



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

## ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

## ANALYTICKÝ NÁSTROJ PRO GENEROVÁNÍ BICÍCH TRIGGERŮ Z DOWNMIX ZÁZNAMU

ANALYSING TOOL FOR GENERATING OF DRUM TRIGGERS FROM DOWNMIX RECORD

### DIPLOMOVÁ PRÁCE

MASTER'S THESIS

### AUTOR PRÁCE

AUTHOR

Bc. Jan Konzal

### VEDOUCÍ PRÁCE

SUPERVISOR

RNDr. Lubor Příklad

BRNO 2020

# Diplomová práce

magisterský navazující studijní obor **Audio inženýrství**

Ústav telekomunikací

**Student:** Bc. Jan Konzal

**ID:** 164745

**Ročník:** 2

**Akademický rok:** 2019/20

## NÁZEV TÉMATU:

### **Analytický nástroj pro generování bicích triggerů z downmix záznamu**

#### POKYNY PRO VYPRACOVÁNÍ:

Navrhněte způsob využití statistické analýzy signálu metodou PCA pro klasifikaci jednotlivých částí bicí soupravy (basový buben, virbl, tom-tomy, hi-hat, crash, ride a další) a navrhněte takovou metodu, která umožní analyzovat smíchanou nahrávku bicí sady a vygeneruje sadu časových značek a příznaků tohoto záznamu. Následně vytvořte aplikaci v programovém prostředí Matlab nebo python jako sekvenci dílčích, laditelných modulů umožňující analýzu smíchané nahrávky bicí sady ve formátu wav a vygenerujte výstup této analýzy jako sekvenci časových značek a bicích příznaků, případně jako MIDI stopu nebo notový záznam.

#### DOPORUČENÁ LITERATURA:

- [1] Mayer, J., Acoustics and the Performance of Music, Springer; 2009, ISBN: 978-0-387-09516-5
- [2] Jan, J., Číslíková filtrace, analýza a resturace signálů, VUTIUM; 2002, ISBN: 80-214-1558-4

**Termín zadání:** 3.2.2020

**Termín odevzdání:** 1.6.2020

**Vedoucí práce:** RNDr. Lubor Přikryl

**Konzultant:** Ing. Jan Kaňka

**prof. Ing. Jiří Mišurec, CSc.**  
předseda oborové rady

#### UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## **Abstrakt**

Tato diplomová práce se zabývá návrhem a realizací nástroje pro generování časových triggerů bicí soupravy z downmix záznamu. V práci je popsáno předzpracování vstupního zvukového signálu a metody pro klasifikaci úderů. Rozeznávání úderů je založeno na podobnosti signálů ve frekvenční oblasti. Pro snížení počtu dimenzí a nalezení charakteristických vlastností vstupních dat byla využita analýza hlavních komponent (PCA). Pro klasifikaci dat do jednotlivých tříd představující části bicí soupravy byla využita metoda podpůrných vektorů (SVM). Program byl realizován v prostředí Matlab. Klasifikační model byl vytrénován na sadě 728 vzorků úderů pro sedm kategorií (velký buben, malý buben, hi-hat, crash, ride, velký buben + hi-hat, malý buben + hi-hat). Systém vykazuje úspěšnost rozeznání úderu 75 %.

## **Klíčová slova**

bicí souprava, analýza hlavních komponent, metoda podpůrných vektorů, klasifikace,

## **Abstract**

This thesis deals with the design and implementation of a tool for generating drums triggers from a downmix record. The work describes the preprocessing of the input audio signal and methods for the classification of strokes. The drum classification is based on the similarity of the signals in the frequency domain. Principal component analysis (PCA) was used to reduce the number of dimensions and to find the characteristic properties of the input data. The method support vector machine (SVM) was used to classify the data into individual classes representing parts of the drum kit. The software was programmed in Matlab. The classification model was trained on a set of 728 drum samples for seven categories (kick, snare, hi-hat, crash, ride, kick + hi-hat, snare + hi-hat). The success of the system in the classification is 75 %.

## **Keywords**

drum kit, principal component analysis, support vector machine, classification

## **Bibliografická citace:**

*KONZAL, Jan. Analytický nástroj pro generování bicích triggerů z downmix záznamu [online]. Brno, 2020 [cit. 2020-05-26]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/127317>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce Lubor Přikryl.*

## **Prohlášení**

*„Prohlašuji, že svou diplomovou práci na téma Analytický nástroj pro generování bicích triggerů z downmix záznamu jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.*

*Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.*

V Brně dne: **1. června 2020**

.....  
podpis autora

## **Poděkování**

*Děkuji vedoucímu této diplomové práce panu RNDr. Luborovi Přikrylovi za vedení práce a přínosné připomínky. Dále děkuji panu konzultantovi Ing. Janovi Kaňkovi za odborné vedení této práce a za cenné rady při častých konzultacích.*

V Brně dne: **1. června 2020**

.....  
podpis autora

# Obsah

1.	Úvod.....	13
2.	Bicí souprava.....	14
3.	Zpracování vstupního signálu .....	15
3.1	Předzpracování .....	15
3.2	Číslicové filtry.....	16
3.2.1	FIR, IIR filtry a jejich vlastnosti .....	18
3.2.2	Návrh IIR filtru .....	19
3.2.3	Banky filtrů .....	21
4.	Analýza hlavních komponent.....	23
4.1	Princip .....	23
4.1.1	Kovarianční matice .....	24
4.1.2	Matice korelačních koeficientů.....	24
4.2	Výběr počtu hlavních komponent .....	25
4.3	Výpočet analýzy hlavních komponent .....	26
4.4	Výstupy analýzy hlavních komponent .....	27
4.5	Příklad výpočtu .....	29
5.	Klasifikace dat.....	36
5.1	Výběr klasifikační metody .....	36
5.2	Metoda podpůrných vektorů .....	40
6.	Realizace .....	43
6.1	Segmentace plovoucím oknem .....	44
6.2	Detekce úderů.....	44
6.3	Banka filtrů.....	45
6.4	Výpočet energií ve frekvenčních pásmech .....	46
6.5	Analýza hlavních komponent.....	47
6.6	Metoda podpůrných vektorů .....	50
6.7	Výstupy a výsledky programu .....	51
7.	Závěr .....	55

# Seznam symbolů a zkratek

## Zkratky:

DAW	...	hudební software (Digital Audio Workstation)
FIR	...	filtr s konečnou impulsní odezvou (Finite Impulse Response)
IIR	...	filtry s nekonečnou impulsní odezvou (Infinite Impulse Response)
PC	...	hlavní komponenta (Principal Component)
PCA	...	analýza hlavních komponent (Principal Component Analysis)
RMS	...	efektivní hodnota ( <i>Root Mean Square</i> )
SVD	...	singulární rozklad matice (Singular Value Decomposition)
SVM	...	metoda podpůrných vektorů (Support Vector Machine)
SW	...	programové vybavení (software)

## Symbols:

$X_{RMS}$	...	efektivní hodnota signálu	[-]
$E$	...	energie signálu	[-]
$x[n]$	...	známá posloupnost	[-]
$y[n]$	...	řešení diferenciální rovnice	[-]
$X(z)$	...	obraz známé posloupnosti	[-]
$Y(z)$	...	obraz partikulárního řešení diferenciální rovnice	[-]
$H(z)$	...	přenosová funkce číslicového filtru	[-]
$b$	...	koeficient přenosové funkce číslicového filtru	[-]
$a$	...	koeficient přenosové funkce číslicového filtru	[-]
$Z$	...	Z transformace	
$h[n]$	...	impulsní odezva číslicového filtru	[-]
$h_n$	...	impulsní charakteristika číslicového filtru	[-]
$z$	...	komplexní rovina	[-]



$w$	...	komplexní rovina	[-]
$\omega_a$	...	kmitočty v analogovém systému	[-]
$\omega_d$	...	kmitočty v diskrétním systému	[-]
$f_c$	...	střední kmitočet	[Hz]
$f_{ref}$	...	referenční kmitočet	[Hz]
$B$	...	šířka pásma	[Hz]
$\varepsilon^2$	...	střední kvadratická odchylka	[-]
$\mathbf{X}$	...	matice vstupních dat PCA	
$\mathbf{X}_C$	...	centrovaná matice vstupních dat PCA	
$\mathbf{x}_k$	...	vektor vzorků úderů	[-]
$\lambda_i$	...	vlastní čísla	[-]
$\mu$	...	průměr	[-]
$\mathbf{C}(\mathbf{x})$	...	kovarianční matice	[-]
$\mathbf{v}_i$	...	vlastní vektor	[-]
$\mathbf{T}$	...	transformační matice	[-]
$\mathbf{Y}$	...	výstupní matice PCA	[-]
$h(\mathbf{x})$	...	hranice SVM	[-]
$\mathbf{w}$	...	vektor udávající orientaci SVM hranice	[-]
$\mathbf{w}_0$	...	vektor udávající polohu hranice SVM	[-]
$d$	...	vzdálenost objektu od hranice SVM	[-]
$J$	...	podmínka pro stanovení hranice	[-]
$y_k$	...	stavová proměnná příslušnosti objektu do třídy	[-]
$\xi_k$	...	toleranční pásmo SVM	[-]

## Seznam obrázků

Obr. 2.1: Bicí souprava [3] .....	14
Obr. 3.1: Obecný rekurzivní systém převzato z [4] .....	17
Obr. 3.2: Aproximace dolní propusti .....	18
Obr. 3.3: Modulová frekvenční charakteristika banky filtrů .....	22
Obr. 4.1: A – objekty v prostoru souřadnic X a Y, B – objekty v novém prostoru $X_2$ a $Y_2$ .....	23
Obr. 4.2: Indexový graf úpatí vlastních čísel .....	25
Obr. 4.3: Graf komponentních vah .....	28
Obr. 4.4: Graf komponentního skóre .....	29
Obr. 4.5: Centrované hodnoty pro Cenu a Citlivost (červené křížky zobrazují jednotlivé typy sluchátek v prostoru prvních dvou parametrů) .....	31
Obr. 4.6: Indexový graf úpatí vlastních čísel .....	32
Obr. 4.7: Graf komponentních vah .....	33
Obr. 4.8: Rozptylový diagram komponentního skóre .....	34
Obr. 4.9: Dvojný graf .....	35
Obr. 5.1: Trénovací nahrávky v prostoru prvních dvou hlavních komponent .....	37
Obr. 5.2: Klasifikace trénovacích nahrávek metodou k – průměrů .....	37
Obr. 5.3: Klasifikace trénovacích nahrávek metodou hierarchického shlukování ...	38
Obr. 5.4: Klasifikace trénovacích nahrávek metodou podpůrných vektorů .....	39
Obr. 5.5 Srovnání reálných tříd a tříd vytvořených metodou podpůrných vektorů ..	39
Obr. 5.6: Příklad množiny hranic u separabilních tříd SVM .....	40
Obr. 5.7: Lineárně neseparovatelné třídy .....	42
Obr. 6.1: Blokové schéma trénovací části programu .....	43
Obr. 6.2: Blokové schéma testovací části programu .....	44
Obr. 6.3: Ukázka funkce detektoru úderů .....	45
Obr. 6.4: Srovnání energií spočtených ve frekvenčních pásmech s amplitudovým spektrem malého bubnu .....	46
Obr. 6.5: Indexový graf úpatí vlastních čísel .....	47
Obr. 6.6: Graf komponentních vah .....	48
Obr. 6.7: Výřez grafu komponentních vah .....	48

Obr. 6.8: Graf komponentního skóre .....	49
Obr. 6.9: Výřez grafu komponentního skóre .....	49
Obr. 6.10: Hranice klasifikačních tříd .....	50
Obr. 6.11: Výřez hranic klasifikačních tříd .....	51
Obr. 6.10: Grafický výstup programu .....	52
Obr. 6.11: Spektra nahrávek úderu na malý a velký buben .....	53
Obr. 6.12: Spektra nahrávek úderu na crash, hi-hat (uzavřená) a ride.....	53
Obr. 6.13: Časový průběh nahrávky se špatnou detekcí úderů.....	54

## Seznam tabulek

Tab. 3.1: Vstupní parametry funkcí pro realizaci IIR filtrů.....	20
Tab. 4.1: Srovnání parametrů bezdrátových sluchátek.....	30
Tab. 4.2: Tabulka hodnot vlastních čísel .....	31
Tab. 4.3: Tabulka komponentních vah .....	32
Tab. 4.4: Tabulka komponentního skóre .....	33
Tab. 6.1: Srovnání úspěšnosti klasifikace s různými bankami filtrů .....	46
Tab. 6.2: Vlastní čísla [%] .....	47
Tab. 6.3: Tabulka záměn rozpoznávacího systému .....	52

# 1. ÚVOD

Tato diplomová práce se zabývá návrhem a realizací softwaru pro rozeznávání úderů na jednotlivé části bicí soupravy a generováním příznaků z downmix záznamu. Hlavním cílem při návrhu a realizaci takového analytického nástroje bylo ověření předpokladů pro vývoj komplexního rozpoznávacího softwaru, na kterém pracuje firma Audified.

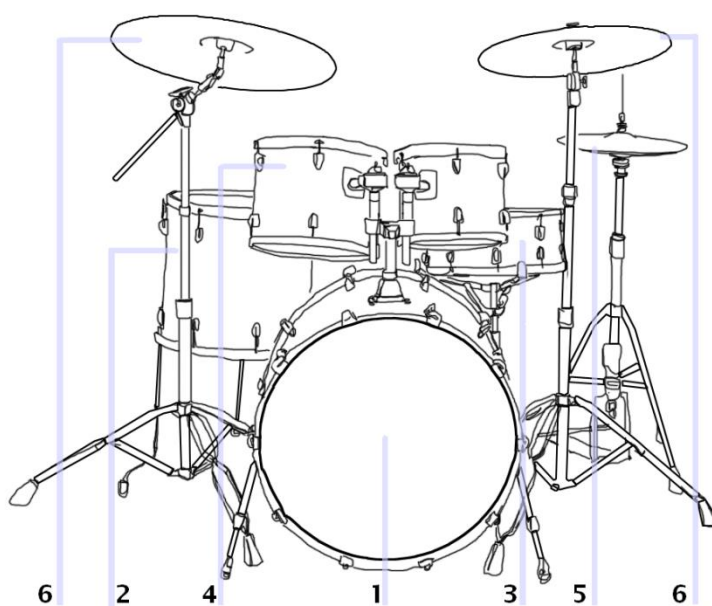
V řadě nahrávek jsou zvuky klasické bicí soupravy nahrazeny různými elektronickými zvuky. Takováto nahrávka může vzniknout přímo v počítači, nebo jsou nahrány klasické bicí a v DAW nahrazeny zvuky elektronickými. Této metodě se říká triggerování nebo také samplování. Při triggerování je zapotřebí mít každou část bicí soupravy ve zvláštní zvukové stopě a dbát na co nejmenší přeslechy mezi mikrofony při nahrávání. Komplexní software pro rozeznávání bicích a generování příznaků z downmixu záznamu, může zásadně zjednodušit práci triggerování bicí soupravy. Dále by mohly být triggerovány bicí v nahrávkách, které nejsou k dispozici v jednotlivých stopách, ale pouze již smíchané v downmixu.

Tato práce je rozdělena do dvou základních částí. První je část teoretická, kde jsou popsány jednotlivé metody zpracování signálů, které byly v softwaru využity. Tato část je rozdělena do tří kapitol. První se zabývá předzpracováním signálu. Další kapitola popisuje princip a využití analýzy hlavních komponent. V poslední kapitole teoretické části je popsán výběr klasifikační metody, a popis vybrané metody podpůrných vektorů.

Druhá část této práce se zabývá realizací softwaru. Jsou zde uvedeny bloková schémata navrženého kódu. U jednotlivých bloků jsou popsány jejich vstupy, výstupy a jejich funkce. Na závěr jsou popsány testy funkčnosti klasifikace a prezentovány výsledky těchto testů.

## 2. BICÍ SOUPRAVA

Bicí souprava je sestava bubnů a činelů, její velikost a složení nejsou dány. Záleží na preferencích hráče a také hudebním žánru. Na obrázku (obr. 2.1) je vidět typická základní bicí souprava. Číslem jedna je označen velký buben (Bass Drum), jako číslo dvě je označen kotel (Floor Tom), číslo tři je malý buben (Snare), číslo čtyři jsou tomy, nebo také přechody (Tom Tom). Zbývají činely, které nemají pro svoje názvy české ekvivalenty, proto se používají názvy anglické, vpravo pod číslem pět se nachází Hi – Hat česky hovorově „hajtká“. Zbývající dva činely jsou vpravo Crash a vlevo Ride. Souprava může být menší nebo i mnohem větší. Záleží na vkusu hráče a hudebním žánru. [1, 2]



Obr. 2.1: Bicí souprava [3]

Počtem bubnů a činelů však rozmanitost zvuků bicí soupravy nekončí. Malý buben může být při hře „zapnut“ nebo „vypnut“. Tím se označuje, zdali jsou nataženy struny na spodní bláně malého bubnu nebo ne. Na malý buben se dá hrát takzvaně „přes ráfek“. To je technika, kdy hráč paličkou neudeří do blány bubnu, ale na jeho okraj. Hi-hat může být otevřená nebo uzavřená, pomocí pedálu. U činelů obecně se může hrát na jejich kraj, nebo střed. [1, 2]

Samostatnou kategorii zvuků tvoří takzvané dvojúdery. Jedná se o úder na buben a činel ve stejnou chvíli. Základními dvojúdery jsou malý buben s hi-hat, velký buben s hi-hat. Hi-hat může být nahrazena například crashem.

Rozmanitost zvuků při hře na bicí soupravu může být velká. Tato práce se zabývá pouze základním rozsahem bicí soupravy a jejími základními zvuky.

## 3. ZPRACOVÁNÍ VSTUPNÍHO SIGNÁLU

### 3.1 Předzpracování

U některých hudebních nahrávek je bicí souprava nahrazena elektronickými zvuky neboli samplý. Jednou z metod pro nahrazení bicích je triggerování nebo také samplování. Nejprve je pomocí mikrofonů nahrána klasická bicí souprava do zvláštních stop, jednotlivé úderý jsou poté v DAW nahrazeny zvuky elektronickými.

Vstupním signálem SW pro generování triggerů je digitální audio nahrávka bicí soupravy. Jedná se tedy o diskretní signál  $f_n = f(t_n) = f(nT)$ , kde T je perioda vzorkování. Omezením takového signálu je, že známe pouze jeho vzorky v určitých časových úsecích. Podle Nyquistova teorému, mohou být zpracovány pouze signály s frekvencí nižší, než je polovina frekvence vzorkovací. Dalším omezením je kvantování, kde je vzorku přiřazena nejbližší hodnota z dané číselné reprezentace, tedy přesnost je dána počtem kvantovacích úrovní. [4]

Nejdříve jsou ze vstupního signálu získávány časové značky začátků a konců jednotlivých úderů. Získávání časových značek je realizováno pomocí časové obálky signálu, kterou počítá funkce prostředí Matlab *envelope*. Výpočet obálky signálu je založen na Hilbertově transformaci. Rozbor Hilbertovy transformace není součástí této práce, a je popsána například v [4]. Při průchodu obálky rozhodovací úrovní je zaznamenána časová značka. Tato metoda vyžaduje stejnou hlasitost vstupních nahrávek. Proto, bylo implementováno normalizování vstupní hlasitosti podle efektivní hodnoty (RMS) (3.1). Nahrávka je normalizována jako celek na hodnotu 0,4 (RMS), která byla určena experimentálně. Matlab v základním nastavení ukládá zvukový soubor ve formátu double v rozsahu hodnot -1 až 1. [9]

$$X_{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x^2(n)} \quad (3.1)$$

Kde N je počet vzorků a  $x(n)$  je vstupním signálem. [5] Dále je signál rozdělen na časové úseky, ve kterých probíhá rozpoznávání úderu. Zde se nabízí dva možné způsoby, a to segmentace plovoucím oknem (kontinuální analýza), nebo dělení na vzorky celých úderů (analýza segmentů).

V dalším kroku předzpracování je provedena frekvenční analýza pomocí banky filtrů. Pro každé frekvenční pásmo je spočítána energie signálu. Tyto energie jsou vstupem pro analýzu hlavních komponent.

Výpočet energie je definován jako součet čtverců absolutních hodnot signálu (3.2). [8]

$$E = \sum_{n=-\infty}^{\infty} |s[n]|^2 \quad (3.2)$$

Kde  $s[n]$  je výstupní signál banky filtrů v daném pásmu.

## 3.2 Číslicové filtry

Číslicové filtry jsou obdobou filtrů analogových pro digitalizované signály, jako transformační funkce, které ze vstupní posloupnosti  $x[n]$  vytvoří požadovanou výstupní posloupnost  $y[n]$ . I přes určitou paralelu k analogovým filtrům mají odlišné vlastnosti. Například:

- strmosti frekvenčních charakteristik, kterých nešlo pomocí analogových obvodů dosáhnout
- vlastnosti filtrů se mění pomocí jejich koeficientů, nemají na ně vliv okolní podmínky, jako například teplota nebo stárnutí kapacitorů
- vlastnosti digitálního filtru lze kdykoli upravit, není třeba změna zapojení
- digitální filtry mohou zpracovávat i nízké kmitočty, na rozdíl od analogových, kde je jejich realizace pro nízké kmitočty náročná
- analogové filtry dosahují běžně útlumu 60 až 70 dB, u číslicových filtrů lze charakteristiku nastavit v širokém rozsahu
- hlavní nevýhodou číslicových filtrů může být delší čas pro zpracování, tedy značné zpoždění
- omezení frekvence zpracovávaného signálu číslicovými filtry do poloviny vzorkovací frekvence
- zařazení analogově-číslcových a číslicově-analogových převodníků

[4, 7]

Funkce číslicových filtrů popisuje diferenční rovnice, která vyjadřuje závislost mezi posloupnostmi a jejich diferencemi. Diferenční rovnice lze řešit pomocí diferenčního počtu nebo jednodušeji pomocí  $Z$  transformace. Transformací  $Z$  (3.3)  $x[n]$  známé posloupnosti a  $y[n]$  hledaného řešení diferenční rovnice, dostaneme obraz  $X(z)$  a obraz partikulárního řešení  $Y(z)$ .

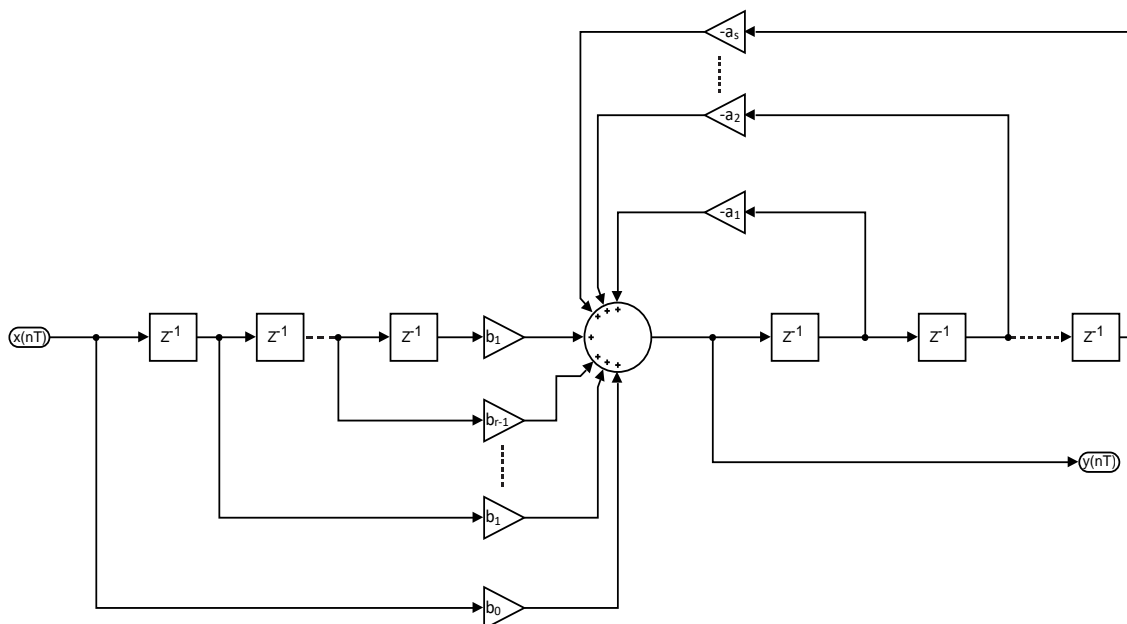
$$Y(z) = \sum_{n=0}^{+\infty} y[n]z^{-n}, \quad X(z) = \sum_{n=0}^{+\infty} x[n]z^{-n} \quad (3.3)$$

Jejich podílem vzniká přenosová funkce číslicového filtru vztah (3.4).

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1z + b_2z^2 + \dots + b_rz^r}{a_0 + a_1z + a_2z^2 + \dots + a_sz^r} \quad (3.4)$$

Kde  $a$  a  $b$  jsou koeficienty přenosové funkce. [8] Na obrázku (obr. 3.1) je znázorněn obecný rekurzivní systém, který je popsán přenosovou funkcí (3.4).





Obr. 3.1: Obecný rekurzivní systém převzato z [4]

Přenosová funkce filtru může být také vyjádřena pomocí nulových bodů a pólů vztah (3.5). [4]

$$H(z) = K \frac{z^{r-s} \prod_{i=0}^r (z - z_{0i})}{\prod_{j=0}^s (z - z_{xj})} \quad (3.5)$$

Nulové body  $z_0$  získáme porovnáním jmenovatele přenosové funkce (3.4) s nulou, tedy řešením rovnice  $Y(z) = 0$ . Póly  $z_x$  získáme, položíme-li jmenovatele přenosové funkce rovno nule, tedy řešením rovnice  $X(z) = 0$ .  $K$  je zesilovací činitel systému. [8]

Impulsní charakteristika je odezva systému na jednotkový impuls, definovaný jako  $\{x_n\} = 1, 0, 0 \dots$ . Obrazem jednotkového impulsu pro jmenovatele přenosové funkce  $X(z) = 1$ , dostáváme rovnost  $H(z) = Y(z)$ . V tuto chvíli je  $H(z)$  obrazem impulsní charakteristiky filtru. Zpětnou  $Z$  transformací (3.6) získáme impulsní odezvu  $h[n]$  (3.7).

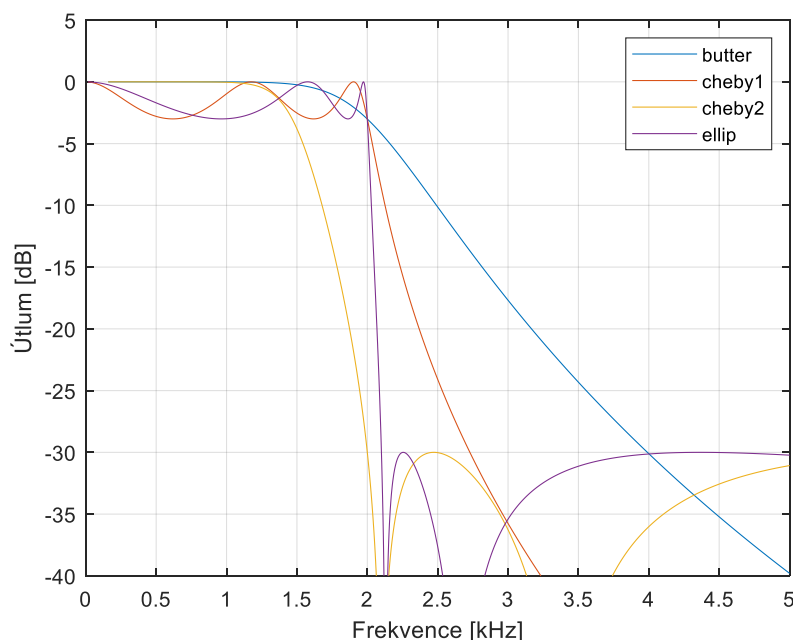
$$x[n] = \frac{1}{2\pi j} \oint_C X(z) z^{k-1} dz \quad (3.6)$$

Kde  $C$  je křivka ležící v oblasti konvergence a obklopující počátek. [4][8]

$$h[n] = Z^{-1}\{H(z)\} \quad (3.7)$$

Filtr, u kterého platí, že jeho impulsní odezva skončí po uplynutí nějaké doby, nazýváme filtr s konečnou impulsní odezvou (FIR, Finite Impulse Response). Pokud je impulsní odezva na výstupu časově neohrazená (nekonečná), filtr se nazývá filtr s nekonečnou impulsní odezvou (IIR, Infinite Impulse Response). [4]

Návrh filtru vychází z ideální analogové dolní propusti, která má v propustném pásmu modul roven jedné a v nepropustném pásmu se hodnota modulu rovná nule. Strmost přechodového pásma ideálního filtru je nekonečná. Aby se takovýto ideální průběh přiblížil skutečnému analogovému filtru, je třeba jeho průběh aproximovat. Tím dojde k rozšíření přechodového pásma a zvlnění modulové frekvenční charakteristiky. Zde budou uvedeny čtyři tyto aproximační funkce. Vybírá se z nich na základě požadavku na zvlnění frekvenční modulové charakteristiky. Butterworthova aproximace obsahuje pouze póly, modulem filtru je monotónně klesající funkce. Kmitočtová charakteristika je hladká, nemá žádné zvlnění, za cenu nejmenší strmosti přechodového pásma. Čebyševova aproximace 1. typu má také pouze póly, dochází ke zvlnění propustného pásma modulové frekvenční charakteristiky. Čebyševova aproximace 2. typu obsahuje nulové body i póly, modulová frekvenční charakteristika se zvlní v nepropustném pásmu. Eliptická (Cauerova) aproximace má nulové body i póly, vyznačuje se největší strmostí přechodového pásma. Ke zvlnění modulové frekvenční charakteristiky dojde jak v propustném, tak v nepropustném pásmu. Na obrázku (obr. 3.2) jsou vidět jednotlivé aproximační funkce pro filtr typu dolní propust. [7, 8]



Obr. 3.2: Aproximace dolní propusti

### 3.2.1 FIR, IIR filtry a jejich vlastnosti

Číslicové filtry se dělí do dvou základních skupin, a to s konečnou impulsní odezvou (FIR) a s nekonečnou impulsní odezvou (IIR). Reprezentace v  $z$ -oblasti je pro FIR dána jen nulovými body, frekvenční charakteristika je tedy periodická funkce s periodou  $2\pi/T$ . Výhodou FIR je lineární fázová charakteristika, je-li dodržena podmínka symetrie nebo antisymetrie impulsní charakteristiky (3.8). [4]

$$h_n = h_{(N-1-n)}, \text{ resp. } h_n = -h_{(N-1-n)} \quad (3.8)$$

Díky ní nemají fázové zkreslení v celém kmitočtovém pásmu. Jsou nerekurzivní, není tedy třeba při výpočtu vyšetřovat jejich stabilitu. Mají obecně menší citlivost na kvantizační šum. Pro dosažení shodných strmostí s IIR filtry mají FIR filtry větší počet koeficientů ( $N$ ), proto mají větší nároky na paměť a způsobují větší zpoždění signálu. Vyšší řád sebou nese vyšší počet koeficientů, kde musí být uloženo více stavových veličin. [4, 8]

Při realizaci IIR systému je nutné použít struktury se zpětnými vazbami. Stabilita systému je dána polohou pólů. Aby byl systém stabilní musí póly ležet uvnitř jednotkové kružnice v rovině  $z$ . IIR filtry jsou citlivé na nepřesnost vzniklou číslicovou realizací, proto se stabilita na konkrétní realizaci ověřuje. [4] Pro IIR filtry lze najít ekvivalentní analogový filtr. Tyto filtry nemají lineární fázovou charakteristiku, proto způsobují fázové zkreslení. Filtry s náročnými frekvenčními charakteristikami realizovanými pomocí IIR, mají řádově nižší řád filtru než při realizaci FIR. To přináší nižší řád přenosové funkce, díky kterému se snižuje potřebné zpoždění signálu. [7] Aby mohl být filtr používán v reálném čase, musí být dosaženo jeho kauzality. Podmínka kauzality je dána vztahem (3.9). [7]

$$h[n] = 0 \text{ pro } n < 0 \quad (3.9)$$

Z uvedených vlastností vyplývá, že FIR filtry se využívají v aplikacích, kde záleží na fázovém zkreslení, není třeba vysokých strmostí pro přechod do nepropustného pásma frekvenční charakteristiky a nezáleží příliš na zpoždění a náročnosti výpočtu. Naopak v aplikacích, které vyžadují vysokou strmost, krátký procesní čas, je vhodné použít filtr typu IIR. IIR filtry nemohou být navrženy s lineární fázovou charakteristikou v celém frekvenčním rozsahu, avšak lze se v úzkém kmitočtovém pásmu lineární fázové charakteristice přiblížit. Proto se používají pro realizaci filtrů s po částech konstantní frekvenční charakteristikou (horní propust, dolní propust, pásmová zadrž, nebo pásmová propust). Což může být pro zpracování zvukových signálů dostačující. [8] Na základě těchto skutečností byla pro frekvenční dělení vstupního signálu vybrána realizace IIR filtrů.

### 3.2.2 Návrh IIR filtru

IIR filtry jsou podobné analogovým filtrům, a jejich návrh lze provést pomocí transformace analogových prototypů do číslicové oblasti. Při transformaci dochází k převodu roviny  $w$  do roviny  $z$ . Použitá transformace musí splňovat tyto požadavky:

- Celá rovina  $w$  se musí transformovat na celou rovinu  $z$  jednojednoznačně.
- Póly přenosové funkce  $z$  levé strany v rovině  $w$  musí být transformovány dovnitř jednotkové kružnice v rovině  $z$ . Tedy stabilní analogový filtr musí být transformován na stabilní číslicový filtr.
- Základní kmitočtové vlastnosti musí být po transformaci zachovány.
- Transformace musí jednojednoznačně transformovat racionální lomenou funkci  $H_a(w)$  na racionální lomenou funkci  $H(z)$ .

Nejčastěji bývá využita bilineární transformace (3.10). [8]

$$w = \frac{z - 1}{z + 1} \quad (3.10)$$

A zpětná bilineární transformace podle vztahu (3.11). [4, 7, 8]

$$z = \frac{1 + w}{1 - w} \quad (3.11)$$

Bilineární transformace zobrazí imaginární osu v komplexní rovině  $w$  do jednotkové kružnice v komplexní rovině  $z$ . Tím je splněna první a druhá podmínka. Bilineární transformace je racionální lomenou funkcí a splňuje čtvrtou podmínku. Splnění třetí podmínky dosáhneme vyšetření zobrazení kmitočtové osy (3.12). [4, 8]

$$z = \frac{1 + j\omega_a}{1 - j\omega_a} = 1 \cdot e^{j2\alpha}, \text{ kde } \alpha = \arctan\left(\frac{\omega_a}{1}\right) \quad (3.12)$$

Kde  $\omega_a$  jsou kmitočty v diskretním systému a  $\omega_a$  jsou kmitočty v analogovém systému. Při transformaci dochází ke zkreslení frekvenční osy, a to z důvodu převodu nekonečné analogové frekvenční osy na konečnou diskretní osu. Vztah mezi kmitočty v analogovém systému  $\omega_a$  a kmitočty v číslicovém systému je dán vztahem (3.13). [4, 8]

$$\omega_d T = 2 \arctan(\omega_a) \quad (3.13)$$

V Matlabu se IIR filtry realizují za pomoci funkcí, které jsou rozdělené podle charakteru filtru. Jsou to funkce butter, cheby1, cheby2 a ellip. Každá z těchto funkcí vrací přímo koeficienty přenosové funkce nebo nulové body, póly a zesílení. Pro filtry vyššího řádu než čtvrtého, je v dokumentaci pro Matlab, [9] doporučeno používat jako výstup rozložení nulových bodů a pólů, nikoli přímo koeficienty filtru. Vzhledem k zaokrouhlení by mohlo dojít k nestabilitě systému. V tabulce (tab. 3.1) jsou uvedeny vstupní parametry pro jednotlivé funkce.

Tab. 3.1: Vstupní parametry funkcí pro realizaci IIR filtrů

vstupní parametry	názvy funkcí			
	butter	cheby1	cheby2	ellip
řád filtru	✓	✓	✓	✓
mezní frekvence	✓	✓	✓	✓
tip filtru	✓	✓	✓	✓
velikost zvlnění [dB]		✓		✓
útlum v nepropustném pásmu [dB]			✓	✓

Pokud je mezní frekvence definována dvouprvkovým vektorem funkce, vytvoří pásmovou propust nebo pásmovou zádrž, kde první prvek vektoru je dolní mezní frekvence, a druhým prvkem je horní mezní frekvence. Mezní kmitočty může nabývat hodnot 0 až 1, kde hodnota 1 představuje polovinu vzorkovacího kmitočtu. [5]

### 3.2.3 Banky filtrů

Pro rozdělení signálu na jednotlivé frekvenční pásma bude použita banka filtrů. Jako nejvhodnější byla vybrána třetino – oktávová banka filtrů. Signál je tedy rozdělen do 30 frekvenčních pásem. Třetino – oktávová banka byla vybrána na základě testu několika bank, který je popsán v kapitole 6.3.

Podle Harvey Fletchera je-li tón maskován bílým šumem, na maskování se podílí jen určité pásmo spektra, které leží v okolí maskovaného tónu. Rozsah tohoto pásma se nazývá šířka kritického pásma. V lidském sluchovém ústrojí šířka kritického pásma odpovídá konstantní vzdálenosti na bazilární membráně. Eberhard Zwicker na základě svých psychoakustických měření stanovil jednotlivým pásmům jejich střední kmitočty a šířky pásma. Pro aproximaci šířky kritického pásma podle Eberhard Zwicker se používá banka třetino-oktávových filtrů. [1, 10, 11]

Střední kmitočty oktávových a zlomko-oktávových filtrů se počítají dle vztahu (3.14) pro lichý počet pásem na oktávu a podle vztahu (3.15) pro sudý počet pásem na oktávu. [10]

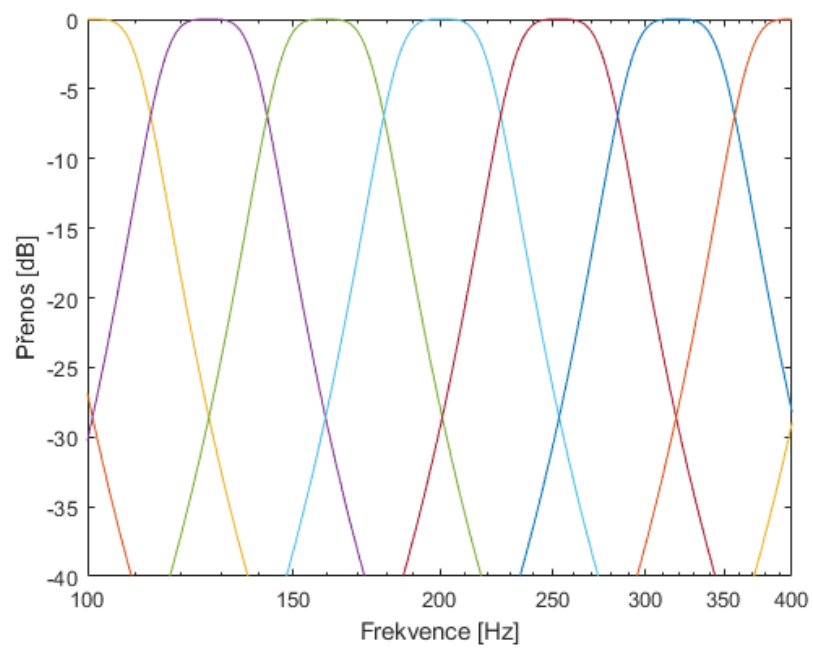
$$f_c = G^{\frac{x}{n}} f_{ref} \quad (3.14)$$

$$f_c = G^{\frac{2x+1}{2n}} f_{ref} \quad (3.15)$$

Šířka pásma se spočítá dle vztahu (3.16).

$$B_{i+1} = B_i = B_0 2^{\frac{x}{n}} \quad (3.16)$$

Kde  $G = 10^{3/10}$ ,  $f_{ref} = 1$  kHz,  $n$  je počet pásem na oktávu a  $x$  je pořadí filtru. [10] Jako počáteční realizace banky filtrů, byla zvolena banka třetino-oktávových filtrů. Na obrázku (obr. 3.3) je znázorněna modulová frekvenční charakteristika banky třetino-oktávových filtrů realizovaná v Matlabu. Pro názornost je zobrazena v rozsahu 100 až 400 Hz.



Obr. 3.3: Modulová frekvenční charakteristika banky filtrů

## 4. ANALÝZA HLAVNÍCH KOMPONENT

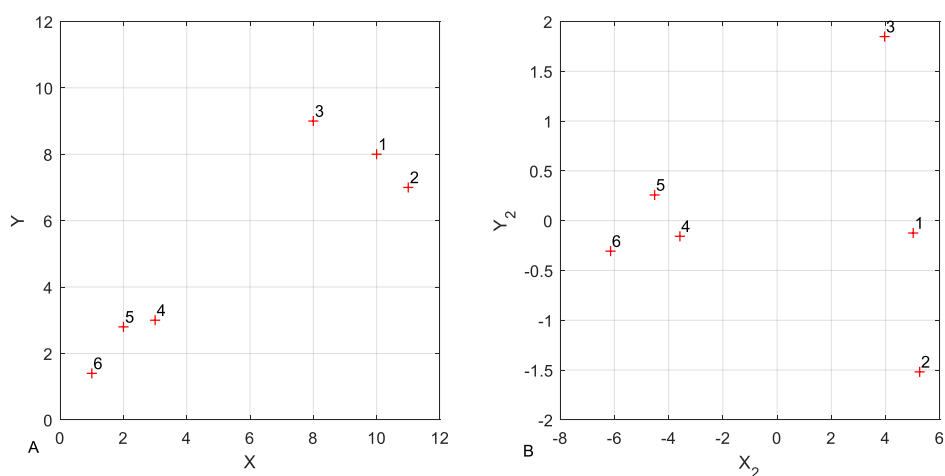
Signál, který prošel bankou filtrů má tolik dimenzí, do kolika je rozdělen pásem. Pro snížení náročnosti následných výpočtů a tím i procesního zpoždění, byla hledána metoda pro snížení počtu dimenzí signálu. Tedy k odstranění redundance. Analýza hlavních komponent (PCA, principal component analysis) je jednou z nejpoužívanějších metod pro analýzu vícerozměrných dat a redukci jejich dimenzionality. Metoda může být popsána jako lineární transformace vstupních proměnných na nové nekorelované, které nazýváme hlavními komponentami. Z geometrického hlediska dochází k hledání nového prostoru, ve kterém lze zobrazit rozložení vstupních dat, pomocí menšího počtu os označovaných jako hlavní komponenty. Redukce dimenzí pomocí PCA také umožňuje hledat významné charakteristiky v datech, čímž je umožněn jejich kvantitativní popis. PCA byla vyvinuta pro data s mnohazměrným normálním rozdělením. Není vhodná pro analýzu vícestavových kvalitativních dat, na které nelze aplikovat euklidovskou metriku. PCA také není vhodná pro data, která obsahují velké množství nul. Je vhodné, aby počet proměnných nepřevyšoval počet prvků. Obecně je doporučeno, aby se počet prvků blížil druhé mocnině počtu proměnných. Často bývá metoda hlavních komponent součástí komplexnější analýzy dat. [12, 13, 14]

### 4.1 Princip

Cílem metody hlavních komponent je zobrazení vstupních, obecně vícerozměrných dat, do nového prostoru s nižší dimenzí. Tento nový prostor je reprezentován sadou latentních výstupních proměnných. Výstupní proměnné označované jako hlavní komponenty jsou vzájemně nekorelované, popisují téměř v neredukované míře přesné rozložení původních znaků a poskytují vhodnější vlastnosti pro další analýzu. [12]

Metoda hlavních komponent transformuje vstupní, obecně vícerozměrná data do nového prostoru s nižší dimenzí.

Pro jednoduchost si představme objekty rozmístěny ve dvourozměrném prostoru popsány souřadnicemi  $X$  a  $Y$  (obr. 4.1 A).



Obr. 4.1: A – objekty v prostoru souřadnic  $X$  a  $Y$ , B – objekty v novém prostoru  $X_2$  a  $Y_2$

Pootáčením souřadného systému je nalezen nový prostor, který lépe popisuje rozložení objektů. Takovýto nový prostor má nové souřadnice  $X_2$  a  $Y_2$ , ty jsou dány lineární kombinací původních souřadnic  $X$ ,  $Y$  (obr. 4.1 B).

Jsou-li objekty rozloženy na přímce, po nalezení nového prostoru dojde k popisu jejich rozložení pomocí pouze jedné souřadnice. Toto je hlavní princip redukce dimenzí prostoru. K popisu rozložení objektů pomocí jedné nové souřadnice zpravidla nedochází. Proto je hledán nový prostor, kde je rozložení objektů co nejpřesněji popsáno, v co nejmenším počtu souřadnic tak, aby souřadnice, které mají malý vliv na rozložení objektů, mohly být zanedbány. Matematický popis tohoto principu je dán jako odchylka všech bodů od redukované reprezentace, popsána minimální střední kvadratickou odchylkou (4.1).

$$\varepsilon_k^2 = \|\mathbf{x}_k - \mathbf{x}'_k\|^2 \quad (4.1)$$

Kde  $\mathbf{x}_k$  je původní reprezentace a  $\mathbf{x}'_k$  redukovaná reprezentace. Analýza dat pomocí metody hlavních komponent se počítá nejčastěji dvěma způsoby, a to kovarianční maticí nebo maticí korelačních koeficientů. [12, 13, 15, 16]

### 4.1.1 Kovarianční matice

Analýza hlavních komponent s kovarianční maticí se někdy nazývá též centrovaná PCA (4.4), (4.5). Počátek nového souřadného systému je posunut do centroidu objektů. Vzdálenosti mezi objekty se s převodem do nového souřadného systému nemění. Součet vlastních hodnot kovarianční matice je roven součtu rozptylů proměnných (4.2).

$$\varepsilon^2 = \frac{1}{K} \sum_{k=1}^K \|\mathbf{x}_k\|^2 = \sum_{i=1}^n \lambda_i \quad (4.2)$$

Kde  $\lambda_i$  jsou vlastní čísla. Tato metoda výpočtu PCA se hodí pro data, která jsou ve stejných jednotkách a ve stejné škále. Nebo pro data v rozdílných jednotkách, které lze přepočítat na stejné číselné reprezentace. [15]

### 4.1.2 Matice korelačních koeficientů

Vstupní data, jsou normalizována na jednotkový rozptyl a nulový průměr. Součet hodnot matice korelačních koeficientů je roven počtu proměnných. Počet proměnných určuje řád korelační matice. Počátek nového souřadného systému je opět posunut do centroidu objektů. Po normalizaci původních proměnných na jednotkový rozptyl, vzdálenosti mezi objekty již nejsou závislé na jednotkách, v kterých byly naměřeny. Normalizací na jednotkový rozptyl získává množina objektů nový kompaktnější zakulacený tvar. Snižují se rozdíly mezi vlivy hlavních komponent, což může vést k výpočetním chybám. Tato metoda se hodí pro data vyjádřená ve zcela rozdílných jednotkách, a nelze je vyjádřit ve stejné číselné škále. [15, 16]



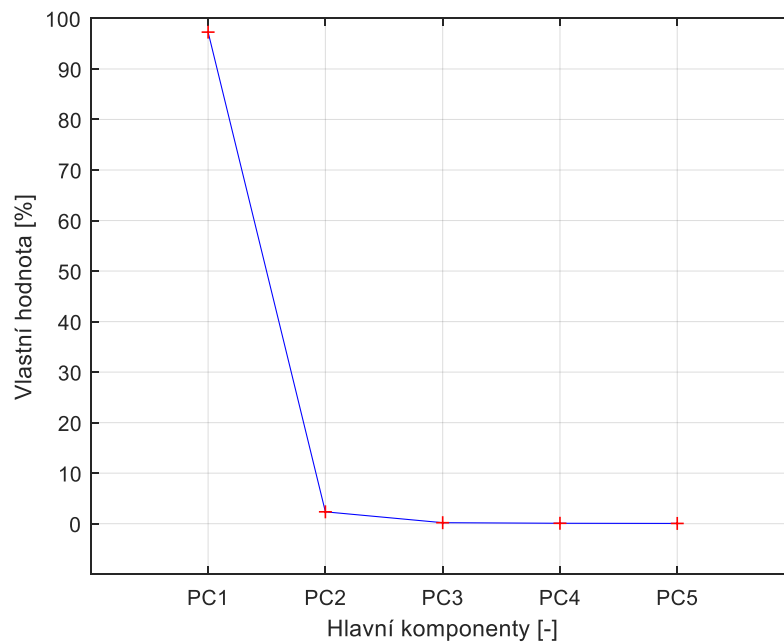
## 4.2 Výběr počtu hlavních komponent

Po výpočtu PCA je vytvořen nový prostor daný hlavními komponentami. Hlavní komponenty jsou seřazeny podle velikosti části rozptylu původních dat, v nich obsaženém. Nyní je třeba z množiny hlavních komponent vybrat několik s nejnižšími pořadovými čísly tak, aby mohlo být dostatečně popsáno rozložení dat. Obecně je dobré brát v úvahu jen komponenty, které mají vlastní hodnoty vyšší, než je průměr všech vlastních hodnot. Vlastní hodnota  $\lambda_i$  udává rozptyl v dané hlavní komponentě. [15]

Pokud je využita metoda s korelační maticí, tak se nejčastěji ke stanovení počtu hlavních komponent používá Kaiserovo kritérium. [15] Toto kritérium předpokládá užití hlavních komponent s vlastní hodnotou (4.8) větší než 1. Popis výpočtu vlastní hodnoty je uveden v kapitole 4.3. Součet vlastních hodnot je roven počtu vstupních proměnných. Nemá tedy smysl uvažovat komponenty, které mají vlastní hodnotu menší než 1. [15]

Je-li použita kovarianční matice, součet vlastních hodnot se rovná součtu rozptylu vstupních proměnných. U této metody můžeme využít hlavní komponenty s vlastní hodnotou vyšší, než je průměr všech vlastních hodnot. [15]

Obecnou možností výběru počtu hlavních komponent může být grafické zobrazení jejich vlastních hodnot. Toto zobrazení se nazývá indexový graf úpatí vlastních čísel (Scree Plot) (obr. 4.2).



Obr. 4.2: Indexový graf úpatí vlastních čísel

Jedná se o graf závislosti vlastních hodnot na hlavních komponentách. Počet hlavních komponent z grafu odečteme tak, že sledujeme pokles vlastních hodnot, když se tento pokles zmírní a křivka se „ohne“, odečteme pořadí příslušné poslední komponenty. V příkladu na (Obr. 4.2) by byly použity první tři hlavní komponenty. [12, 14, 15]

Další metodou může být výběr komponent, které dohromady vyjadřují 90 až 99 % celkového rozptylu. [14]

### 4.3 Výpočet analýzy hlavních komponent

Nejpoužívanějším algoritmem pro výpočet hlavních komponent je singulární rozklad matice (SVD – Singular Value Decomposition). Tento algoritmus je založen na Karhunenově – Loèevově transformaci, a zaručuje globální optimum. [13] Dále je uveden postup výpočtu analýzy hlavních komponent pomocí algoritmu SVD.

Nejdříve je potřeba vstupní data uspořádat do matice  $\mathbf{X}$  (4.3),

$$\mathbf{X}_{i,j} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,j} \\ x_{2,1} & x_{2,2} & \dots & x_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i,1} & x_{i,2} & \dots & x_{i,j} \end{bmatrix} \quad (4.3)$$

Kde řádky představují jednotlivé vzorky úderů a sloupce frekvenční pásma. [17], [18]

V dalším bodě probíhá centrování dat. Nejprve se z vektorů  $\mathbf{x}_k = \mathbf{x}_{k,1} \dots \mathbf{x}_{k,j}$  vypočte pomocí vztahu (4.4) průměr. [16]

$$\mu = \frac{1}{K} \sum_{k=1}^K \mathbf{x}_k \quad (4.4)$$

Centrovaný vektor z matice  $\mathbf{X}$  se vypočítá pomocí vztahu (4.5), z těchto vektorů je složena centrovaná matice  $\mathbf{X}_c$ . [16]

$$\tilde{\mathbf{x}}_K = \mathbf{x}_K - \mu \quad (4.5)$$

Následně se vypočítá kovarianční matice  $\mathbf{C}$ . Obsahem této matice jsou prvky s indexy  $i, j$  což jsou kovariance  $i$ -té a  $j$ -té složky původních dat  $\mathbf{X}$ . Kovariance se spočítá na základě vztahu (4.6), jednotlivé kovariance jsou uspořádány do kovarianční matice  $\mathbf{C}$ . [12, 14, 16]

$$\mathbf{C}(\mathbf{x}) = \frac{1}{K} \sum_{k=1}^K \tilde{\mathbf{x}}_k \cdot \tilde{\mathbf{x}}_k^T \quad (4.6)$$

Nyní se vypočítají vlastní čísla (eigenvalues) a vlastní vektory (eigenvectors) kovarianční matice  $\mathbf{C}$ . Vlastní vektory  $\mathbf{v}_i$  jsou získány řešením rovnice (4.7).

$$(\mathbf{C} - \lambda_i \mathbf{I})\mathbf{v}_i = 0 \quad (4.7)$$

Kde  $(\mathbf{C} - \lambda_i \mathbf{I})$  je charakteristická rovnice pro výpočet vlastních hodnot  $\lambda_i$ . Vlastní hodnoty jsou získány řešením rovnice (4.8). [14, 16, 17, 18]

$$\det|\mathbf{C} - \lambda_i \mathbf{I}| = 0 \quad (4.8)$$

Kde  $\det|\mathbf{C} - \lambda_i \mathbf{I}|$  je determinanem charakteristické rovnice. Vlastní hodnoty jsou rozptylem odpovídající příslušným hlavním komponentám. [14, 16]

Vlastní čísla  $\lambda_i$  jsou spolu s příslušnými vlastními vektory  $\mathbf{v}_i$  seřazeny podle velikosti od největších k nejmenším. Z vlastních vektorů se složí transformační matice  $\mathbf{T}$ . Jednotlivé vlastní vektory  $\mathbf{v}_i$  jsou v matici  $\mathbf{T}$  řazeny sestupně podle vlastních čísel  $\lambda_i$  ve sloupcích. Prvky vlastních vektorů jsou váhy původních proměnných. Tyto váhy udávají pozici objektů v novém souřadném systému.

[12, 14, 16]

Nyní se vybírá počet hlavních komponent (vlastních vektorů), podle některého z kritérií uvedených v kapitole 4.2. Počet vybraných komponent  $n$  se vloží do nové transformační matice  $\mathbf{T}_n$ . Podle vybraných hlavních komponent jsou původní data promítnuta do nového prostoru. [17, 18]

Na závěr je vypočtena matice  $\mathbf{Y}$  podle vztahu (4.9), která obsahuje původní data promítnutá do nového prostoru s  $n$  dimenzemi. [14, 16, 17, 18]

$$\mathbf{Y} = \mathbf{X}_c \cdot \mathbf{T}_n \quad (4.9)$$

Kde  $\mathbf{X}_c$  je centrovaná matice vstupních dat.

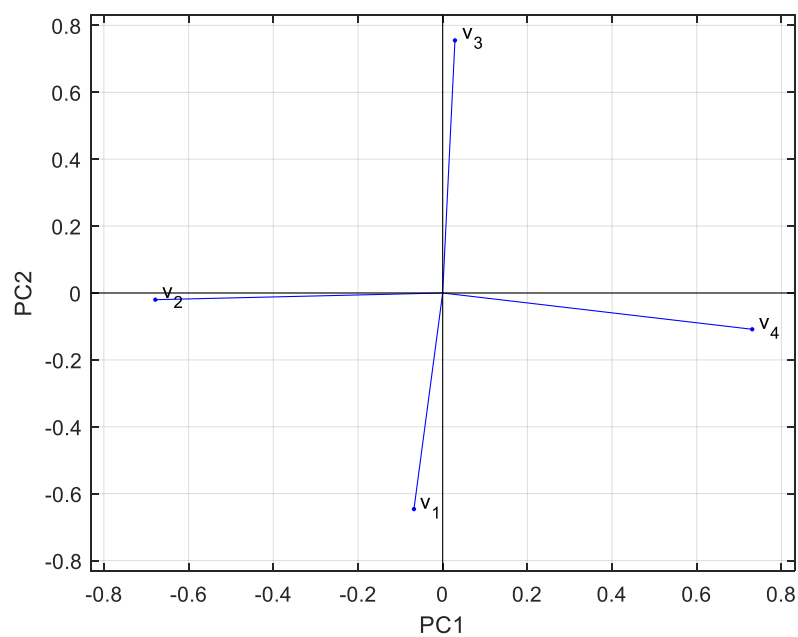
## 4.4 Výstupy analýzy hlavních komponent

Výstupy důležité pro interpretaci výsledků PCA – jsou vlastní hodnoty, komponentní váhy, komponentní skóre a grafy vyjadřující tyto hodnoty. [15]

**Vlastní hodnoty** vyjadřují podíl rozptylu rozložení původních dat v hlavní komponentě, udávají tedy význam dané hlavní komponenty pro vyjádření rozložení dat. Číselná hodnota není důležitá, důležitý je procentuální podíl v dané hlavní komponentě z celkového součtu vlastních hodnot všech hlavních komponent. [14, 15]

**Komponentní váhy** představují míru vzájemné korelace mezi původními proměnnými a hlavními komponentami. Znázorňují, jak velký vliv mají původní proměnné na hlavní komponenty. [14]

**Graf komponentních vah (Plot Components Weights)** zobrazuje původní proměnné v ordinačním prostoru (obr. 4.3). Komponentní váhy (v1,2,3,4) představující původní proměnné mohou být zobrazené pomocí vektorů vycházejících z počátku souřadného systému. Délka vektoru znázorňuje velikost vlivu proměnné. Úhel mezi proměnnou a hlavní komponentou vyjadřuje jejich vzájemnou korelaci. [14, 15]

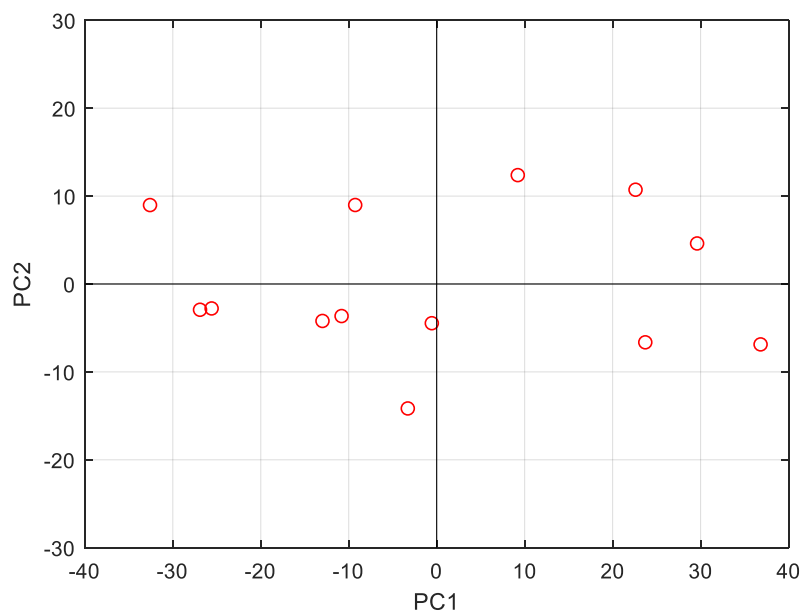


Obr. 4.3: Graf komponentních vah

Čím menší úhel mezi hlavní komponentou a vektorem proměnné je, tím více proměnná komponentu ovlivňuje. Korelace mezi proměnnými je dána kosinem úhlu mezi vektory daných proměnných. Vzdálenost mezi proměnnými zobrazuje jejich vzájemnou korelaci. [14, 15]

**Komponentní skóre** obsahuje souřadnice původních objektů v prostoru hlavních komponent. [14]

**Graf komponentního skóre (Scatterplot)** zobrazuje původní objekty v ordinačním prostoru. Objekty jsou znázorněny jako body, jejichž pozici určuje komponentní skóre. Na obrázku je vidět příklad ordinačního diagramu objektů (obr. 4.4). [14, 15]



Obr. 4.4: Graf komponentního skóre

Tento graf slouží k nalezení podobných objektů, které v prostoru hlavních komponent leží blízko sebe. Nebo také objektů silně odlišných, tedy odlehlých od ostatních. Objekty vzdálené od počátku jsou extrémní, naopak objekty umístěné blízko počátku, jsou objekty obvyklými. Nejdůležitější vlastností výstupu PCA je, že rozložení objektů v prostoru hlavních komponent je vhodné pro třídění dat pomocí shlukovacích metod. [14]

**Dvojný graf (Biplot)** je grafem, který spojuje graf komponentních vah a komponentního skóre. Existují dva typy biplotů. Dělí se podle standardizace vlastních vektorů. Prvním je biplot vzdáleností (distance biplot). Délky vlastních vektorů jsou standardizovány na jednotkovou délku. Pozice objektů v grafu mají rozptyl roven vlastnímu číslu. Euklidovské vzdálenosti v PCA prostoru jsou aproximací euklidovských vzdáleností v původním prostoru, euklidovské vzdálenosti tedy lze interpretovat. Délka vektorů původních proměnných určuje jejich příspěvek k definici daného prostoru. Úhly mezi vektory znázorňující původní proměnné, nelze nijak interpretovat. Druhou variantou je biplot korelací (correlation biplot). Délky vlastních vektorů standardizuje na druhou mocninu z vlastních čísel a pozice objektů mají jednotkový rozptyl. Euklidovské vzdálenosti v prostoru PCA nelze interpretovat, nejsou totiž aproximací euklidovských vzdáleností v původním prostoru. Délky vektorů původních proměnných popisují jejich směrodatnou odchylku. Úhly mezi vektory původních proměnných zobrazují jejich vzájemnou korelaci. [14, 15]

## 4.5 Příklad výpočtu

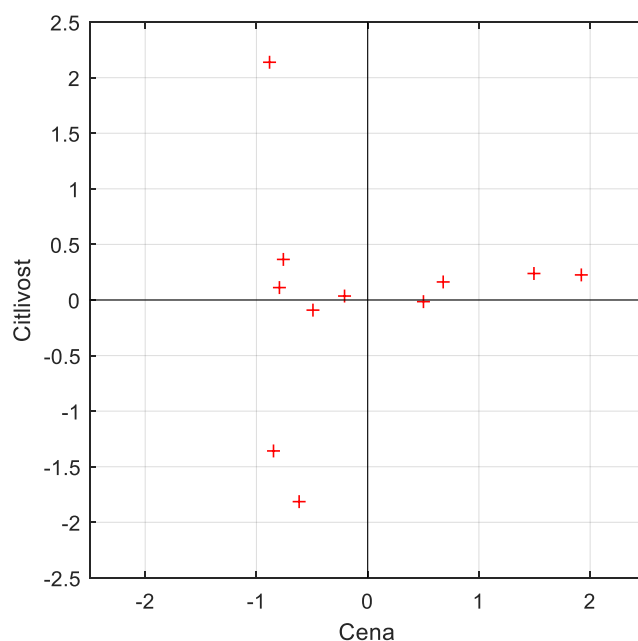
Pro názornost je v této kapitole uveden příklad analýzy hlavních komponent. Jedná se o srovnání parametrů bezdrátových sluchátek. Z článku o testu bezdrátových sluchátek [20], byly převzaty typy a parametry sluchátek, na kterých byla vypočtena

PCA. I když není soubor dat příliš rozsáhlý, bylo by obtížné z něj vyvodit nějaké poznatky, například podobnosti nebo rozdílnosti jednotlivých sluchátek, a to hlavně proto, že data obsahují šest proměnných (šest parametrů). Předpokladem je, že po výpočtu PCA bude možné vytvořit skupiny sluchátek s podobnými vlastnostmi, nebo najít sluchátka, která svými vlastnostmi vybočují. Případně, které vlastnosti to způsobují. Jednotlivé typy sluchátek budeme považovat za objekty a jejich parametry za proměnné. Takovýto soubor dat má tedy šest dimenzí a nelze graficky vynést. Ke zřehlednění dat využijeme analýzu hlavních komponent.

Tab. 4.1: Srovnání parametrů bezdrátových sluchátek

Sluchátka	cena Kč	citlivost dB/mW	impedance $\Omega$	výdrž baterie hod.	nabíjení hod.	hmotnost g
Sony WH-1000XM3	8500	104,5	47	38	3	255
Marshall Major III	2500	97	32	30	3	178
Niceboy HIVE	850	100	32	13	2,5	115
Beats by Dr.Dre Solo3	5000	102	16	40	2	215
JBL Tune 500BT	1350	24	32	16	2	155
Apple AirPods PRO	7300	105	32	4,5	0,3	11
Niceboy HIVE Podsie	950	110	16	3,5	2	9
QCY T1C	600	180	32	4	2	70
Apple AirPods	4500	95	16	5	0,4	16
Niceboy HIVE Pods	1700	92	32	3	2	80
Fixed Steel	700	42	32	6	1,5	33

Nejprve jsou data vložena do matice vstupních dat a následně vypočteny průměry jednotlivých parametrů. (4.4) Pomocí nich je vypočtena centrovaná matice. (4.5) Na obrázku (obr. 4.5) jsou vidět centrované hodnoty pro Cenu a Citlivost.



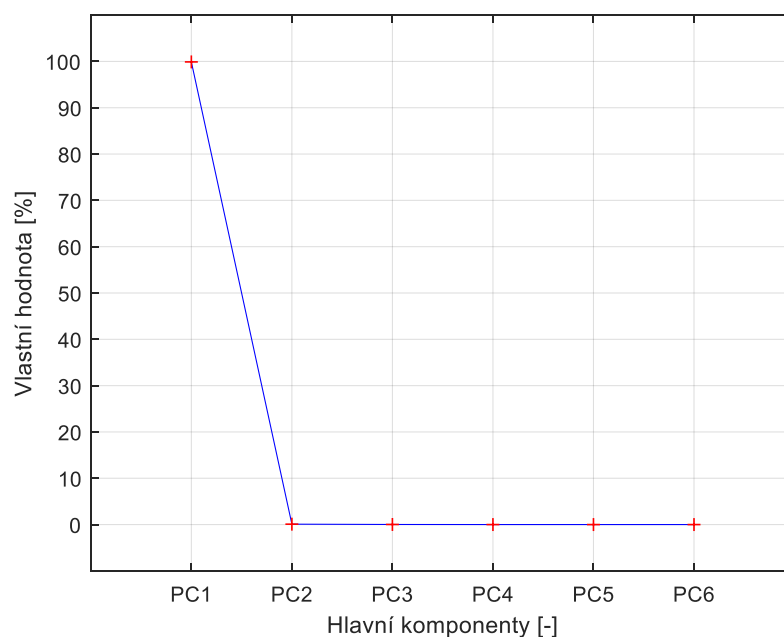
Obr. 4.5: Centrované hodnoty pro Cenu a Citlivost (červené křížky zobrazují jednotlivé typy sluchátek v prostoru prvních dvou parametrů)

Dále se vypočte kovarianční matice (4.6), ze které jsou získány vlastní vektory a vlastní hodnoty, tedy hlavní komponenty a jejich váhy.

Tab. 4.2: Tabulka hodnot vlastních čísel

Hlavní komponenta	Vlastní číslo $\lambda_i$ [-]	Vlastní číslo [%]
PC1	$7,946 \cdot 10^6$	99,89
PC2	$6,824 \cdot 10^3$	0,086
PC3	$1,516 \cdot 10^3$	0,019
PC4	$9,005 \cdot 10^1$	0,001
PC5	9,978	$1,255 \cdot 10^{-4}$
PC6	0,126	$1,578 \cdot 10^{-6}$

Vlastní čísla slouží k určení počtu hlavních komponent, které budou využity pro interpretaci dat. Na grafu úpatí vlastních čísel (obr. 4.6) je vidět, že první hlavní komponenta PC1, obsahuje téměř všechny rozptyly původních dat. Pro zobrazení výsledků budou využity první dvě hlavní komponenty PC1 a PC2.



Obr. 4.6: Indexový graf úpatí vlastních čísel

Hlavní komponenty PC2, PC3, PC4, PC5 a PC6 budou zanedbány, protože jejich popis rozložení objektů má velmi malý význam. V tabulce (tab. 4.3) jsou uvedeny vypočtené vlastní vektory. Hodnoty, v nich obsažené, jsou váhy určující korelaci mezi původními proměnnými a hlavními komponentami.

Tab. 4.3: Tabulka komponentních vah

proměnná	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$
cena	0,1	-0,011	-0,002	$-1,841 \cdot 10^{-5}$	-0,001	$2,077 \cdot 10^{-4}$
citlivost	0,001	-0,072	0,998	0,009	0,009	-0,005
impedance	$8,633 \cdot 10^{-4}$	0,044	-0,008	0,928	0,369	-0,029
výdrž baterie	0,003	0,142	0,005	-0,372	0,917	-0,033
nabíjení	$-5,684 \cdot 10^{-5}$	0,009	0,005	0,015	0,041	0,1
hmotnost	0,011	0,986	0,072	0,013	-0,148	-0,003

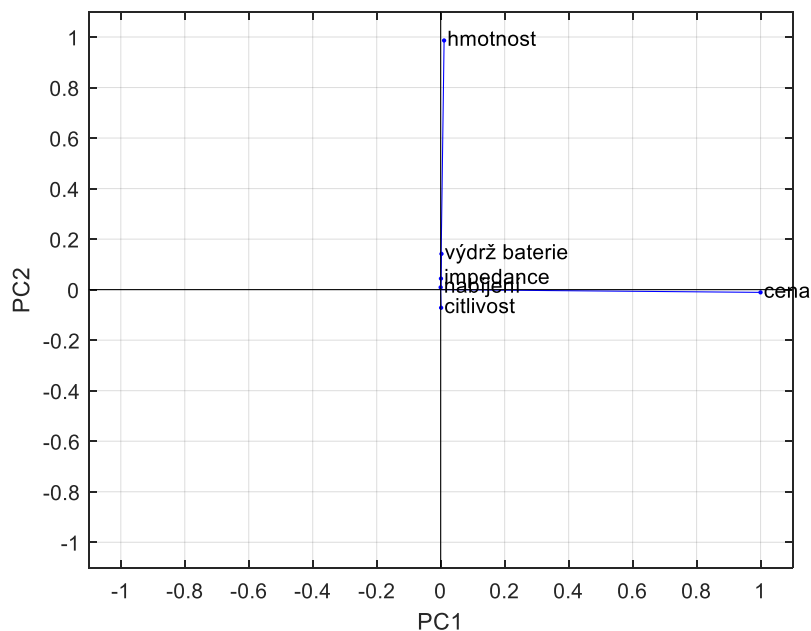
V další tabulce (tab. 4.4) jsou hodnoty koeficientů hlavních komponent, neboli komponentní skóre jednotlivých typů sluchátek.



Tab. 4.4: Tabulka komponentního skóre

Index	Sluchátka	PC1	PC2	PC3	PC4	PC5	PC6
1	Sony WH-1000XM3	5 415	94	8	10	0,1	0,4
2	Marshall Major III	-585	82	8	-2	5	0,1
3	Niceboy HIVE	-2 236	35	10	4	0,2	0,1
4	Beats by Dr.Dre Solo3	1 915	92	10	-20	-0,04	-0,4
5	JBL Tune 500BT	-1 736	75	-64	3	-4	-0,2
6	Apple AirPods PRO	4 212	-139	-6	6	1	-0,2
7	Niceboy HIVE Podsie	-2 137	-73	12	-11	0,1	0,8
8	QCY T1C	-2 487	-13	87	7	-0,5	-0,4
9	Apple AirPods	1 413	-103	-10	-9	-2	-0,2
10	Niceboy HIVE Pods	-1 387	-9	-2	7	-5	0,2
11	Fixed Steel	-2 387	-41	-54	5	6	-0,2

Z uvedených hodnot v (tab. 4.3) je vynesena graf komponentních vah (obr. 4.7). Graf zobrazuje původní proměnné v prostoru první a druhé hlavní komponenty (PC1 a PC2). Z uvedeného grafu je možné vyčíst, že hmotnost a cena se zásadně liší od zbývajících parametrů.

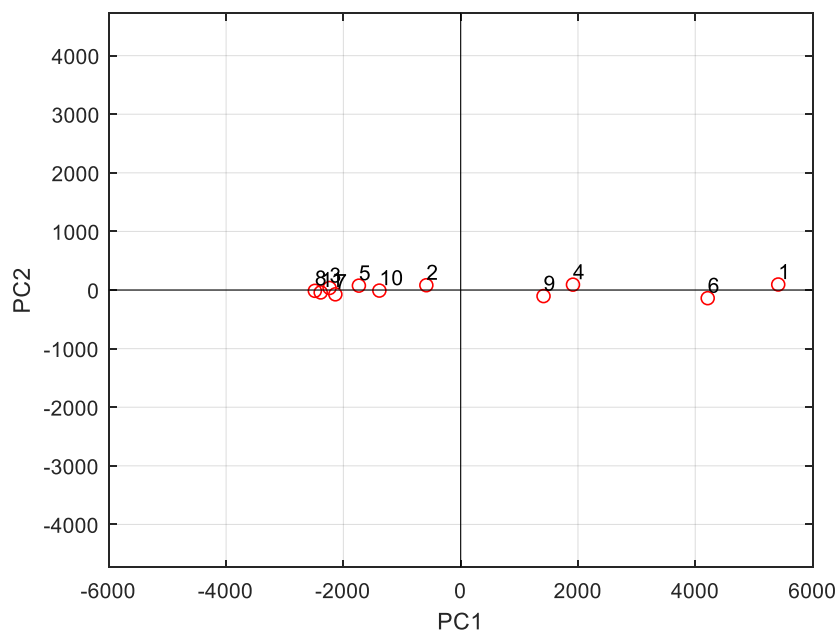


Obr. 4.7: Graf komponentních vah

Je možné říci, že cena koreluje s první hlavní komponentou, kdežto zbývající parametry více korelují s druhou hlavní komponentou. Největší vliv na první hlavní komponentu měla cena, na druhou hlavní komponentu měla nejvyšší vliv hmotnost.

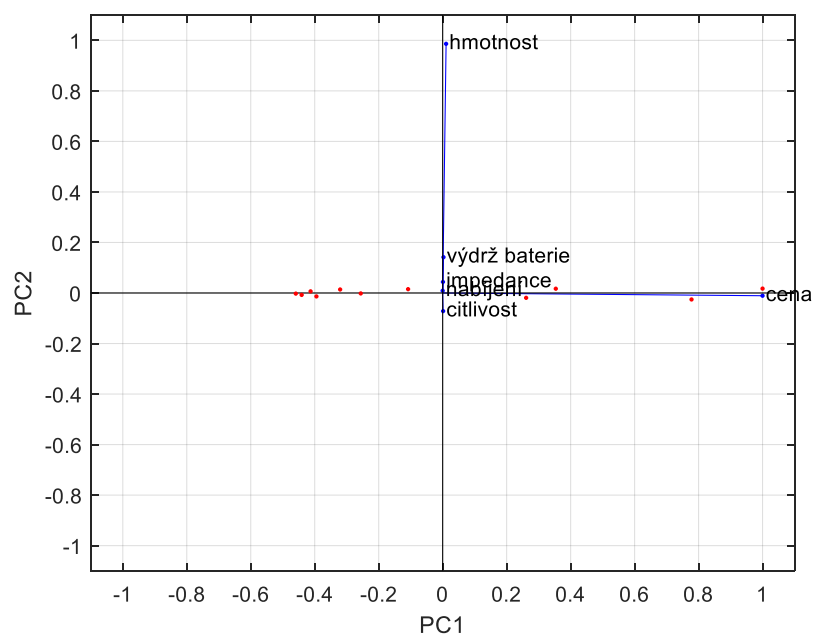
Rozptylový diagram komponentního skóre je vidět na obrázku (obr. 4.8). Body zobrazují jednotlivé objekty (sluchátka) v prostoru první a druhé komponenty. Čísla

odpovídají indexům sluchátek v tabulce (tab. 4.4). Sluchátka značek Marshall, Niceboy, JBL, QCY, Fixed si jsou podobná a tvoří shluk. Osamocená sluchátka jako Apple AirPods PRO a Sony WH-1000XM3 se vymykají. Můžeme také říci, že jejich vlastnosti nejsou v uvedeném výčtu typické.



Obr. 4.8: Rozptylový diagram komponentního skóre

Posledním zobrazením výstupu analýzy hlavních komponent je dvojgraf (biplot). Spojuje zobrazení komponentních vah a komponentního skóre do jednoho zobrazení (obr. 4.9). U dvojgrafu se sleduje vzdálenost mezi proměnnými a objekty. Jedná se o dvojgraf korelací, pozice objektů (typů sluchátek) mají jednotkový rozptyl.



Obr. 4.9: Dvojný graf

Kdyby byl objekt ve stejném bodě jako proměnná (parametr sluchátek) nebo byl v její blízkosti, znamenalo by to vzájemnou interakci mezi danou proměnnou a blízkým objektem. Interakce může sloužit interpretaci objektů. [14] V tomto případě si jsou velice blízko sluchátka Sony a cena. To může naznačovat, že u sluchátek Sony je cena zásadním odlišujícím parametrem. Tato sluchátka jsou v uvedeném přehledu nejdražší (tab. 4.1). Výsledky PCA, nemusí být využity pouze pro grafické zobrazení, ale mohou být podrobeny dalším analýzám pro třídění a klasifikaci dat tak, aby sluchátka mohla být rozdělena do skupin s podobnými vlastnostmi.

## 5. KLASIFIKACE DAT

Dále bude na data analyzovaná pomocí PCA, aplikována některá z klasifikačních metod, která údery rozdělí do jednotlivých klasifikačních tříd. Tyto klasifikační třídy představují jednotlivé bubny nebo činely. Klasifikační metody hledají vzájemnou podobnost mezi objekty. Klasifikace dat je jedna z hlavních úloh strojového učení (machine learning). Pomocí klasifikačních metod je možné jeden objekt přiřadit do existující třídy, nebo množinu objektů rozdělit na několik klasifikačních tříd. [14] Rozdělení objektů do tříd může být provedeno pomocí následujících způsobů:

- Klasifikace pomocí diskriminačních funkcí
- Klasifikace pomocí minimální vzdálenosti od etalonů tříd
- Klasifikace pomocí hraničních ploch

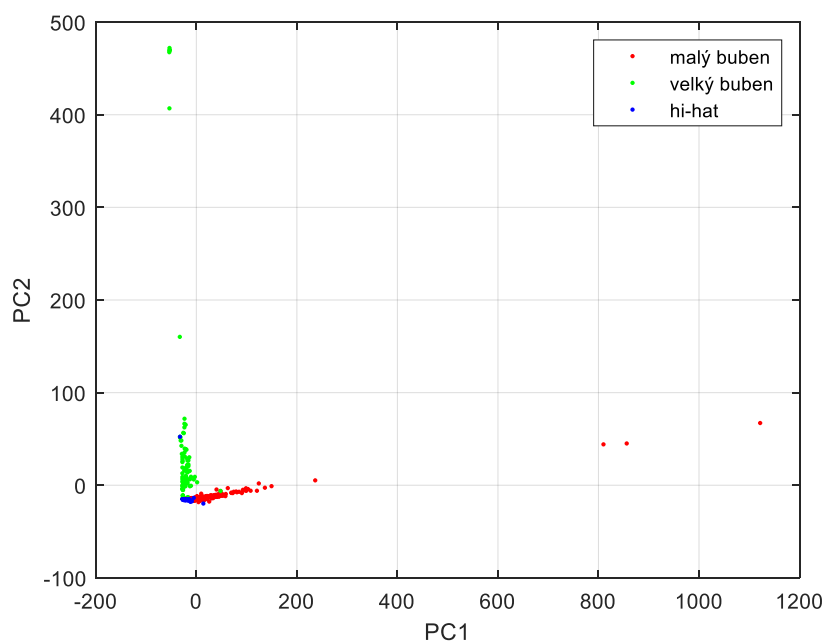
**Klasifikace pomocí diskriminačních funkcí** je založena na výpočtu míry příslušnosti daného objektu k dané klasifikační třídě. Objekt je přiřazen do té klasifikační třídy, pro kterou byla vypočtena největší míra příslušnosti. Do této skupiny patří například metoda rozhodovacího stromu. [13, 16]

**Klasifikace pomocí minimální vzdálenosti od etalonů tříd** přiřazuje objekty do klasifikačních tříd na základě výpočtu vzdálenosti daného objektu od etalonu dané třídy. Etalon je reprezentativní objekt dané třídy. Počet etalonů klasifikační třídy není dán, záleží na zvolené klasifikační metodě. Například u centroidové metody je etalon pouze jeden (centroid), oproti metodě průměrné vazby, kde je etalonem každý prvek dané třídy. Takto probíhá výpočet například u metody  $k$  – průměrů. [13, 16]

**Klasifikace pomocí hraničních ploch** využívá hranic definovaných tak, aby v prostoru oddělovaly jednotlivé třídy. Objekt umístěný v daném hraničním prostoru je přidělen dané třídě. Typickým představitelem této skupiny může být metoda podpůrných vektorů (SVM). [13, 16]

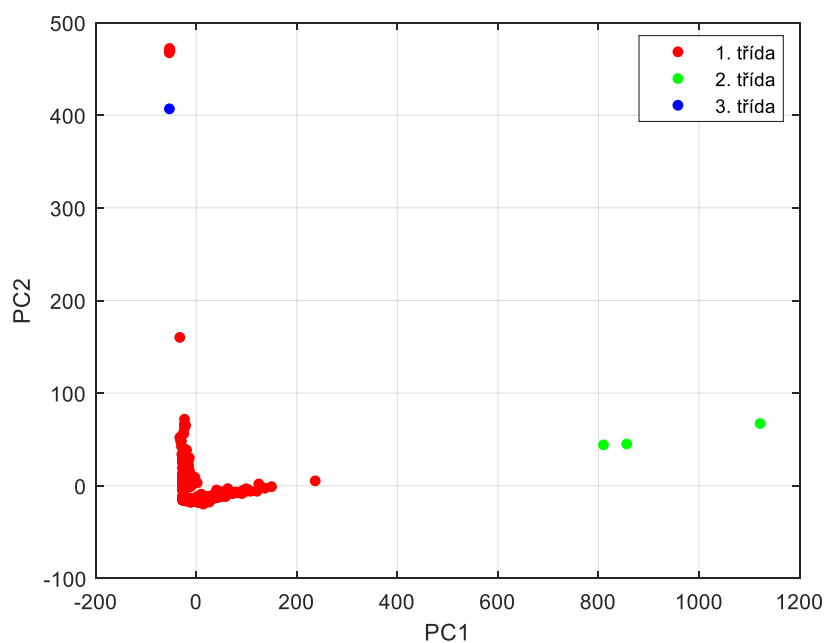
### 5.1 Výběr klasifikační metody

Pro výběr vhodné klasifikační metody byl proveden pokus na vzorku trénovacích dat. Bylo použito celkem 427 nahrávek. Konkrétně 221 nahrávek malého bubnu, 102 nahrávek velkého bubnu a 104 nahrávek hi-hat. Tyto nahrávky byly filtrovány třetino-oktávovou bankou filtrů, a následně byly vypočteny energie v jednotlivých pásmech dle vztahu (3.2). Dále byla provedena analýza hlavních komponent. Na data v novém prostoru pěti prvních hlavních komponent byly aplikovány klasifikační algoritmy. Na obrázku (obr. 5.1) je znázorněno rozložení trénovacích nahrávek v prostoru prvních dvou hlavních komponent.

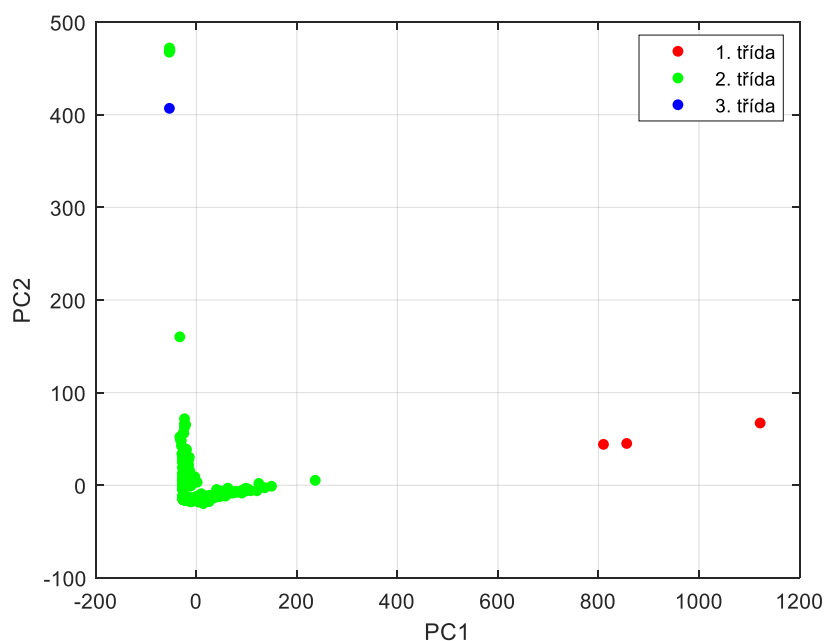


Obr. 5.1: Trénovací nahrávky v prostoru prvních dvou hlavních komponent

Pro test byly vybrány tři klasifikační metody a to k-průměrů, hierarchické shlukování a metoda podpůrných vektorů. První testovanou metodou byla k-průměrů. Tato metoda má předem daný počet shluků. Na začátku se provede náhodné rozdělení do shluků. Určí se centroidy těchto shluků a jednotlivé prvky jsou přiřazeny k nejbližším centroidům. Centroidy se znovu přepočítají. Tyto operace se opakují, dokud se přesuny neustálí. [15]



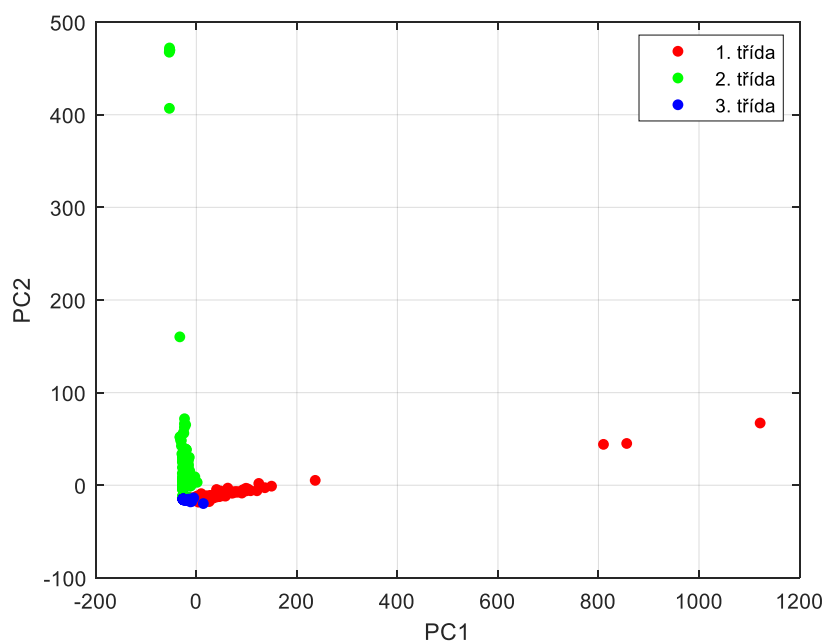
Obr. 5.2: Klasifikace trénovacích nahrávek metodou k – průměrů



Obr. 5.3: Klasifikace trénovacích nahrávek metodou hierarchického shlukování

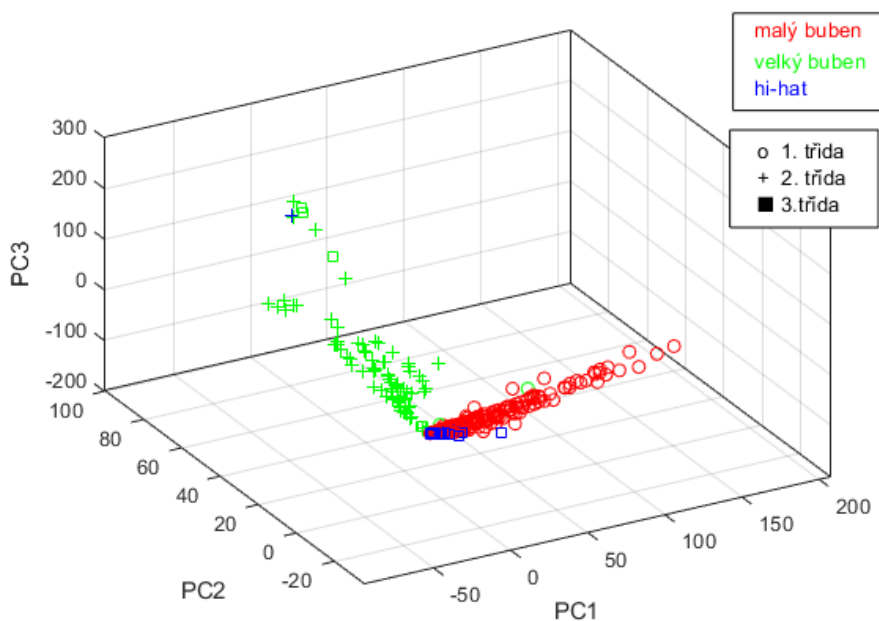
Na obrázku (obr. 5.2) je znázorněn výsledek klasifikace pomocí metody k-průměrů a na obrázku (obr. 5.3) je znázorněn výsledek klasifikace pomocí metody hierarchického shlukování. Hierarchické shlukování nemá předem daný počet tříd. Každá třída se rozdělí na dvě poloviny dělení, končí, když v třídě zůstane poslední prvek. Tento proces může probíhat i obráceně, tedy jednotlivé prvky se seskupují, dokud nejsou v jedné třídě. [15] Z obrázků vyplývá, že oběma metodám činí potíže osamocené objekty a z objektů v levém dolním rohu vytvoří jeden shluk.

Lepších výsledků bylo dosaženo pomocí metody podpůrných vektorů, které je možné vidět na (obr. 5.4).



Obr. 5.4: Klasifikace trénovacích nahrávek metodou podpůrných vektorů

Na obrázku (obr. 5.5) je přiblížený shluk nacházející se poblíž počátku. Barevně jsou rozlišeny reálné třídy podle vstupních dat. Rozdělení do tříd pomocí metody podpůrných vektorů je rozlišeno pomocí symbolů. Zde je vidět, že opravdová většina dat je zařazena do správných tříd.



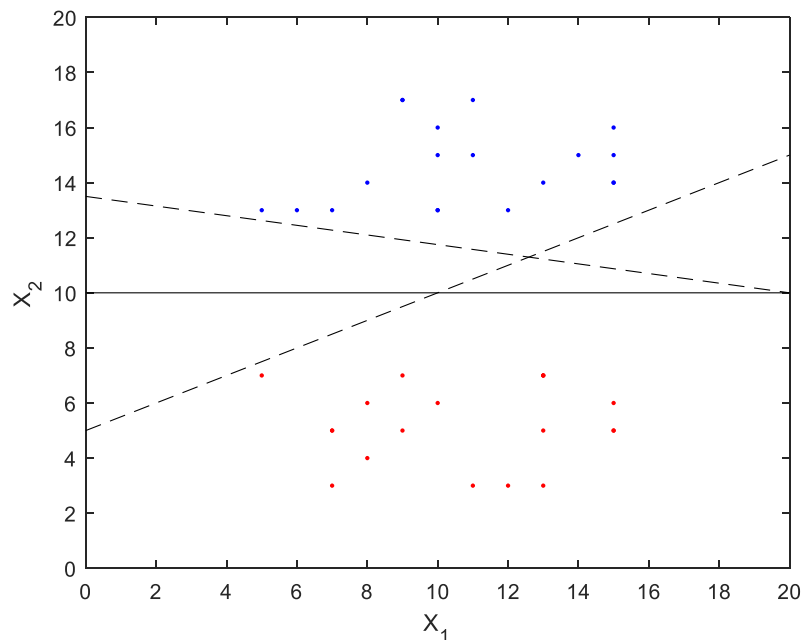
Obr. 5.5 Srovnání reálných tříd a tříd vytvořených metodou podpůrných vektorů

Na základě výsledků tohoto testu byla vybrána metoda podpůrných vektorů.

## 5.2 Metoda podpůrných vektorů

Metoda podpůrných vektorů (SVM, support vector machine) je jedním z nejpoužívanějších klasifikačních algoritmů. [12] Jedná se o typ strojového učení s učitelem (supervised learning), tedy data jsou klasifikována na základě trénovací množiny známých dat. SVM má mnoho variant použitelných pro různé klasifikační metody. Pro pochopení problematiky je zde vysvětlena lineární separace do klasifikačních tříd. [16]

Jsou-li třídy lineárně separovatelné existuje množina hranic, podle kterých může být prostor rozdělen tak, aby na jedné straně hranice byly pouze objekty patřící do jedné třídy a na druhé straně hranice pouze objekty patřící do druhé třídy. Na obrázku (obr. 5.6) je uveden příklad několika hranic. Algoritmus podpůrných vektorů hledá hranici, která představuje nejrobustnější rozdělení nových testovacích objektů, tedy hranici, která je stejně vzdálená od objektů první i druhé třídy. Na obrázku (obr. 5.6) je tato hranice zobrazena plnou čarou. [13, 16]



Obr. 5.6: Příklad množiny hranic u separabilních tříd SVM

Kritériem pro nalezení optimální hranice je vytvoření co nejširšího tolerančního pásma mezi hranicí a oběma třídami v množině trénovacích dat. Hranice je definovaná vztahem (5.1). [13]

$$h(x) = \mathbf{w}^T \mathbf{x}_k + \mathbf{w}_0 \quad (5.1)$$

Kde vektor  $\mathbf{w}$  udává orientaci hranice a  $\mathbf{w}_0$  její polohu,  $\mathbf{x}$  je prvkem, který má být klasifikován.



$$\mathbf{w}^T x_k + \mathbf{w}_0 \begin{cases} \geq 1 & \text{pro všechna } x \text{ v první třídě} \\ \leq 1 & \text{pro všechna } x \text{ v druhé třídě} \end{cases} \quad (5.2)$$

Objekt  $\mathbf{x}$  bude přiřazen do jedné ze dvou tříd na základě výsledku vztahu (5.2). [16, 19]

$$d = \frac{|h(x)|}{\|\mathbf{w}\|} \quad (5.3)$$

Vzdálenost bodu od hranice je dána vztahem (5.3). Souřadnice vektoru  $\mathbf{w}$  určují orientaci hranice. Je-li výsledek funkce  $h(x)$  v nejbližším bodě první klasifikační třídy roven +1 a -1 v nejbližším bodě druhé klasifikační třídy, je šířka tolerančního pásma definována dle vztahu (5.4).

$$\frac{1}{\|\mathbf{w}\|} + \frac{1}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|} \quad (5.4)$$

Z toho vyplývá definice kritéria pro stanovení hranice (5.5), kde je hledáno minimum funkce.

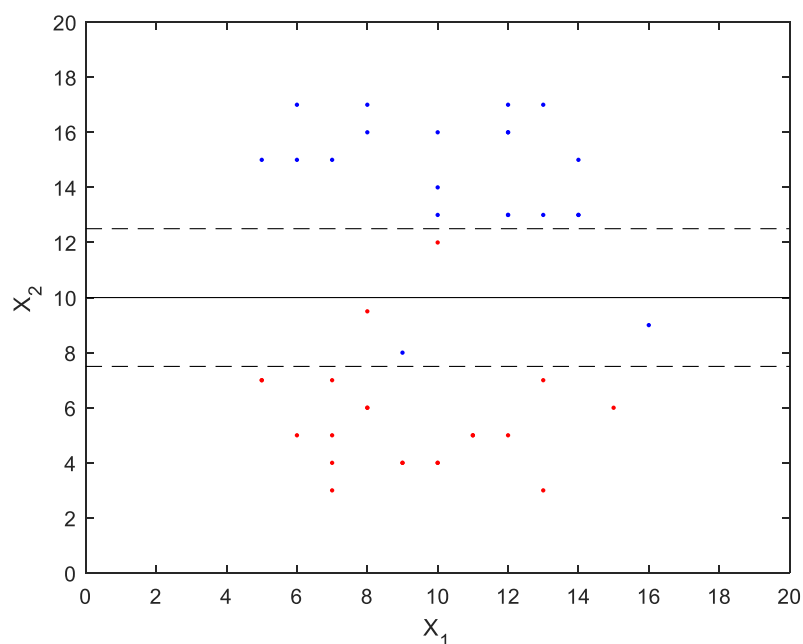
$$J(\mathbf{w}, \mathbf{w}_0) = \frac{\|\mathbf{w}\|^2}{2} \quad (5.5)$$

Za podmínky (5.6).

$$y_k \cdot (\mathbf{w}^T x_k + \mathbf{w}_0) \geq 1, \text{ pro všechna } x \text{ trénovací množiny} \quad (5.6)$$

Kde  $y_k = 1$  pro první třídu a  $y_k = -1$  pro druhou třídu. [16, 13, 19]

Jsou-li klasifikační třídy lineárně neseparovatelné, podmínka (5.6) nemůže platit pro všechny objekty trénovací množiny. Příklad takovýchto tříd je uveden na obrázku (obr. 5.7). Je vidět, že některé objekty jsou již v tolerančním pásmu, a dokonce i za hranicí a budou špatně klasifikovány.



Obr. 5.7: Lineárně neseparovatelné třídy

Někdy také může být výhodnější vytvořit robustnější klasifikátor s širším tolerančním pásmem za cenu špatné klasifikace několika málo objektů. Klasifikátory s úzkým tolerančním pásmem nemusí správně třídit nové testovací objekty. [13]

Kritérium (5.6) je tedy upraveno. Zavádí se nová proměnná  $\xi_k$ , nazývaná relaxační (slack variable), pro kterou platí:

- je-li objekt mimo toleranční pásma  $\xi_k = 0$
- leží-li objekt uvnitř tolerančního pásma a je správně klasifikován  $0 < \xi_k \leq 1$
- nachází-li se objekt na druhé straně hranice a je špatně vyhodnocen, pak  $\xi_k > 1$

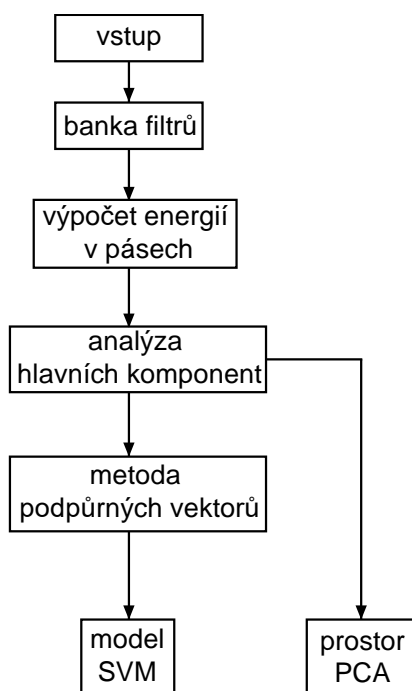
Podmínka je tedy nově definována jako (5.7).

$$y_k \cdot (\mathbf{w}^T x_k + \mathbf{w}_0) \geq 1 - \xi_k \quad (5.7)$$

Při optimalizaci není teď bráno v úvahu jen co nejširší toleranční pásmo, ale také co nejmenší počet objektů, pro které platí  $\xi_k > 0$ . [13, 16]

## 6. REALIZACE

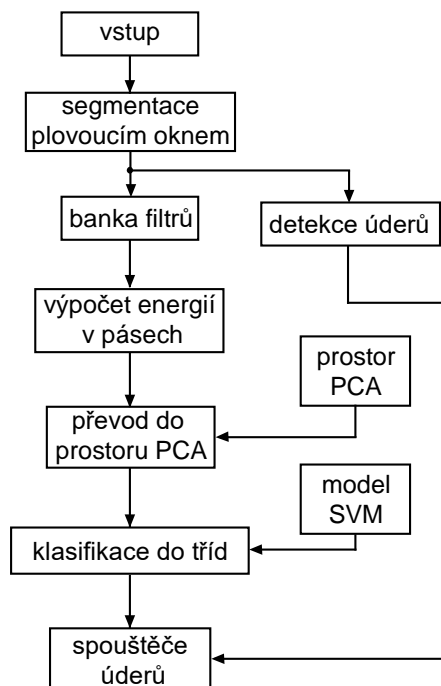
Program pro triggerování bicí soupravy byl realizován v prostředí Matlab. Program je rozdělen na dvě základní části, a to trénovací s testovací. Hlavním souborem pro spuštění trénovací části je *uceni.m*. Slouží k výpočtu prostoru hlavních komponent a rozdělení clusterovacích tříd na základě trénovací sady dat. Na obrázku (obr. 6.1) je zobrazeno blokové schéma této části programu.



Obr. 6.1: Blokové schéma trénovací části programu

Testovací část s hlavním souborem *main.m*, analyzuje nahrávku a jednotlivé údery přiřazuje do daných tříd. Blokové schéma této části programu je zobrazeno na obrázku (obr. 6.2).

V této kapitole budou popsány jednotlivé bloky obou částí programu. Některé bloky jsou využity v obou větvích programu, proto je většina kódu naprogramovaná jako knihovny (funkce), které využívají obě větve.



Obr. 6.2: Blokové schéma testovací části programu

## 6.1 Segmentace plovoucím oknem

V teoretickém úvodu byly zmíněny dvě možnosti časové segmentace vstupního signálu. První možností bylo rozdělit nahrávku po jednotlivých úderech a ty následně analyzovat. Druhou možností byla segmentace plovoucím oknem (kontinuální režim). Segmentace plovoucím oknem přináší možnost, že by systém v budoucnu mohl pracovat v reálném čase.

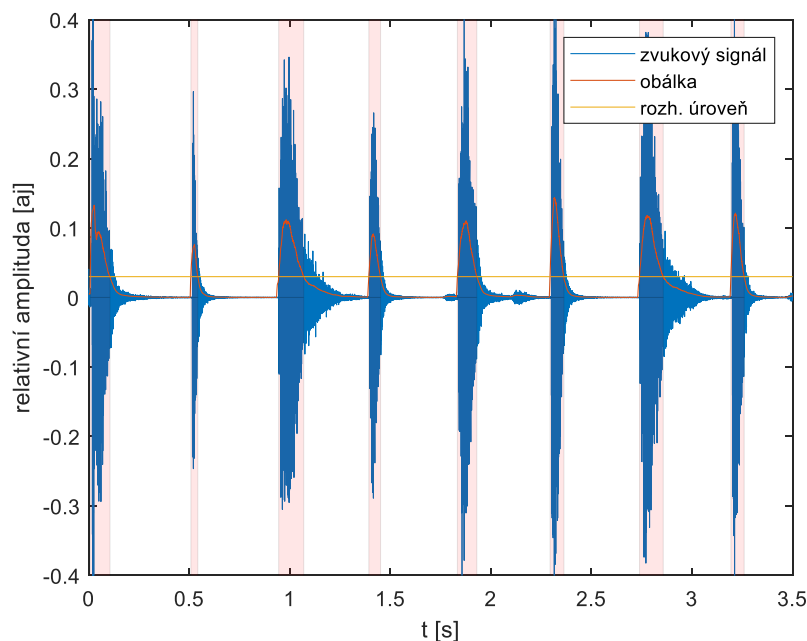
Testy obou variant bylo zjištěno, že obě jsou funkční. Implementována byla segmentace plovoucím oknem, díky tomu má rozpoznávací systém předpoklad pro budoucí fungování v reálném čase. Experimentálně bylo zjištěno, že bezpečnou délkou okna pro rozpoznání úderu je 3 500 vzorků. Pro přesnější detekci začátků a konců úderů byl zaveden přesah oken na 2/3.

Segmentaci do jednotlivých oken zajišťuje funkce *okno.m*. Vstupem této funkce je samotná nahrávka, délka okna a přesah. Výstupy jsou časové značky oken, počet oken, časové značky úderů a počet úderů v nahrávce.

## 6.2 Detekce úderů

Detekce začátků a konců úderů je také začleněna do funkce *okno.m*. Aby detekce správně fungovala, je potřeba normalizovat hlasitost nahrávky. Ve většině programů pro zpracování zvuku, se pod normalizací hlasitosti myslí zesílení signálu tak, že jeho nejhlasitější špička dosáhne úrovně přenosu 1. Takovýto přístup nezaručuje stejnou nebo alespoň podobnou hlasitost signálů. Proto zde není prováděno měření špičkové hodnoty, ale hodnoty efektivní (RMS). Všechny vstupní nahrávky jsou tedy normalizovány na hodnotu 0.4 (RMS). Pro zvýraznění tranzientů v signálu byl použit diferenční filtr. Jedná se o výpočet rozdílu dvou sousedních vzorků.

Pro takto upravený signál je vypočtena jeho obálka pomocí funkce *envelop* obsažené v Matlabu. Průchod obálky signálu rozhodovací úrovní je zaznamenán jako začátek nebo konec úderu; podle toho, jestli křivka obálky po průtnutí rozhodovací úrovně stoupá nebo klesá. Tento proces probíhá v rámci časového okna, kde není znám průběh zbytku signálu, který by mohl mít vliv na výpočet obálky. Proto bylo využito přesahu mezi okny. Díky přesahu se jeden průchod rozhodovací úrovní promítne do více oken. Časy těchto průchodů se průměrují.



Obr. 6.3: Ukázka funkce detektoru úderů

Na obrázku (obr. 6.3) je znázorněna funkce detektoru úderů. Signál je nahrávkou bicí soupravy, která již prošla předzpracováním popsáním výše. Dále je na obrázku vidět obálka počítaná z RMS hodnoty signálu a rozhodovací úroveň. Růžovou barvou jsou zvýrazněny úseky signálu, které detektor označil jako úder. Výstupem tohoto bloku je počet nalezených úderů a časové značky začátků a konců úderů.

### 6.3 Banka filtrů

Banka filtrů je v této práci využita pro frekvenční analýzu vstupního signálu. Tuto analýzu provádí knihovna *banka\_filtru.m*. Vstupem tohoto bloku je neupravený vstupní signál. Výstupem je vyfiltrovaný signál a střední kmitočty filtrů.

Banku tvoří řada IIR číslicových filtrů, jedná se o filtry typu pásmová propust druhého řádu. Jako aproximace byla zvolena Butterworthova, zejména protože nezpůsobuje zvlnění frekvenční charakteristiky v propustném pásmu filtru.

Rozlišení spektrální analýzy je dáno počtem pásem. Na základě podobnosti kritických pásem lidského sluchu byla jako výchozí zvolena třetino – oktávová banka filtrů. Pro ověření její vhodnosti byly ještě otestovány oktávová a šestino – oktávová banka filtrů. Výsledky tohoto testu jsou zaznamenány v tabulce (tab. 6.1).

Tab. 6.1: Srovnání úspěšnosti klasifikace s různými bankami filtrů

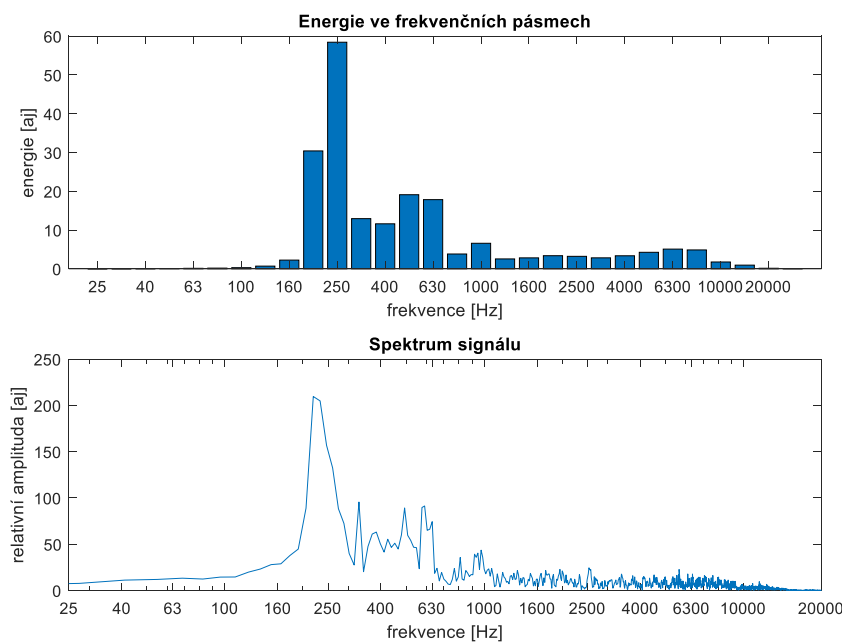
	oktávová banka [%]	1/3 oktávová banka [%]	1/6 oktávová banka [%]
malý buben	88	75	78
velký buben	91	94	91
hi-hat	47	53	53
crash	78	84	67
ride	38	56	63
malý + hi-hat	72	84	72
velký + hi-hat	62	81	100
průměrně	68	75	75

Z tabulky vyplývá, že třetino – oktávová banka filtrů je vhodným řešením. Oktávová banka filtrů má o 7 % nižší úspěšnost správného rozeznání úderu než třetino – oktávová banka. Šestino – oktávová banka nevykazuje žádné zlepšení v celkové úspěšnosti, jen prodlužuje procesní čas potřebný na výpočet. Podrobné výsledky testů jsou v tabulkách v příloze (Příloha 1, Příloha 1 a Příloha 3).

## 6.4 Výpočet energií ve frekvenčních pásmech

Tato část programu je obsažena v knihovně *energie.m*. Vstupem funkce je výstup z banky filtrů. Výstupem je vektor energií, pro jednotlivá frekvenční pásma. Výpočet energie je definován jako součet čtverců absolutních hodnot signálu (3.2).

Na obrázku (obr. 6.4) je zobrazeno srovnání energií spočítaných ve frekvenčních pásmech s amplitudovým spektrem stejného signálu. Jako ukázkový signál byla použita nahrávka úderu na malý buben.



Obr. 6.4: Srovnání energií spočtených ve frekvenčních pásmech s amplitudovým spektrem malého bubnu

## 6.5 Analýza hlavních komponent

Pro výpočet PCA je využita funkce *pca.m* prostředí Matlab. Funkce využívá přednastavený algoritmus SVD viz kapitola 4.3.

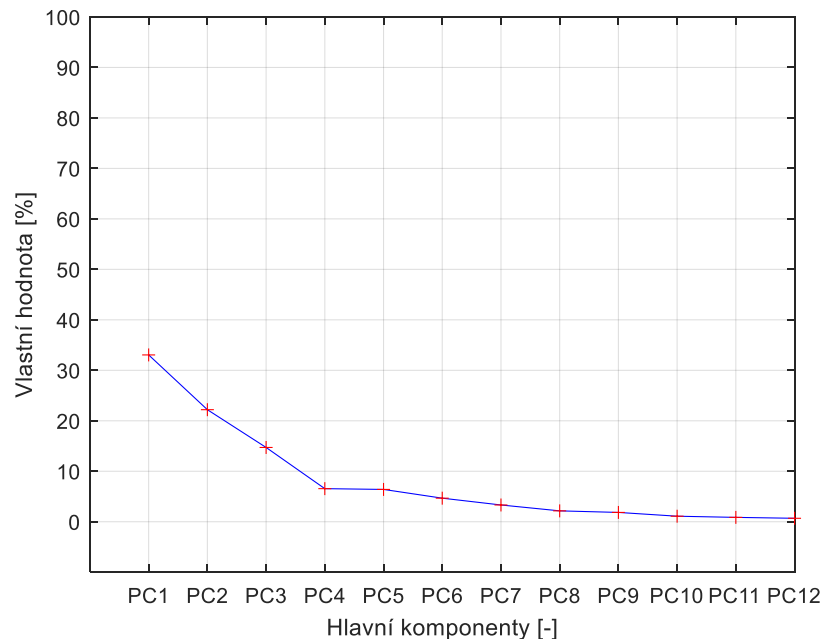
Vstupem této funkce je matice vypočtených energií, kde řádky jsou frekvenčními pásmy a sloupce představují jednotlivé nahrávky úderů. Dalším vstupem je počet hlavních komponent, které budou vypočteny.

Prvním důležitým výstupem jsou vlastní čísla. Vektor s vlastními čísly nese název *latent*. Vlastní čísla vyjádřená v procentech jsou ve vektoru *explained*. Vlastní čísla v procentech (tab. 6.2) poslouží k určení počtu hlavních komponent.

Tab. 6.2: Vlastní čísla [%]

PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
33,05	22,20	16,71	6,56	6,40	4,67	3,32	2,16	1,84	1,10
PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
0,86	0,68	0,49	0,37	0,24	0,24	0,22	0,18	0,16	0,13
PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30
0,11	0,1	0,07	0,04	0,04	0,02	0,01	-	-	-

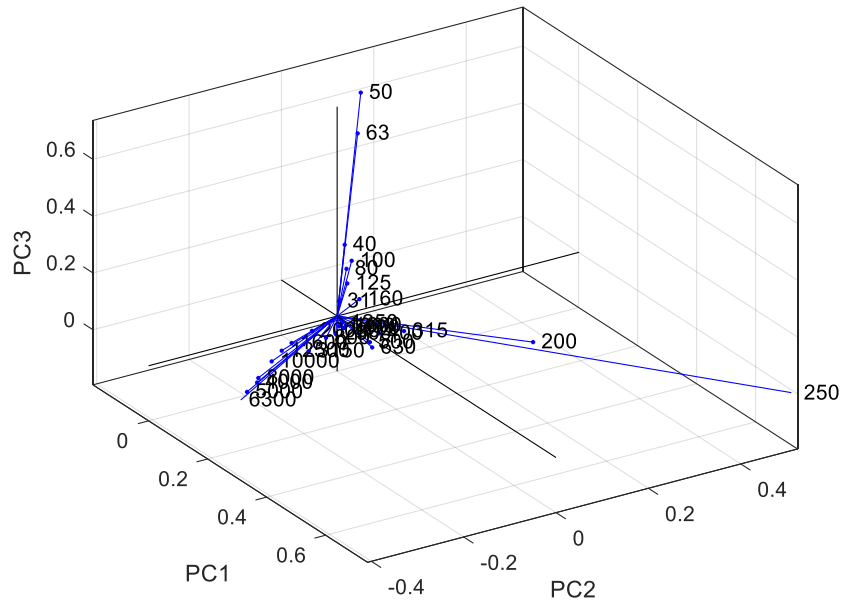
Pro lepší představu o vlivu jednotlivých hlavních komponent je na (obr. 6.5) uveden indexový graf úpatí vlastních čísel.



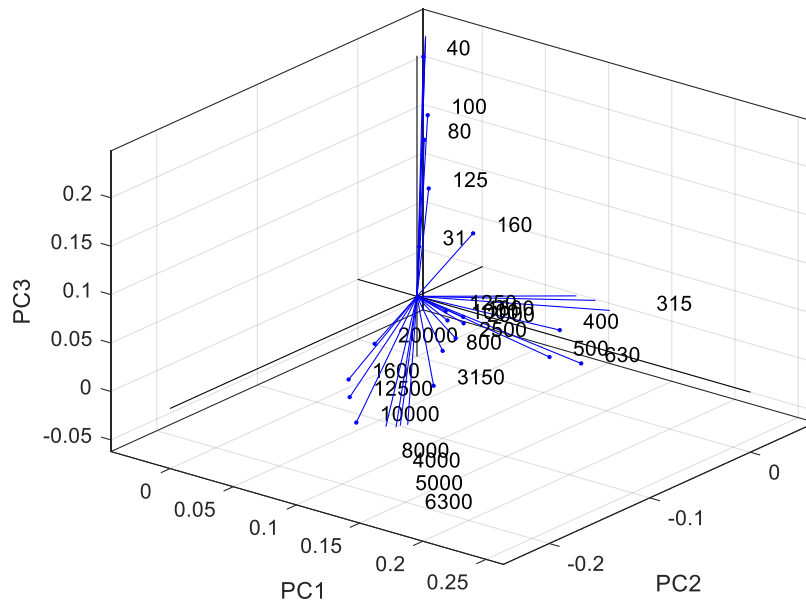
Obr. 6.5: Indexový graf úpatí vlastních čísel

Prvních devět hlavních komponent obsahuje 96,82 % rozptylu původních dat před výpočtem PCA. PCA bude dále počítána pouze pro 9 hlavních komponent. Počet dimenzí prostoru, ve kterém se data nachází se snížil z 30 na 9. Výsledky PCA mohou být graficky zobrazeny pouze v prostoru prvních tří hlavních komponent. Takové zobrazení obsahuje pouze 71,96 % rozptylu původních dat.

Druhým výstupním parametrem je matice komponentních vah označená jako *coeff*. Tento výstup lze reprezentovat grafem komponentních vah, který zobrazuje obrázek (obr. 6.6).



Obr. 6.6: Graf komponentních vah



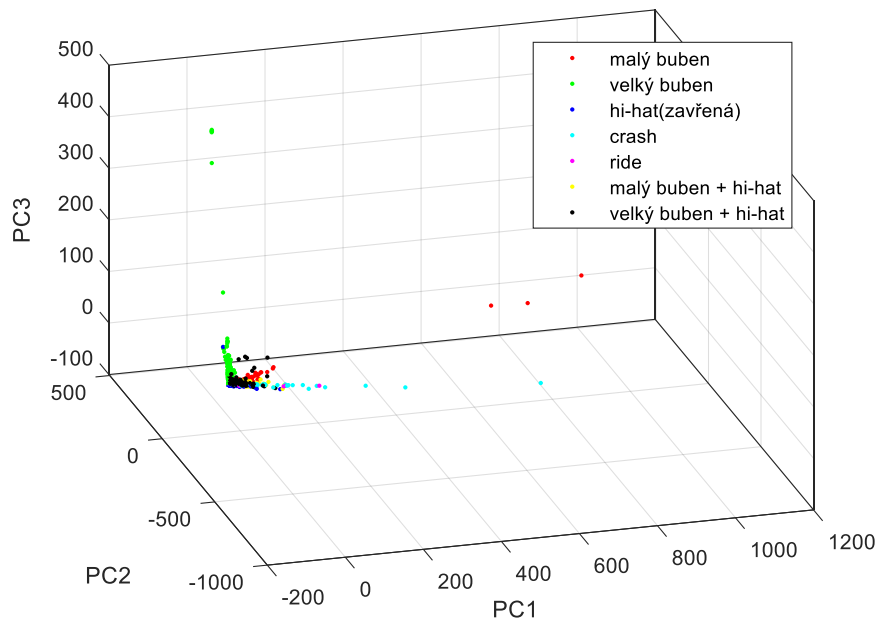
Obr. 6.7: Výřez grafu komponentních vah

Z grafu se dá usoudit, že tři hlavní skupiny frekvencí, které vystupují ze shluku, budou zásadní pro tři základní kategorie. Frekvenční pásma s  $f_c = 50$  Hz a 60 Hz pro velký buben, 200 a 250 Hz pro malý buben a frekvenční pásma s  $f_c = 4\ 000$



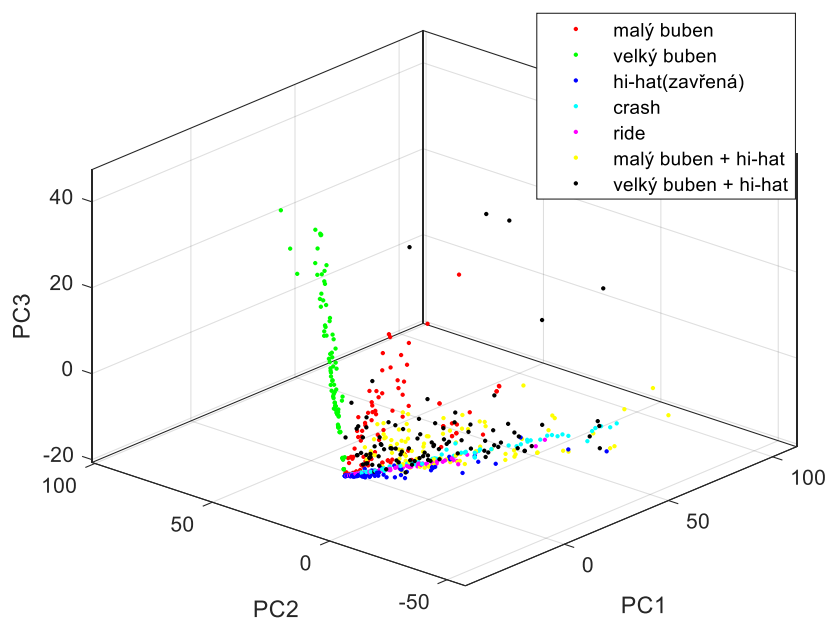
až 8 000 Hz pro činely. Zbylé frekvence ve shluku se mohou kategoriemi prolínat, a proto nejsou jejich typickými představiteli.

Dalším výstupním parametrem je matice komponentního skóre. Nese označení *score*. Graf komponentního skóre je vynesena na obrázku (obr. 6.8).



Obr. 6.8: Graf komponentního skóre

Na obrázku (obr. 6.9) se nachází výřez z grafu komponentního skóre pro lepší zobrazení shluku, nacházejícího se v levém dolním rohu.



Obr. 6.9: Výřez grafu komponentního skóre

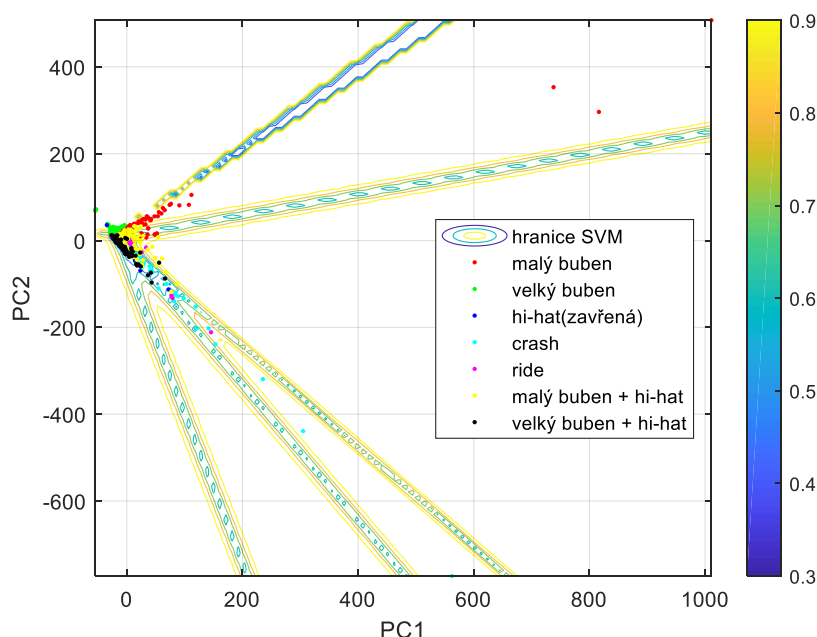
Z grafu vyplývá, že nejlépe separovatelnou třídou by měl být velký buben. Další dobře separovatelnou třídou by mohl být crash. Komplikovaně vypadá separace třídy pro hi-hat. Protože není možné data zobrazit v prostoru všech 9 hlavních komponent jedná se o pouhý odhad.

K převodu testovacích dat do prostoru PCA vytvořeného na trénovací sadě dat, slouží knihovna *data2pca.m*. Vstupem jsou komponentní váhy, odhadované průměry trénovacích dat a nová testovací data. Nejprve jsou od testovacích dat odečteny odhadované průměry, čímž jsou testovací data centrovány podle dat trénovacích. Následně jsou testovací data násobena komponentními váhami, a tím převedena do vytrénovaného PCA prostoru. Výstupem jsou komponentní skóre testovacích dat, převedených do vytrénovaného PCA prostoru.

## 6.6 Metoda podpůrných vektorů

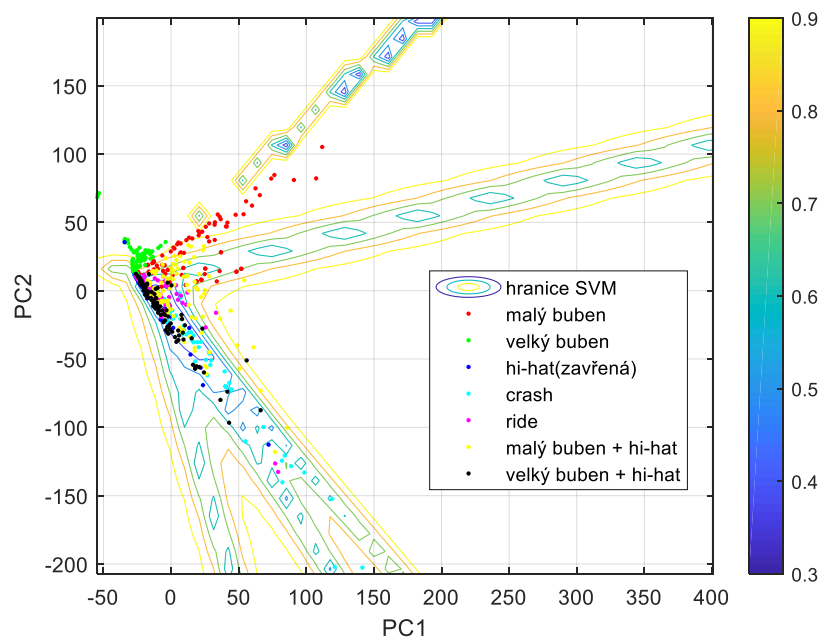
Pro výpočet podpůrných vektorů byla využita funkce prostředí Matlab *fitcecoc*. Jedná se o obecnou funkci pro výpočet různých klasifikátorů. Jedním z nich je i výpočet SVM pro více než dvě třídy. Vstupy této funkce jsou trénovací data a vektor štítků klasifikačních tříd, který označuje klasifikační třídy vstupních dat. Výstupem je model podpůrných vektorů připravený pro klasifikaci testovacích dat.

Na obrázku (obr. 6.10) jsou zobrazeny hranice klasifikačních tříd SVM. Barevnou škálou je zobrazena pravděpodobnost příslušnosti do dané třídy. Body zobrazují trénovací data klasifikačního modelu. Graf je vynesena v prostoru prvních dvou hlavních komponent.



Obr. 6.10: Hranice klasifikačních tříd

Na obrázku (obr. 6.11) je opět uveden výřez shluku. Z těchto obrázků se může zdát, že hraniční prostory klasifikačních tříd úplně neodpovídají rozložení trénovacích dat. To je způsobeno vynesemím v prostoru pouze dvou hlavních komponent. Takovéto zobrazení vysvětluje pouze 55 % rozptylu trénovacích dat před výpočtem PCA.



Obr. 6.11: Výřez hranic klasifikačních tříd

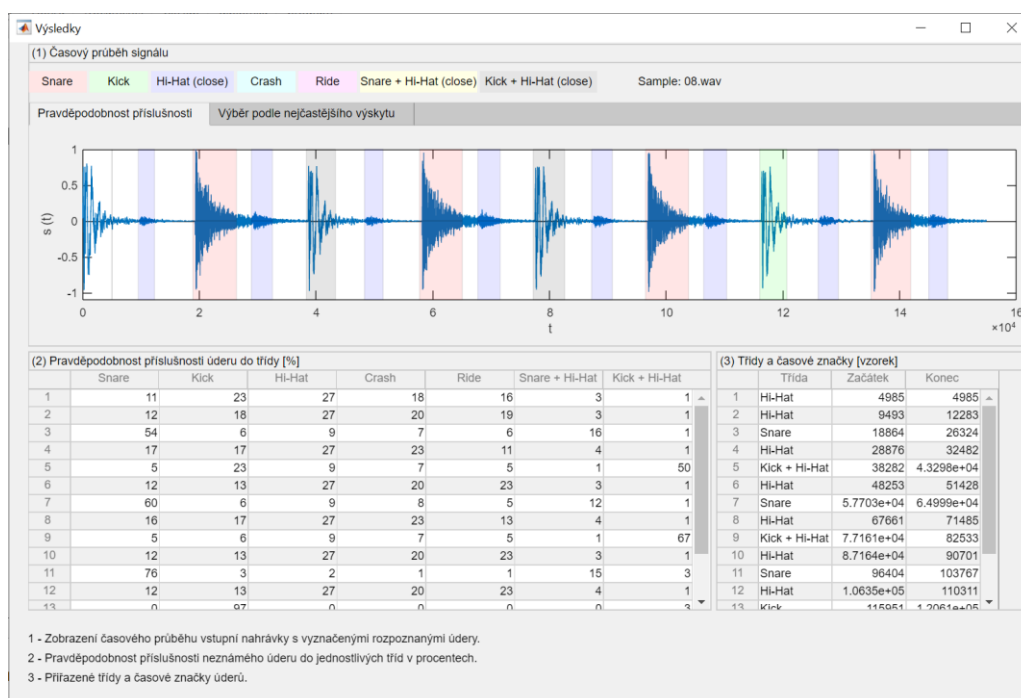
Pro klasifikaci testovacích dat podle vytrénovaného modelu, byla využita funkce Matlabu *predict*. Vstupy této funkce jsou model SVM a testovací data převedená do prostoru PCA. Výstupy jsou vektor s predikovanými štítky tříd a pravděpodobnostní matice příslušnosti dané třídy. Predikce klasifikační třídy je založena na minimalizaci relativní entropie (Kullbackova–Leiblerova divergence), která je definována například zde [21]. Podrobný popis algoritmu predikce je uveden zde [22].

## 6.7 Výstupy a výsledky programu

Výstupy trénovací části programu tvoří tři modely. Jsou jimi *trenovaci.mat*, *PCAModel.mat* a *SVMMModel.mat*. V trénovacím modelu jsou uloženy energie vzorků trénovací sady a informace o jejich počtech a kategoriích. Tento model lze využít k přetrénování systému bez nutnosti mít k dispozici trénovací sadu vzorků úderů na bicí soupravu. Model PCA slouží k převodu testovacích dat do vytrénovaného prostoru PCA. Model SVM slouží ke kategorizaci dat, jsou v něm uloženy hraniční prostory jednotlivých tříd.

Systém byl vytrénován na sadě 728 vzorků úderů, pro tyto kategorie: malý buben, velký buben, uzavřená hi-hat, činel crash, činel ride, dvojúder malého bubnu s uzavřenou hi-hat a dvojúder velkého bubnu s uzavřenou hi-hat. Vzorky tomů nebyly do trénovací sady zařazeny, protože volně dostupné banky úderů obsahují velmi málo těchto vzorků.

Na obrázku (obr. 6.12) se nachází zachycen grafický výstup z programu. Tento výstup je rozdělen do tří bloků. V prvním se nachází časový průběh testované nahrávky, ve kterém jsou barevně vyznačeny rozeznané údery. V druhém bloku se nachází tabulka s procentuální pravděpodobností příslušnosti neznámého úderu do jednotlivých tříd. Například rozeznáný úder číslo 3 má nejvyšší pravděpodobnost příslušnosti do třídy malého bubnu a to 54 %. Systém tedy tento úder vyhodnotil jako úder na malý buben.



Obr. 6.12: Grafický výstup programu

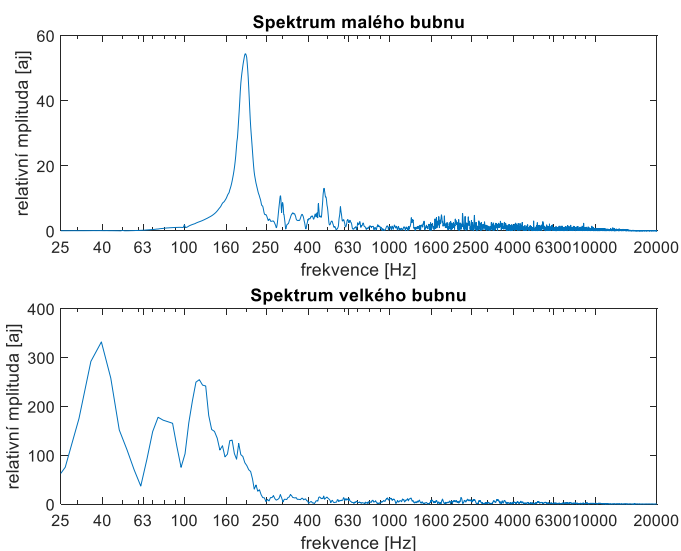
Posledním blokem je tabulka, ve které jsou uvedeny názvy tříd, kam byly jednotlivé úderu přiřazeny a časové značky začátků a konců těchto úderů.

Systém pro rozeznávání úderů na bicí soupravu, byl otestován na sadě 224 vzorků rozličných od trénovací sady. Výsledky tohoto testu jsou uvedeny v tabulce záměn (Confusion matrix) (tab. 6.3).

Tab. 6.3: Tabulka záměn rozpoznávacího systému

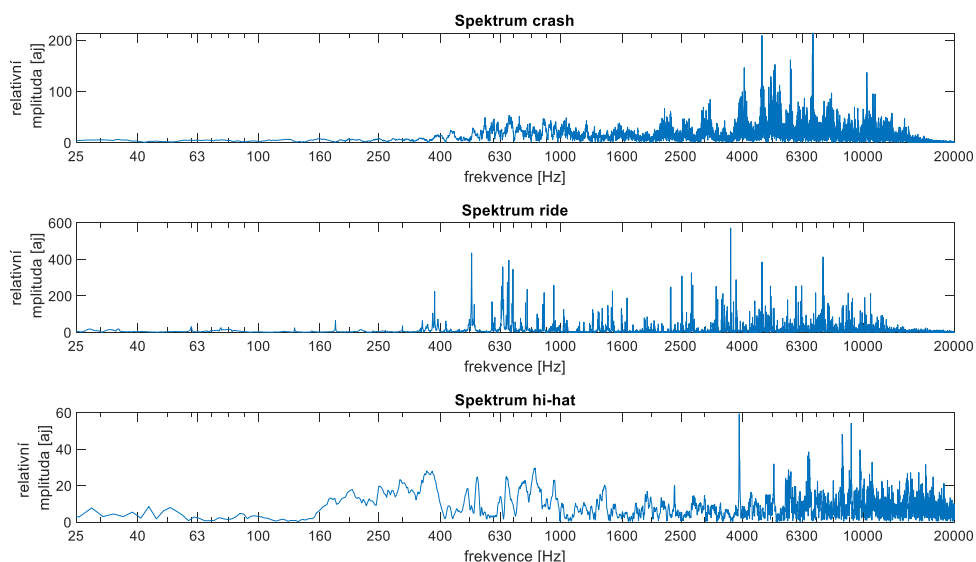
		predikce						
		malý buben	velký buben	hi-hat	crash	ride	malý b. + hi-hat	velký b. + hi-hat
reálné třídy	malý buben	75	0	0	0	0	19	6
	velký buben	0	94	0	0	0	0	6
	hi-hat	0	0	53	38	6	3	0
	crash	0	0	4	84	12	0	0
	ride	0	0	16	22	56	6	0
	malý b. + hi-hat	12	0	0	0	0	84	4
	velký b. + hi-hat	0	19	0	0	0	0	81

Kompletní výsledky tohoto testu jsou uvedeny v příloze (Příloha 1). Hodnoty jsou uváděny v procentech. Zelenou barvou jsou označeny správně vyhodnocené třídy. Žlutou barvou jsou označeny nesprávně vyhodnocené třídy, které jsou k dané třídě podobné. Červenou barvou jsou označeny zbylé nesprávně vyhodnocené třídy. Celková pravděpodobnost na úspěšnou klasifikaci úderu je 75 %. Z tabulky záměn vyplývá, že nejobtížněji klasifikovatelné třídy jsou hi-hat a ride. Obě tyto třídy jsou nejčastěji zaměňovány za crash.



Obr. 6.13: Spektra nahrávek úderu na malý a velký buben

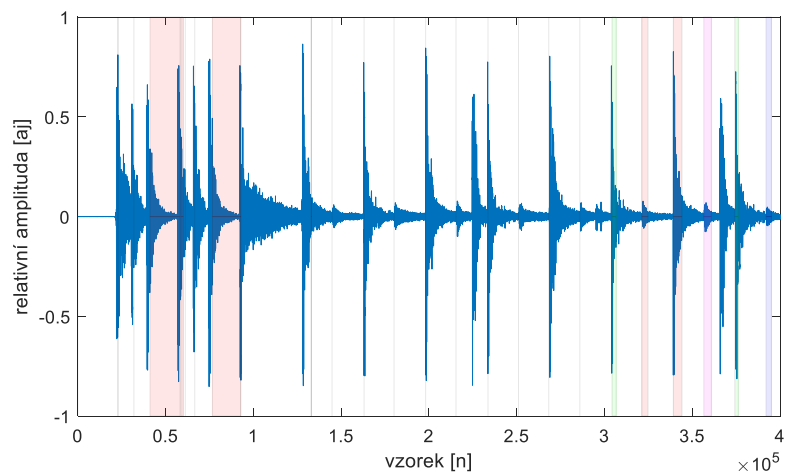
Na obrázku (obr. 6.13) jsou zobrazeny spektra nahrávek malého a velkého bubnu, na obrázku (obr. 6.14) se nachází spektra činelů.



Obr. 6.14: Spektra nahrávek úderu na crash, hi-hat (uzavřená) a ride

Na těchto obrázcích je vidět, že spektra činelů si jsou navzájem podobná, na rozdíl od spekter bubnů. Tato podobnost spekter vede k záměně některých nahrávek činelů.

Nejslabším článkem celého programu je detektor začátků a konců úderů. I přes normalizaci hlasitosti signálu dochází u některých nahrávek ke špatné detekci úderů. Špatná detekce je ukázána na časovém průběhu nahrávky sóla na bicí soupravu (obr. 6.15).



Obr. 6.15: Časový průběh nahrávky se špatnou detekcí úderů

Detekce začátků a konců úderů na základě prostého odhadu obálky se jeví jako nedostatečná. Další výzkum by se pro to měl zaměřit na preciznější detekci úderů, ať už optimalizací detekce zvlášť pro bubny a zvlášť pro činely nebo využitím některé z pokročilejších metod (například on-set detektor).

## 7. ZÁVĚR

V rámci této práce byl navržen a realizován analytický nástroj pro generování bicích triggerů z downmix záznamu bicí soupravy. Klasifikace jednotlivých úderů je založena na podobnosti signálů ve spektrální oblasti. Spektrální vlastnosti signálu jsou analyzovány pomocí banky filtrů. Na základě testů se nejvíce osvědčila třetino – oktávová banka filtrů. Signál je tedy rozdělen na 30 frekvenčních pásem. Pro jednotlivá frekvenční pásma jsou vypočteny energie signálu, které jsou vstupem pro následnou analýzu. Převod signálu na energie ve frekvenčních pásmech zásadním způsobem redukuje množství dat popisujících zvukovou nahrávku, a tím přispívá k rychlejšímu výpočtu dalších analýz.

Pro další odstranění redundance byla využita analýza hlavních komponent. Pomocí PCA byl snížen počet 30 dimenzí na 9, které obsahují 96,82 % rozptylu původních dat. Výpočet PCA také umožňuje v datech nalézt jejich charakteristické vlastnosti, na jejichž základě je možné data třídit.

Metoda pro klasifikaci dat byla vybrána na základě testu, kde byly porovnány tři klasifikační metody, a to metoda k – průměrů, hierarchické shlukování a metoda podpurných vektorů. Nejlepších výsledků bylo dosaženo za pomoci metody podpurných vektorů.

Rozpoznávací nástroj byl naprogramován v prostředí Matlab. Program se dělí na dvě základní části, a to trénovací a testovací. Trénovací část slouží k trénování klasifikačního modelu na základě trénovací množiny dat. Část testovací slouží k analýze nahrávky bicí soupravy. Za pomoci trénovací sady, která obsahovala 728 vzorků úderů byl vytrénován klasifikační model. Tento model rozlišuje sedm klasifikačních tříd odpovídajících úderům na různé části bicí soupravy (velký buben, malý buben, hi-hat, crash, ride, velký buben + hi-hat, malý buben + hi-hat). Vytrénovaný klasifikační model byl otestován 32 vzorky úderů pro každou klasifikační třídu. Z této testovací sady bylo správně rozpoznáno 75 % úderů.

Jako nejslabší část systému se ukázal detektor začátků a konců úderu, založený na sledování obálky signálu. U některých nahrávek nedokáže správně detekovat začátky a konce úderů, zejména u nahrávek s přeznívajícím činely. Je otázkou dalšího výzkumu, zdali půjde tuto metodu lépe optimalizovat nebo bude třeba vybrat vhodnější metodu.

Další vývoj navazující na tuto práci by se měl nejprve zabývat zlepšením funkce detektoru začátků a konců úderů. Například aplikací detektoru optimalizovaného zvlášť pro činely a zvlášť pro bubny. Následovalo by rozšíření klasifikačního modelu o další třídy. Rozšíření by se týkalo tomů, více kombinací dvojúderů, nebo rozlišení druhu úderu na malý buben. Klasifikační model by měl být otestován, jak obstojí bude-li testovaná nahrávka zarušena dalšími nástroji.

# Literatura

- [1] SYROVÝ, Václav. *Hudební akustika*. 3., dopl. vyd. V Praze: Akademie múzických umění, 2013. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-297-8.
- [2] MASTROIANY, Petr. *Výuka hry na bicí* [online]. [cit. 2020-05-19]. Dostupné z: <https://bici.mypage.cz/>
- [3] Drum kit illustration. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2020-02-05]. Dostupné z: [https://commons.wikimedia.org/wiki/File:Drum\\_kit\\_illustration\\_edit.png](https://commons.wikimedia.org/wiki/File:Drum_kit_illustration_edit.png)
- [4] JAN, Jiří. *Číslicová filtrace, analýza a restaurace signálů*. 2. upr. a rozš. vyd. Brno: VUTIUM, 2002. ISBN 80-214-1558-4.
- [5] NOVOTNÝ, Martin a Miloš SEDLÁČEK. *Měření efektivní hodnoty s využitím algoritmu DSP v prostředí MATLAB* [online]. In: . Praha: České vysoké učení technické v Praze, Fakulta elektrotechnická, katedra měření [cit. 2020-05-13]. Dostupné z: <https://www2.humusoft.cz/www/papers/tcp04/novotny.pdf>
- [6] VLACHÝ, Václav. *Praxe zvukové techniky*. 3., aktualiz. a dopl. vyd. Praha: Muzikus, c2008. ISBN 978-80-86253-46-6.
- [7] SMÉKAL, Z. (2009). *Číslicové zpracování signálů*. Vysoké Učení Technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 208.
- [8] SMÉKAL, Z. a SYSEL, P. *Číslicové filtry*. Elektronická skripta FEKT VUT v Brně. 2004, 130 s.
- [9] MATLAB Documentation. *MathWorks* [online]. The MathWorks, 2020 [cit. 2020-02-03]. Dostupné z: <https://www.mathworks.com/help/matlab/index.html>
- [10] SCHIMMEL, Jiří. *Akustika a zvukové systémy.: Učební text* [online]. Brno, 2018 [cit. 2020-02-03]. Dostupné z: [https://www.vutbr.cz/www\\_base/priloha.php?dpid=185282](https://www.vutbr.cz/www_base/priloha.php?dpid=185282)
- [11] SCHIMMEL, Jiří. *Elektroakustika*. Brno: Vysoké učení technické v Brně, 2013. ISBN 978-80-214-4716-5.
- [12] MELOUN, Milan a Jiří MILITKÝ. *Statistická analýza experimentálních dat*. Vyd. 2., upr. a rozš. Praha: Academia, 2004. ISBN 80-200-1254-0.
- [13] HOLČÍK, Jiří, KOMENDA, Martin (eds.) a kol. *Matematická biologie: e-learningová učebnice* [online]. 1. vydání. Brno: Masarykova univerzita, 2015. ISBN 978-80-210-8095-9.
- [14] MELOUN, Milan a Jiří MILITKÝ. *Kompendium statistického zpracování dat*. Praha: Karolinum, 2012. ISBN 978-80-246-2196-8.
- [15] HARUŠTIAKOVÁ, Danka. *Vícerozměrné statistické metody v biologii*. Brno: Akademické nakladatelství CERM, 2012. ISBN 978-80-7204-791-8.
- [16] HOLČÍK, Jiří. *Analýza a klasifikace dat*. Brno: Akademické nakladatelství CERM, 2012. ISBN 978-80-7204-793-2.



- [17] SMITH, Lindsay. *A tutorial on Principal Components Analysis* [online]. 2002, 27 [cit. 2020-01-31]. Dostupné z: <https://ourarchive.otago.ac.nz/bitstream/handle/10523/7534/OUCS-2002-12.pdf>
- [18] ŠEMBERA, Jan. *Aplikace analýzy hlavních komponent pro redukci dimenze transportne-reakčního problému* [online]. 13 [cit. 2020-01-31]. Dostupné z: [https://math.fce.vutbr.cz/~pribyl/workshop\\_2007/prispevky/Sembera.pdf](https://math.fce.vutbr.cz/~pribyl/workshop_2007/prispevky/Sembera.pdf)
- [19] ABE, Shigeo. *Support vector machines for pattern classification*. 2nd ed. London: Springer, 2010. ISBN 18-499-6097-6.
- [20] HRBOLKA, Daniel. *Test 10 nejlepších bezdrátových sluchátek 2020* [online]. [cit. 2020-04-10]. Dostupné z: <https://www.arecenze.cz/bezdratova-sluchatka/>
- [21] Kullback–Leibler divergence. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2020-05-29]. Dostupné z: [https://en.wikipedia.org/wiki/Kullback-Leibler\\_divergence](https://en.wikipedia.org/wiki/Kullback-Leibler_divergence)
- [22] MATLAB predict: Classify observations using multiclass error-correcting output codes (ECOC) model. *MathWorks* [online]. The MathWorks, 2020 [cit. 2020-05-29]. Dostupné z: <https://www.mathworks.com/help/stats/classificationecoc.predict.html>

## **Seznam příloh**

Příloha 1 - Tabulka úspěšnosti klasifikace s oktávovou bankou filtrů .....	59
Příloha 2 - Tabulka úspěšnosti klasifikace s třetino – oktávovou bankou filtrů .....	60
Příloha 3 - Tabulka úspěšnosti klasifikace s šestino – oktávovou bankou filtrů .....	61

## Příloha 1 - Tabulka úspěšnosti klasifikace s oktávovou bankou filtrů

malý buben	velký buben	hi-hat	crash	ride	malý + hi-hat	velký + hi-hat
malý	velký	hi-hat	crash	hi-hat	malý	velký+hh
malý	velký	crash	crash	hi-hat	malý+hh	velký+hh
malý	velký	crash	crash	crash	malý	velký
malý+hh	velký	hi-hat	crash	ride	malý+hh	velký+hh
velký+hh	velký	hi-hat	crash	ride	malý+hh	velký+hh
velký+hh	velký+hh	crash	crash	malý+hh	malý+hh	velký+hh
malý	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký+hh	ride	ride	crash	malý+hh	velký
malý	velký	malý+hh	crash	hi-hat	malý+hh	velký+hh
malý	velký	ride	crash	ride	malý	velký+hh
malý	velký	crash	crash	hi-hat	malý+hh	velký+hh
malý+hh	velký	crash	ride	ride	malý	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký+hh
malý+hh	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	malý+hh	malý+hh	velký
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý+hh	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký+hh
malý+hh	velký	hi-hat	crash	hi-hat	malý+hh	velký
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký
malý	velký	hi-hat	ride	ride	malý+hh	velký
malý+hh	velký	hi-hat	ride	ride	velký+hh	velký+hh
malý	velký	hi-hat	hi-hat	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
75 %	94 %	53 %	84 %	56 %	84 %	81 %

## Příloha 2 - Tabulka úspěšnosti klasifikace s třetino – oktávovou bankou filtrů

malý buben	velký buben	hi-hat	crash	ride	malý + hi-hat	velký + hi-hat
malý	velký	hi-hat	crash	hi-hat	malý+hh	velký
malý	velký	crash	crash	hi-hat	malý+hh	velký+hh
malý	velký	malý+hh	crash	crash	ride	velký
malý+hh	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	hi-hat	ride	malý+hh	velký+hh
malý	velký+hh	crash	crash	malý+hh	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký
malý	velký	ride	crash	crash	malý+hh	velký+hh
malý	velký+hh	malý+hh	ride	crash	malý	velký
malý	velký	malý+hh	crash	hi-hat	malý	velký
ride	velký	malý+hh	malý	hi-hat	malý+hh	velký
malý	velký	malý+hh	crash	ride	malý	velký+hh
malý	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký+hh
malý+hh	velký	crash	crash	hi-hat	malý+hh	velký+hh
malý	velký	malý+hh	malý+hh	crash	malý+hh	velký+hh
malý	velký	malý+hh	crash	malý+hh	malý+hh	velký
malý	velký	ride	crash	malý+hh	ride	velký+hh
malý	velký	hi-hat	crash	ride	malý	velký+hh
malý	velký	malý+hh	crash	ride	malý+hh	velký+hh
malý	velký	crash	malý	hi-hat	malý+hh	velký+hh
malý	velký	crash	crash	crash	ride	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký
malý	velký	malý+hh	crash	hi-hat	ride	velký
malý	velký	crash	crash	hi-hat	malý+hh	velký
malý	velký	hi-hat	malý	ride	malý+hh	velký
malý+hh	velký	hi-hat	crash	ride	malý+hh	velký
malý	velký+hh	hi-hat	ride	ride	malý	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
78 %	91 %	53 %	67 %	63 %	72 %	100 %

## Příloha 3 - Tabulka úspěšnosti klasifikace s šestino – oktávovou bankou filtrů

malý buben	velký buben	hi-hat	crash	ride	malý + hi-hat	velký + hi-hat
malý	velký	hi-hat	crash	ride	malý	velký+hh
malý	velký	crash	crash	hi-hat	malý+hh	velký+hh
malý	velký	crash	crash	crash	malý+hh	velký+hh
malý+hh	velký	hi-hat	ride	ride	malý+hh	velký+hh
velký+hh	velký	hi-hat	ride	ride	malý+hh	velký+hh
velký+hh	velký+hh	crash	crash	malý+hh	malý+hh	velký+hh
malý	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký+hh	ride	ride	crash	malý+hh	velký+hh
malý	velký	ride	crash	valký+hh	malý+hh	velký+hh
malý	velký	ride	crash	ride	malý+hh	velký+hh
malý	velký+hh	sn+hh	crash	hi-hat	malý	velký+hh
malý	velký	crash	ride	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý	velký+hh
malý+hh	velký	crash	ride	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	crash	malý+hh	velký+hh
malý	velký	hi-hat	ride	hi-hat	malý	velký+hh
malý	velký	hi-hat	crash	ride	malý	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	kick+hh	velký+hh
malý+hh	velký	crash	crash	crash	malý+hh	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký	ride	crash	crash	malý+hh	velký+hh
malý+hh	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	crash	ride	ride	malý	velký+hh
malý	velký	crash	crash	ride	malý+hh	velký+hh
malý	velký	kick+hh	ride	ride	malý+hh	velký+hh
malý+hh	velký	hi-hat	ride	ride	malý+hh	velký+hh
malý	velký	ride	hi-hat	ride	kick+hh	velký+hh
malý	velký	hi-hat	ride	ride	malý	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
malý	velký	hi-hat	crash	ride	malý+hh	velký+hh
78 %	91 %	53 %	67 %	63 %	72 %	100 %