnávány s frekvencí odražené vlny. V případě, že dojde ke shodě těchto frekvencí se vzorky, je vyhlášen poplach. [2]

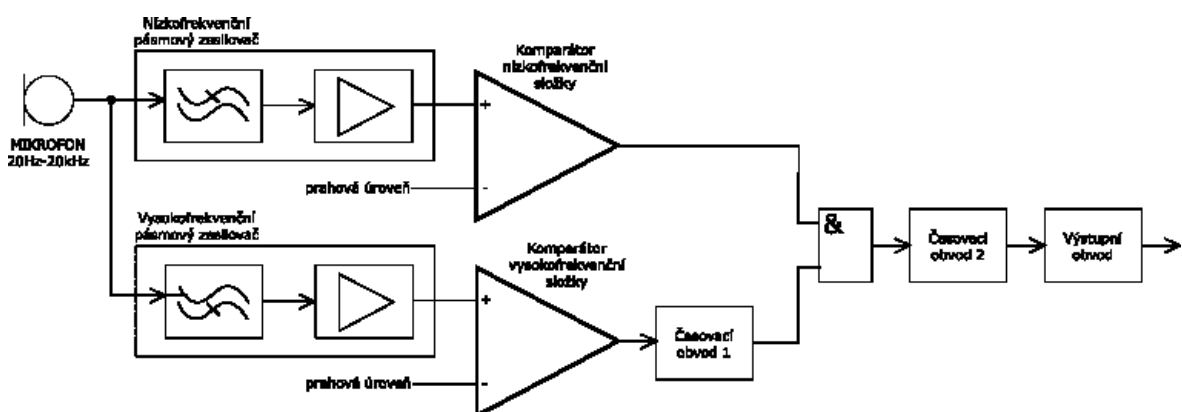
1.1.2.1 Zpracování akustického signálu detektorem

Zpracování akustického signálu může být realizováno dvěma způsoby. Signál může být zpracován analogově nebo digitálně. Využívají se buď jednopásmové nebo vícepásmové systémy. V případě jednopásmových systémů se analyzuje akustický signál pouze na základě frekvenčního spektra a je porovnáván se vzorky uloženými v paměti detektoru. Nevýhodou těchto systémů je vysoké riziko vzniku planých poplachů. Poplach zde mohou vyvolat

i zvuky, které jsou z hlediska frekvence podobné rozbití skla, jako zvonek nebo provoz na ulici. Více pásmové systémy nejsou tolik náchylné na vznik falešných poplachů. Akustický signál se nezkoumá jen na základě frekvenčního spektra, ale také je měřena jeho intenzita a časová skladba signálu. Skladba signálu je analyzována z hlediska přítomnosti úderu předmětu na skleněnou plochu následovaném jejím rozbitím. K vyhlášení poplachu je potřeba přítomnost obou zvukových projevů charakteristických pro rozbití skleněné výplně s danými vlastnostmi.

Algoritmus pro vyhodnocení akustického signálu detektorem musí být schopen bezchybně odlišit rozbití skleněné plochy od dalších podobných zvuků. Zvuk, který je charakteristický pro rozbití skla závisí na typu skleněné výplně, její tloušťce, předmětu použitým pro její rozbití a akustickém prostředí. Algoritmy používané pro vyhodnocení zvuku analyzují akustický signál z hlediska časové a frekvenční skladby. [3]

Analogové zpracování vlnění je založeno na dvoufázovém zpracování. Nejprve je pomocí mikrofonu zachycen krátký nízkofrekvenční impuls typický pro úder předmětu na skleněnou výplň. Tento signál je zesílen pomocí pásmového zesilovače a přiveden do komparátoru, kde je porovnán s prahovou hodnotou frekvence typickou pro úder na skleněnou plochu. Následně je očekáváno rozbití skleněné plochy charakteristické vysokofrekvenčním signálem. Tento signál je také zesílen a porovnán s prahovými hodnotami pomocí komparátoru. Pokud je přítomna nízkofrekvenční i vysokofrekvenční složka je na výstupu generován impuls k spuštění poplachu. [2]



Obr. 3 Schéma zpracování akustického signálu analogovým obvodem.

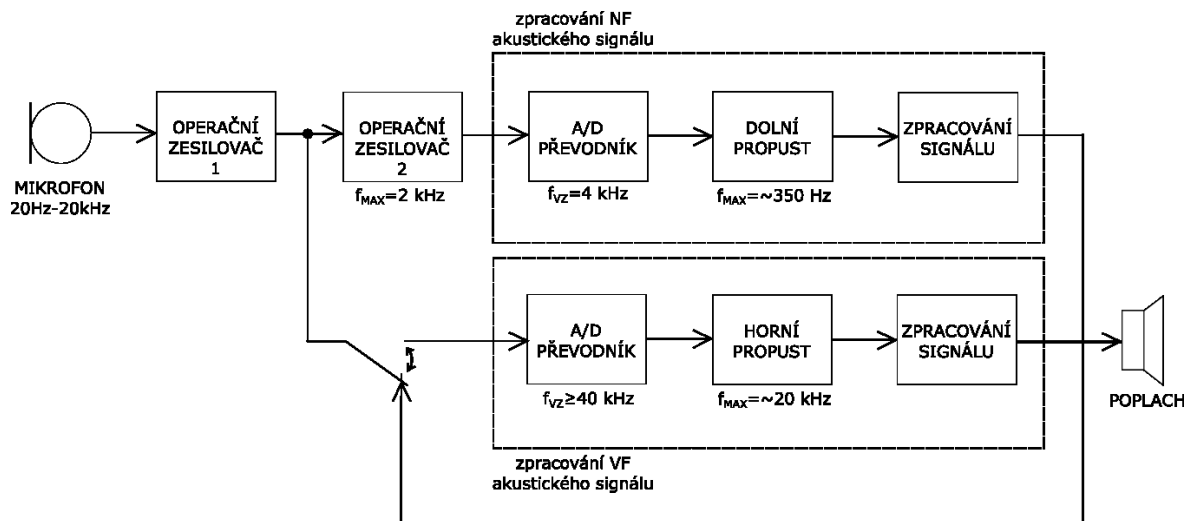
Nevýhodou analogového zpracování signálu je, že zvuky podobné rozbití skla mohou způsobit falešný poplach. To je způsobeno pásmovými zesilovači, které mají neměnnou frekvenční útlumovou charakteristiku. Proto mohou být zvuky jako zvonění domovního zvonku,

pohyb vozidel v okolí objektu nebo například vhazování sklenic do kontejneru vyhodnoceny jako narušení skleněné výplně střeženého prostoru. [2]

Naproti tomu digitální zpracování akustického signálu vykazuje daleko větší odolnost proti vzniku falešných poplachů. Princip zpracování akustického signálu je v tomto případě založen na jeho převodu do digitální formy. Digitální zpracování signálu je graficky zobrazeno na obr. 4. Ke zpracování digitálního signálu se používají mikrokontrolery (MCU). MCU jsou levné procesory využívající se v jednoduchých aplikacích např. digitální hodiny nebo měřicí přístroje. Jejich výhodou je nízká cena a malé nároky na spotřebu.

Vstupním zařízením pro snímání akustického signálu je mikrofon schopný zaznamenat zvuk v rozsahu od 20 Hz do 20 kHz. Nejčastěji používaným typem je elektretový mikrofon. Analogový signál zaznamenaný mikrofonem je pak zesílen pomocí operačního zesilovače. V první fázi je na vstupu očekáván intenzivní nízkofrekvenční akustický signál charakteristický pro úder předmětu na skleněnou plochu. Nízkofrekvenční signál je nejprve zesílen pomocí druhého operačního zesilovače, kterým též slouží jako dolní propust. Dolní propust v tomto případě filtruje signály o frekvenci větší než 2 kHz. Následně je signál převeden do digitální formy pomocí A/D převodníku se vzorkovací frekvencí minimálně 4 kHz. Musí zde být dodržen Shannonův teorém, kdy vzorkovací frekvence musí být vyšší než dvojnásobek nejvyšší harmonické složky vzorkovaného signálu. Náráz předmětu na plochu má frekvenci okolo 300 Hz a nachází se vždy na začátku celého jevu rozbití skleněné plochy. Proto je za A/D převodních zařazen dolnoproputný filtr, který k dalšímu zpracování propustí jen frekvence nižší než 350 Hz. Takto získané vzorky jsou pomocí MCU analyzovány z hlediska intenzity daného signálu. Vzorky jsou porovnány s prahovými hodnotami intenzity a v případě jejich překročení je spuštěna druhá část algoritmu pro analýzu vysokofrekvenčního signálu. Tento signál je charakteristický pro rozbití skleněné plochy, dopad střepů na podlahu a jejich lámání. Tato fáze následuje bezprostředně po nárazu předmětu na skleněnou plochu. Je pro ni charakteristická nižší intenzita signálu, frekvence se pohybuje mezi 12–15 kHz a větší délka tohoto jevu. Signál je opět převeden do digitální formy pomocí A/D převodníku. Vzorkovací frekvence zde musí být větší nebo rovna 40 kHz. Horní propust zajišťuje zpracování těchto signálů, které mohou obsahovat rozbití skleněné výplně. Samotné vyhodnocení signálu pak zahrnuje průměrování signálu, detekci průchodu nulou a detekci špiček. Algoritmus pracuje v každém cyklu s přibližně 2400 vzorky signálů. Nejprve se vyhodnocuje poměr energie signálu představující tříštění skla s energií celého zpracovávaného signálu. Tento poměr se pohybuje mezi 1,75 a 14. Druhým krokem je zjištění počtu špiček

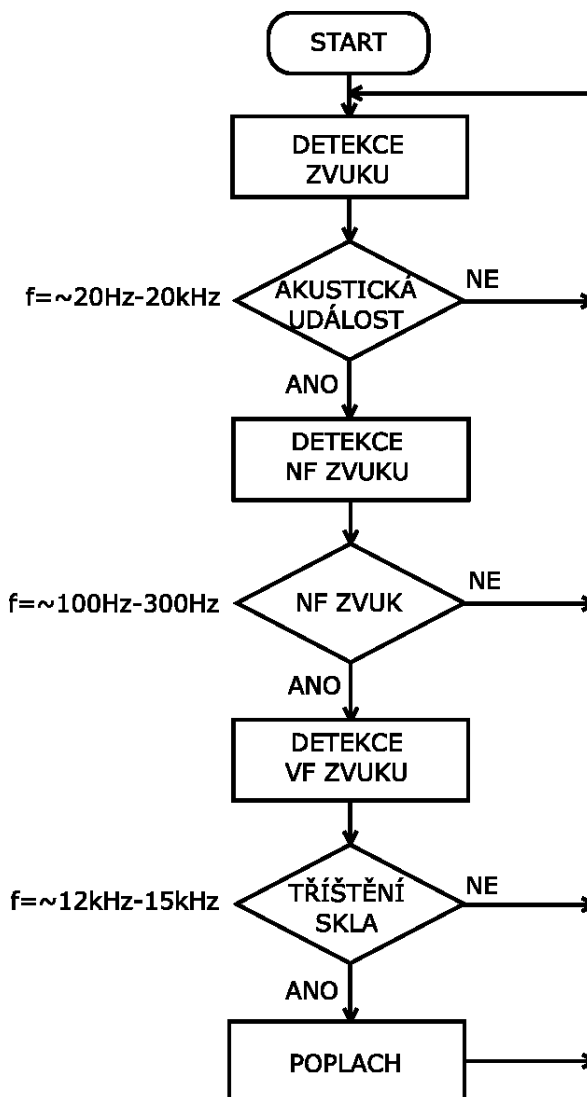
signálu, který se porovnává s prahovou hodnotou. Tento počet by se měl pohybovat mezi 160 a 320. Nakonec se zkoumá počet průchodů signálu nulou. Ten by se měl v případě tříštění skla pohybovat mezi 95 a 300. Pokud jsou všechny tyto tři podmínky splněny můžeme říct, že daný analyzovaný signál zachycuje rozbití skleněné výplně a je vyhlášen poplach. [3], [4]



Obr. 4 Průběh digitálního zpracování akustického signálu detektorem. [5]

1.1.2.2 Algoritmus vyhodnocení akustické události

Při detekce a vyhodnocení rozbití skla probíhá podle následujícího vývojového diagramu viz obr. 5. V první části mikrofon neustále sleduje akustické pozadí v místnosti. Pokud je zaznamenána aktivita je zvuk vyhodnocen nejprve na přítomnost úderu předmětu na skleněnou plochu. Pro tento zvuk je typická vysoká intenzita a frekvence mezi 100–300 Hz. Pokud je identifikace pozitivní je aktivována druhá část algoritmu a zvukový signál je zkoumán z hlediska přítomnosti vyšších frekvencí charakteristických pro rozbití skleněné plochy. Pro tento zvuk je typická nižší intenzita, větší délka akustického efektu a frekvence mezi 12–15 kHz. Pokud je i tento zvuk identifikován, pak je vyhlášen poplach. [4]



Obr. 5 Vývojový diagram pro vyhodnocení akustické události detektorem. [5]

1.2 Biometrické systémy využívající rozpoznávání hlasu

Systémy využívající ověření identity člověka na základě nějakého identifikačního prvku usilují o co největší bezpečnost a zároveň komfort uživatele. Uživatel je sám o sobě nositelem identifikačního prvku a odpadá tedy nutnost pamatovat si přístupový pin nebo se starat o čipovou kartu. Bezpečnost je jako u všech identifikačních systémů závislá hlavně na obezřetnosti uživatele. Typickou aplikací biometrických metod jsou systémy kontroly vstupů.

Systémy kontroly vstupu řídí přístup osob, resp. vozidel do chráněných prostorů nebo ke chráněným zařízením, případně informacím, na základě přidělených přístupových práv. Tato zařízení umožňují sledovat pohyb osob v definovaných prostorových zónách. [6]

1.2.1 Biometrické systémy

Biometrické systémy slouží k rozpoznávání identity člověka na základě jeho charakteristických anatomických a behaviorálních rysů.

Výhodami biometrie je nepřenositelnost rysů, pohodlí, uživatel nemůže ztratit identifikační prvek. Identifikační prvky je složitější zcizit.

Mezi anatomické rysy člověka používané pro biometrickou identifikaci patří:

- Otisk prstu,
- obličej,
- oční duhovka,
- oční sítnice,
- geometrie ruky,
- dlaň,
- termogram obličeje,
- termogram ruky,
- dentální obraz,
- podpis,
- tvar ucha,
- snímek nehtu,
- DNA.

Mezi behaviorální rysy člověka používané pro biometrickou identifikaci patří:

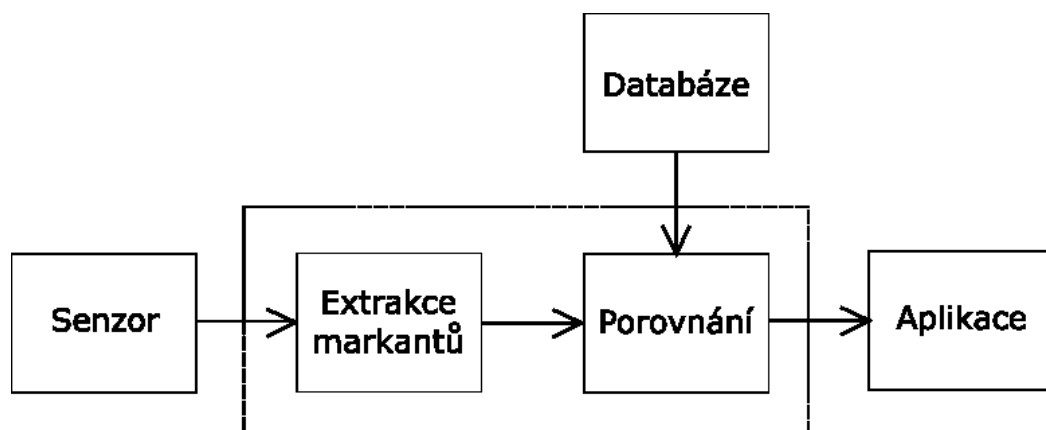
- Hlas,
- mimika obličeje a pohyb rtů,
- dynamika podpisu
- chůze.

Anatomické vlastnosti člověka jsou pevně dané už od narození. Zůstávají neměnné po celou délku života. Analýza anatomických vlastnosti člověka se také označuje jako statická metoda. Behaviorální vlastnosti vyžadují provedení nějaké akce. Tyto vlastnosti mohou být ovlivněny aktuálním stavem daného člověka (nemoc, zranění). Analýza behaviorálních vlastnosti člověka se označuje jako dynamická metoda. V praxi se používají unimodální nebo modální biometrické systémy. Unimodální systémy používají k autentizaci uživatele jen jednu biometrickou vlastnost. Tyto systémy jsou nejvíce rozšířené. Jejich výhodou jsou

nízké pořizovací náklady. Na druhou stranu jsou méně spolehlivé a méně odolné proti útokům. Multimodální systémy vykazují vyšší spolehlivost, protože k autentizaci uživatele využívají více než jednu biometrickou vlastnost. Nevýhodou jsou vyšší pořizovací náklady, jsou ale odolnější proti útokům.

Použité biometrické vlastnosti musí být:

- Univerzální,
- jedinečné,
- konstantní,
- získatelné,
- výkonné,
- akceptovatelné,
- odolné proti falšování. [7]



Obr. 6 Schéma biometrického systému.

1.2.1.1 Identifikace

Na vstupu uživatel poskytne biometrickému systému požadovaný biometrický prvek. Tento prvek je pak porovnán s celou databází vzorků. Systém provede rozpoznání identity uživatele. Výstupem je tedy pozitivní nebo negativní ověření identity uživatele za pomoci poskytnutého biometrického údaje a vzorků uložených v databázi. [7]

1.2.1.2 Verifikace

V případě verifikace uživatel zadá na vstupu požadovaný biometrický prvek, a navíc i identitu s kterou má být tento prvek porovnán. Vstupní biometrický údaj je porovnán s konkrétním vzorkem v databázi a na základě srovnání je buď potvrzena nebo zamítnuta identita uživatele. [7]

1.2.2 Hlasové rozpoznávání

Hlasové rozpoznávání je technologie, která přináší možnost ovládní přístrojů pomocí řeči. Hlavním cílem je tedy usnadnění života lidí. Hlasové ovládní může značně usnadnit ovládní přístrojů zdravotně postiženým lidem. Nejdostupnější aplikací ovládní pomocí hlasu představují hlasový asistenti integrovaní do mobilních technologií. V bezpečnostních technologiích se hlasové rozpoznávání nejčastěji využívá v systémech kontroly vstupů.

Rozpoznávání řeči je založeno na použití matematického aparátu. Uplatňují se zde statistické a pravděpodobnostní metody. Nejobtížnější je výběr správných parametrů řečového signálu, které by jednoznačně dokázali rozpoznat každý signál. Lidská řeč je velmi různorodá a každý jazyk má svá specifika. Navíc každý mluvčí interpretuje řečový signál unikátním způsobem. Často ani prostředí, ve kterém se provádí rozpoznání hlasu není ideální. Můžou se zde nacházet různé rušivé vlivy, které mohou analýzu hlasu znesnadnit.

Rozpoznávání hlasu se dělí do tří oblastí:

- Rozpoznávání řeči,
- rozpoznávání mluvčího,
- speciální rozpoznávání. [7]

1.2.2.1 Rozpoznávání řeči

Rozpoznávání řeči je zaměřeno na obsahovou stránku řečového signálu. Zabývá se rozpoznáváním plynulé řeči a rozpoznáváním s porozuměním významu. Úkolem je rozeznat jednotlivé hlásky, slova, věty a určit jejich význam. Rozpoznávání řeči se používá v aplikacích schopných na základě hlasových pokynů provádět různé úkony. Nejjednodušším příkladem může být hlasové vytáčení. Nejtěžším úkolem je porozumění významu řeči. Klíčové je, aby daná aplikace byla schopná obsáhnout gramatiku, syntaxi a sémantiku daného jazyka. [7]

1.2.2.2 Rozpoznávání mluvího

V závislosti na aplikaci je požadována verifikace nebo identifikace uživatele. Verifikace je založena na ověření identity uživatele na základě mluveného slova. Ověření identity může být textově závislé nebo nezávislé. Textově závislé ověření vyžaduje, aby uživatel vyslovil předem dané slovo nebo více slov. Při textově nezávislém ověření není vyžadován konkrétní obsah. V praxi se z verifikací uživatele na základě hlasu setkáme například v aplikacích, kde není možný jiný způsob identifikace uživatele nebo se využívá v kombinaci s jiným způsobem ověření. Příkladem je ověření identity v telefonickém bankovníctví. Pro zákazníka je pohodlné mít možnost provádět bankovní transakce pomocí mobilního telefonu. V tomto případě je nutné zajistit ověření, že osoba, která chce provést daný bankovní úkon má opravdu potřebné oprávnění. Chybná identifikace zákazníka by v tomto případě mohla vést k peněžním ztrátám. Dalším příkladem je použití v systémech kontroly vstupů do střežených objektů a prostor.

Identifikace vyžaduje, aby byl hlasový projev uživatele porovnán s celou databází hlasových vzorů. Cílem je identifikovat daného člověka. Tento požadavek je často kladen v případech využití hlasové identifikace při forenzní audio analýze. Může jít například o identifikaci vyděrače z telefonického záznamu. Případný pachatel se může snažit měnit svůj hlasový projev, a tak ztížit jeho identifikaci. Při analýze záznamu je třeba zaměřit se na ty parametry audio nahrávky, které nemůže člověk změnit. [7]

1.2.2.3 Speciální rozpoznávání

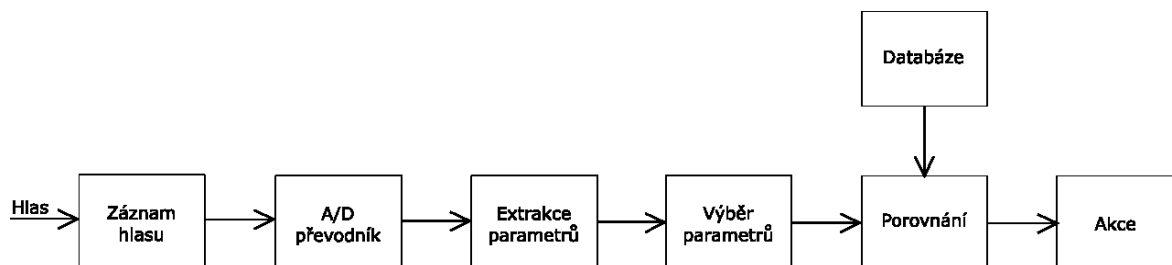
Úkolem speciálního rozpoznávání na základě hlasu daného jedince je zjistit jeho aktuální stav. Jedná se o charakteristické znaky hlasového projevu, které mohou vypovídat o jeho psychickém a fyzickém stavu. Analýzou hlasu je možné zjistit, zda byl jedinec pod vlivem omamné látky. Při vyšším množství alkoholu v krvi dochází ke změnám hlasového projevu, které nemůže člověk ovlivnit. Z psychologického hlediska lze u daného jedince zjistit náladu nebo zda je pod vlivem stresu. [7]

Na základě řečových projevů lze také zjistit jeho národnost, vzdělání, inteligenci, sociální prostředí nebo probíhající nemoc.

1.2.2.4 Postup zpracování řečového signálu a rozpoznání uživatele

Hlasový signál se šíří prostorem jako změna akustického tlaku. Tyto změny je třeba zachytit a převést z analogové do digitální formy. Rozpoznání hlasu se provádí na základě extrakce

parametrů z digitalizovaného signálu. Musí být vybrány takové parametry, které umožňují jednoznačně charakterizovat daný vzorek. Zvolené parametry jsou pak porovnány se vzorky uloženými v databázi.



Obr. 7 Schéma zpracování hlasového signálu.

1.2.2.5 Charakteristiky hlasového rozpoznávání

Rozhodujícím faktorem každé biometrické technologie je její akceptace uživateli. To znamená, zda je veřejnost ochotná danou technologii využívat. Hlavními hledisky jsou její bezpečnost, hygienické požadavky, časová náročnost a uživatelská přívětivost. Hlas je pro člověka přirozený komunikační nástroj, a proto je tato vysoce akceptovatelná.

Z hlediska spolehlivosti patří systémy založené na rozpoznávání hlasu k těm méně spolehlivým. To je způsobeno charakterem řečových signálů. Hlas konkrétního člověka není vždy stejný. Na hlasový projev může mít vliv například aktuální psychické rozpoložení nebo nemoc. [7]

1.3 Autentizace na základě rozpoznávání hlasu v bankovním sektoru

Hlasová biometrie představuje jeden ze způsobů, kterým lze ověřit identitu uživatele prostřednictvím telefonu. Potřeba ověřování identity zákazníka je důležitá pro bankovní sektor. Komunikace zákazníka s bankou probíhá běžně po telefonu a je potřeba zajistit, aby s bankovním účtem mohl nakládat pouze jeho majitel. Běžně se k ověření uživatele po telefonu používá pin kód zákazníka případně ověření znalosti určité informace. Problém je, že tímto nelze zcela spolehlivě ověřit identitu klienta ale prakticky se zjišťuje jen zdali má daná osoba k dispozici požadované informace. Pokud se k těmto informacím dostane někdo jiný než majitel účtu, může provádět pomocí zákaznické linky bankovní úkony. Řešením tohoto nedostatku je využití systému rozpoznávání hlasu „speaker recognition system“. Přínosem je zvýšení pohodlí klienta a rychlosti ověření jeho identity. [8]

1.3.1 Průběh hlasového ověření

V prvé řadě se provede záznam hlasu z prvního telefonického rozhovoru s operátorem, z kterého se vytvoří hlasový otisk (voiceprint). Tento otisk je uložen do systému a přiřazen k danému klientovi. Při další komunikaci s klientem je tento otisk porovnán s otiskem získaným při hovoru. Pokud odpovídají parametry hlasového otisku uloženého v databázi s parametry extrahovanými z aktuálního hovoru, je klientovi umožněn přístup do systému. Ověření může probíhat dvěma způsoby. V případě identifikace klienta se hlasový záznam musí porovnat se všemi otisky uloženými v databázi. Nevýhodou je větší časová náročnost. Druhou možností je autentizace řečníka, kdy se na základě sdělené identity ověří shoda hlasu klienta s konkrétním vzorkem z databáze. Ověření může být založeno na znalosti konkrétního hesla, na zopakování náhodných slov vygenerovaných systémem nebo může být hlas vyhodnocován v průběhu hovoru. [8]

1.3.2 Bezpečnost hlasového ověření

Pro každého klienta je důležité, aby jeho informace a přístupy k bankovním účtům nemohly být zneužity. Banka si zase chce být jistá, že komunikuje s oprávněným vlastníkem daného účtu. Systém založený na ověření pomocí konkrétního hesla by mohl být teoreticky napadnutelný. Pokud útočník získá nahrávku daného hesla. Člověk však není schopen vyslovit každou frázi vždy stejně. Na základě toho systém pozná, zda daná fráze pochází od člověka nebo jde jen o nahrávku. Bezpečnějším způsobem je ověření pomocí náhodně vygenerovaných frází v průběhu hovoru, které musí klient zopakovat. Pro útočníka je v tomhle případě nemožné připravit si předem dané fráze. Klíčové fráze mohou také obsahovat slovní spojení, které se v běžné mluvě nevyskytují. Uživatel může být ověřen na začátku hovoru pomocí klíčových frází a následně také v průběhu celého hovoru. Systémy jsou schopné detailně charakterizovat hlasový projev klienta. Každý člověk má unikátní hlasové ústrojí, které imitátor nemůže dokonale napodobit. Důležité je, aby systém fungoval spolehlivě za všech okolností. Správnou funkci systému může ovlivnit aktuální zdravotní stav klienta nebo okolní podmínky narušující hovor (hluk). Z toho důvodu se využívá vícefaktorové ověření klienta. Používá se kombinace hlasového ověřování se znalostí další informace. Může se jednat například o zadání pin kódu. Dalším problémem je, že i klient si musí být jistý, že jedná s bankou. V jiném případě by se útočník mohl vydávat za operátora banky. Útok by také mohl probíhat tak, že by útočník ve stejném čase komunikoval jak s operátorem banky, tak

s klientem. Útočník by přehrával požadavky automatického systému hlasového ověření klientovy. Klient by takto komunikoval s bankou prostřednictvím útočníka, kterému by byl umožněn přístup do systému. Z toho důvodu je důležité kontrolovat klienta i v průběhu hovoru. [8]

Úspěšnost vyhodnocení je závislá na nastavení citlivosti systému. Systém musí zohlednit zařízení používané uživatelem, vlivy okolního prostředí a omezení, které ovlivňují hlasový projev. Citlivost se nastavuje tak, aby se úspěšnost autentizace uživatele pohybovala okolo 95 %. Chybou systému může dojít k odmítnutí přístupu uživateli, který má potřebná oprávnění (FRR – False Rejection Rate). To lze omezit využitím vícefaktorové autentizace. Větším problémem je chybné umožnění přístupu uživateli, který nemá potřebná oprávnění (FAR – False Acceptance Rate). [9]

Česká republika patří mezi špičkám v oblasti autentizace pomocí hlasu. Zpracováním hlasu se zabývá například výzkumná skupina Speech@FIT. Produkty pro použití ve státním a komerčním sektoru působí firma Phonexia s.r.o.

1.4 Rozpoznávání zvuku ve forenzním vyšetřování

Rozpoznávání zvuku pro potřeby kriminalistického vyšetřování se označuje jako forenzní audio analýza. Úkoly forenzní audio analýzy jsou získávání, analýza a vyhodnocování zvukových záznamů pro potřeby kriminálního vyšetřování. Hlavní cíle audio analýzy jsou:

- Ověření integrity audio záznamu,
- zpracování audio nahrávky za účelem zlepšení její srozumitelnosti,
- identifikace osoby, která je zachycena na nahrávce.

Techniky používané při audio analýze slouží ke zlepšení srozumitelnosti nahrávky. Patří zde zlepšení srozumitelnosti mluvího nebo extrakce zvuků z pozadí nahrávky. Některé metody audio analýzy nemusí být při důkazním šetření přijaty soudem. Proto je vždy nutné, aby bylo možné dokázat, že při práci s nahrávkou nemohlo dojít ke změně jejího obsahu. [10]

Analýza hlasu je založena na faktu, že žádné dvě osoby nemají stejný hlas. Každý člověk má naprosto jiné uspořádání orgánů sloužících k tvorbě hlasu. Audio analýza v kriminalistické praxi slouží k objasňování trestných činů, kde má zvukový záznam přímou souvislost s daným činem. Jde například o záznamy týkající se výhrůžek, pomluv, plánování trestného činu atd. [11]

1.4.1 Objekty kriminalistické audio analýzy

Objektem zkoumání pomocí kriminalistické audio analýzy jsou zvukové nahrávky. Cílem zkoumání je identifikace osoby na nahrávce, zkoumání obsahu nahrávky nebo ověření pravosti nahrávky. Případně mohou být objektem zkoumání i zařízení za jejíž pomoci byly nahrávky vytvořeny. Objekty kriminalistické audio analýzy jsou:

- Sporné záznamy zvuku,
- srovnávací záznamy zvuku,
- pomocné materiály.

Sporné záznamy zvuku

Sporné záznamy zvuku jsou hlasové nahrávky mající souvislost s vyšetřovaným trestným činem. Typicky se jedná o záznamy hovorů na tísňových linkách, hovory zaznamenané na soukromých telefonních přístrojích, média zajištěná na místě činu nebo získaná při domovních prohlídkách. Záznamy bývají uchovávány na různých datových nosičích. Klíčovým je ověřit integritu audiozáznamu. Také je důležité, aby nahrávka měla, co největší vypovídající hodnotu. V případě výhružného telefonátu by se osoba přijímající hovor měla snažit, co nejvíce prodloužit čas, kdy vyděrač mluví. Případně aby byl hovor orientován k získání co nejvíce informací o pachateli (věk, sociální postavení, státní příslušnost atd.). Je tak získán kvalitní vzorek pro následnou audio analýzu.

Srovnávací záznamy zvuku

Srovnávací záznamy zvuku jsou takové záznamy, kdy známe jejich původce. Provádění záznamu je dělíme na:

- Ukázky řeči,
- zkoušky řeči,
- zkoušky hlasu.

Ukázky řeči jsou záznamy provedení bez vědomí nahrávané osoby. Jedná se především o záznamy telefonické komunikace nebo záznamy rozhovorů.

Zkoušky řeči se provádí za přítomnosti vyšetřované osoby. Je zaznamenáván mluvený projev dané osoby. Jedná se o řízený rozhovor, kdy vyšetřovaný odpovídá na připravené otázky nebo čte daný text.

Zkoušky hlasu se provádí s vyšetřovanou osobou. Vyšetřovaný přeríkává slova nebo celé věty, které jsou obsahem sporného materiálu.

Pomocné materiály

Pomocné materiály popisují okolnosti vzniku všech zvukových záznamů využitelných k objasnění trestného činu. Patří sem protokoly z ohledání místa činu, protokoly o provedené audio analýze, protokoly o výslechu osob. [11]

2 ZPRACOVÁNÍ SIGNÁLŮ

Zpracování signálu je proces, při němž jsou ze signálu získávány informace. Informace jsou získávány na základě poznatků z širokého spektra vědních oborů. Hlavně se jedná o matematiku a elektrotechniku. Tyto poznatky dovolují signály transformovat a analyzovat. Pro akustické signály je typické před dalším zpracováním provést jejich převod z analogové do digitální podoby. Výhodami digitálního signálu jsou větší možnosti zpracování pomocí výpočetní techniky. Jedná se například o možnosti filtrace, komprese, práci se signálem v časové i frekvenční doméně, extrakce parametrů signálu a následné klasifikace. [7]

Signál vyjadřuje stav nějaké fyzikální veličiny, která se mění. Bývá vyjádřena jako závislostí dvou a více proměnných.

2.1 Záznam zvuku

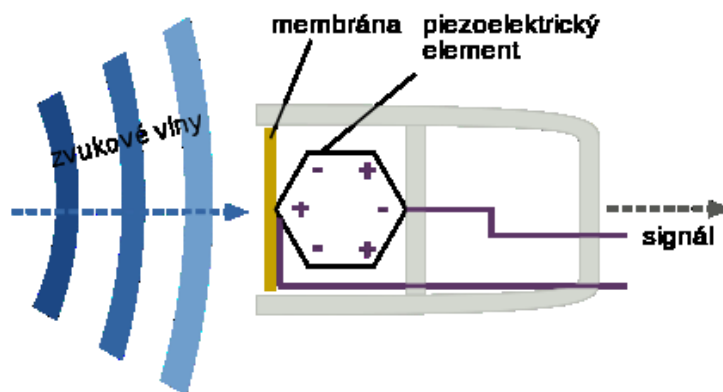
Při zpracování audio signálu je v první řadě potřeba jeho zaznamenání. K tomuto účelu slouží mikrofony, které jsou schopny přeměnit akustické vlnění na elektrický signál. Základními druhy mikrofonů jsou:

- Kondenzátorový,
- piezoelektrický,
- elektretový,
- dynamický,
- páskový,
- uhlíkový.

V praxi se nejčastěji využívají piezoelektrické a elektretové mikrofony.

2.1.1 Piezoelektrický mikrofon

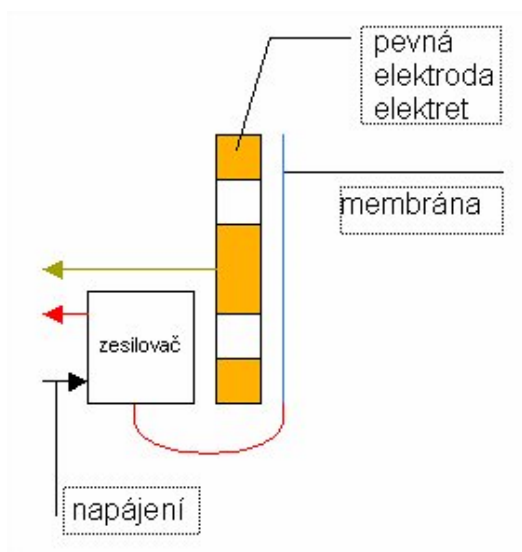
Piezoelektrický mikrofon je tvořen membránou a piezoelektrickým elementem. Piezokrystal je materiál, který při deformaci generuje elektrické napětí. Mezi piezoelektrické materiály patří křemík a křišťál. Membrána mikrofonu je s tímto piezoelektrickým elementem propojena. Při pohybu membrány dochází ke stlačování piezokrystalu a generování elektrického napětí.



Obr. 8 Piezoelektrický mikrofon. [12]

2.1.2 Elektretový mikrofon

Elektretový mikrofon je typem kondenzátorového mikrofonu. Elektretový mikrofon je tvořen dvěma elektrodami. Jedna tvoří membránu, která se vlivem akustického signálu pohybuje. Druhá elektroda je pokryta elektretem. Elektret je nevodivý materiál, která je permanentně nabitý. V rytmu pohybu membrány se mění kapacita kondenzátoru a napětí mezi elektrodami. Tyto změny napětí jsou zesíleny pomocí předzesilovače. Pro zesílení se používá se FET tranzistory. Typický frekvenční rozsah tohoto typu mikrofonu se pohybuje mezi 20 Hz a 20 kHz. Mezi typické vlastnosti elektretových mikrofonů patří velká citlivost, malé zkreslení, velká výstupní impedance a vyrovnaná frekvenční charakteristika. Tento typ mikrofonu se používají pro měřicí účely, mobilní telefony, spotřební elektroniku atd. [13] [14]



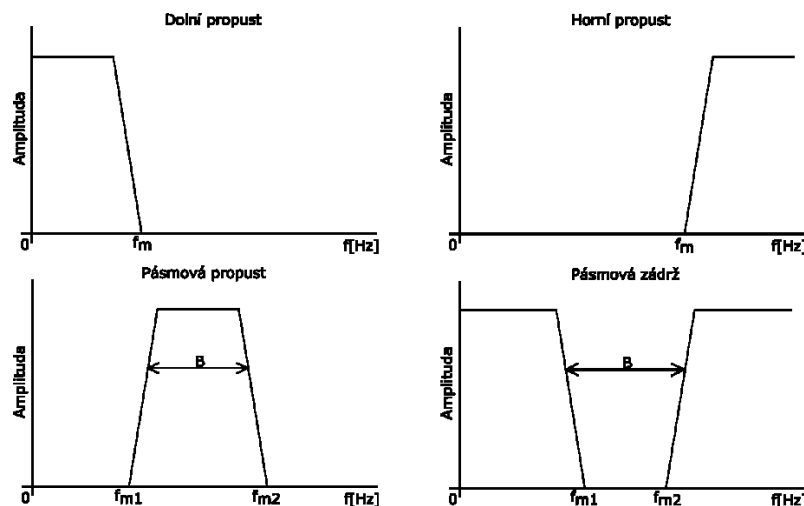
Obr. 9 Elektretový mikrofon. [15]

2.2 Filtrace signálu

Filtrace je proces získávání požadovaného frekvenčního spektra pro další zpracování. Jedná se o selekci informací ze vstupního signálu. Filtrace signálu slouží k oddělení užitečné a rušivé složky. Cílem je určitou složku signálu potlačit a jinou zesílit. Například může být použita k odstranění šumu nebo potlačení části spektra. Podle typu zpracovávaného signálu se filtry dělí na analogové a digitální.

2.2.1 Frekvenčně selektivní filtry

Frekvenčně selektivní filtry slouží k propuštění jedné části spektra a potlačení jiné části. Podle části filtrovaného spektra rozlišujeme filtry typu dolní propust, horní propust, pásmová propust a pásmová zadrž.



Obr. 10 Typy filtrů podle funkce.

Dolní propust

Filtr typu dolní propust propouští spodní část frekvenčního spektra a utlumuje vysoké frekvence. Typicky se využívá v A/D převodnicích jako antialiasing filtr. Filtr propouští frekvence od 0 Hz do mezní frekvence f_m . [16]

Horní propust

Filtr typu horní propust blokuje nízké frekvence a propouští frekvenční spektrum od f_m do nekonečna.

Pásmová propust

Pásmová propust propouští část spektra mezi dvěma mezními hodnotami F_{m1} a F_{m2} . Signály o frekvenci menší, než F_{m1} jsou blokovány, stejně tak část spektra větší než F_{m2} .

Pásmová zádrž

Pásmová zádrž blokuje část spektra mezi dvěma mezními hodnotami F_{m1} a F_{m2} . Signály o frekvenci menší, než F_{m1} a větší, než F_{m2} jsou propuštěny.

2.2.2 Analogové filtry

Analogové filtry se typicky používají pro předzpracování signálu. Zpracovávaným signálem může být zvuk zachycený mikrofonom. Analogové filtry se realizují pomocí pasivních elektronických součástek, jakými jsou rezistory, kondenzátory a cívky. Jedná se o pasivní filtry. Aktivní filtry jsou realizovány pomocí aktivních elektronických součástek, jakými jsou tranzistory nebo operační zesilovače. Umožňují propustit určité frekvence analogového signálu a jiné potlačit. Jednoduché filtry se konstruují jako kombinace pasivních elektronických součástek. Zapojením prvků RLC lze realizovat dolní propust, horní propust, pásmovou propust a pásmovou zádrž. [17]

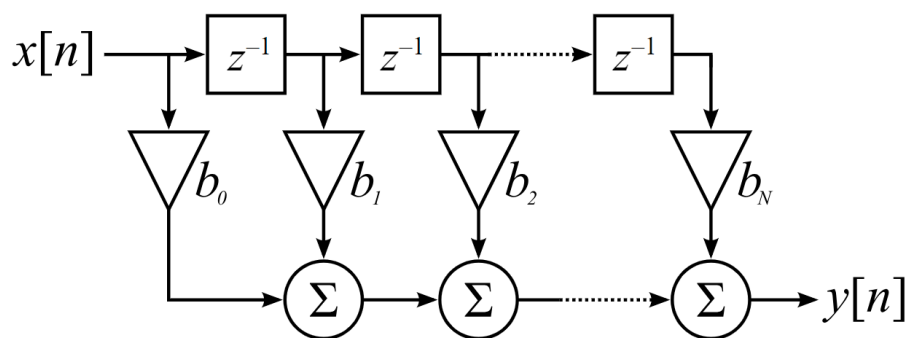
2.2.3 Číslicové filtry

Číslicové filtry využívají pro zpracování signálu procesor, který provádí výpočty se vzorky signálu. Analogový signál musí být před zpracování pomocí číslicového filtru digitalizován. Číslicová filtrace se využívá v systémech pracujících se signálem v reálném čase. Jedná například o systémy rozpoznávání hlasu. Výhodami oproti analogovým filtru jsou neměnné parametry filtrů z důvodu parazitních jevů elektrotechnických součástek, umožňují adaptaci filtru na právě zpracovávaný signál, číslicová filtrace je mnohem rychlejší. Uplatňuje se v aplikacích, kde je potřeba provést veškeré zpracování signálu za periodu jednoho vzorku. Základní typy číslicových filtrů jsou filtr s konečnou impulsní odezvou (FIR) a filtr s nekonečnou impulsní odezvou (IIR). [18]

Filtr s konečnou impulsní odezvou (FIR)

Tento typ filtru se vyznačuje konečnou impulsní odezvou. To znamená, že po průchodu impulsem filtrem je na výstupu konečný počet nenulových hodnot. FIR filtry se vyznačují velkou stabilitou, lineárním fázovým posuvem pro všechny frekvence a jednoduchou realizací. Mezi nevýhody patří vyšší výpočetní náročnost. Používají se například pro odstranění šumu z řečového signálu. [7] [19]

$$y[n] = \sum_{i=0}^N b_i \cdot x[n - i]$$

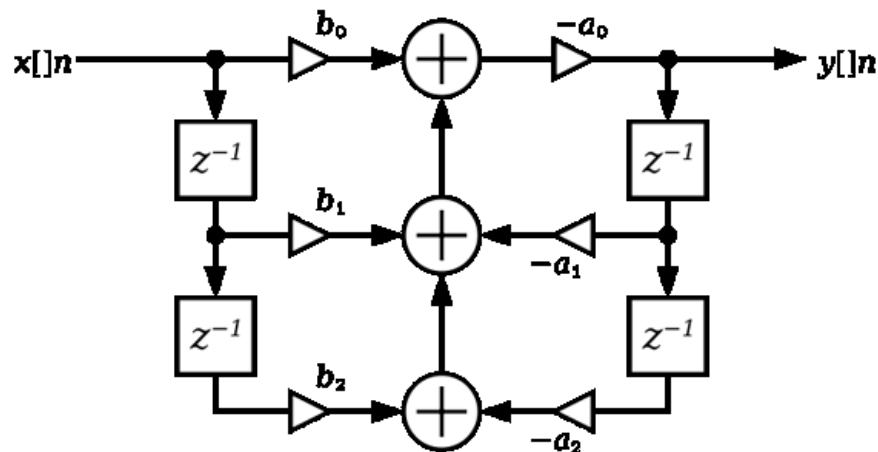


Obr. 11 Funkce filtru s konečnou impulsní odezvou. [20]

Filtr s nekonečnou impulsní odezvou (IIR)

Filtr s nekonečnou impulsní odezvou se vyznačuje zpětnou vazbou. Výhodou oproti FIR filtrům je nižší výpočtová náročnost. Nevýhodou je, integrační zapojení se zpětnou vazbou může způsobit nestabilitu filtru. Také není možné jako u FIR zaručit lineární fázový posuv pro všechny frekvence. [7] [21]

$$y[n] = \frac{1}{a_0} \cdot \sum_{i=0}^P b_i \cdot x[n - i] \cdot \sum_{j=1}^Q a_j \cdot y[n - j] \quad (2.1)$$



Obr. 12 Funkce filtru s nekonečnou impulsní odezvou. [22]

2.2.4 Antialiasing filtr

Aliasing je jev, ke kterému dochází při převodu analogového signálu na digitální. Při vzorkování signálu musí být dodržen Shannonův teorém, kdy vzorkovací frekvence musí být minimálně dvakrát větší, než je frekvence nejvyšší harmonické složky signálu. K předcházení vzniku aliasingu se používá antialiasingový filtr. Jedná se o filtr typu dolní propust. [16]

2.3 Digitalizace

Digitalizace analogových signálů je jejich převod z analogové do digitální formy. Signál je převáděn ve dvou krocích. Nejprve se provádí vzorkování a následně je těmto vzorkům přiřazena hodnota. Druhý krok je označován jako kvantování. Výsledkem jsou diskrétní hodnoty signálu.

2.3.1 Vzorkování signálu

Vzorkování je proces, při kterém jsou z signálu vybírány jeho úrovně po předem definovaných časových úsecích. Jednotlivé vzorky jsou vyjádřeny jako posloupnost Diracových impulzů. [23]

$$i(t) = \sum_{n=-\infty}^{\infty} \delta \cdot (t - nT) \quad (2.2)$$

kde $i(t)$ – vzorkovací funkce, δ – Diracova funkce, T – vzorkovací perioda.

Vzorkovaný signál lze potom vyjádřit jako:

$$s(n) = x(t) \cdot i(t) \quad (2.3)$$

$$s(n) = \sum_{n=-\infty}^{\infty} x \cdot (n \cdot T) \cdot \delta \cdot (t - n \cdot T) \quad (2.4)$$

kde $x(t)$ – vstupní analogový signál, $i(t)$ – vzorkovací funkce, $s(n)$ – vzorkovaný signál.

Při vzorkování je nutné dodržet Shannonův teorém, kdy vzorkovací frekvence musí být vyšší nebo rovna dvojnásobku nejvyšší harmonické složky vzorkovaného signálu.

$$f_s \geq 2 \cdot f_{max} \quad (2.5)$$

Pokud by nebyl dodržen Shannonův teorém vzniklo by zkreslení signálu označované jako aliasing. Například v telefonii se používají akustické signály o frekvenci 300–3400 Hz. Minimální vzorkovací frekvence by tedy měla být $f_s=6800$ Hz. V praxi se používá vzorkovací frekvence 8 kHz. Vzorkovací perioda pak:

$$T_{vz} = \frac{1}{f_{vz}} = \frac{1}{8000} = 125 \mu s \quad (2.6)$$

2.3.2 Kvantování signálu

Kvantování je proces přiřazení velikosti jednomu vzorku pomocí kvant. Velikost je danému vzorku přiřazena pomocí kvantovacích hladin. Počet kvantovacích hladin závisí na délce kódovaného slova. Například pokud je délka kódovaného slova 8 bitů. Máme k dispozici 256 kvantovacích hladin. Při kvantování vzniká tzv. kvantizační chyba. Je to chyba, která vznikne zaokrouhlením velikosti vzorku k nejbližší vyšší kvantovací hladině. Maximální kvantizační chyba:

$$\Delta_{max} = \frac{1}{2^{n+1}} \quad (2.7)$$

Uniformní kvantování

Uniformní kvantování se vyznačuje stejnou vzdáleností kvantizačních hladin. [7]

Neuniformní kvantování

Tento typ kvantování umožňuje snížit kvantizační chybu v oblasti signálu, která je považována za důležitou. Kvantizační hladiny mezi sebou nemají stejnou vzdálenost. Používá se například logaritmické nebo exponenciální rozdělení hladin. [7]

2.4 Fourierova transformace

Fourierova transformace umožňuje převádět signály z časové do frekvenční domény a na zpět. Frekvenční doménu označujeme jako spektrum. Spektrum je časově nezávislé zobrazení frekvencí obsažených v signálu. Rozlišujeme spojitou a diskrétní Fourierovu transformaci. Spojitá Fourierova transformace pracuje se spojitým signálem. Signál v časové doméně je převáděn na spojité spektrum. Diskrétní Fourierova transformace, jak už z názvu vyplývá pracuje s diskrétními hodnotami signálu. [7]

2.4.1 Diskrétní Fourierova transformace

Při práci s digitálním signálem, kde již proběhlo vzorkování se využívá DFT (Discrete Fourier Transform). Využívá se v případech číslicového zpracování signálu. Například tedy všechny biometrické systémy využívají DFT, kde se zpracovává vzorkovaný signál.

$$x[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi \frac{k}{N}n} \quad (2.8)$$

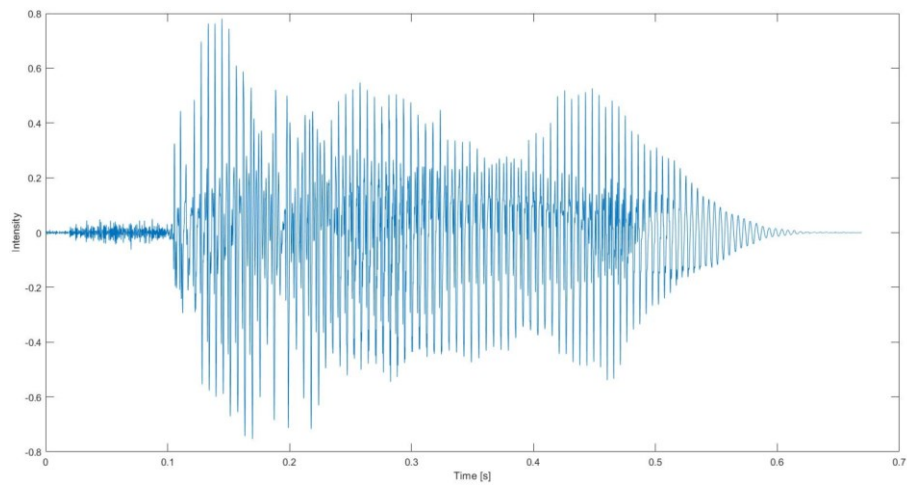
kde $x[k]$ – výsledek analýzy, k -velikost frekvence, $x[n]$ – analyzovaný signál, n – počítadlo vzorků, k/N -normovaná frekvence.

Normovaná frekvence je frekvence vztažená k vzorkovací frekvenci. Jedná se o bezrozměrnou veličinu.

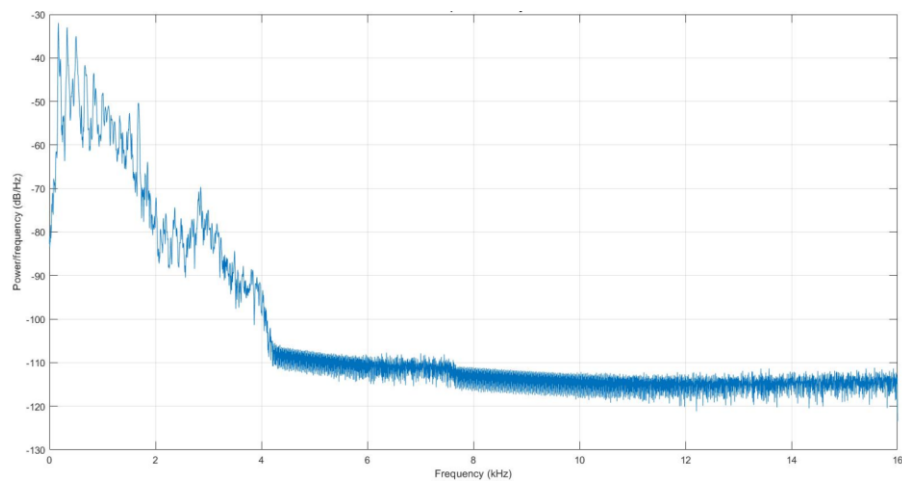
$$f = \frac{f_{skutečná}}{F_s} \quad (2.9)$$

kde f – normovaná frekvence [-], F_s – vzorkovací frekvence [Hz].

V softwarech pro zpracování signálu se používá FFT (Fast Fourier Transform). FFT je algoritmus pro provedení výpočtu DFT. Vzorkovací frekvence pro FFT musí být mocninou čísla dva, $N = 2^m$. [7]



Obr. 13 Vstupní signál v časové doméně.



Obr. 14 Frekvenční spektrum signálu.

3 ZÁKLADNÍ PARAMETRY ZVUKU PRO AUDIO ANALÝZU

Pro popis a následnou klasifikaci akustického signálu je potřeba z daného signálu extrahovat jeho charakteristické parametry. Ty pak mohou být využity pro odlišení několika vzorků akustického signálu. Získávání příznaků daného signálu označujeme jako parametrizace.

3.1 Parametrizace

Úkolem parametrizace je vyhledání parametrů akustického signálu schopných popsat charakteristické příznaky daného signálu. Tyto parametry slouží k rozpoznání daného signálu při porovnání s jinými akustickými signály. Rozpoznávání řeči je založeno na extrakci dostatečného množství parametrů vhodných pro jednoznačnou identifikaci mluvčího. Využívají se parametry získané, jak z časové, tak frekvenční domény daného akustického signálu. Parametrizaci dělíme podle popisu signálu na:

Neparametrický popis signálu – je založeny na poznacích o zpracování signálu (filtrace, Fourierova transformace, korelace atd.).

Parametrický popis signálu – je založených na vlastnostech akustického signálu vztažených k mechanismu vzniku daného signálu. [24]

3.2 Odstranění stejnosměrné složky

Průchodem signálů prvky jako zesilovače a převodníky může vznikat rušivý signál, který je přidán ke zpracovávanému signálu. Vliv tohoto rušivého signálu může být odbourán odstraněním stejnosměrné složky akustického signálu. Stejnosměrná složka je střední hodnota všech vzorků signálu. Střední hodnota se pak od celého signálu odečte.

Střední hodnota:

$$\bar{s} = \frac{1}{N} \cdot \sum_{n=1}^N s[n] \quad (3.1)$$

kde \bar{s} – střední hodnota, N – počet vzorků signálu, $s[n]$ – n -tý vzorek signálu.

Toto platí v případě off-line záznamu, ke kterému se můžeme kdykoliv vrátit. V případě potřeby zpracování signálu v reálném čase se využívá IIR filtr. [24]

3.3 Rozdělení na rámce

Při zpracování akustických signálů je vhodné analyzovaný signál rozdělit na rámce. Akustický signál je považován za náhodný. Pro potřeby jeho analýzy je výhodné získat jeho krátké časové úseky, které lze považovat za stacionární. Rámce jsou definovány jejich délkou, překrytím, a posunem. Délka rámce pro potřeby zpracování řečového signálu se používá 20–25 ms. Odpovídá 160–200 vzorkům signálu při vzorkovací frekvenci $F_s = 8000$ Hz. Tato délka je vztažena k velikosti setrvačnosti hlasového ústrojí. [24]

3.3.1 Okénkové funkce

Okénkové funkce se využívají pro selekci analyzovaného signálu. Využívají se také pro výběr rámců signálu. Úkolem okénkové funkce je vybrat část analyzovaného signálu tak, aby bylo jeho spektrum co nejpoužitelnější pro následnou extrakci parametrů. [26]

Příkladem okénkové funkce je pravoúhle okno, pro které je charakteristické ostré vykrojení části spektra. Pro zpracování akustického signálu se nejčastěji používá Hammingovo okno, které se vyznačuje utlumením spektra na jeho okrajích. Maximum signálu je orientováno uprostřed. [24]

3.4 Spektrogram

Spektrogram je grafické zobrazení rozložení frekvencí v čase. Pro zobrazení se používají okénkové funkce. Energie analyzovaného signálu je v grafu vyjádřena pomocí barevných odstínů. [27]

3.5 Střední krátkodobá energie

Střední krátkodobá energie je parametr sloužící ke zjištění energie vyskytující se v rámci. V oblasti rozpoznávání řeči umožňuje odlišit znělé a neznělé hlásky.

$$E = \frac{1}{l_{ram}} \cdot \sum_{n=0}^{l_{ram}-1} x^2[n] \quad (3.2)$$

kde E – střední krátkodobá energie, l_{ram} – délka rámce, $x[n]$ – vzorek signálu. [24]

3.6 Korelace

Slouží k určování podobnosti signálů. Pomocí korelační funkce lze v neznámém signálu nalézt známý vzor. Korelace mezi dvěma různými signály se nazývá vzájemná korelace (cross-

corelation). Pro určení periody signálu se využívá autokorelace. Je možné tedy určit, zda se signál periodicky neopakuje. [25]

Korelační funkce:

$$c = \sum_{n=0}^{N-1} x[n] \cdot a[n] \quad (3.3)$$

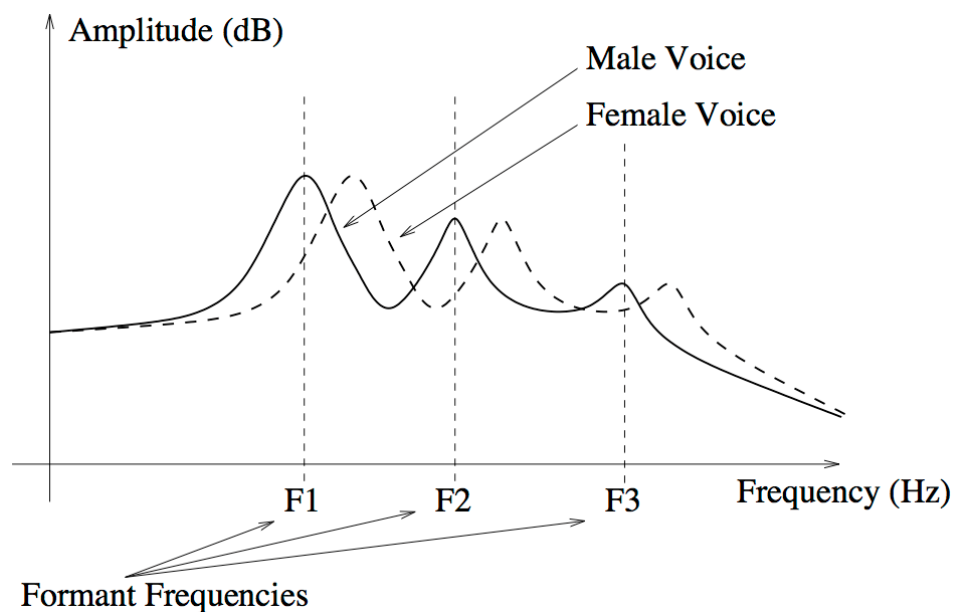
kde c – míra podobnosti dvou signálů, $x[n]$ – vzorky analyzovaného (porovnávaného) signálu, $a[n]$ – analyzační signál (vyhledávaný vzor).

3.7 Počet průchodů nulou

Parametr vyjadřující počet překročení nuly signálem. Periodické signály periodicky překračují nulu. Určuje se počet průchodů za rámeček. Funkce je citlivá na výskyt šumu v analyzovaném signálu. [27]

3.8 Formantové frekvence

Jde o rezonanční frekvence základního tónu hlasu. Základní tón f_0 se šíří hlasovým ústrojím (hltan, dutina ústní), kde dochází ke vzniku rezonancí.



Obr. 15 Formantové frekvence. [28]

3.9 Cepstrum

Cepstrální analýza se využívá v rozpoznávání řečového signálu. Slouží k oddělení hlasové složky vzniklé buzením a modifikační složky hlasu. Buzení vzniká proudem vzduchu z plic

skrz průdušnici a hrtan, vytváří se zde základní tón. Ten je se liší podle pohlaví. Výška základního tónu muže se pohybuje mezi 80 a 160 Hz, u žen mezi 150 a 300 Hz, u dětí mezi 200 a 600 Hz. Modifikační složka hlasu vzniká průchodem základního tónu těmi částmi hlasového ústrojí, které jsou pro každého člověka unikátní. Jedná se o hlasivky, dutinu ústní, jazyk a zuby. Pohybem těchto částí dochází k modulaci základního tónu. Modifikační složka hlasu je pak využívána v systémech rozpoznávání řeči. Výsledkem cepstrální analýzy jsou cepstrální koeficienty. Pro účely rozpoznávání hlasu se využívá mocninné cepstrum (power cepstrum).

Mocninné cepstrum je definováno jako inverzní FFT přirozeného logaritmu spektra analyzovaného signálu. [24]

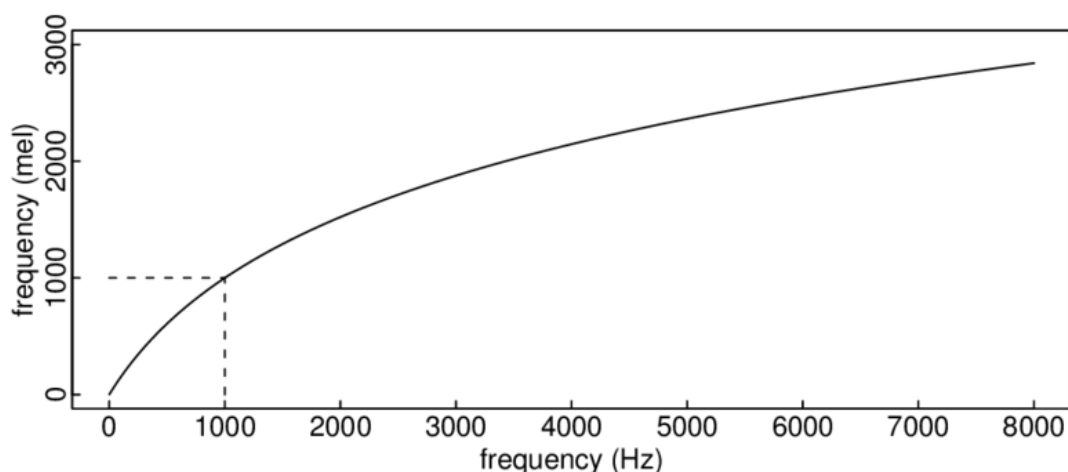
$$c_p(n) = F^{-1}\{\ln|x(f)|^2\}$$

3.10 Mel-frekvenční spektrální koeficienty

Mel-frekvenční spektrální koeficienty (MFCC) na rozdíl od cepstra respektují fyziologické vlastnosti lidského sluchu. Člověk má mnohem větší rozlišovací schopnosti pro nízké frekvence než pro vysoké. Různého rozlišení na frekvenční ose můžeme dosáhnout použitím nelineárních filtrů. Daná nelineární úprava využívá převodu Herzů na Mely. [24]

Převod z Herzů na Mely:

$$F_{Mel} = 2959 \cdot \log_{10}\left(1 + \frac{F_{Hz}}{700}\right) \quad (3.4)$$



Obr. 16 Grafické znázornění nelineární závislosti převodu z Herzů na Mely. [29]

Následným použitím zpětné Fourierovy transformace získáme Mel-frekvenční cepstrální koeficienty (MFCC). [24]

3.11 Lineární predikce

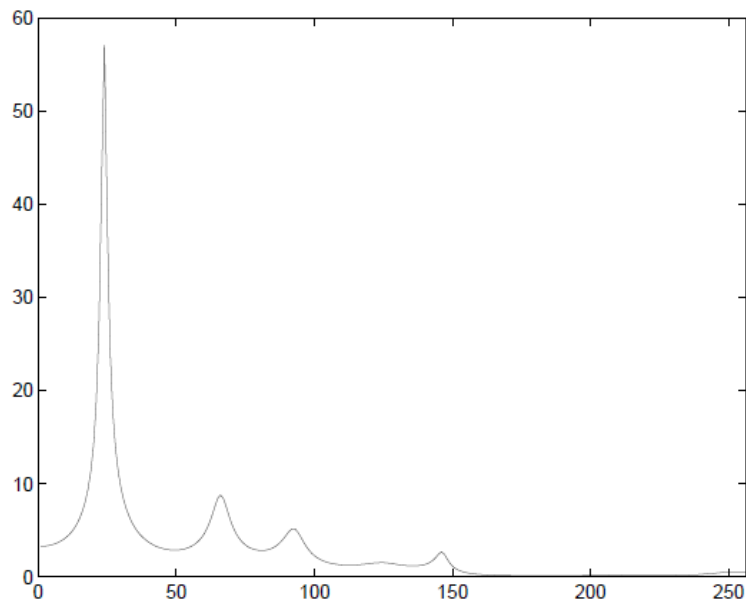
Analýza lineární predikcí (LP) je jednou z nejdůležitějších technik používaných pro rozpoznávání hlasu. Cílem je vytvoření filtrů, které jsou schopné namodelovat lidský hlasový trakt. Za pomoci LP je možné předpovědět následující vzorky signálu z předcházejících vzorků. Predikované následující vzorky signálu lze získat jako váženou sumu předcházejících vzorků. [27]

$$\hat{s}[n] = - \sum_{i=1}^p a_i \cdot s[n-i] \quad (3.5)$$

kde $\hat{s}[n]$ – predikovaný vzorek, a_i – koeficienty filtru, $s[n-i]$ – předchozí vzorky signálu.

LP spektrum

Z koeficientu filtru a_i lze pomocí Fourierovy transformace a jejich následném invertování získat LP spektrum. LP spektrum zobrazuje velmi ostře formantové frekvence, je zde odstraněn vliv základního tónu. [27]



Obr. 17 LP spektrum. [27]

Pomocí LPC lze získat spektrální hustotu každého rámce signálu. To se využívá pro odlišení znělých a neznělých řečových rámců. Tedy rámce obsahující znělé a neznělé hlásky. [24]

LP cepstrum

Dalším parametrem odvozeným od lineární predikce je LP cepstrum (LPCC), které se využívá v systémech rozpoznávání hlasu založených na skrytých Markovových modelech. [24]

LP cepstrální koeficienty:

$$c(n) = -a_n - \frac{1}{n} \cdot \sum_{k=1}^{n-1} k \cdot c_k \cdot a_{n-k} \quad (3.6)$$

3.12 Další parametry vhodné pro audio analýzu

Maximum

Velikost maximální energie obsažené v analyzovaném akustickém signálu.

Minimum

Minimální energie obsažené v analyzovaném akustickém signálu.

Medián

Medián je střední hodnota statistického souboru, který je seřazen podle velikosti.

$$\tilde{x} = \frac{n+1}{2} \quad (3.7)$$

Kde \tilde{x} – medián, n – počet prvků statistického souboru. [30]

Peak to Peak

Velikost amplitudy peak to peak je rozdíl mezi hodnotou největší amplitudy analyzovaného signálu a jeho nejmenší hodnoty.

Efektivní hodnota

Efektivní hodnota RMS (Root mean square) je definovaná odmocnina aritmetického průměru druhých mocnin daného souboru hodnot.

Aritmetický průměr

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (3.8)$$

Standartní deviace

Vyjadřuje rozptyl hodnot kolem aritmetického průměru celého souboru hodnot.

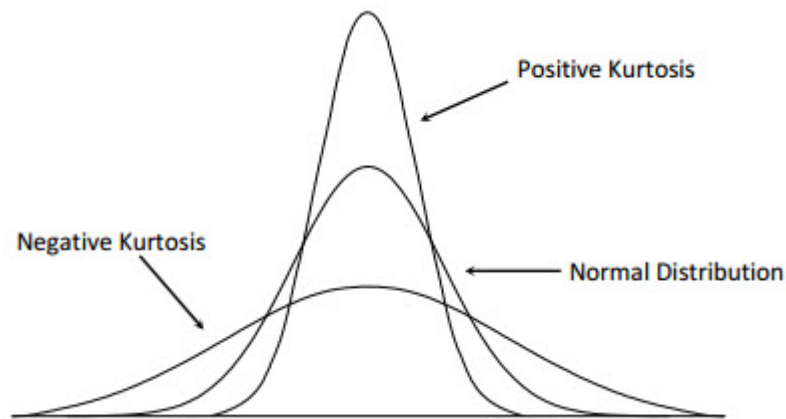
$$\sigma = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (x_i - \bar{x})^2} \quad (3.9)$$

Efektivní délka

Efektivní délka je časový úsek, kdy energie signálu je vyšší než předem stanovená prahová hodnota energie. [31]

Kurtosis

Parametr kurtosis vyjadřuje plochost rozdělení hodnot analyzovaného signálu vzhledem k jeho střední hodnotě. Parametr lze získat jak z časové, tak frekvenční domény. [31]



Obr. 18 Hodnota parametru kurtosis v závislosti na plochosti distribuce signálu. [32]

Podle hodnot, jakých parametr kurtosis (K) nabývá rozlišujeme:

- $K = 3$ – normální,
- $K < 3$ – plošší rozložení,
- $K > 3$ – špičatější rozložení. [31]

Spektrální centroid (spectral centroid)

Spektrální centroid je parametr vyjadřující pozici těžiště spektra. Nachází v oblasti s největší částí spektra. Matematicky lze vyjádřit jako vážený průměr frekvencí vztažených k amplitudě signálu. [31]

$$\mu = \frac{\sum_{n=0}^{N-1} f[n] \cdot x[n]}{\sum_{n=0}^{N-1} x[n]} \quad (3.10)$$

kde μ – spektrální centroid, $f[n]$ – střední frekvence vztažená k odpovídající amplitudě, $x[n]$ – amplituda signálu.

Spektrální rozptyl (spectral spread)

Spektrální rozptyl je definován jako rozptyl spektra okolo jeho střední hodnoty. [31]

$$\sigma = \frac{\sum_{n=0}^{N-1} (f[n] - \mu)^2 \cdot x[n]}{\sum_{n=0}^{N-1} x[n]} \quad (3.11)$$

kde σ – spektrální rozptyl, $f[n]$ – střední frekvence vztažená k odpovídající amplitudě, μ – střední hodnota, $x[n]$ – amplituda signálu.

Spektrální šikmost (spectral skewness)

Vyjadřuje hodnotu asymetrie daného spektra okolo jeho střední hodnoty. Počítá se z centrálního momentu třetího řádu. [31]

Moment třetího řádu:

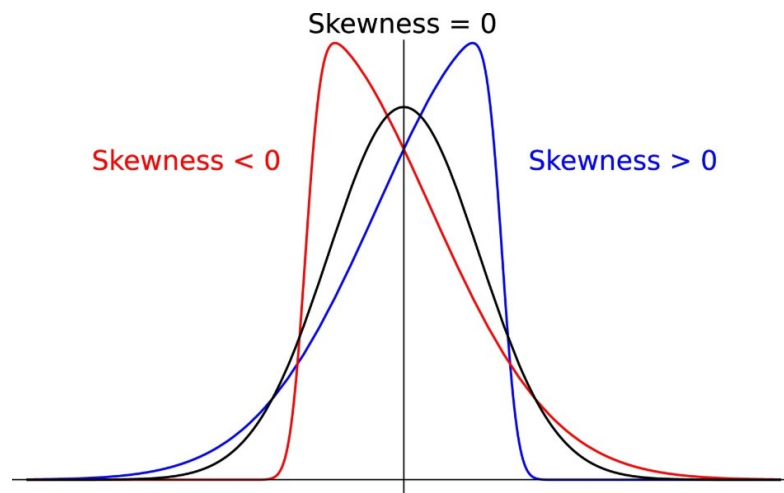
$$m_3 = \frac{\sum_{n=0}^{N-1} (f[n] - \mu)^3 \cdot x[n]}{\sum_{n=0}^{N-1} x[n]} \quad (3.12)$$

kde m_3 – centrální moment třetího řádu, $f[n]$ – střední frekvence vztažená k odpovídající amplitudě, μ – střední hodnota, $x[n]$ – amplituda signálu.

pak spektrální šikmost:

$$\gamma_1 = \frac{m_3}{\sigma^3} \quad (3.13)$$

kde γ_1 – spektrální šikmost, m_3 – moment třetího řádu, σ – standartní deviace.



Obr. 19 Hodnoty spektrální šikmosti. [33]

Spektrální sklon (spectral slope)

Spektrální sklon udává počet sestupů spektrální amplitudy. Matematicky je vyjádřen pomocí lineární regrese spektrální amplitudy. [31]

$$S = \frac{1}{\sum_{n=0}^{N-1} x[n]} \cdot \frac{\sum_{n=0}^{N-1} f[n] \cdot x[n] - \sum_{n=0}^{N-1} f[n] \cdot \sum_{n=0}^{N-1} x[n]}{\sum_{n=0}^{N-1} f^2[n] - (\sum_{n=0}^{N-1} f[n])^2} \quad (3.14)$$

kde S – spektrální sklon, $f[n]$ – střední frekvence vztažená k odpovídající amplitudě, $x[n]$ – amplituda signálu.

4 STROJOVÉ UČENÍ

Strojové učení označuje výpočetní modely dovolující počítačovým systémům učit se. Výhodou těchto modelů je adaptace na změny vnějších podmínek. Metody strojového učení se využívají v oblastech jako například robotika, biomedicína, a informační technologie. Používá se také v systémech rozpoznávání řeči. Strojové učení má za cíl, aby počítačový systém byl schopen dělat rozhodnutí na základě učení se z velkého množství dat. [34]

Oblasti strojového učení lze rozdělit vzhledem k účelu na:

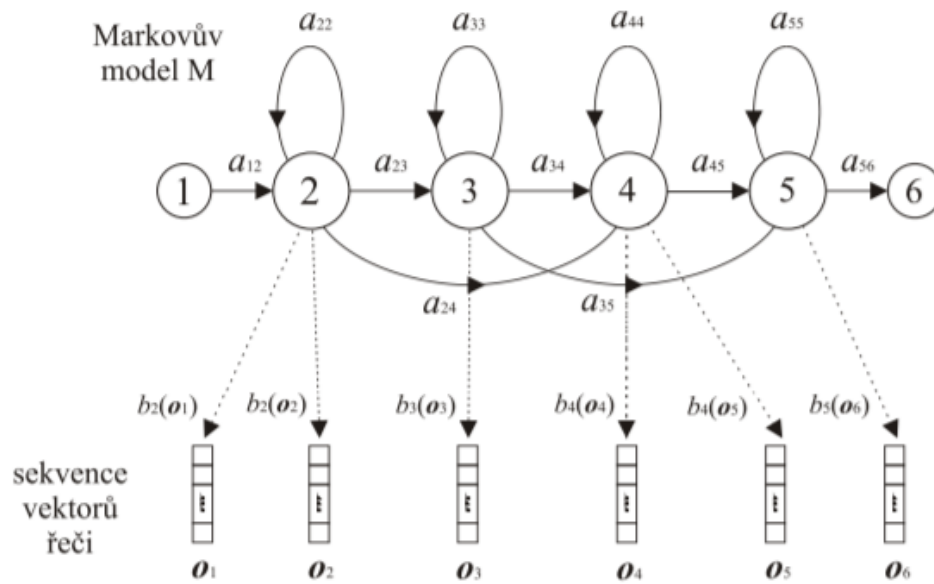
- Klasifikace – zařazení dat do určité skupiny,
- Regrese – analýza a ohodnocení za účelem odhadu budoucího vývoje dat,
- clusterování – zjišťování podobnosti dat,
- asociace – hledání pravidel v analyzovaných datech. [35]

Mezi algoritmy strojového učení patří:

- Skryté Markovovy modely,
- support vector machine,
- neuronové sítě,
- bayesovské sítě atd.

4.1 Skryté Markovovy modely

Skryté Markovovy modely jsou určeny pro modelování statických procesů s konečným počtem stavů. Mezi statické procesy patří i řeč. Model mění v každém okamžiku t svůj stav na základě pravděpodobnosti a_i . Při každé změně stavu modelu se generuje vektor pozorování o_t s výstupní pravděpodobností b_j . [36] [37]



Obr. 20 Příklad Markovova modelu se šesti stavy. [36]

Na příkladu viz Obr. 20 je zobrazen příklad Markovova modelu. Při zpracování řečového signálu jsou výsledkem přechodu každého stavu parametrické vektory θ_i , které obsahují parametry popisující analyzovaný signál. Zpracovávaný akustický signál je pak vyjádřen posloupností těchto parametrických vektorů. Aby byl rozpoznávač schopen plnit svoji funkci je ho potřeba tzv. natrénovat. Trénování probíhá prostřednictvím databáze slov a jejich variant vyjádření. Dalším důležitým krokem je zajištění kompenzace šumu. Řečový signál zaznamenávaný v reálném prostředí obsahuje často různé úrovně šumu. [36] [37]

4.2 Support vector machine

Support Vector Machine je algoritmus strojového učení, který se využívá pro klasifikaci a regresní analýzu. Algoritmus je založen na rozdělení různých tréninkových dat pomocí nadroviny (Hyperplane). Cílem je, co nevíce oddělit data. Nadrovina by měla být umístěna v co největší vzdálenosti od krajních bodů skupin dat.

Nadrovina je definována jako:

$$w^T \cdot x + b = 0 \quad (4.1)$$

kde w – váhový vektor, x – vstupní vektor, b – absolutní člen.

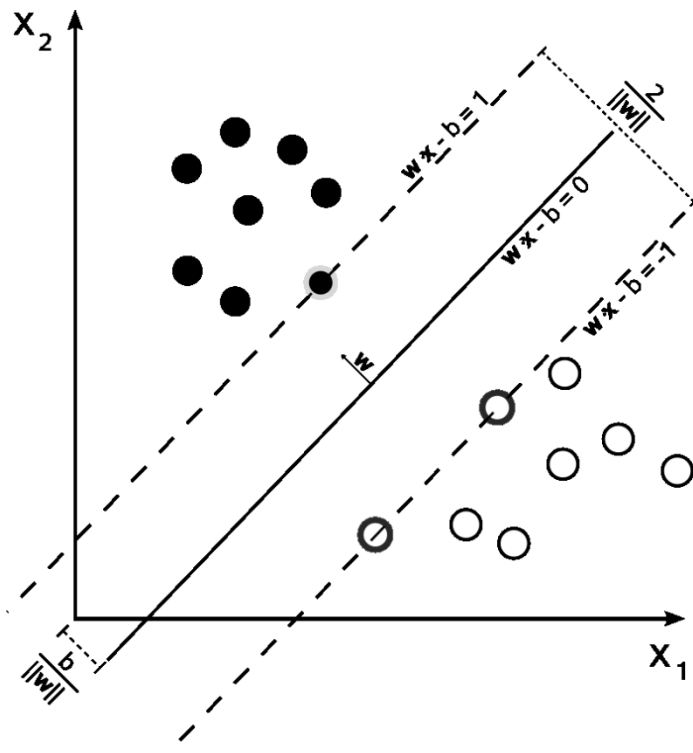
Support vector machine se dělí podle rozdělení prostoru na lineární a nelineární rozdělení. [38]

4.2.1 Lineární rozdělení

Lineární rozdělení se používá, pokud je možné data rozdělit lineárně, pomocí jedné přímky.

4.2.2 Nelineární rozdělení

Nelineární rozdělení se využívá v případech, kdy není možné data rozdělit lineárně. Používá se operace jádrová transformace, rozšíří původní dvojrozměrné zobrazení dat o třetí dimenzi. Trojrozměrné zobrazení je už lineárně rozdělitelné.



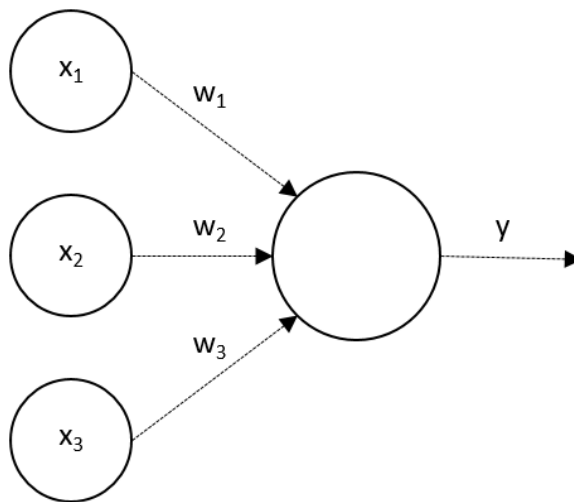
Obr. 21 Grafické zobrazení algoritmu Support Vector Machine. [38]

4.3 Neuronové sítě

Algoritmus umělé neuronové sítě je inspirován funkcí lidského mozku. Základním prvkem mozku jsou mozkové buňky neurony. Stejně tak je to i u umělých neuronových sítí. Každý neuron může mít několik vstupů. Může se jednat o informace zvenčí nebo výstup z jiného neuronu. Každý vstup má svoji specifickou váhu. Výstup z neuronu je vždy jen jeden. Všechny přijaté vstupy jsou v neuronu vynásobeny se svými vahami a sečteny. Následně jsou porovnány s prahovou hodnotou. [39]

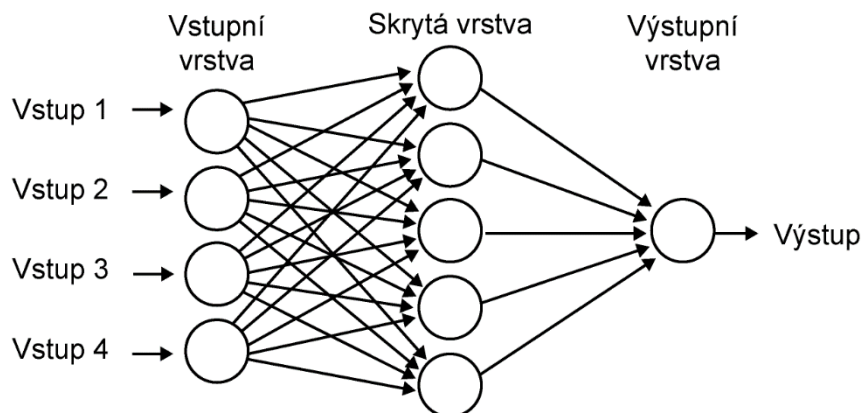
$$y = f\left(\sum_{i=1}^n w_i \cdot x_i - \theta\right) \quad (4.2)$$

kde x_i – vstupy do neuronu, w_i – váhy vstupů, θ – prahová hodnota, f – přenosová funkce, y – hodnota výstupu.



Obr. 22 Model neuronu. [39]

V praxi se využívají neuronové sítě složené z velkého počtu neuronů. Tyto sítě se skládají se vstupní vrstvy, která má za úkol přijmout informace z vnějšku. Ty jsou pak předány všem neuronům ve skryté vrstvě. Zde probíhají výpočty. Používané neuronové sítě mívají typicky dvě skryté vrstvy. Výsledky výpočtů jsou pak předány výstupní vrstvě. Aby byla neuronová síť schopná plnit svůj účel, je potřeba přizpůsobit pomocí trénovacího algoritmu. Ten modifikuje parametry i topologii celé sítě. [40]



Obr. 23 Model neuronové sítě. [40]

5 FYZIKÁLNÍ PRINCIP PRO ZPRACOVÁNÍ AKUSTICKÝCH SIGNÁLŮ

5.1 Akustika

Akustika je obor zabývající se vznikem a vlastnostmi zvuku. Zvuk je popsán jako mechanické podélné vlnění šířící se prostředím. Podle frekvence rozlišujeme tři typy zvuků. Lidské ucho je schopné vnímat zvuk o frekvenci 16 Hz–20 kHz. Tuto část spektra označujeme jako slyšitelný zvuk. Frekvence, které je člověk schopen slyšet jsou ovlivněny věkem. S přibývajícím věkem člověk přestává vnímat vyšší frekvence. Zvuky mimo pásmo slyšitelnosti neslyšíme ale dokážou ovlivnit zdravý člověka a jeho psychický stav. Zvuk pod hranicí 16 Hz označujeme jako infrazvuk. Zvuk nad hranicí 20 kHz se nazývá ultrazvuk. Zvuk vzniká kmitáním bodů a bodových soustav. Při kmitavém pohybu se v čase mění vlastnosti vlny jako amplituda, rychlost, vlnová délka atd. [41]

Harmonický pohyb lze popsat vztahem:

$$a = A \cdot \sin\left(\frac{2\pi}{T} \cdot t\right) \quad (5.1)$$

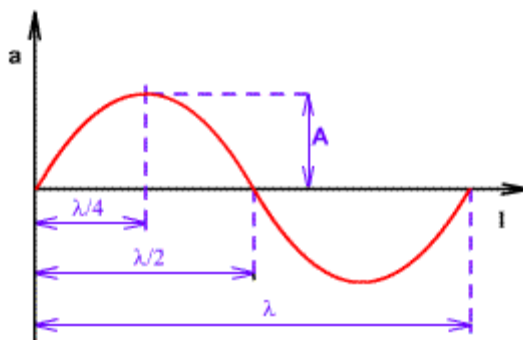
Po dosazení $T = 1/f$:

$$a = A \cdot \sin(2\pi \cdot f \cdot t) \quad (5.2)$$

kde A – amplituda signálu, f – frekvence [Hz], t – čas [s].

$$\omega = 2\pi \cdot f \quad (5.3)$$

ω – úhlová frekvence [$\text{rad}\cdot\text{s}^{-1}$]. Úhlová frekvence je definována jako změna fáze za jednotku času.



Obr. 24 Harmonická vlna. [41]

Harmonická vlna je dále charakterizována vlnovou délkou – λ [m], která udává vzdálenost dvou nejbližších bodů kmitajících ve stejné fázi. [41]

5.1.1 Šíření zvuku v prostoru

Zvuk se šíří ve volném prostoru všemi směry. Šíření zvuku je popsáno pomocí vlnoploch, které mají v případě bodového zdroje tvar koule. Pokud je tvar zdroje zvuku rovná deska, jsou vlnoplochy rovinné. V plynném prostředí se zvuk šíří jako postupné podélné vlnění. Ve vakuu se zvuk nešíří. Ve vzduchu se zvuk šíří rychlostí 340 m.s^{-1} . Rychlost šíření zvuku je závislá na prostředí. Dále je ovlivněna tlakem, teplotou a vlhkostí. Vlivem mechanického kmitání zdroje zvuku se částice vzduchu v některých místech přibližují a v jiných vzdalují, tím dochází ke změnám tlaku. Lidské ucho vnímá změny tlaku jako měnící se hlasitost zvuku. [41]

Hlasitost zvuku je popisována veličinou intenzita zvuku I [W.m^{-2}]. Je definována jako podíl výkonu P [W] a plochy S [m^2], kterou vlnění prochází. [42]

$$I = \frac{P}{S} \quad (5.4)$$

Lidské ucho je nejvíce citlivé na frekvence 700–6000 Hz. Práh slyšitelnosti lidského ucha je dán nejnižší intenzitou zvuku, kterou je schopno zaznamenat. Nejnižší intenzita zvuku $I_0 = 10^{-12} \text{ W.m}^{-2}$ a odpovídá hladině akustického tlaku $p_0 = 2 \cdot 10^{-5} \text{ Pa}$. Naopak hodnoty intenzity zvuku vyšší než $I = 1 \text{ W.m}^{-2}$ odpovídající akustickému tlaku $P = 130 \text{ Pa}$ jsou označovány jako práh bolesti a mohou poškodit sluch. Hlasitost zvuku lze také vyjádřit logaritmičticky jako hladinu intenzity zvuku a hladinu intenzity akustického tlaku. [42]

Hladinu intenzity zvuku lze vyjádřit:

$$L = 10 \cdot \log. \frac{I}{I_0} \quad (5.5)$$

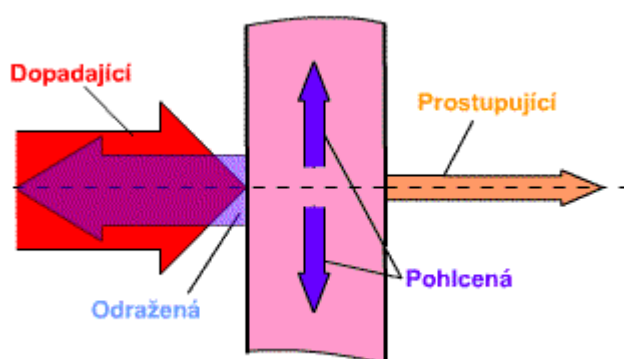
kde L – hladina intenzity zvuku [dB], I – intenzita zvuku [W.m^{-2}], I_0 – prahová hodnota intenzity zvuku [W.m^{-2}], $I_0 = 10^{-12} \text{ W.m}^{-2}$.

Hladina intenzity akustického tlaku:

$$L = 20 \cdot \log. \frac{p}{p_0} \quad (5.6)$$

kde L – hladina intenzity akustického tlaku [dB], p – akustický tlak [Pa], p_0 – prahová hodnota akustického tlaku [Pa], $p_0 = 2 \cdot 10^{-5} \text{ Pa}$.

Zvuková vlna je při šíření uzavřeným prostorem odráží od překážek. V případě dopadu vlny na překážku se část energie odráží a část je pohlcena. Intenzita drážené vlny je vždy nižší než intenzita dopadající vlny. Vlna se odráží vždy pod stejným úhlem, pod jakým dopadla na překážku. Množství pohlcené energie vlny je závislá na materiálu překážky. Pokud je překážka pevná dojde při odrazu ke změně fáze akustické vlny. U měkkých materiálů ke změně fáze nedochází. V případě, že je překážka několikanásobně větší, než je vlnová délka akustické vlny, vzniká ozvěna. Při šíření zvuku prostorem může dojít k ohybu vlny za překážku. K tomu dochází, pokud má předmět nebo otvor menší rozměry, než je vlnová délka akustické vlny. [43]



Obr. 25 Vlna dopadající na překážku. [41]

5.1.2 Šíření zvuku v pevných látkách a kapalinách

V pevných látkách se zvuk může šířit nejen jako podélné vlnění ale i jako příčné. Rychlost šíření zvuku v pevných látkách závisí na hustotě a pružnosti prostředí. V látkách, které mají velmi silné vazby mezi částicemi se zvuk šíří vyšší rychlostí. Rychlost šíření se naopak snižuje se zvyšující se hustotou látky. Například rychlost šíření zvuku oceli je $5941 \text{ m}\cdot\text{s}^{-1}$.

V kapalinách se zvuková vlna šíří jako podélné vlnění. Rychlost šíření zvuku v kapalinách je vyšší než v prostoru a zároveň nižší než v pevných látkách. Například rychlost zvuku ve vodě je $1481 \text{ m}\cdot\text{s}^{-1}$. Hustota vody je sice větší, než vzduchu ale na druhou stranu je také méně stlačitelná. Proto je rychlost zvuku vyšší než ve vzduchu. [44]

Rychlost šíření zvuku v závislosti na prostředí lze vyjádřit:

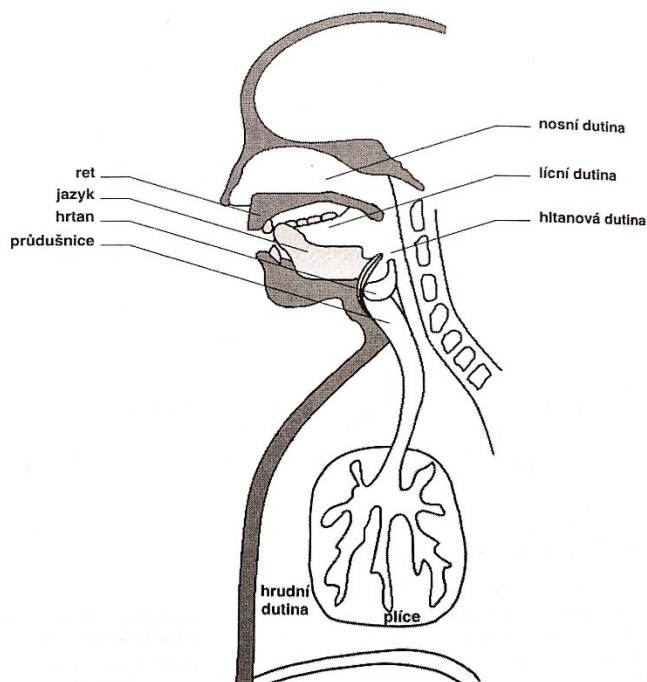
$$v = \sqrt{\frac{B}{\rho}} \quad (5.7)$$

kde v – rychlost zvuku [$\text{m}\cdot\text{s}^{-1}$], B – modul objemové pružnosti [Pa], ρ – hustota [$\text{kg}\cdot\text{m}^{-3}$].

5.2 Vznik mluveného slova

Člověk je schopný vytvářet hlasový signál pomocí hlasového ústrojí. Hlasové ústrojí se skládá z hlasivek a hrtanu. Hlasivky se skládají z dvou pružných vazů napnutých mezi štítnou chrupavku a hlasivkové chrupavky. Při řeči proudí vydechovaný vzduch přes štěrbinu a rozechvívá hlasivky. Štěrbina se stahuje a roztahuje prostřednictvím hrtanových svalů. Hlas pak vzniká souhrou hlasového ústrojí, horních dýchacích cest, jazyka, dutiny ústní, zubů a rtů. Jazyk umožňuje artikulaci, určuje kvalitu hlásek a tím srozumitelnost řeči. [45]

Plíce jsou zdrojem energie pro hlasové ústrojí. Kmitáním hlasivek se vytváří základní ton řeči, který je u mužů 90–120 Hz, u žen 150–300 Hz a u dětí 350–400 Hz. Jazyk, dutina ústní, nosní dutina, zuby a rty umožňují artikulaci, určuje kvalitu hlásek a tím srozumitelnost řeči. [46]



Obr. 26 Hlasové ústrojí. [47]

Neznělá hláska vzniká průchodem vzduchu zúženou hlasovou štěrbinou, kde se proudění vzduchu stává turbulentní. Znělá hláska vzniká, pokud je hlasová štěrbinu zcela uzavřená. Vzduch proudící hrtanem zde vytváří tlak a při náhlém uvolnění štěrbinu vznikne znělá hláska. [47]

Unikátnost každého hlasu je dána vlastnostmi hlasového ústrojí, které se u každého jedince liší. Jednotlivé části hlasového ústrojí mají u každého člověka jiné charakteristiky. Liší se

například délka hlasivek, postavení hlasových chrupavek, napětí hlasivek a tlak vydechovaného vzduchu. Barvu získává hlas až při průchodu nadhrtanovými prostory, které fungují jako rezonátory. Barvu dále určuje postavení jazyka, rtů a hrdla. [48]

II. PRAKTICKÁ ČÁST

6 ALGORITMUS PRO ANALÝZU AUDIOSIGNÁLU

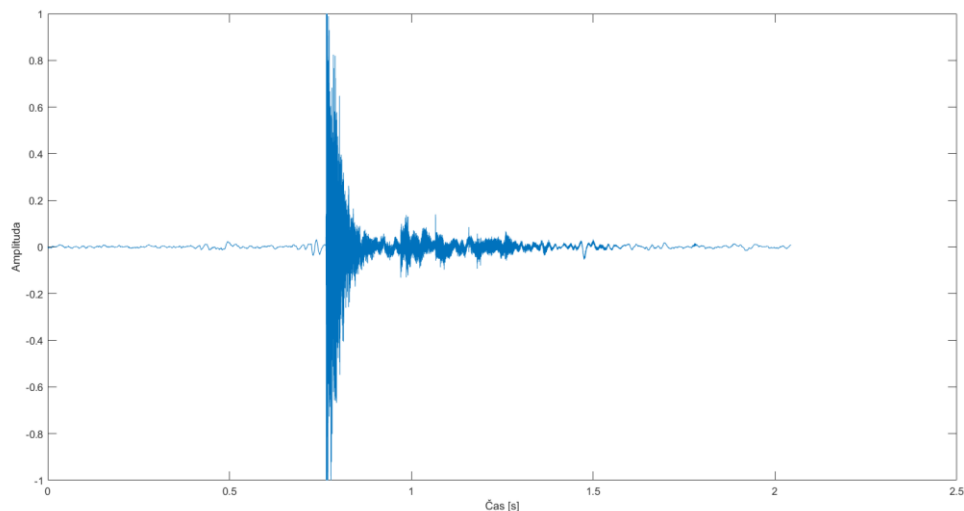
Cílem praktické části diplomové práce je navrhnout algoritmus pro rozpoznání zvuků typických pro neoprávněné vniknutí pachatele do objektu. Jedná se o zvuky jako rozbití skla, různé nárazy, řezání bruskou, řezání řetězovou pilou a kroky po objektu. Pro vytvoření algoritmu je využité vývojové prostředí Matlab.

6.1 Matlab

Matlab je nástroj pro vědecké a technické výpočty. Využívá se pro aplikace zahrnující zpracování signálu, modelování a simulace, analýzu dat, měření atd. Obsahuje také nástroje pro vývoj vlastního grafického uživatelského prostředí. Výhodou je, že vytvořené algoritmy jsou dále přenositelné a mohou být použity v jiných aplikacích. Matlab je založen na programovacím jazyku Fortran. Pro Matlab je dostupné velké množství knihoven funkcí, které rozšiřují jeho možnosti využití. [49]

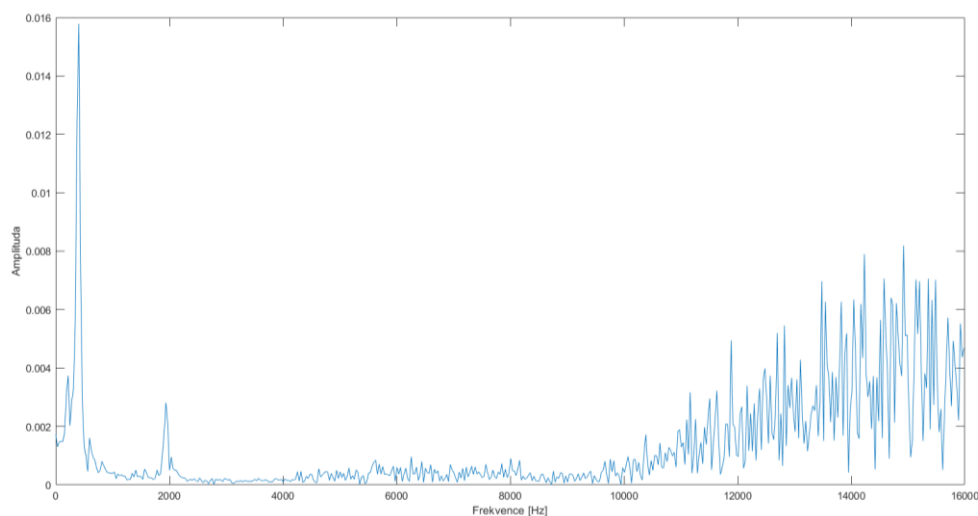
6.2 Analýza audiosignálu

Cílem analýzy audiosignálu je získat pohled na základní charakteristiku zvuků. Zvukový signál je charakterizován v časové a frekvenční doméně. Analyzované záznamy jsou ve formátu wav. Pro zvuky vznikající při vloupání do objektu je charakteristická jejich velká intenzita. Obvykle jsou to z pohledu času krátké děje o velké intenzitě. Může docházet k jejich opakování. Pro příklad je uveden časový průběh rozbití skleněné plochy viz Obr. 27. Na obrázku lze vidět, že celý děj trval okolo 1 vteřiny. Na začátku děje je vidět vysoká intenzita úderu na skleněnou plochu a její tříštění. Dále dochází k dopadání největších kusů skla na podlahu. Intenzita klesá při dopadání menších kusu na podlahu a na již ležící sklo na podlaze.



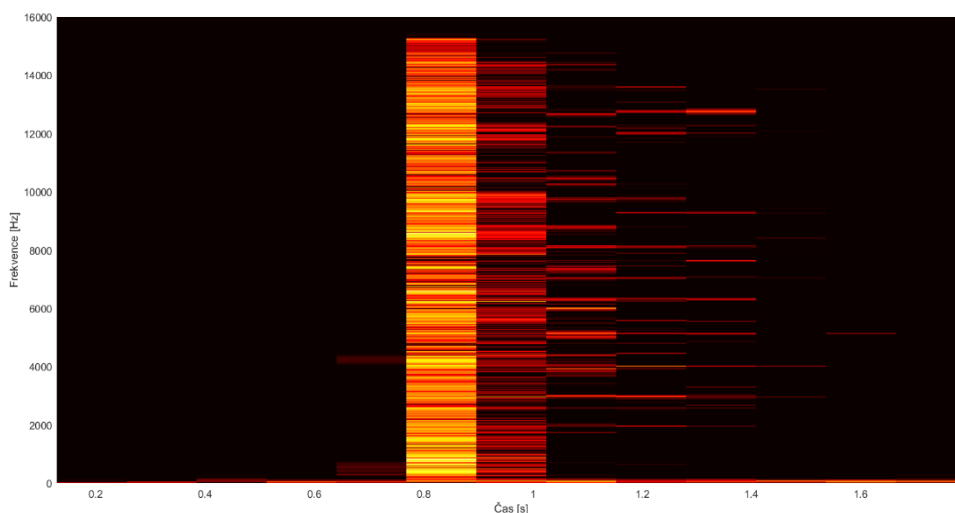
Obr. 27 Průběh rozbití skleněné plochy v časové doméně.

Dále byl audiosignál analyzován ve frekvenční doméně. Spektrum zobrazuje rozbití skleněné plochy viz Obr. 28. Na průběhu je vidět, že nejvyšší energie je obsažena na frekvencích do 400 Hz. Nižší frekvence je charakteristická pro dopad předmětu na skleněnou plochu a vyšší frekvence pro její rozbití.



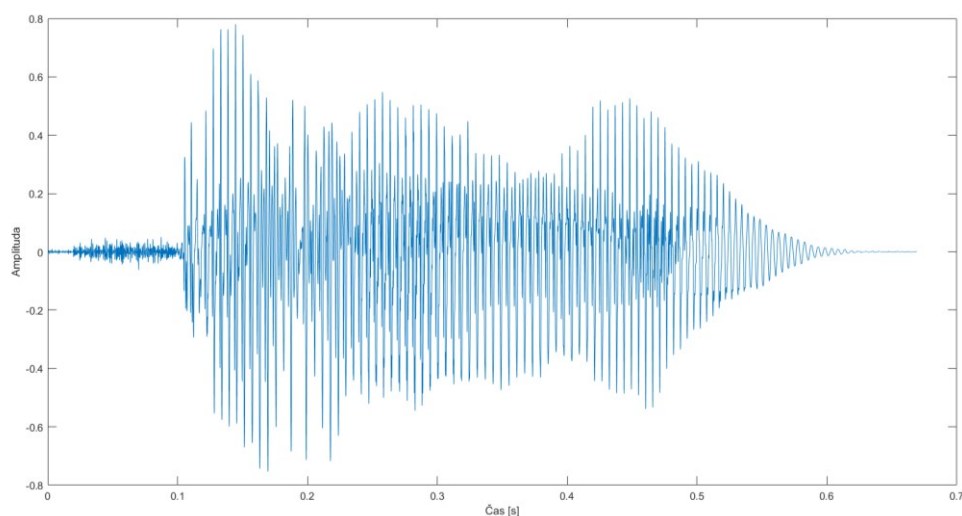
Obr. 28 Průběh rozbití skleněné plochy ve frekvenční doméně

Rozložení energie vznikající při rozbití skleněné plochy je dále zobrazena pomocí histogramu. Histogram zobrazuje, jak jsou v audio signálu zastoupeny jednotlivé frekvence. Na Obr. 29 je vidět, že nejvíce energie bylo uvolněno mezi 0,7 a 0,9 vteřinou. Frekvence s největším zastoupením jsou vyobrazeny žlutou barvou.



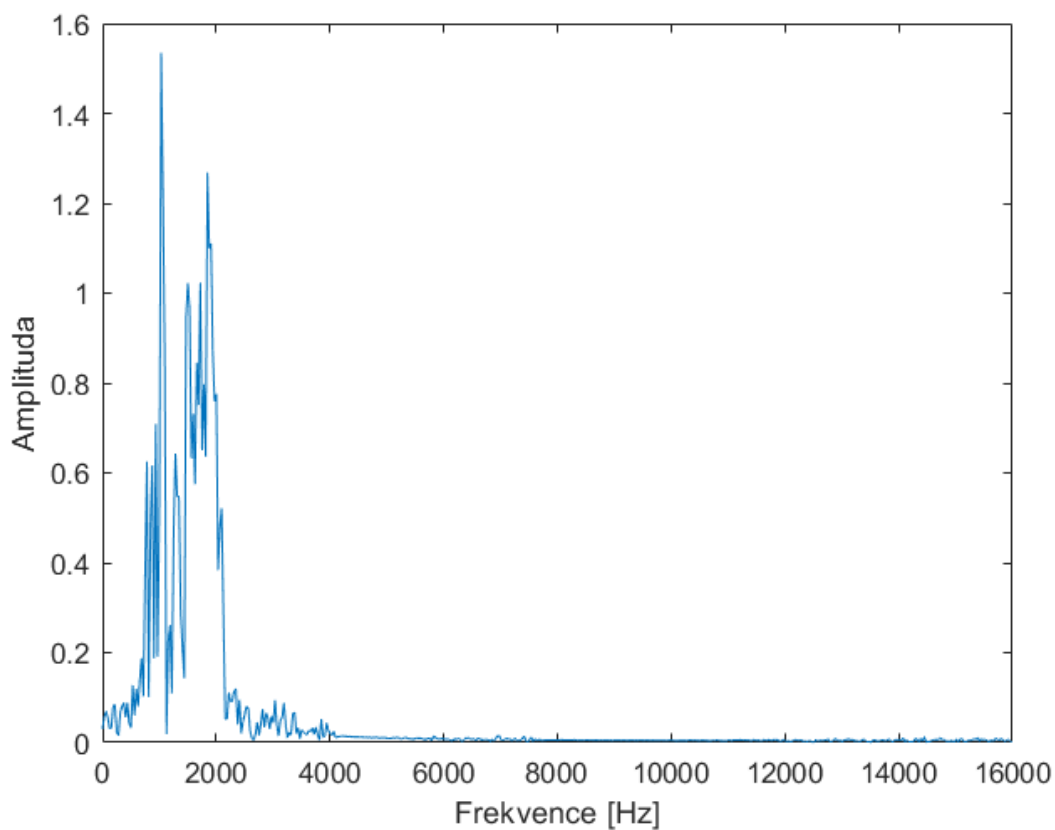
Obr. 29 Histogram rozbití skleněné plochy

Pro srovnání byl analyzován lidský hlas. Konkrétně ženský hlas vyslovující slovo „hello“. Opět je audiosignál zobrazen v časové a frekvenční doméně. Audiosignál nemá v časové doméně tak velkou amplitudu jako rozbití skleněné plochy. Energie je rovnoměrněji rozprostřena.



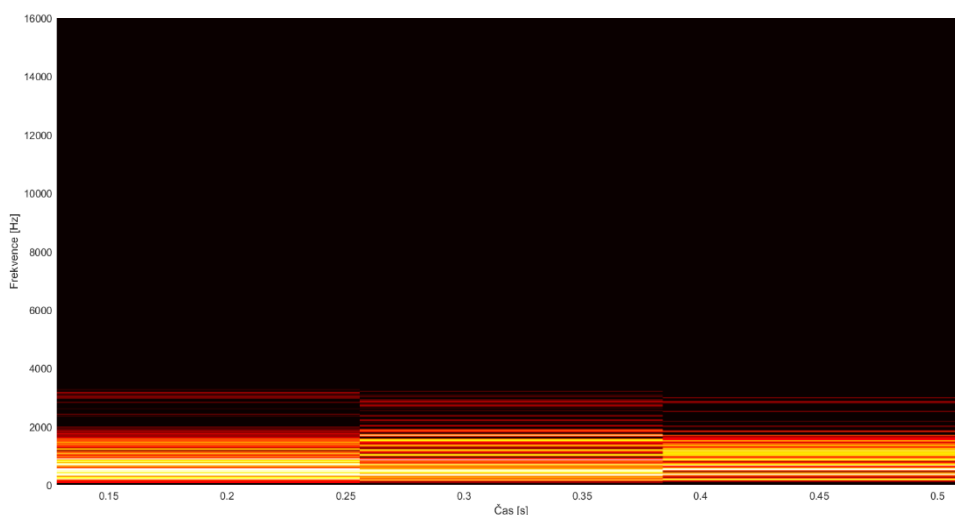
Obr. 30 Hlas v časové doméně („hello“)

Ve frekvenční spektru jde vidět, že největší energii mají frekvence mezi 1000 a 2000 Hz.



Obr. 31 Hlas v časové doméně („hello“)

Histogram zobrazuje rovnoměrné zastoupení frekvencí s maximální hodnotou do 2000 Hz. Zde je vidět rozdíl v porovnání s rozbitím skleněné plochy, kde jsou zastoupeny i vyšší frekvence.

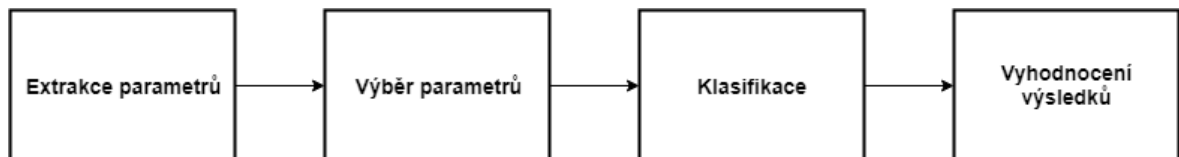


Obr. 32 Histogram hlasu („hello“)

Celý algoritmus analýzy audiosignálu je uveden v příloze P I: Algoritmus pro analýzu audiosignálu

7 ALGORITMUS PRO ROZPOZNÁVÁNÍ ZVUKŮ

Cílem algoritmu je rozpoznání zvuků typických pro vloupání do objektu. Postup vytvoření algoritmu se skládá z následujících kroků viz Obr. 33.



Obr. 33 Postup vytváření algoritmu pro rozpoznávání zvuku

Prvním krokem je extrakce parametrů audio signálu v časové a frekvenční doméně. Dále jsou vybrány ty parametry, které jsou schopné, co nejlépe odlišit zvuky vloupání od ostatních zvuků. Tyto parametry jsou získány z referenčních záznamů obsahující daný jev. Jsou tak vytvořena referenční data, která slouží pro porovnání s parametry analyzovaných souborů. Porovnáním dostatečného množství parametrů lze určit o jaký zvuk se jedná. Nakonec jsou získané výsledky statisticky zpracovány a je určena správnost detekce zvuků algoritmem.

7.1 Extrakce a výběr referenčních parametrů

Každý signál může být charakterizován množstvím parametrů. Jedná se o parametry charakterizující jak časovou, tak frekvenční stránku audio signálu. Parametry v časové doméně zobrazují, jak se signál mění v čase. Parametry ve frekvenční doméně ukazují, jaká část signálu leží na každé frekvenci.

Použité parametry v časové doméně:

- Maximum,
- medián,
- aritmetický průměr,
- peak to peak,
- peak to RMS,
- RMS,
- standartní deviace,
- variace,
- kurtosis.

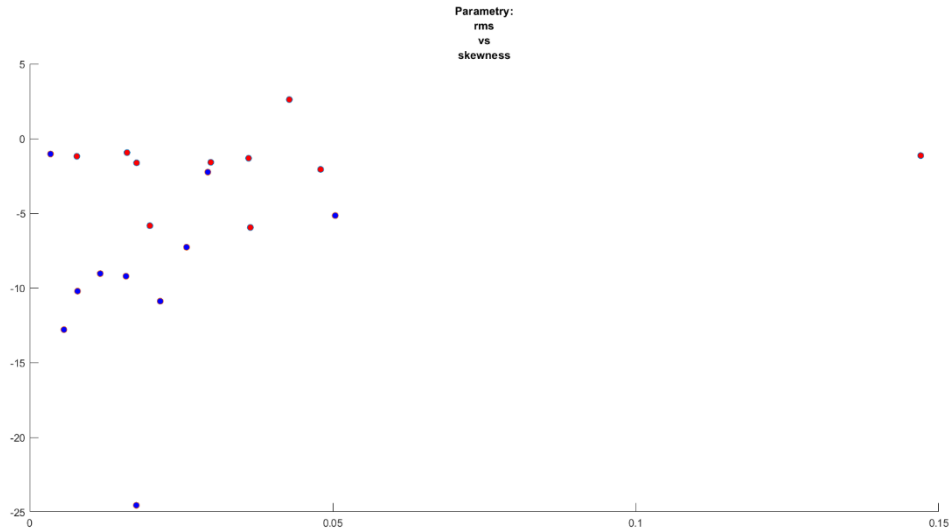
Použité parametry ve frekvenční doméně:

- Centroid,
- skewness,
- kurtosis.

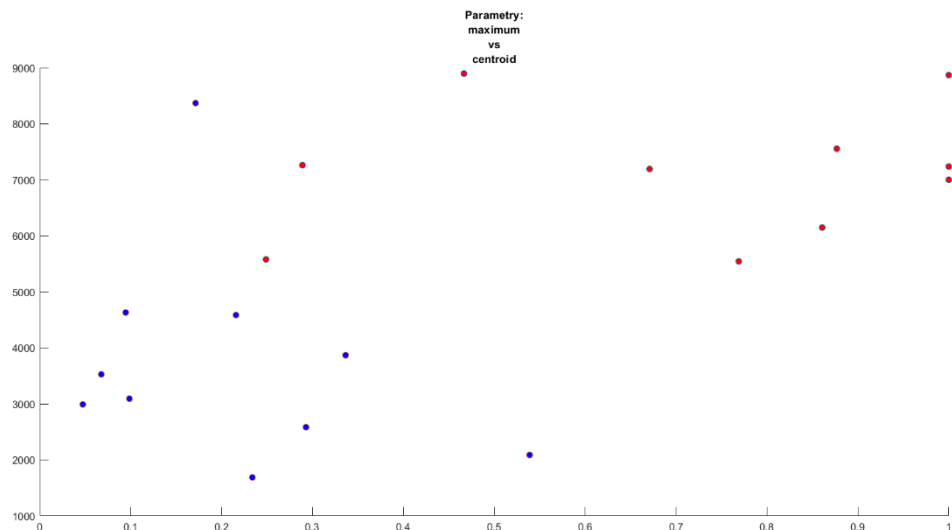
Parametry získané z referenčních audio záznamu byly zobrazeny v grafu po párech. Pro extrakci referenčních parametrů byly použity zvukové záznamy ve formátu wav. Tyto záznamy byly získány z volně dostupných zdrojů. Bylo tak možné určit, jaké parametry jsou nejlepší pro identifikaci daného typu zvuku. Celkově bylo použito 12 parametrů, které umožňují 66 kombinací zobrazených do grafů.

$$C_k(n) = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{12}{2} = \frac{12!}{2!(12-2)!} = 66 \quad (7.1)$$

Cílem výběr parametrů bylo získat 10 kombinací, které dokážou co nejlépe odlišit zvuky vznikající při vloupání od zvuků běžného provozu objektu. Porovnání parametrů je zobrazeno viz Obr. 34 a 35. Tyto grafy zobrazují porovnání parametrů pro odlišení zvuku rozbití skla od ostatních zvuků vznikajících při běžné činnosti v objektu a jeho okolí. Červenou barvou jsou v grafech zobrazeny záznamy zvuků obsahující rozbití skla a modrou ostatní jevy. Každý bod představuje jeden zvukový záznam. Pro odlišení obou jevů je důležité, aby body představující rozbití skla byly, co nejdále od ostatních bodů. Zároveň musí být body představující rozbití skla, co nejbližší u sebe. Na Obr. 34 je vidět oddělení bodů představujících rozbití skla od ostatních zvuků.

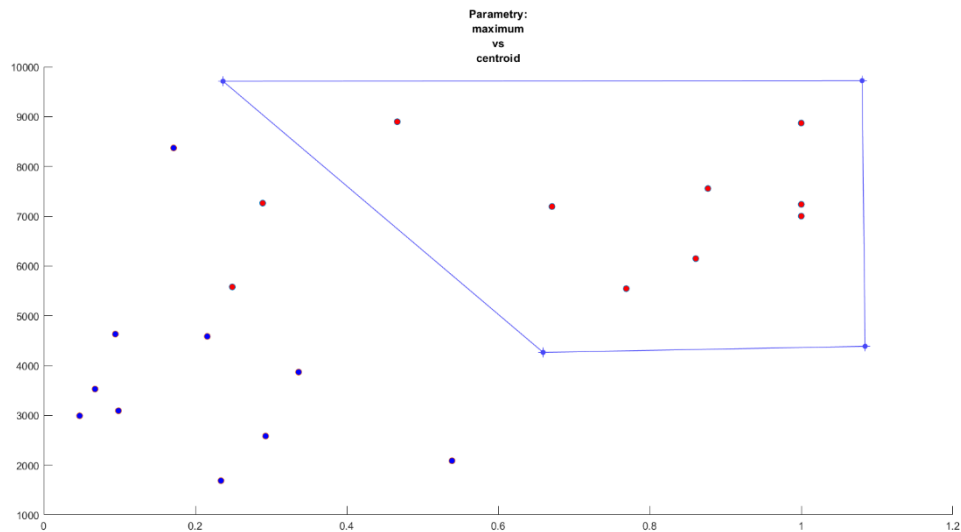


Obr. 34 Ukázka zobrazení porovnání parametrů nevhodných pro klasifikaci audiosignálu.



Obr. 35 Ukázka zobrazení porovnání parametrů vhodných pro klasifikaci audiosignálu.

Po zvolení nejlepší kombinace parametrů ve vybrána oblast referenčních dat. Výběr referenční oblasti je realizován prostřednictvím polygonu. Polygon je tvořen čtyřmi body, které umožňují přizpůsobit jeho tvar referenčním datům. Oblast výběru je nakonec uložena do souboru MS excel. Celý algoritmus extrakce referenčních dat v Matlabu je uveden v příloze P II: Algoritmus pro získání referenčních dat.



Obr. 36 Ukázka výběru referenční oblasti

Tabulka 1 Ukázka parametrů referenčních oblastí pro zvuk rozbití skla uložených v XLS

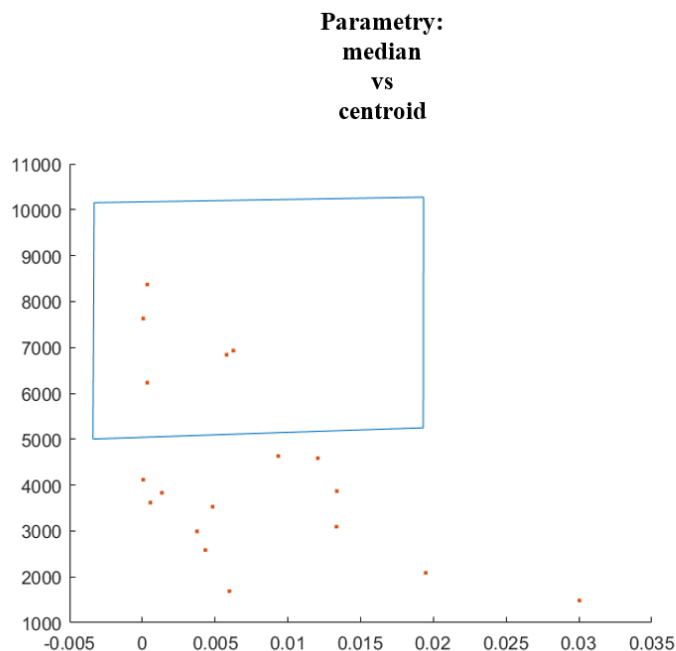
Maximum	Medián	Peak-to-Peak	Kurtosis_TD	Centroid	Kurtosis_FD
0,367339		1,215164		9652,218	5,020492
1,090726		2,633197		5109,879	5,498634
1,180242		1,979167		5123,656	309,597
0,431855		0,395492		9652,218	215,403
0,407661				3561,475	
1,114113				6293,716	
0,990726				10433,06	
0,218145				7646,175	
	0,000689			5002,049	
	0,046119			4965,164	
	0,046052			9563,525	
	0,000823			9502,049	
			180,9476	3968,579	
			1,108871	5116,12	
			1,310484	12173,5	
			180,746	7258,197	

7.2 Klasifikace

Cílem klasifikace je vyhodnocení analyzovaného záznamu. Algoritmus je schopný říct o jaký zvuk se jedná. V rámci praktické části byly vytvořeny dva algoritmy. První algoritmus má za cíl vyhodnotit, zda analyzovaný záznam neobsahuje jeden ze zvuků typických pro vloupání do objektu. Druhý algoritmus simuluje funkci detektoru rozbití skla.

7.2.1 Algoritmus pro rozpoznání zvuků typických pro vloupání do objektu

Prvním krokem algoritmu je načtení složky obsahující záznamy k analýze. Všechny záznamy musí být ve formátu wav. Na základě načtení složky je zjištěno, kolik obsahuje záznamů. Na základě počtu záznamů dojde přečtení jmen souborů a jejich uložení do proměnné „playlist“. Dále je v každém cyklu načtený jeden záznam. Z každého záznamu jsou extrahovány parametry analyzovaného záznamu v časové a frekvenční doméně. Jde o stejné parametry na základě, kterých byly získány hodnoty referenční oblasti. Referenční hodnoty jsou načteny z XLS souboru a porovnány s hodnotami získanými z analyzovaného záznamu viz Obr.37. Jde o porovnání dvojice parametrů ve dvojrozměrné matici. Body nacházející se uvnitř modrého polygonu představují parametry analyzovaných záznamů, které spadají do referenční oblasti.



Obr. 37 Porovnání hodnot referenční oblasti s parametry analyzovaných záznamů.

Po zjištění, které záznamy mají parametry v rozsahu referenčních oblastí je pomocí bodovacího systému rozhodnuto, jaký typ zvuku je na nahrávce. Za každou dvojici parametrů nacházející se v referenční oblasti jsou parametry určené pro daný typ zvuku obodovány jedním bodem. Počet bodů je vztažen k počtu dvojic parametrů pro každý zvuk. Maximální počet bodů je 5. Pro pozitivní rozpoznání daného zvuku je potřeba dosáhnout minimálního počtu alespoň 3 body. V případě, že ani jeden typ parametrů nenasbírá potřebný počet bodu je zvuk obsažený na nahrávce označen jako „Neznámý“. V případě že více typů parametrů dosáhne stejného počtu bodů je nahrávka označena jako „Nelze rozhodnout“.

7.2.1.1 Použité parametry pro analýzu záznamů

Klasifikace záznamů je založena na použití pěti dvojic parametrů v časové a frekvenční doméně pro každý typ zvuku.

Zvuk rozbití skla:

- Maximum : peak-to-peak,
- maximum : centroid,
- medián : centroid,
- kurtosis v časové doméně : centroid,
- kurtosis ve frekvenční doméně : centroid.

Zvuk úderů:

- Maximum : centroid,
- aritmetický průměr : peak-to-peak,
- RMS : kurtosis v časové doméně,
- medián : peak-to-peak,
- peak-to-peak : centroid.

Zvuk kroků:

- Maximum : peak-to-rms,
- maximum : RMS,
- kurtosis ve frekvenční doméně : skewness,
- peak-to-peak : peak-to-RMS,
- peak-to-RMS : skewness.

Zvuk brusky:

- Aritmetický průměr : centroid,
- medián : peak-to-RMS,
- peak-to-RMS : standartní deviace,
- standartní deviace : centroid,
- variace : centroid.

Zvuk řetězové pily:

- Maximum : aritmetický průměr,
- maximum : kurtosis v časové doméně,
- aritmetický průměr : peak-to-peak,
- peak-to-peak : kurtosis v časové doméně,
- peak-to-peak : centroid.

7.2.1.2 Výstup z algoritmu pro rozpoznání zvuků typických pro vloupání do objektu

Výstup z algoritmu je v textové podobě. Skládá se ze jména posuzovaného záznamu a typu zvuku, který obsahuje.

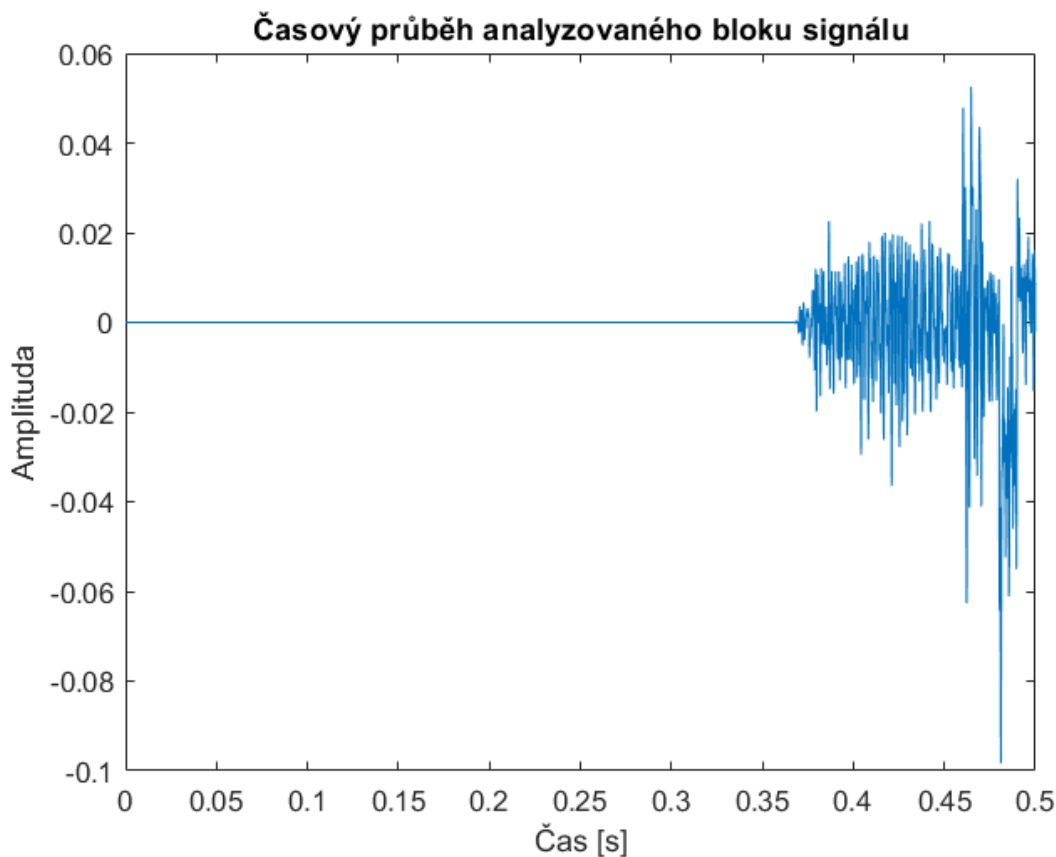
```
Soubor: test 1.wav Nelze rozhodnout
Soubor: test 2.wav Bruska!!!
Soubor: test 3.wav Kroky!!!
Soubor: test 4.wav Kroky!!!
Soubor: test 5.wav Kroky!!!
Soubor: test 6.wav Neznámy zvuk!!!
Soubor: test 7.wav Kroky!!!
Soubor: test 8.wav Tříštění skla!!!
Soubor: test 9.wav Tříštění skla!!!
Soubor: test10.wav Kroky!!!
Soubor: test11.wav Neznámy zvuk!!!
```

Obr. 38 Textový výstup do příkazového řádku

7.2.2 Simulace detektoru rozbití skla

Algoritmus simulující funkci detektoru rozbití skla je založen na snímání zvuku pomocí mikrofonu počítače. V první části jsou zaznamenávány 500 ms dlouhé bloky záznamu. Zde

jsou nastaveny hodnoty vzorkovací frekvence, počet bitů na vzorek a počet kanálů. Vzorkovací frekvence byla nastavena $F_s = 32\,000$ Hz, počet bitů na vzorek na 24 b a počet kanálů na 1. Hodnoty bloku jsou uloženy do proměnné. Aktuálně zpracovávaný blok je zobrazovaný pomocí grafu v časové doméně.



Obr. 39 Časový průběh analyzovaného bloku signálu

Každý blok je následně rozdělen na 250 ms dlouhé rámce. Pro analýzu jsou vybrány rámce, které neobsahují ticho. Rámce, které obsahují ticho nejsou na základě hraniční hodnoty amplitudy analyzovány. Z každého rámce jsou extrahovány parametry v časové a frekvenční doméně. Navíc je z každého bloku získán počet průchodů nulou. Tento parametr dokáže spolehlivě odlišit zvuk rozbití skla od ostatních běžných zvuků. Hraniční hodnota pro počet průchodů nulou byla stanovena na 1000. Pokud je tato hodnota překročena, je pravděpodobné, že v daném bloku záznamu se nachází zvuk rozbití skla. Zvolené dvojice parametrů jsou následně porovnány s hodnotami referenčních oblastí. Referenční parametry jsou uloženy v XLS souboru. Referenční parametry jsou extrahovány z 250 ms dlouhých záznamů přesně zachycujících rozpoznávaný jev rozbití skla. Následně je zjištěno, které dvojice pa-

parametrů získaných z analyzovaného rámce odpovídá referenčním oblastem a jsou obodovány. Výsledné skóre celého rámce může mít maximální hodnotu 13 bodů. Pro rozhodnutí, zda daný rámeček obsahuje zvuk rozbití skla je potřeba minimálně 9 bodů. Překročení hranice počtu průchodu akustického signálu nulou je ohodnocena třemi body. V případě, že bodové ohodnocení je nižší než 5, je rozhodnuto, že daný rámeček neobsahuje rozbití skla.

7.2.2.1 Použité parametry pro simulaci detektoru rozbití skla

Pro rozpoznání zvuku rozbití skla zaznamenaného pomocí mikrofону bylo použito 10 dvojic parametrů v časové a frekvenční doméně.

- Maximum : peak-to-peak,
- maximum : centroid,
- medián : centroid,
- kurtosis v časové doméně : centroid,
- kurtosis ve frekvenční doméně : centroid,
- maximum : skewness,
- maximum : kurtosis ve frekvenční doméně,
- aritmetický průměr : kurtosis v časové doméně,
- peak-to-peak : P2rms.

7.2.2.2 Výstup z algoritmu pro simulaci detektoru rozbití skla

Výstup z algoritmu je v textové podobě. Skládá se z typu zvuku a bodového ohodnocení analyzovaného rámce.

```
Recording...
Neznámý zvuk
score: 3
Neznámý zvuk
score: 2
Recording...
Tříštění skla!!!
score: 6
Tříštění skla!!!
score: 8
Recording...
Neznámý zvuk
score: 3
```

Obr. 40 Textový výstup do příkazového řádku

8 ZHODNOCENÍ EFEKTIVITY VYTVOŘENÝCH ALGORITMŮ

Efektivita obou vytvořených algoritmů byla posouzena z hlediska úspěšnosti, s jakou jsou schopny rozpoznat dané zvuky. Byly vytvořeny dva algoritmy. První algoritmus slouží k odlišení zvuků typických pro pokusy o vniknutí do objektu a druhý má za úkol simulovat funkci detektoru rozbití skla.

8.1.1 Efektivita algoritmus pro rozpoznání zvuků typických pro vloupání do objektu

Pro testování algoritmu byly použity záznamy obsahující jednotlivé typy zvuků, na které byl algoritmus vytvořen. Rozpoznání každého typu zvuků bylo ověřeno na osmi záznamech. Celkem bylo k testování použito 40 záznamů.

Tabulka 2 Výsledky testování algoritmu pro rozpoznání zvuků typických pro pokusy o vniknutí do objektu

Soubor	Výsledek	Vyhodnocení
Soubor: test 1_rozbití skla.wav	Nelze rozhodnout	Vyhodnoceno jako jedna z událostí
Soubor: test 2_rozbití skla.wav	Údery	Vyhodnoceno jako jedna z událostí
Soubor: test 3_rozbití skla.wav	Nelze rozhodnout	Vyhodnoceno jako jedna z událostí
Soubor: test 4_rozbití skla.wav	Údery	Vyhodnoceno jako jedna z událostí
Soubor: test 5_rozbití skla.wav	Rozbití skla	Správně vyhodnoceno
Soubor: test 6_rozbití skla.wav	Údery	Vyhodnoceno jako jedna z událostí
Soubor: test 7_rozbití skla.wav	Údery	Vyhodnoceno jako jedna z událostí
Soubor: test 8_rozbití skla.wav	Řetězová pila	Vyhodnoceno jako jedna z událostí
Soubor: test11_údery.wav	Údery	Správně vyhodnoceno
Soubor: test12_údery.wav	Údery	Správně vyhodnoceno
Soubor: test13_údery.wav	Údery	Správně vyhodnoceno
Soubor: test14_údery.wav	Nelze rozhodnout	Vyhodnoceno jako jedna z událostí
Soubor: test15_údery.wav	Údery	Správně vyhodnoceno
Soubor: test16_údery.wav	Údery	Správně vyhodnoceno
Soubor: test17_údery.wav	Údery	Správně vyhodnoceno
Soubor: test18_údery.wav	Údery	Správně vyhodnoceno
Soubor: test21_bruska.wav	Řetězová pila	Vyhodnoceno jako jedna z událostí

Soubor: test22_bruska.wav	Nelze rozhodnout	Vyhodnoceno jako jedna z událostí
Soubor: test23_bruska.wav	Údery	Vyhodnoceno jako jedna z událostí
Soubor: test24_bruska.wav	Bruska	Správně vyhodnoceno
Soubor: test25_bruska.wav	Bruska	Správně vyhodnoceno
Soubor: test26_bruska.wav	Bruska	Správně vyhodnoceno
Soubor: test27_bruska.wav	Bruska	Správně vyhodnoceno
Soubor: test28_bruska.wav	Nelze rozhodnout	Vyhodnoceno jako jedna z událostí
Soubor: test31_řetězová pila.wav	Neznámý zvuk	Vyhodnoceno jako jedna z událostí
Soubor: test32_řetězová pila.wav	Neznámý zvuk	Vyhodnoceno jako jednaz událostí
Soubor: test33_řetězová pila.wav	Řetězová pila	Správně vyhodnoceno
Soubor: test34_řetězová pila.wav	Řetězová pila	Správně vyhodnoceno
Soubor: test35_řetězová pila.wav	Řetězová pila	Správně vyhodnoceno
Soubor: test36_řetězová pila.wav	Řetězová pila	Správně vyhodnoceno
Soubor: test37_řetězová pila.wav	Řetězová pila	Správně vyhodnoceno
Soubor: test38_řetězová pila.wav	Bruska	Vyhodnoceno jako jedna z událostí
Soubor: test41_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test42_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test43_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test44_kroky.wav	Neznámý zvuk	Chybné vyhodnocení
Soubor: test45_kroky.wav	Neznámý zvuk	Chybné vyhodnocení
Soubor: test46_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test47_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test48_kroky.wav	Kroky	Správně vyhodnoceno
Soubor: test51_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno
Soubor: test52_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno
Soubor: test53_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno
Soubor: test54_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno
Soubor: test55_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno

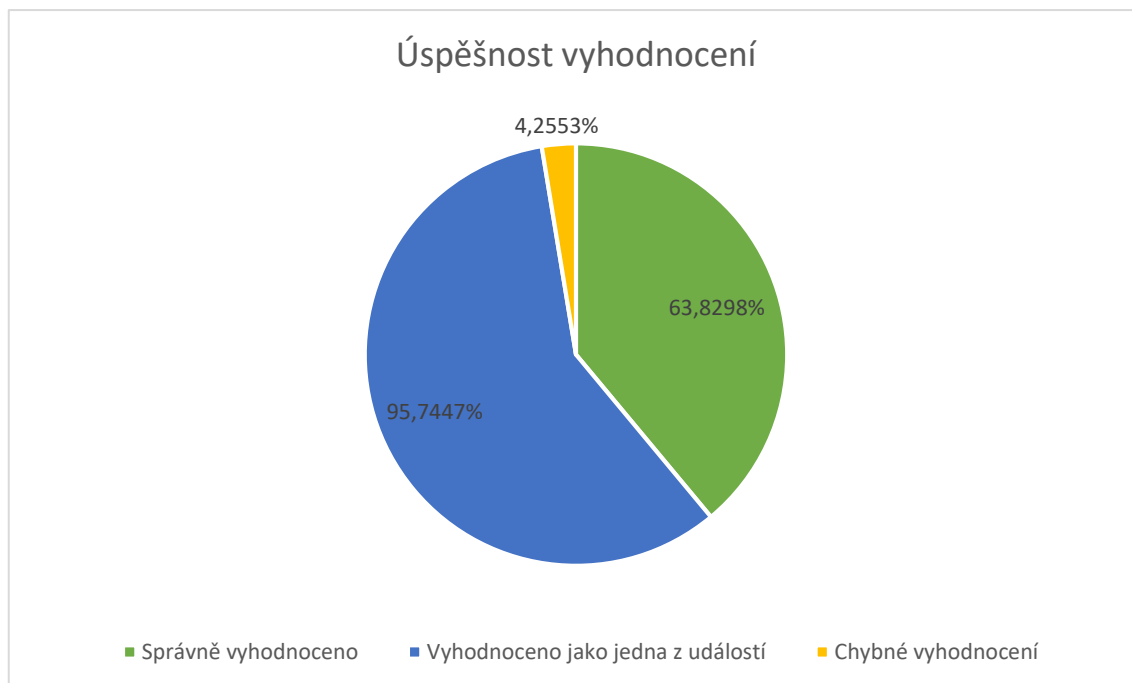
Soubor: test56_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno
Soubor: test57_ostatní zvuky.wav	Bruska	Chybné vyhodnocení
Soubor: test58_ostatní zvuky.wav	Neznámý zvuk	Správně vyhodnoceno

Z výsledků testování algoritmu vyplývá, že nejlepších výsledků dosahuje algoritmus při rozpoznávání zvuků úderů a kroků. Dále je méně úspěšný u zvuků řetězové pily a brusky. Velmi malou úspěšnost zaznamenal v případě rozpoznávání zvuků rozbití skla. Celkové je algoritmus schopný velmi spolehlivě odlišit zvuky typické pro pokusy o vniknutí do objektu od ostatních zvuků běžně přítomných v prostředí. Zvuky běžně vyskytující se v okolí objektu jsou ruch dopravy, ptáci a jiná zvířata, řeka, lidský hlas, déšť a vítr. Algoritmus je velmi úspěšný při odlišení zvuků typických pro vniknutí do objektu od ostatních zvuků. Při určování konkrétního typu zvuku je úspěšnost nižší. Procentuální úspěšnost algoritmu je zobrazena viz Tabulka 3 a Graf 1.

Tabulka 3 Procentuální vyhodnocení úspěšnosti

Vyhodnoceno	Úspěšnost
Správně vyhodnoceno	63,8298 %
Vyhodnoceno jako jedna z událostí	95,7447 %
Chybné vyhodnocení	4,2553 %

Graf 1 Grafické zobrazení úspěšnosti algoritmu



8.2 Efektivita algoritmu pro simulaci detektoru rozbití skla

Efektivita algoritmu byla testována z hlediska vzniku falešných poplachů a z hlediska úspěšnosti detekce akustického jevu rozbití skla.

8.2.1 Testování vzniku falešných poplachů při běžném provozu

V prvním případě byl algoritmus spuštěn po dobu jedné hodiny v prostředí, kde se vyskytovali akustické jevy běžného provozu objektu. Testovacím objektem byly vnitřní prostory rodinného domu. V testovacím prostoru se vyskytovali zvuky jako lidský hlas, zvuky vznikající při manipulaci s vybavením objektu a vnější zvuky z ulice, jako dopravní ruch. Algoritmus při testování nevykazoval žádnou reakci na zvuky vznikající při běžném provozu objektu. Lze tedy říct, že algoritmus je odolný proti vzniku falešných poplachů (NAR). Při pokusech vyvolat falešný poplach jiným typem zvuku, než je rozbití skla bylo nutné dosáhnout velmi vysoké intenzity akustického jevu. Poplach byl vyvolán silným nárazem dvou dřevěných desek a také silným úderem do papírové krabice. Stejným způsobem reaguje reálná detektor rozbití skla. Stejně jako u reálného detektoru je i v případě simulačního algoritmu možné přizpůsobit citlivost detekce v závislosti na prostředí.

8.2.2 Testování reakce na zvuky rozbití skla

Testování reakce na zvuky rozbití skla bylo provedeno prostřednictvím nahrávky obsahující zvuky typické pro městské prostředí. Nahrávka obsahuje lidský hlas, dopravní ruch, zvuky tramvajové dopravy a ruch budovy železniční stanice. Tato nahrávka byla doplněna o 40 akustických jevů rozbití skla. Šlo o několik různých záznamů, které byly do nahrávky přimíchány. Následně byla nahrávka reprodukována pomocí reproduktoru v místnosti se spuštěným algoritmem. Algoritmus byl nejprve testován při nastavení hraniční hodnoty bodového ohodnocení na 6 bodů. Při tomto nastavení vykazuje algoritmus vyšší citlivost na zvuky rozbití skla. Je ale také více citlivý na okolní zvuky. Ve výsledku dokázal algoritmus detekovat všech 40 záznamů. Úspěšnost byla tedy 100 %. Druhé testování bylo provedeno při nastavení bodové hranice citlivosti na 9 bodů. Algoritmus byl schopný detekovat 35 ze 40 jevů rozbití skla. To znamená úspěšnost detekce 87,5 %.

Nastavení citlivosti detekce je nutné přizpůsobit provozním podmínkám v objektu.

Úspěšnost algoritmu je na základě provedených testů velmi vysoká. Algoritmus může konkurovat reálnému detektoru rozbití skla.

ZÁVĚR

Diplomová práce byla zaměřena na systémy využívající zpracování akustických signálů v bezpečnostních technologiích. Zvláště se jednalo detektory rozbití skla a biometrické systémy založené na rozpoznávání hlasu. V rámci teoretické části byly uvedeny systémů rozpoznávání zvuku v bezpečnostních technologiích. Ty byly zpracovány od těch nejjednodušší, jako jsou detektory rozbití skla až po složitější aplikace představující systémy rozpoznávání hlasu. Hlavním tématem teoretické části bylo digitální zpracování zvukových signálů. Je zde popsán princip digitalizace analogového akustického signálu. Po digitalizaci následuje extrakce parametrů využitelných pro rozpoznávání zvuku. Jsou zde popsány charakteristické parametry, které se využívají v systémech rozpoznávání zvuku. Jedná se o parametry jak v časové, tak frekvenční doméně akustického signálu. V systémech rozpoznávání hlasu se hojně využívá algoritmů strojového učení. V této kapitole jsou popsány nejvyužívanější algoritmy strojového učení. Závěrem teoretické části byl popsán fyzikální princip šíření zvuku v prostoru a pevných látkách. Je zde také popsán princip vzniku lidského hlasu.

Výsledkem praktické části jsou dva algoritmy. Algoritmy byly vytvořeny v prostředí softwarového nástroje Matlab. První algoritmus má za cíl rozpoznat zvuky, které jsou typické pro pokusy o vniknutí do objektu. Jedná se o zvuky provázené rozbitím skla, řezání bruskou a motorovou pilou, kroky v objektu a různé nárazy do konstrukcí budovy a vnitřního vybavení. Druhý algoritmus je zaměřen na simulaci detektoru rozbití skla. Algoritmus je na základě snímání zvukového pozadí prostřednictvím mikrofону schopen rozpoznat zvuky rozbití skla. Oba algoritmy jsou založeny na extrakci parametrů typických pro rozpoznávané zvuky jak v časové, tak frekvenční doméně. Na základě těchto parametrů je vytvořena referenční oblast, která umožňuje porovnávat parametry referenčních a analyzovaných zvuků. Prostřednictvím srovnání parametrů jsou algoritmy schopny odlišit analyzované akustické jevy. Nakonec byla vyhodnocena úspěšnost vytvořených algoritmů. Algoritmus určený pro rozpoznávání zvuků typických pro vniknutí do objektu nebyl tolik nebyl z hlediska určení konkrétního zvuku tolik úspěšný. Byl však schopný odhalit, že analyzovaný zvuk je ten, který je typický pro vloupání. Na druhou stranu nevykazoval reakci na zvuky vznikající běžnou činností v objektu. Algoritmus pro simulaci detektoru rozbití skla měl velmi vysokou úspěšnost. Z hlediska reakce na zvuky běžné v testovaném objektu nevykazoval rušení danými zvuky. Nedošlo tak ke vzniku falešných poplachů. Při testování reakce na zvuky obsahující rozbití skla dokázal odhalit všechny úseky testovací nahrávky obsahující zvuk rozbití skla.

Přínosem mé diplomové práce je vytvoření algoritmů určených k rozpoznávání zvuku na základě parametrů charakterizujících audio signál v časové a frekvenční doméně. Ověření spolehlivosti vytvořených algoritmu a tím i jejich využitelnosti v praxi. Pro mě je hlavním přínosem získání znalostí z oboru zpracování signálu a zkušenosti vývojem algoritmů v softwarovém nástroji Matlab.

SEZNAM POUŽITÉ LITERATURY

- [1] UHLÁŘ, Jan. *Technická ochrana objektů*. 2. vyd. Praha: Policejní akademie České republiky v Praze, 2009, 229 s. ISBN 978-80-7251-313-0.
- [2] LUKÁŠ, L. *Bezpečnostní technologie, systémy a management I*. 1. vyd. Zlín: VeRBuM, 2011, 316 s. ISBN 978-80-87500-05-7.
- [3] Navrhujeme odolný detektor tříštění skla – 1. část. *HW.cz* [online]. 22 prosinec 2010. Dostupné z: <http://vyvoj.hw.cz/teorie-a-praxe/navrhujeme-odolny-detektor-tristení-skla-1cast.html>
- [4] Navrhujeme odolný detektor tříštění skla – 2. část. *HW.cz* [online]. 22 prosinec 2010. Dostupné z: <http://vyvoj.hw.cz/teorie-a-praxe/navrhujeme-odolny-detektor-tristení-skla-2cast.html>
- [5] MLČOCH, Martin. *Elektroakustické detektory v zabezpečovací technice*. Zlín, 2017. Bakalářská práce. Univerzita Tomáše Bati ve Zlíně.
- [6] Systémy kontroly a řízení vstupu v bezpečnostních aplikacích (ACS / Acces Control Systems). *Security Technologies* [online]. [cit. 2019-02-04]. Dostupné z: <https://www.security.cz/systemy-kontroly-a-rizeni-vstupu--2431.html>
- [7] DRAHANSKÝ, Martin a Filip ORSÁG. *Biometrie*. [Brno: M. Drahanský], 2011. ISBN 978-80-254-8979-6.
- [8] ČERMÁK, Miroslav. Autentizace: hlasová biometrie. *Clever and Smart* [online]. [cit. 2019-02-04]. Dostupné z: <https://www.cleverandsmart.cz/autentizace-hlasova-biometrie/>
- [9] RŮŽIČKA, Jan. Ověřování totožnosti na základě hlasu a obsahu: Hlasová biometrie nadchla banky, pomáhá klientovi i operátorovi. *CFO World.cz* [online]. [cit. 2019-02-05]. Dostupné z: <https://cfoworld.cz/trendy/overovani-totoznosti-na-zaklade-hlasu-a-obsahu-hlasova-biometrie-nadchla-banky-pomaha-klientovi-i-operatorovi-3184>
- [10] Zakariah, M., Khan, M.K. & Malik, H. *Multimed Tools Appl* (2018) 77: 1009. <https://doi.org/10.1007/s11042-016-4277-2>
- [11] VICHLENDÁ, Milan. *Kriminalistika* [online]. Zlín: Střední odborná škola ochrany osob a majetku, 2011 [cit. 2019-02-04]. Dostupné z: <http://www.sosoom-zlin.cz/media/skripta/kriminalistika.pdf>
- [12] Mikrofon. In: *Wikiwand* [online]. [cit. 2019-02-21]. Dostupné z: <http://www.wikiwand.com/sk/Mikrof%C3%B3n>
- [13] PRŮCHA, Petr. Začínáme s mikrofony. *Frontman* [online]. [cit. 2019-02-21]. Dostupné z: <http://frontman.cz/zaciname-s-mikrofony>
- [14] PRŮCHA, Petr. Elektrostatický mikrofon. *Encyklopedie fyziky* [online]. [cit. 2019-02-21]. Dostupné z: <http://fyzika.jreichl.com/main.article/view/388-elektrostaticky-mikrofon>
- [15] Princip funkce mikrofónů. In: *Haaf* [online]. [cit. 2019-02-21]. Dostupné z: <http://haaf.cz/view.php?cislocianku=2007060305>

- [16] Filtering: How to Remove Interference from your Signal. *Windmill Software Ltd* [online]. [cit. 2019-02-27]. Dostupné z: <http://www.windmill.co.uk/filter.html>
- [17] What is Analog Filter? – Different Types of Analog Filters. *Windmill Software Ltd* [online]. [cit. 2019-02-27]. Dostupné z: <https://www.elprocus.com/types-of-analog-filters/>
- [18] *Digital Filters Part 1* [online]. 21. 2. 2011 [cit. 2019-02-27]. Dostupné z: <https://www.youtube.com/watch?v=loHy8v9A8LY>
- [19] VOJÁČEK, Antonín. Použití filtrů FIR v digitálním zpracování signálů. *Automatizace.hw.cz* [online]. 18.11.2004 [cit. 2019-02-27]. Dostupné z: <https://automatizace.hw.cz/clanek/2005110801>
- [20] IIR filter what does it mean?. *Automatizace.hw.cz* [online]. 10.6.2011 [cit. 2019-02-27]. Dostupné z: <https://electronics.stackexchange.com/questions/15206/iir-filters-what-does-infinite-mean>
- [21] Filtry s nekonečnou impulsní charakteristikou. *Matematická biologie* [online]. [cit. 2019-02-27]. Dostupné z: <http://portal.matematickabiologie.cz/index.php?pg=analyza-a-modelovani-dynamicky-biologicky-dat--linearni-a-adaptivni-zpracovani-dat--linearni-filtrace-filtry-1--filtry-s-nekonecnou-impulsni-charakteristikou>
- [22] IIR Filter. *Super Secret Website* [online]. [cit. 2019-02-27]. Dostupné z: <https://sites.google.com/site/reallysupersecretwebsite/home/iir-filter>
- [23] KADLEC, František. Zpracování akustických signálů. Praha: Vydavatelství ČVUT, 2005, 189 s. ISBN 8001025888.
- [24] ČERNOCKÝ, Jan. *Zpracování řečových signálů — studijní opora* [online]. 2006 [cit. 2019-03-08]. Dostupné z: http://www.fit.vutbr.cz/study/courses/ZRE/public/opora/zre_opora.pdf
- [25] HORÁK, Karel. Úvod do zpracování signálů. *Multimediální interaktivní didaktický systém* [online]. [cit. 2019-02-27]. Dostupné z: http://midas.uamt.feec.vutbr.cz/ZVS/Lectures/01_Uvod_do_zpracovani_signalu.pdf
- [26] SVOBODA, Viktor. Teoretické minimum o signálech pro systémové inženýry – 3. část. *Music store* [online]. [cit. 2019-03-08]. Dostupné z: <http://www.music-store.cz/clanky/teoreticke-minimum-o-signalech-pro-systemove-inzenyry-3-cast>
- [27] ROBINSON, Tony. *Speech Analysis* [online]. 1998 [cit. 2018-11-22]. Dostupné z: https://www.researchgate.net/publication/2443731_Speech_Analysis
- [28] Ultra Analog VA – manual. *Applied Acoustics Systems* [online]. [cit. 2019-03-08]. Dostupné z: <https://www.applied-acoustics.com/ultra-analog-va-2/manual/>
- [29] The mel scale. *Research Gate* [online]. [cit. 2019-03-08]. Dostupné z: https://www.researchgate.net/figure/The-mel-scale_fig7_236646001
- [30] Medián. *Statistika* [online]. [cit. 2019-03-08]. Dostupné z: <http://www.ekovyp.cz/median/>

- [31] PEETERS, Geoffroy. *A Large Set of Audio Features for Sound Description* [online]. 2004, , 1-25 [cit. 2018-11-22]. Dostupné z: http://recherche.ir-cam.fr/equipes/analyse-synthese/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf?fbclid=IwAR3iETsatY5uVYtk0iLjwExgrody0TC2DYY4OH1MMrfXcjkRGGH7dIXkdn4
- [32] How is the kurtosis of a distribution related to the geometry of the density function?. *Stack Exchange* [online]. [cit. 2019-03-08]. Dostupné z: <https://stats.stackexchange.com/questions/84158/how-is-the-kurtosis-of-a-distribution-related-to-the-geometry-of-the-density-fun>
- [33] The statistical nature of turbulent barotropic ocean jets. *Science Direct* [online]. 2017 [cit. 2019-03-10]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1463500317300392>
- [34] BRANSCOMB, Mary. Strojové učení ve firemním IT. *Computerworld* [online]. 2018 [cit. 2019-03-10]. Dostupné z: <https://computerworld.cz/software/strojove-uceni-ve-firemnim-it-55114>
- [35] KURFÜRSTOVÁ, Jana. Strojové učení kouzla zbavené. *EDTECH KISK* [online]. 2018 [cit. 2019-03-10]. Dostupné z: <https://medium.com/edtech-kisk/strojov%C3%A9-u%C4%8Den%C3%AD-kouzla-zbaven%C3%A9-e066d79ebe51>
- [36] SMÉKAL, Zdeněk, Hicham ATASSI, Vojtěch STEJSKAL a Jiří MEKYSKA. Soubor programů pro práci se skrytými Markovovými modely (HTK). *Elektrorevue* [online]. 2009(11) [cit. 2019-03-12]. Dostupné z: <http://www.elektrorevue.cz/cz/download/soubor-programu-pro-praci-se-skrytymi-markovovymi-modely--htk-/>
- [37] Metody kompenzace šumu pro rozpoznávače řeči využívajících skrytých Markovových modelů. *Laboratoř zpracování řeči a analýzy signálů* [online]. [cit. 2019-03-12]. Dostupné z: <http://noel.feld.cvut.cz/vyu/tss/hmm9.htm>
- [38] Support Vector Machines. *Strojové učení* [online]. [cit. 2019-03-12]. Dostupné z: <http://smnd.sk/mcibula/alg/SVM.html>
- [39] DURČÁK, Pavel. Neuronové sítě a princip jejich fungování. *Strojové učení* [online]. 2017 [cit. 2019-03-12]. Dostupné z: <https://www.napocitaci.cz/33/neuronove-site-a-princip-jejich-fungovani-uniqueid-gOke4NvrWuNY54vrLeM670eFNQh552VdDDulZX7UDBY/>
- [40] Koncept umělé neuronové sítě. *Matematická biologie* [online]. [cit. 2019-03-12]. Dostupné z: <http://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickych-dat--umela-inteligence--neuronove-site-jednotlivy-neuron--uvod-do-neuronovych-siti--koncept-umele-neuronove-site>
- [41] Akustika, vznik a šíření zvuku, frekvenční analýza a syntéza, sluchový vjem zvukového signálu. *Vysoká škola báňská — Technická univerzita Ostrava* [online]. 13 duben 2005. Dostupné z: http://homen.vsb.cz/~ber30/texty/varhany/anatomie/pistaly_akustika.htm
- [42] RŮŽIČKA, Jan. Hlasitost a intenzita zvuku. *Encyklopedie fyziky* [online]. [cit. 2019-02-07]. Dostupné z: <http://fyzika.jreichl.com/main.article/view/208-zakladni-definice>

- [43] Zvuk v uzavřeném prostoru. *Encyklopedie fyziky* [online]. [cit. 2019-02-12]. Dostupné z: <http://fyzika.jreichl.com/main.article/view/1180-zvuk-v-uzavrenem-prostoru>
- [44] Relative speed of sound in solids, liquids and gases. *Khanova škola* [online]. [cit. 2019-02-12]. Dostupné z: <https://khanovaskola.cz/video/8/27/2605-relativni-rychlost-zvuku-v-pevnych-latkach-kapalinach-a-plynech>
- [45] Vznik lidského hlasu. *Encyklopedie fyziky* [online]. [cit. 2019-02-12]. Dostupné z: <http://fyzika.jreichl.com/main.article/view/188-vznik-lidskeho-hlasu>
- [46] ČERNOCKÝ, Jan. *Předzpracování řeči, tvorba řeči, cepstrum* [online]. [cit. 2019-03-12]. Dostupné z: https://www.fit.vutbr.cz/study/courses/ZRE/public/pred/03_prepro_model_ceps/03_prepro_model_ceps.pdf
- [47] RAK, Roman, Vašek MATYÁŠ a Zdeněk ŘÍHA. *Biometrie a identita člověka ve forezních a komerčních aplikacích*. Praha: Grada, 2008, 631 s., 32 s. barev. obr. příl. Profesionál. ISBN 978-80-247-2365-5.
- [48] KRČMOVÁ, Marie. Ústrojí hlasové (fonační). *Fonetika a fonologie* [online]. [cit. 2019-02-12]. Dostupné z: <https://is.muni.cz/elportal/estud/ff/js08/fonetika/ucebnice/ch05s02s02.html>
- [49] MATLAB. *Humusoft* [online]. [cit. 2019-04-10]. Dostupné z: <https://www.humusoft.cz/matlab/details/>

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

MCU	Microcontroller unit
FRR	False Rejection Rate
FAR	False Acceptance Rate
FET	Field-Effect Transistor
FIR	Finite impulse response
IIR	Infinite impulse response
DFR	Discrete Fourier Transform
FFT	Fast Fourier Transform
MFFC	Mel Frequency Cepstral Coefficient
LP	Linear Prediction
LPCC	Linear Prediction Cepstral Coefficients
RMS	Root Mean Square
NAR	Nuisance Alarm Rate

SEZNAM OBRÁZKŮ

Obr. 1 Časový průběh akustického signálu rozbití skleněné výplně. [3]	12
Obr. 2 Frekvenční spektrum akustického signálu rozbití skleněné výplně. [3].....	13
Obr. 3 Schéma zpracování akustického signálu analogovým obvodem.....	14
Obr. 4 Průběh digitálního zpracování akustického signálu detektorem. [5].....	16
Obr. 5 Vývojový diagram pro vyhodnocení akustické události detektorem. [5].....	17
Obr. 6 Schéma biometrického systému.	19
Obr. 7 Schéma zpracování hlasového signálu.	22
Obr. 8 Piezoelektrický mikrofon. [12].....	28
Obr. 9 Elektretový mikrofon. [15].....	28
Obr. 10 Typy filtrů podle funkce.	29
Obr. 11 Funkce filtru s konečnou impulsní odezvou. [20]	31
Obr. 12 Funkce filtru s nekonečnou impulsní odezvou. [22]	32
Obr. 13 Vstupní signál v časové doméně.	35
Obr. 14 Frekvenční spektrum signálu.....	35
Obr. 15 Formantové frekvence. [28]	38
Obr. 16 Grafické znázornění nelineární závislosti převodu z Herzů na Mely. [29] ...	39
Obr. 17 LP spektrum. [27].....	40
Obr. 18 Hodnota parametru kurtosis v závislosti na plochosti distribuce signálu. [32]	42
Obr. 19 Hodnoty spektrální šikmosti. [33]	43
Obr. 20 Příklad Markovova modelu se šesti stavy. [36].....	46
Obr. 21 Grafické zobrazení algoritmu Support Vector Machine. [38].....	47
Obr. 22 Model neuronu. [39].....	48
Obr. 23 Model neuronové sítě. [40].....	48
Obr. 24 Harmonická vlna. [41].....	49
Obr. 25 Vlna dopadající na překážku. [41].....	51
Obr. 26 Hlasové ústrojí. [47]	52
Obr. 27 Průběh rozbití skleněné plochy v časové doméně.	56
Obr. 28 Průběh rozbití skleněné plochy ve frekvenční doméně	56
Obr. 29 Histogram rozbití skleněné plochy	57
Obr. 30 Hlas v časové doméně („hello“)	57
Obr. 31 Hlas v časové doméně („hello“)	58

Obr. 32 Histogram hlasu („hello“)	58
Obr. 33 Postup vytváření algoritmu pro rozpoznávání zvuku	59
Obr. 34 Ukázka zobrazení porovnání parametrů nevhodných pro klasifikaci audiosignálu	61
Obr. 35 Ukázka zobrazení porovnání parametrů vhodných pro klasifikaci audiosignálu	61
Obr. 36 Ukázka výběru referenční oblasti	62
Obr. 37 Porovnání hodnot referenční oblasti s parametry	63
Obr. 38 Textový výstup do příkazového řádku	65
Obr. 39 Časový průběh analyzovaného bloku signálu	66
Obr. 40 Textový výstup do příkazového řádku	67

SEZNAM TABULEK

Tabulka 1 Ukázka parametrů referenčních oblastí pro zvuk rozbití skla uložených v XLS	62
Tabulka 2 Výsledky testování algoritmu pro rozpoznání zvuků typických pro pokusy o vniknutí do objektu.....	68
Tabulka 3 Procentuální vyhodnocení úspěšnosti	70

SEZNAM PŘÍLOH

P I Dokumentační CD obsahující:

- Algoritmus pro analýzu audiosignálu,
- algoritmus pro získání referenčních dat,
- algoritmus pro klasifikaci většího počtu zvuků,
- algoritmus pro simulaci detektoru rozbití skla,
- referenční data ve formátu XLS,
- výsledky vyhodnocení efektivity algoritmů.