



TECHNICKÁ UNIVERZITA V LIBERCI  
Fakulta mechatroniky, informatiky  
a mezioborových studií ■

# INTELIGENTNÍ OVLADAČ KAMERY NA MINIPOČÍTAČI RASPBERRY PI

## Bakalářská práce

*Studijní program:* B2646 – Informační technologie  
*Studijní obor:* 1802R007 – Informační technologie

*Autor práce:* **Vojtěch Starý**  
*Vedoucí práce:* doc. Ing. Zbyněk Koldovský, Ph.D.



## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Vojtěch Starý**  
Osobní číslo: **M11000121**  
Studijní program: **B2646 Informační technologie**  
Studijní obor: **Informační technologie**  
Název tématu: **Inteligentní ovladač kamery na minipočítači Raspberry Pi**  
Zadávající katedra: **Ústav informačních technologií a elektroniky**

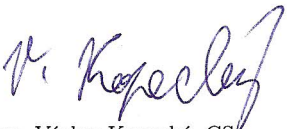
### Z á s a d y p r o v y p r a c o v á n í :

1. Sestavte zařízení s kamerou, natáčecím zařízením pro ovládání kameru, dvěma mikrofony a minipočítačem Raspberry Pi.
2. Vytvořte aplikaci pro Raspberry Pi, která bude kontinuálně sledovat zvuk ze dvou mikrofonů, vyhodnocovat směr příchodu akustického signálu a podle toho vyšle příkaz k natočení kamery směrem ke zdroji.
3. Vyřešte detaily jako je optimalizace způsobu vyhodnocování směru příchodu zvuku, úspora energie, inteligentní sledování zdroje (Kalmanův filtr) a vyřešte problémy kalibrace.
4. Zařízení ověřte v praxi.

Rozsah grafických prací: **Dle potřeby dokumentace**  
Rozsah pracovní zprávy: **cca 30 stran**  
Forma zpracování bakalářské práce: **tištěná/elektronická**  
Seznam odborné literatury:

- [1] I. Tashev, "Sound Capture and Processing" Practical Approaches", Wiley, 2008.
- [2] B. Porat, "A Course in Digital Signal Processing", John Wiley & Sons, 1997

Vedoucí bakalářské práce: **doc. Ing. Zbyněk Koldovský, Ph.D.**  
Ústav informačních technologií a elektroniky  
Konzultant bakalářské práce: **Ing. Jiří Málek, Ph.D.**  
Ústav informačních technologií a elektroniky  
Datum zadání bakalářské práce: **12. září 2014**  
Termín odevzdání bakalářské práce: **15. května 2015**

  
prof. Ing. Václav Kopecký, CSc.  
děkan



  
prof. Ing. Zdeněk Pliva, Ph.D.  
vedoucí ústavu

V Liberci dne 12. září 2014

## Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.


Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum: 7. 5. 2015

Podpis: 

## Poděkování

Mé poděkování patří vedoucímu této práce doc. Ing. Zbyňkovi Koldovskému, Ph.D. za odborné konzultace a cenné rady. Velké poděkování patří také mé rodině za trpělivost a podporu během studia a vypracovávání této bakalářské práce.

## Abstrakt

Tento dokument popisuje použité metody a realizaci funkčního prototypu inteligentního ovladače kamery s využitím počítače Raspberry Pi model B. Takové zařízení na základě analýzy zaznamenaných zvukových signálů rotuje kamerou ve směru zvukového zdroje. Lze jej použít při pořizování videozáznamu pohybujícího se mluvčího bez nutnosti obsluhy kamery člověkem. Úhel natočení servomotoru kamery je stanoven podle hodnoty fázového posunu signálů zaznamenaných dvěma mikrofony. Ke zjištění tohoto posunutí je použita metoda křížové korelace, na jejíž výsledky je aplikována kvadratická interpolace, jejíž maximální výsledek poskytuje přesnější hodnotu fázového posunu. Dosažení optimálního řízení servomotoru zajišťuje aplikace Kalmanova filtru na výsledné hodnoty fázového posunu. Celý program je implementován v jazyce Python verze 2.7. Součástí této práce jsou rozборы provedených experimentů, které prokazují krátkou časovou odezvu zařízení.

### **Klíčová slova:**

Raspberry Pi, stereofonní záznam zvuku, analýza zvukového signálu, sledování zdroje zvuku kamerou, zařízení pracující v reálném čase, Python

## Abstract

This document describes methods and an implementation of functional prototype of intelligent camera controller based on Raspberry Pi Model B. The device rotates camera in the direction of a sound source depending on the analysis of recorded sound signals. It can be used for taking a video of a moving speaker without the need of human operator. The angle of servo rotation is determined by phase shift of signals recorded with two microphones. The phase shift value is given by the cross-correlation method, on whose results the method of quadratic interpolation is applied. The maximum value provides more accurate value of the phase shift. The best result of servo controlling provides Kalman filter application on the solved phase shift values. The whole code is implemented in programming language Python 2.7. Part of this document describes experiments whose analysis proves a fast response of the device.

### **Keywords:**

Raspberry Pi, stereophonic sound recording, sound signal analysis, sound source camera tracking, real-time device, Python

## Obsah

Úvod.....	11
1 Současný stav problematiky.....	12
2 Teoretická příprava.....	13
2.1 Záznam zvukového signálu.....	13
2.1.1 Zvuk.....	13
2.1.2 Digitální signál.....	13
2.1.3 Senzory.....	14
2.1.4 Analogově digitální převodník.....	15
2.2 Předzpracování zaznamenaných dat.....	16
2.2.1 Výkon signálů.....	16
2.2.2 Butterworthův filtr.....	16
2.3 Analýza zvukového signálu.....	17
2.3.1 Křížová korelace.....	17
2.3.2 Kvadratická interpolace .....	18
2.4 Zpracování výsledků analýzy.....	18
2.4.1 Kalmanův filtr.....	19
2.5 Použitý hardware.....	20
2.5.1 Minipočítač Raspberry Pi model B 512 RAM.....	20
2.5.2 Tříbarevná dioda emitující světlo.....	21
2.5.3 Digitální teplotní čidlo DALLAS DS18B20.....	21
2.5.4 Servomotor TowerPro MG90.....	21
2.5.5 Dynamický mikrofon Behringer Ultravoice XM1800S.....	21
2.5.6 Předzesilovač M-Audio Audio Buddy.....	22
2.5.7 Zvuková karta Behringer UCA222.....	23
2.6 Použitý software.....	23
2.6.1 Operační systém Raspbian.....	23
2.6.2 Programovací jazyk Python 2.7.....	24
2.6.3 Advanced Linux Sound Architecture.....	24
2.6.4 Knihovna NumPy.....	24
2.6.5 Knihovna SciPy.....	25
2.6.6 ServoBlaster.....	25
3 Praktické zpracování.....	26
3.1 Konstrukce.....	26

3.1.1	Propojení komponent.....	27
3.1.2	Zapojení vstupně výstupních pinů Raspberry Pi.....	27
3.2	Algoritmus.....	28
3.2.1	Automatické spuštění.....	28
3.2.2	Hlavní program.....	29
3.2.3	Hlavní cyklus.....	30
4	Analýza experimentů.....	34
4.1	Kalibrace.....	35
4.2	Vliv vzdálenosti zdroje.....	36
4.3	Vliv prostředí.....	37
4.4	Časová náročnost.....	39
4.5	Plynulost změny směru.....	40
5	Závěr.....	41
	Seznam použité literatury.....	42
	Přílohy.....	44
A	Obsah přiloženého CD.....	44
B	Fotografie zařízení.....	45
C	Zobrazení zařízení v praktickém využití.....	46



## Seznam ilustrací

Graf 1: Obecné znázornění směrových charakteristik mikrofonů.....	14
Graf 2: Vliv řádu filtru na frekvenční charakteristiku filtru typu horní propust.....	17
Graf 3: Znázornění frekvenční charakteristiky mikrofonu XM1800S [10].....	22
Graf 4: Znázornění směrové charakteristiky mikrofonu XM1800S [10].....	22
Obrázek 5: Schéma zapojení modulů do vstupně výstupních pinů desky Raspberry Pi.....	27
Graf 6: Navržený filtr podle Butterworthovy aproximace .....	30
Graf 7: Znázornění vlivu nedůvěryhodnosti na Kalmanovo zesílení.....	32
Graf 8: Výše podílu nedůvěryhodnosti dle chyby měření podle úhlu příchodu zvuku .....	33
Graf 9: Výše podílu nedůvěryhodnosti dle dvou úhlů příchodu zvuku.....	33
Obrázek 10: Náčrt rozmístění prvků v prostředí během experimentování.....	34
Graf 11: Hodnoty výkonu signálů včetně středních hodnot a prahové hodnoty.....	35
Graf 12: Rozptyly hodnot směrů ovlivněné vzdáleností akustického zdroje .....	36
Graf 13: Rozptyly hodnot směrů ovlivněné charakteristikou prostředí.....	37
Graf 14: Vliv prostředí a vzdálenosti na směrodatnou odchylku naměřených hodnot.....	38
Graf 15: Hodnoty dob trvání procedur včetně jejich středních hodnot.....	39
Graf 16: Ukázka regulace naměřených hodnot implementovaným Kalmanovým filtrem....	40
Graf 17: Ukázka regulace hodnot Kalmanovým filtrem – případ střídajících se mluvčích...	40
Obrázek 18: Fotografie inteligentního ovladače kamery s připevněnou kamerou.....	45
Obrázek 19: Použití inteligentního ovladače k záznamu divadelního představení.....	46

## Seznam použitých zkratek

AD	analogově digitální
API	Application Programming Interface, aplikační rozhraní pro programování
DA	digitálně analogový
GND	Ground, uzemnění – záporný pól napájení
GUI	Graphical User Interface, grafické uživatelské rozhraní
HDMI	High-Definition Multimedia Interface, rozhraní pro přenos vysoce kvalitního audiovizuálního digitálního signálu
IDLE	Integrated Development Environment, integrované vývojové prostředí
LED	Light-Emitting Diode, světlo emitující dioda
MATLAB	Matrix Laboratory, programovací prostředí a skriptovací jazyk
PWM	Pulse Width Modulation, pulzně šířková modulace
RAM	Random-Access Memory, operační paměť s libovolným přístupem
RGB	Red Green Blue, aditivní způsob míchání barev za použití červené, zelené a modré barvy
SD	Secure Digital, typ nonvolatilní paměťové karty
SDHC	Secure Digital High Capacity, formát SD karty s vysokou kapacitou
SNR	Signal to Noise Ratio, poměr výkonu signálu a šumu
SPI	Serial Peripheral Interface, sériové periferní rozhraní pro komunikaci mezi mikroprocesory integrovaných obvodů
UART	Universal Asynchronous Receiver and Transmitter, univerzální asynchronní sériové rozhraní
USB	Universal Serial Bus, univerzální sériová sběrnice
XLR	typ konektoru audio kabelu

## Seznam použitých symbolů

$c$	rychlost zvuku [ $\text{ms}^{-1}$ ]
$f, f_s, f_{max}$	frekvence, vzorkovací frekvence, nejvyšší sledovaná frekvence [Hz]
$i, j, k$	proměnná nesoucí index prvku v poli
$K_k$	Kalmanovo zesílení
$N$	počet vzorků signálu
$P$	výkon signálu [dB]
$P_k$	aktuální odhad chyby
$P_{k-1}$	předchozí odhad chyby
$P_k^-$	predikovaný odhad chyby
$R$	nedůvěryhodnost
$s_{mic}$	vzdálenost mikrofonů [m]
$t$	teplota [ $^{\circ}\text{C}$ ]
$x, x_1, x_2$	datový typ pole, numerická podoba zaznamenaného signálu
$x_{2zeros}$	datový typ pole, data signálu doplněna o nulové signály zleva i zprava o délce původního pole bez jedné, celková délka tohoto pole je 3 délky bez 2
$x_k$	aktuální odhad měření
$x_{k-1}$	předchozí odhad měření
$x_k^-$	predikovaný odhad měření
$z_k$	aktuální měření
$\lambda$	řecké písmeno lambda, vlnová délka [m]
$\pi$	řecké písmeno pí, matematická konstanta, Ludolfovo číslo
$\Sigma$	řecké písmeno sigma, matematická operace součet řady
$\mathbb{Z}$	množina celých čísel
$\mathbb{R}$	množina reálných čísel
$\in$	prvek množiny
$\subset$	vlastní podmnožina

## Úvod

Tato práce se zabývá konstrukcí inteligentního ovladače kamery, který zaznamená zvuk dvěma mikrofony a implementací algoritmu v jazyce Python na minipočítači Raspberry Pi, který analyzuje signál. Výsledek analýzy bude využit k vypočítání směru příchodu zvukového signálu a následné obsluze servomotoru, jenž zajišťuje horizontální rotaci stativové hlavy s upevněnou videokamerou.

Takové zařízení najde využití například při videokonferenci či záznamech výkladu, kde se mluvčí během záznamu pohybuje v širokém prostoru před kamerou, a přesto je požadován detail na jeho osobu po celou dobu záznamu. Další možností uplatnění je pořízení sekundárního videozáznamu nezvučeného jevištního výkonu, ať už se jedná o mluvené slovo nebo hudební interpretaci. Sekundárního proto, že při sledování záznamu (například divadelní hry) snímá primární kamera celou scénu staticky – používá se tedy pro orientaci diváka v dění, zatímco druhá (sekundární) kamera vždy zabírá detail mluvčího. Pro zvýšení atraktivity výsledného videozáznamu je tedy vhodné využít prostřihů statických celků a dynamických detailů.

V teoretické části práce jsou shrnuty poznatky a metody potřebné k dosažení kvalitního výsledku lokalizace akustického zdroje. V rámci optimalizace fungování systému v reálném čase jsou zaznamenaná data předzpracovávána a rozhoduje se, zda obsahují hodnotné informace. Z důvodu chyb způsobených prostředím či ruchy je před obslužením servomotoru na výsledek analýzy aplikován Kalmanův filtr, jehož parametry jsou nastaveny tak, aby byly ojedinělé anomálie eliminovány.

Praktická část práce se zaměřuje na konstrukci celého zařízení, včetně přesného popisu použitých komponent a implementaci algoritmu v jazyce Python.

Závěrem jsou uvedeny výsledky experimentů v reálném prostředí s různou vzdáleností akustického zdroje od zařízení, včetně analýzy časové náročnosti cyklu hlavního algoritmu.

# 1 Současný stav problematiky

Cílem práce je realizace funkčního prototypu zařízení rotujícího kamerou ve směru příchodu zvukového signálu z akustického zdroje. Takové zařízení je využitelné v případě pořizování videozáznamu mluvčího či umělce v jinak tichém prostředí. Na rozdíl od existujících systémů pro ovládání kamery bez asistence člověka by tento inteligentní ovladač (kromě spuštění a prvotní kalibrace podle prostředí) neměl vyžadovat další instalaci. Požadavkem k jednoduchému používání je také určena vysoká míra integrace všech hardwarových součástí do jednoho celku tak, aby bylo zařízení mobilní.

Například obdobný systém, primárně vyvinutý ke sledování diskuze několika statických mluvčích, kombinuje použití funkce rozpoznávání tváří z obrazů více kamer a rozpoznávání aktuálního mluvčího podle zaznamenaných signálů z mikrofonů. Takové řešení vyžaduje vysoké množství techniky rozmístěné tak, aby vyhodnocení poskytovalo kvalitní informace. Při rozpoznávání tváře v obrazu může dojít k chybnému vyhodnocení pozice tváře mluvčího, a rotace kamery tedy nemusí být správná. Navíc každý mluvčí je vybaven mikrofonem zapojeným do systému, kde je podle hlasitosti získaných signálů jeden vyhodnocen jako aktuální mluvčí. Tento systém není příliš mobilní a je náročný na kalibraci při změně prostředí [1].

Dalším existujícím systémem pro sledování řečníků je radiová lokalizace. Takové zařízení má nízké požadavky na kalibraci a mobilitu, avšak hardwarové nároky zařízení jsou vysoké. Řečník musí být vybaven radiovým vysílačem, který vysílá signál. Tento signál je zachycen několika přijímači a výpočetní jednotka podle času příchodu vyhodnotí pozici vysílače – mluvčího. Vyhodnocování pozice radiolokací je robustní, avšak nelze ho využít při sledování více řečníků, kteří se střídají v projevu [2].

Inteligentní ovladač kamery na minipočítači Raspberry Pi bude vyhodnocovat pozici sledovaného objektu podle zvukových vln, a tak nebude použitelný při ozvučeném projevu či v hlučném prostředí. Na rozdíl od výše uvedených systémů však nebude chybovat vlivem vizuální podobnosti sledovaných objektů a nebude nutné vybavit mluvčího sledovacím zařízením.

## 2 Teoretická příprava

### 2.1 Záznam zvukového signálu

Účelem inteligentního ovladače kamery je rotace kamery směrem ke zvukovému zdroji. Ideální zdroj má bodovou charakteristiku. Z takového zdroje je šířen zvuk v kulových vlnách všemi směry. Toto vlnění je zachyceno senzory, vzájemně umístěnými v krátké vzdálenosti tak, aby bylo možné považovat příchozí vlnění za rovinné. Poté lze předpokládat, že směr příchodu vln je totožný na všech senzorech. Podle rozdílu fázového posunu jednotlivých záznamů je možné vypočítat úhel dopadu vlnění na senzory, a tím určit směr zdroje vlnění.

#### 2.1.1 Zvuk

Mechanické vlnění v intervalu frekvencí od 16 Hz až do 20 kHz je považováno za zvuk slyšitelný lidským uchem. Jelikož audiovizuální záznam pořizovaný kamerou, umístěnou na vyvíjeném zařízení, je určen především lidským divákům, není potřeba směřovat kameru za zdroji vlnění s frekvencí z oboru infrazvuku ani ultrazvuku.

Rychlost zvuku ve vzduchu je nezávislá na barometrickém tlaku, ale závisí na teplotě prostředí (viz rovnici (1)) [3]. Po změření aktuální teploty vzduchu lze spočítat rychlost šíření vlnění, kterou lze dosadit do výpočtu potřebné délky záznamu a výpočtu velikosti fázového posunu.

$$c = 331,8 + 0,61 t \quad (1)$$

Přesné zjištění odpovídající rychlosti zvuku je důležité pro analýzu zaznamenaných signálů. Tato hodnota určuje poměr při přepočtu vypočítaného fázového posunu na výsledný úhel natočení, jenž udává směr akustického zdroje.

#### 2.1.2 Digitální signál

Převod analogového signálu získaného ze sensorů, či předzesíleného analogového signálu umožňuje AD převodník. Diskrétní hodnoty spojitého signálu se získávají vzorkováním a kvantováním.

Vzorkovací frekvence určuje počet pořízených vzorků ze signálu za sekundu. Podle Shannon-Nyquistova teorému (viz rovnici (3) na str. 14) musí být vzorkovací frekvence větší než dvojnásobek maximální požadované frekvence [4]. Má-li nejvyšší kmitočet zvuku

lidského hlasu hodnotu 5 kHz, musí být hodnota vzorkovací frekvence pro snímání lidského hlasu více než 10 kHz.

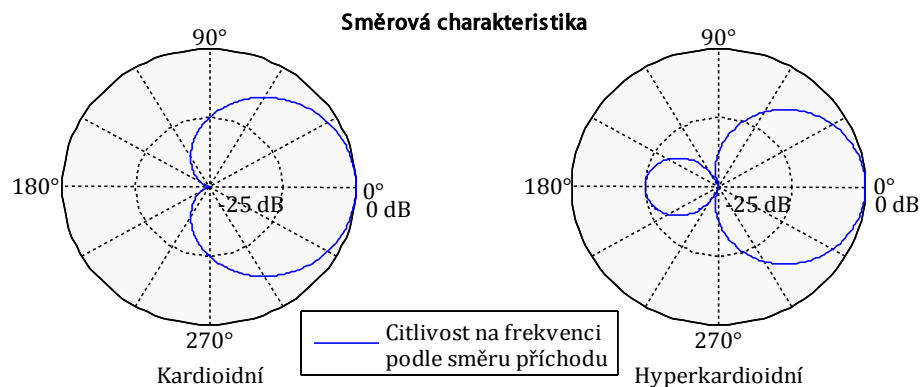
$$f_s > 2 f_{max} \quad (3)$$

Kvantování určuje bitovou hloubku jednoho vzorku, tedy kolika úrovní hodnot může jeden vzorek nabývat. Běžně je velikost jednoho vzorku rovna násobku 8 bitů (1 bajtu). Pro ukládání získaných diskrétních hodnot se volí datový typ o velikosti minimálně bitové hloubky jednoho vzorku. U datového typu se také rozhoduje o oboru funkčních hodnot (signed – unsigned).

### 2.1.3 Senzory

Pro záznam slyšitelného zvuku jsou vhodné dynamické mikrofony, nebo kondenzátorové mikrofony s vyšší citlivostí. Pro získání plného výkonu kondenzátorového mikrofonu je potřeba zapojení phantomového napáječe mezi mikrofon a zvukovou kartu. V případě nedostačujícího výkonu senzorů lze získaný signál zesílit nízkofrekvenčním zesilovačem k získání signálu, jehož kvantizace analogově digitálním převodníkem bude mít vyšší bitovou hloubku, než kvantizace signálu nezesíleného – zaznamenaná data obsahují více úrovní a jejich analýza poskytuje přesnější výsledky.

Volba senzorů pro snímání zvuku závisí na předpokládané vzdálenosti od zdroje a na kmitočtu předpokládaného signálu. Pokud se bude zdroj zvuku pohybovat na krátkou vzdálenost pouze před zařízením, kupříkladu 1 až 3 metry, je vhodná směrová charakteristika mikrofonů kardioidní (potlačuje příjem signálu zpoza senzorů). V případě velké vzdálenosti zdroje od senzorů je vhodnější hyperkardioidní charakteristika mikrofonů (více směrová než kardioidní). Frekvenční charakteristika mikrofonu udává míru zesílení jednotlivých kmitočtů v závislosti na vzdálenosti zdroje zvuku.



**Graf 1: Obecné znázornění směrových charakteristik mikrofonů**

Vzájemná vzdálenost senzorů je dána vlnovou délkou sledovaného vlnění. V případě sledování hlasu mluvčího je požadovaný rozsah frekvencí mezi 500 Hz a 5 kHz. Vlnová délka harmonického signálu závisí na rychlosti zvuku a kmitočtu (viz rovnici (2)) [3]. V mezních kmitočtech lidského hlasu je při teplotě 20 °C vlnová délka rovna 0,688 m a 0,0688 m.

$$\lambda = \frac{c}{f} \quad (2)$$

K zamezení snímání různých, ale podobných period signálu, je tedy nutné mít umístěné senzory maximálně na poloviční vzdálenost vlnové délky nejvyššího kmitočtu. Podle předpokladu, že snímaný signál nemá vlastnosti harmonického signálu a neobsahuje pouze maximální frekvenci, je možné umístit senzory i na vzdálenost delší, než je polovina vlnové délky nejvyšší zastoupené frekvence.

#### **2.1.4 Analogově digitální převodník**

Pro převod analogového signálu z mikrofonů do binární podoby v počítači slouží zvuková karta. Zvuková karta je zařízení složené z konektorů, AD a DA převodníků, které zajišťují konverzi vstupních a výstupních signálů, a rozhraní pro připojení k počítači (u externích dvoukanálových karet typu Plug&Play je to nejčastěji USB, u vícekanálových je to například rozhraní FireWire).

Výběr zvukové karty závisí na její frekvenční charakteristice, vzorkovací frekvenci, typu převodníku a výši SNR. Požadované hodnoty těchto vlastností se odvíjí podle požadovaného výstupu. AD převodník charakterizuje bitová hloubka, jelikož kvantizuje analogový signál. Frekvenční charakteristika definuje rozsah zaznamenaných kmitočtů. Zvukové karty mají přednastavené nejběžnější hodnoty vzorkovacích frekvencí, ze kterých lze vybrat vyhovující frekvenci. Neméně významným údajem je odstup signálu od šumu. Udává poměr výkonu signálu a výkonu šumu, uvádí se v decibelech.

Kvalitní záznam zvukového signálu velmi závisí na kvalitě zpracování konektorů a kabelů použitých k propojení jednotlivých komponent zařízení. Získání nezarušeného signálu zajistí čisté nezoxidované konektory z teplotně stálého materiálu, nejlépe pozlacené. Vzhledem k integraci zařízení musí být zvukové kabely dobře stíněné, jelikož pokud by byly nestíněné kabely vedeny v těsné blízkosti jiných kabelů (včetně napájecích), mohlo by dojít k rušení vedeného signálu vlivem elektromagnetického pole.



## 2.2 Předzpracování zaznamenaných dat

Analogový signál z mikrofonů může být předzpracován předzesilovačem ještě před zpracováním zvukovou kartou, jde však jen o zvýšení napětí signálu. Digitální signál získaný ze zvukové karty však může obsahovat šum způsobený prostředím, předzesilovačem či zvukovou kartou. Před použitím takového nezpracovaného digitálního signálu je vhodné provést analýzu vedoucí ke zjištění, zda jsou obsažená data vhodná k použití v dalších výpočtech. Nepotřebné zaznamenané hodnoty je možné eliminovat filtrací.

### 2.2.1 Výkon signálů

Na základě vypočítání výkonu signálu pro každý kanál lze rozhodnout, zda byl zaznamenaný zvuk dostatečně hlasitý pro další zpracování, anebo zda bylo zaznamenáno ticho – šum, který nenesou žádnou významnou informaci pro další výpočty. Takové rozhodování lze využít například při zaznamenávání projevu mluvčího, který se čas od času odmlčí. V takovém případě je analýza během pomlček nežádoucí.

Výkonu signálu je definován aritmetickým průměrem kvadrátů diskretních hodnot (viz rovnici (4)) [5]. Výsledek výpočtu tedy patří do oboru kladných reálných čísel včetně nuly. V rámci rozhodovacího algoritmu tedy lze určit prahovou hodnotu, podle které bude rozhodnuto, zda jsou data na základě výkonu užitečná či nikoli.

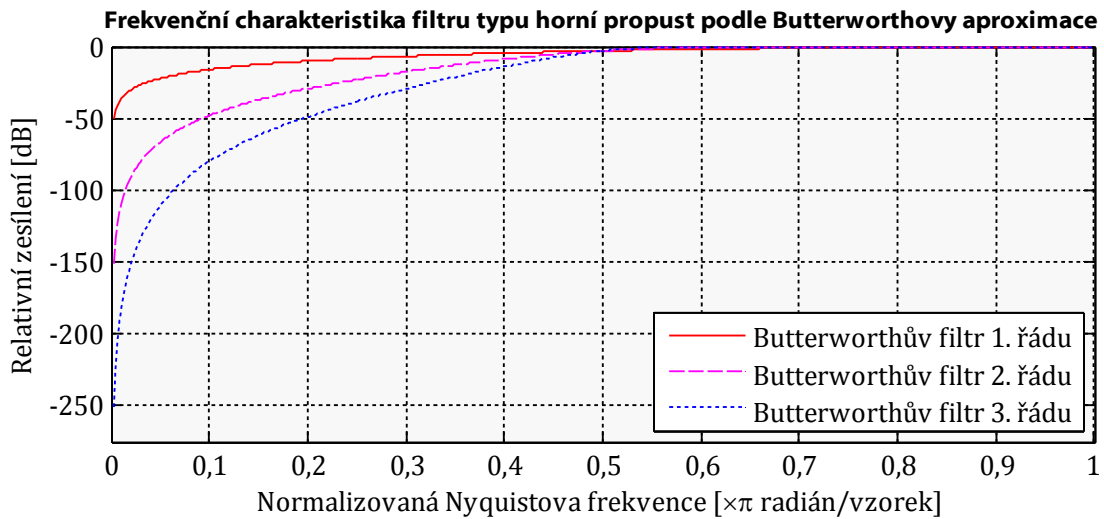
$$P = \frac{1}{N} \sum_{i=0}^N x[i]^2 \quad (4)$$

### 2.2.2 Butterworthův filtr

Eliminací data představujících hluk zaznamenaný mikrofony může dojít ke zlepšení výsledku další analýzy signálu. Filtr s co nejméně zvlněnou frekvenční charakteristikou a mezní frekvencí blíží se nule navrhl v roce 1930 britský inženýr Stephen Butterworth. Původně navrženou dolní propust lze modifikovat i pro horní propust, jež lze využít právě k filtraci brumů elektrické sítě a dalších nízkofrekvenčních hluků s minimálním zkreslením vyšších frekvencí [6].

Návrh filtru vyžaduje udání řádu filtru, jež určuje strmost relativního zesílení (frekvenční charakteristiky filtru), a mezní frekvence, která je udávána v poměru k Nyquistově frekvenci (polovina vzorkovací frekvence). Graf 2 na straně 17 znázorňuje vliv řádu filtru na frekvenční charakteristiky tří filtrů typu horní propust, s hodnotou mezní frekvence

rovné jedné polovině Nyquistovy frekvence (pro názornost), navržených podle Butterworthovy aproximace.



*Graf 2: Vliv řádu filtru na frekvenční charakteristiku filtru typu horní propust*

## 2.3 Analýza zvukového signálu

Cílem analýzy předzpracovaných dat je zjištění vzájemného fázového posunu zaznamenaných signálů. Počet vzorků, o který jsou signály posunuty, je dán stavem posunutí signálů při výpočtu křížové korelace, kde je hodnota křížové korelace maximální. Hodnota takto vypočteného posunutí náleží oboru celých čísel s intervalem, jenž je definován rychlostí zvuku, vzdáleností mikrofonů a vzorkovací frekvencí (viz rovnici (5)).

$$bandwidth = f_s \cdot \frac{S_{mic}}{c}$$

$$phaseShift \in \langle -bandwidth, bandwidth \rangle \subset \mathbb{R} \quad (5)$$

Protože fázový posun náleží oboru celých čísel, je počet výsledků omezený. Získání výsledku z oboru reálných čísel je možné interpolací. Proložení tří bodů je možné kvadratickou funkcí, jejíž parametry lze získat dosazením hodnoty bodu maxima korelace a hodnot jeho dvou sousedních bodů. Hodnota maxima této funkce pak náleží oboru reálných čísel a lépe odhaduje skutečnou hodnotu fázového posunu.

### 2.3.1 Křížová korelace

Křížová korelace poskytuje počet výsledků o jednu menší než je dvojnásobný počet vzorků jednoho signálu. Výsledky odpovídají sumě součinů odpovídajících vzájemně posunutých hodnot signálů [4]. Algoritmicky jsou data posunována doplněním jednoho ze

signálů nulami v první polovině výpočtů zleva, v druhé polovině výpočtů zprava (viz rovnici (6)). Součet součinů hodnot vzorků je maximální, jsou-li si data podobná. Proto je možné podle indexu maximální hodnoty výsledků korelace určit fázový posun. K získání relativního fázového posunutí je od výsledného posunutí nutné odečíst délku jednoho signálu (viz rovnici (7)). V případě vzorkovaných diskretních hodnot vzorkovaných z analogového signálu poskytuje lepší výsledek vyšší vzorkovací frekvence.

$$xcorr[i] = \sum_{j=0}^N x_1[j] \cdot x_{2\_zeros}[i+j]$$

$$i \in \langle 0, 2N-1 \rangle \subset \mathbb{Z} \quad (6)$$

$$phaseShift = \left( \underset{k}{argmax} (xcorr(x_1, x_2)[k]) \right) - count(x_1) \quad (7)$$

Pro kvalitní výsledek je nutné zvolit dostatečnou délku signálů. Minimální délka signálu musí pokrývat počet vzorků z intervalu, kterému náleží fázový posun.

### 2.3.2 Kvadratická interpolace

Kvadratická interpolace slouží k nalezení požadovaných hodnot proložením známých bodů kvadratickou funkcí – na rozdíl od lineární interpolace, kde jsou známé body prokládány prostou lineární funkcí. K určení neznámých parametrů kvadratické rovnice (viz rovnici (8)) je třeba vyřešit soustavu rovnic se třemi neznámými. K vyřešení slouží znalost tří bodů, kterými funkce prochází.

$$ax^2 + bx + c = 0 \quad (8)$$

V případě hledání bodu, ve kterém funkce nabývá maximální hodnoty, je nutné nalézt lokální extrémy. Derivace funkce rovna nule s dosazenými hodnotami parametrů řeší hodnotu bodu v maximu. Pro zjištění funkční hodnoty maxima stačí dosadit parametry a hodnotu vypočítanou z derivace.

### 2.4 Zpracování výsledků analýzy

Výsledkem analýzy signálů je reálné číslo vyjadřující fázový posun. Toto číslo náleží známému intervalu (viz rovnici (7)), avšak již v oboru reálných čísel. Díky znalosti krajních mezí tohoto intervalu je možné poměrem přepočítat výsledný fázový posun na hodnotu úhlu z rozsahu možných úhlů. Hodnota úhlu je potřebná pro vyslání příkazu k natočení servomotoru.

Před vysláním příkazu do motoru je však vhodné provést regulaci hodnoty pro případ, že došlo k nesprávnému odhadu směru zdroje či záznamu jiného zvukového zdroje, který je jiným směrem než zdroj aktuálně sledovaný. Regulaci měřených hodnot podle znalosti předchozí hodnoty včetně zohlednění důvěryhodnosti měření umí provést správně nastavený Kalmanův filtr.

### 2.4.1 Kalmanův filtr

Kalmanův filtr je nazván po vývojáři R. E. Kálmánovi. Pracuje na základě predikce (zohlednění posledního stavu) a korekce – odhadu hodnoty měření se známou chybou měření [4]. V případě neznámé chyby měření, jako například šumu prostředí při záznamu zvuku, lze použít zjednodušený model filtru, který ve stavu predikce nijak nemění hodnotu předchozího odhadu měření a odhadu předpokládané chyby (viz rovnici (9)). Zvolený zjednodušený model ve stavu korekce řeší hodnotu Kalmanova zesílení  $K_k$  na základě predikovaného odhadu chyby a zvolené nedůvěryhodnosti  $R$  – přednastavená hodnota (viz rovnici (10)). Kalmanovo zesílení je dále během stavu korekce využito k výpočtu odhadu aktuálního měření (viz rovnici (11)) a výpočtu aktuálního odhadu chyby (viz rovnici (12)). Hodnota Kalmanova  $K_k$  zesílení je závislá na velikosti nedůvěryhodnosti  $R$  – čím blíže nule, tím vyšší hodnota zesílení a tím nižší hodnota aktuálního odhadu chyby pro příští výpočet.

$$\begin{aligned}x_k^- &= x_{k-1} \\ P_k^- &= P_{k-1}\end{aligned}\tag{9}$$

$$\begin{aligned}K_k &= \frac{P_k^-}{P_k^- + R} \\ K_k &\in \langle 0, 1 \rangle \subset \mathbb{R} \\ R &\in \langle 0, +\infty \rangle \subset \mathbb{R}\end{aligned}\tag{10}$$

$$x_k = x_k^- + K_k(z_k - x_k^-)\tag{11}$$

$$P_k = P_k^- (1 - K_k)\tag{12}$$

Ke zlepšení výsledku činnosti zařízení a eliminace chyb je tento filtr zapojen mezi výpočet úhlu rotace a zpracování příkazu k rotaci servomotoru. Parametr  $R$  je nastavován dynamicky v takovém rozmezí, aby byla hodnota Kalmanova zesílení v intervalu od 0 do 1.

## 2.5 Použitý hardware

Prvky, ze kterých je celé zařízení sestaveno, byly zvoleny podle vhodnosti jejich charakteristiky vzhledem k charakteristice celku. Kvůli požadavku na mobilitu byl za výpočetní jednotku vybrán minipočítač Raspberry Pi, jenž disponuje mnohými užitečnými vlastnostmi. Nízkonapěťový servomotor MG90 lze zapojit přímo k základní desce počítače, bez nutnosti použití regulátoru či napájecího adaptéru. Díky podpoře Plug&Play je s Raspberry Pi možné použít širokou škálu kvalitních externích USB zvukových karet, Behringer UCA 222 disponuje dvěma jednobokanálními mikrofonními vstupy. Pro záznam zvuku byly podle frekvenční i směrové charakteristiky zvoleny dynamické mikrofony Behringer Ultravoice XM1800S.

### 2.5.1 Minipočítač Raspberry Pi model B 512 RAM

Počítač Raspberry Pi je integrován na jedné desce o velikosti kreditní karty. Původním záměrem vývojářů ze Spojeného království bylo vytvoření výukového prvku, který by studentům přiblížil vývoj výpočetních technologií. Z důvodu spolehlivosti, vysokého výkonu, variability a nízkých pořizovacích a provozních nákladů se stalo Raspberry Pi oblíbenou výpočetní jednotkou i v komerčním využití. Prvotní model A se tak dočkal mnoha revizí a modifikací, včetně vývoje bohatého příslušenství.

Model B 512 RAM je stejně jako starší modely vybaven výstupem kompozitního videa (pro připojení náhledového monitoru), HDMI výstupem (v případě potřeby zobrazování obrazu ve vysokém rozlišení) a audio výstupem. Zpracování obrazu zajišťuje grafický procesor VideoCore IV vyvinutý korporací Broadcom. Na rozdíl od modelu A a modelu B 256 RAM tento model disponuje síťovým adaptérem Ethernet s přenosovou rychlostí 10 Mbps a 100 Mbps, dvěma USB A porty verze 2.0 a operační pamětí RAM o velikosti 512 MB. Centrální výpočetní jednotkou tohoto minipočítače je 32 bitový procesor s označením ARM1176JZF-S s frekvencí 700 MHz, avšak lze přetaktovat až na 1 GHz. Při zvýšení frekvence procesoru je doporučeno opatřit čip alespoň pasivním chladičem. Pevný disk Raspberry Pi je zprostředkován vyjímatelnou SD kartou, která je jedinou nonvolatilní pamětí počítače. Pro fungování operačního systému je doporučený typ SDHC s třídou nejméně Class 6 (přenosové rychlosti se mohou lišit podle výrobce). Nedílnou součástí základní desky Raspberry Pi je řada vstupně výstupních konektorů (model B poskytuje 26 pinů) pracujících s napětím 3,3 V a 5 V. Konektory jsou rozděleny podle jejich speciálních vlastností, nastavení jejich módu lze změnit programovacím jazykem v operačním

systemu. Některé konektory mohou být použity pro přenos sběrnicí I<sup>2</sup>C, jiné SPI či UART, 4 piny poskytují napájení periferií, 5 pinů je vyčleněno k uzemnění [7].

Základní deska Raspberry Pi je připravena pro nasazení a provoz operačních systémů s Linuxovým jádrem. Operační systém musí být nainstalován na vložené paměťové kartě. Distribucí systémů vyvinutých na míru Raspberry Pi existuje mnoho, liší se podle způsobu využití počítače (multimediální centrum, osobní počítač, stanice pro sběr dat). Tyto systémy jsou často klony existujících linuxových systémů pro osobní počítače.

### **2.5.2 Tříbarevná dioda emitující světlo**

V rámci samostatného fungování zařízení bude stav procesu signalizován pouze svitem vícebarevné světlo emitující diody. Barva vyzařovaného světla je regulována 3,3–5V PWM signály zvlášť pro všechny tři barevné složky – červenou, zelenou a modrou.

### **2.5.3 Digitální teplotní čidlo DALLAS DS18B20**

Digitální čidlo k měření teploty od -55 °C do 125 °C vyvinula společnost Dallas Semiconductor. Při teplotách od -10 °C do 80 °C je přesnost měření 0,5 °C. Napájen musí být 3 až 5 volty. Datovým vodičem je již přímo přenášén digitální signál, a tak není nutné vést vodič na analogově digitální převodník ke konverzi [8].

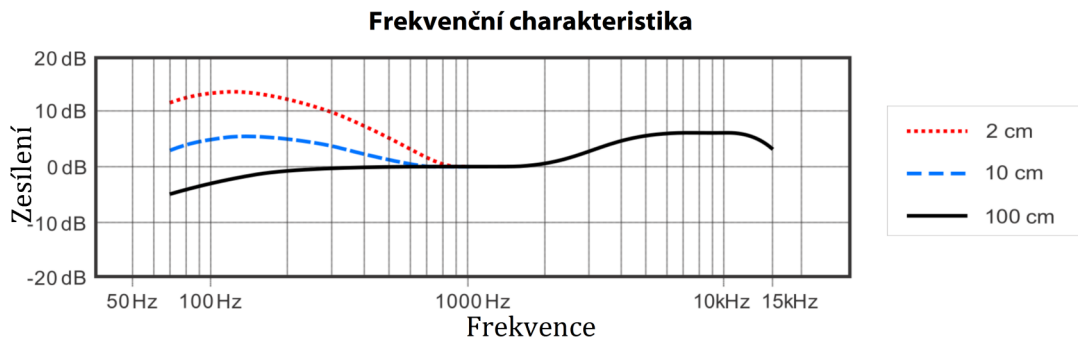
### **2.5.4 Servomotor TowerPro MG90**

Produkt MG90 od společnosti TowerPro je servomotor ovládaný PWM signály. Oproti modelu SG90 tento motor obsahuje kovové převody, které zajišťují delší životnost a vyšší sílu. Operační napětí je uváděno od 4,8 do 6 voltů, točivý moment pak odpovídá přibližně 17,5 až 21,5 Ncm a úhlová rychlost odpovídá 60 stupňů za 0,1 sekundy. K řízení servomotoru slouží tři vodiče – napájecí, uzemňující a řídicí. Řídicí vodič přenáší pulzně šířkově modulovaný signál, který délkou trvání pulzů udává směr natočení motoru [9].

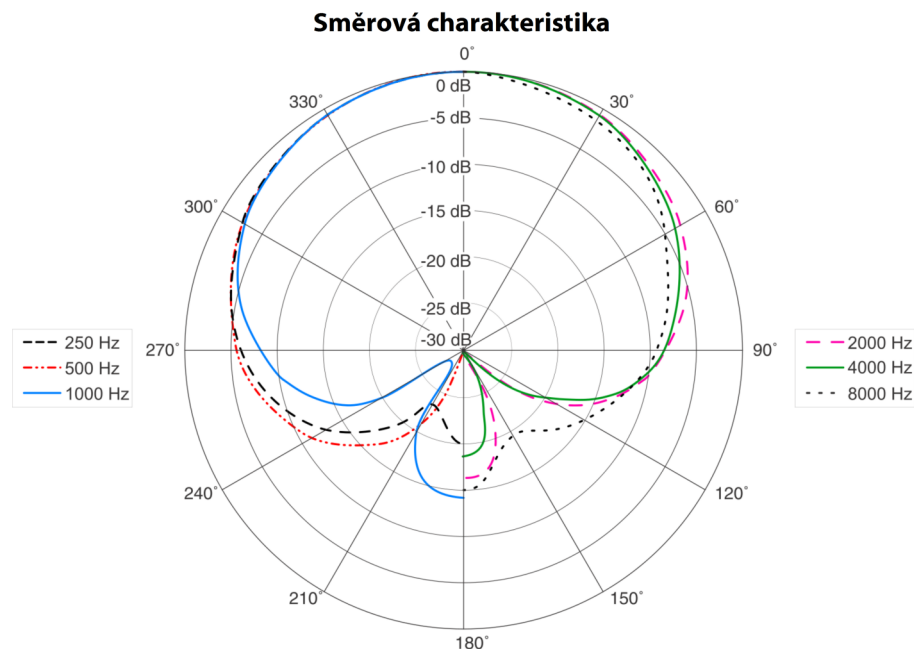
### **2.5.5 Dynamický mikrofon Behringer Ultravoice XM1800S**

Dynamický mikrofon XM1800S byl vyvinut především pro snímání lidského hlasu. Rozsah kmitočtů je uveden od 80 Hz do 15 kHz s tím že, vnitřní konstrukce mikrofonu mění jeho frekvenční charakteristiku podle vzdálenosti od zvukového zdroje (viz graf 3 na str. 22). Citlivost mikrofonu je 2,5 mV/Pa. Směrová charakteristika je superkardioidní, to znamená, že mikrofon přijímá zvuk v širším kuželu zředu (ne tak širokém jako kardioidní

a ne v tak úzkém jako hyperkardioidní) a částečně také zezadu (více než hyperkardioidní). Znázornění citlivosti vybraných frekvencí ve všech směrech je znázorněno v grafu 4 [10].



*Graf 3: Znázornění frekvenční charakteristiky mikrofonu XM1800S [10]*



*Graf 4: Znázornění směrové charakteristiky mikrofonu XM1800S [10]*

### 2.5.6 Předzesilovač M-Audio Audio Buddy

Předzesilovač Audio Buddy s nízkým šumem disponuje dvěma nezávislými vyváženými vstupy XLR a dvěma nevyváženými vstupy typu mono jack 6,3 mm pro zařízení s vysokou impedancí. Výstupní konektory na zadní straně předzesilovače jsou také typu mono jack 6,3 mm – jeden pro každý kanál. U zabudovaných nezávislých zesilovačů (pro každý kanál jeden), lze zesílení regulovat jednotlivě. Při použití XLR vstupů je možné zapnout vestavěný phantomový napáječ (napájí oba kanály). Frekvenční rozsah je uváděn od 5 Hz do 50 kHz. Maximální hodnota zesílení zadních XLR vstupů je 60 dB, u předních jack vstupů je rovna 40 dB [11].

## 2.5.7 Zvuková karta Behringer UCA222

Externí zvuková karta připojitelná přes USB verze 1.1 od výrobce audio techniky Behringer je typu Plug&Play a její použití nevyžaduje instalaci ovladačů. Mezi systémové požadavky patří minimálně 128 MB operační paměti RAM a frekvence procesoru nejméně 400 MHz. Tím je zaručena kompatibilita se systémem Raspbian na Raspberry Pi. Kromě rozhraní USB má karta UCA222 dva jednobanňové mikrofonní vstupy typu cinch, dva jednobanňové výstupy typu cinch a monitorovací stereo výstup typu jack 3,5 mm, na který je přiveden stejný signál jako na primární výstup zvukové karty.

Podle výrobce je vhodné na vstup přivádět signál z mixážního pultu, či předzesílený signál z hudebních nástrojů či mikrofónů. Tato zvuková karta je vybavena 16 bitovým AD převodníkem s použitelnými vzorkovacími frekvencemi 32 kHz, 44,1 kHz a 48 kHz. Frekvenční charakteristika karty doporučuje používat vzorkovací frekvenci 48 kHz pro záznam spektra od 10 Hz do 22 kHz. Uvedené SNR AD převodníku má hodnotu 89 dB [12].

## 2.6 Použitý software

K vytvoření funkčního prototypu zařízení je nutné zvolit vhodný operační systém pro Raspberry Pi, na kterém bude aplikace spuštěna. Nejpodporovanějším operačním systémem je doporučovaný Raspbian – upravená distribuce systému Debian. Raspbian umožňuje použití grafického uživatelského rozhraní a přímé ovládání vstupně výstupních pinů. Mimo jiné obsahuje integrované prostředí pro vývoj v programovacím jazyce Python. Jelikož bude aplikace získávat a zpracovávat data ze zvukové karty, byly vybrány rozšiřující moduly pro Python, které podporují práci se vstupními zařízeními, či implementují širokou škálu matematických operací. Protože Raspberry Pi Model B revize 2 podporuje PWM signály pouze na dvou výstupních konektorech, k ovládání více externích komponent pomocí pulzně šířkově modulovaných signálů musí být použit externí modul.

### 2.6.1 Operační systém Raspbian

Operační systém Raspbian byl vyvinut v roce 2012 jako volně dostupný systém s linuxovým jádrem. Je založen distribuci Debian Wheezy, optimalizován pro hardwarové možnosti minipočítače Raspberry Pi. Instalační obraz obsahuje velké množství instalačních balíčků programů a předinstalovaných aplikací. Mezi ně patří především vývojové prostředí pro Python IDLE, ovladač audio zařízení ALSA a ovladač vstupně výstupních konektorů pro jejich řízení [13].



Po prvním spuštění systému je zobrazena konzole, ze které je možné provést základní nastavení, jako například automatické spouštění grafického uživatelského rozhraní, výběr typu vysílaného grafického signálu, nastavení komunikačních vlastností vstupních výstupních pinů, nebo přetaktování procesoru.

Vzhledem k tomu, že základní deska Raspberry Pi disponuje dvěma USB porty verze 2.0, podporuje operační systém Raspbian zařízení typu Plug&Play a zařízení, která ke svému provozu nevyžadují speciální ovladače.

### **2.6.2 Programovací jazyk Python 2.7**

Programovací jazyk Python verze 2.7 je interpretovaný, objektově orientovaný jazyk. Plně podporuje používání externích modulů (včetně modulů kompilovaných z jazyků C a C++), výjimek, tříd a velké množství datových typů. Výhodou programování v Pythonu je dynamické přetypování objektů, díky čemuž není nutné objekty deklarovat [14].

Volba verze 2.7 je založena na možnosti použití dříve vyvíjených modulů, jelikož Python verze 3 není zpětně kompatibilní. Python 2.7 byl vybrán díky vysoké podpoře v operačním systému Raspbian a vysokému počtu dostupných a využitelných modulů.

### **2.6.3 Advanced Linux Sound Architecture**

Používaný zkrácený název ALSA (název modulu *alsaaudio*) reprezentuje balík ovladačů linuxových operačních systémů pro práci se zvukovými zařízeními. Kromě možnosti konfigurace zvukových vstupních a výstupních streamů z příkazové řádky či GUI obsahuje tento balík také API pro využití v programovacím jazyce Python [15]. Umožňuje vytvářet objekty reprezentující zvuková zařízení a následně, zaznamenávat či přehrávat zvuk. Požadované parametry objektů lze nastavit stručným zápisem kódu. Při záznamu jsou data ze streamu získávána v binární podobě jako textový řetězec, a je tedy nutné je před použitím konvertovat.

### **2.6.4 Knihovna NumPy**

Knihovna NumPy je součástí balíku SciPy. Umožňuje základní matematické operace nejen s čísly, ale také s vlastními objekty v podobě vícedimenzionálních polí – zavádí objektový datový typ *ndarray*. Díky této knihovně je zápis výpočtů v Pythonu velice podobný zápisu v programu MATLAB a jemu podobných. V podstatě obsahuje téměř stejné funkce jako SciPy, avšak výsledky funkcí se mohou lišit (např. vráceným datovým typem).

### **2.6.5 Knihovna SciPy**

Jak je uvedeno výše, knihovna SciPy je mateřskou knihovnou knihovny NumPy. Tato knihovna však obsahuje velké množství funkcí, které provádí složité matematické operace, například v oboru analýzy signálů – návrhy filtrů a filtrace. Je doporučováno pracovat se zavedeným datovým typem *ndarray*. Konverze na standardní pole či seznam je možná [16].

### **2.6.6 ServoBlaster**

ServoBlaster je softwarové rozhraní pro vstupně výstupní piny Raspberry Pi. Slouží k řízení až 8 výstupních konektorů (servomotorů) PWM signály. Pulzně šířkově modulované signály lze ale také využít k ovládní svitu RGB diody, kde lze na každou barevnou složku přivést jinak modulovaný signál. Po zavedení ovladačů do systému je možné šířku pulzů kontrolovat příkazem v terminálu zadáním čísla výstupního pinu a hodnoty pulzu v počtu kroků, procent či mikrosekund [17].

## 3 Praktické zpracování

Jelikož cílem této práce není jen vytvoření návrhu inteligentního ovladače kamery, ale především praktická realizace, jsou následující kapitoly věnovány konkrétní konstrukci, zapojení a implementaci programu celého zařízení.

### 3.1 Konstrukce

S ohledem na mobilitu zařízení jsou všechny použité komponenty umístěny uvnitř či vně plastového boxu. Celý box je uchycen na stativu, kterým je umožněna regulace výšky a náklonu. Změna pozice zařízení je vhodná zejména pro nastavení záběru kamery.

Kamera umístěná na zařízení má sledovat objekt, jenž se bude pohybovat před zařízením. Rozsah rotace kamery, a tedy rozsah sledování zvukového zdroje, činí 180°. Proto je ke snímání zvuku dostačující použití dvou mikrofónů směřujících před zařízení a na základě dvou zaznamenaných signálů lze určit směr zdroje vůči zařízení. Mikrofony jsou umístěny ve vzdálenosti 11 cm od sebe (6,5 cm od osy kamerové točny). Při teplotě 20 °C odpovídá vzdálenost mikrofónů přibližně vlnové délce frekvence 3127 Hz, tzn., že v případě harmonického zvukového signálu této a nižší frekvence nemůže na zaznamenaných signálech dojít k posunu o celou fázi, a tedy k nesprávnému vyhodnocení směru.

Zjištění aktuální teploty vzduchu, potřebné k výpočtu rychlosti zvuku a na ni závislých proměnných, provádí senzor DS18B20 umístěný na přední straně vně zařízení.

Osa otáčení stativové hlavy, ke které lze připevnit kameru, je umístěna ve středu mezi mikrofony na úrovni předních okrajů hlav mikrofónů. Tím je zajištěna shoda vyhodnoceného úhlu příchodu signálu a úhlu natočení kamery.

Z důvodu absence tlačítka, které by zajistilo bezpečné ukončení operačního systému a vypnutí Raspberry Pi, je k vstupně výstupním pinům připojeno externí tlačítko vyvedené na zadní straně zařízení. V hlavním programu stisk tlačítka způsobuje přerušování procedury a vyvolá příkaz k řádnému ukončení systému.

Kromě diagnostiky zařízení není nutné mít připojený monitor ke grafickým výstupům Raspberry Pi. K oznámení stavu je využita tříbarevná LED dioda, jež je napájena podle aktuálního stavu programu. Po načtení systému je rozsvícena bíle, po dokončení kompilace kódu programu je rozsvícena zeleně, během záznamu zvuku svítí červeně, během analýzy signálů je zhasnuta a při vyslání příkazu k vypnutí svítí modře.

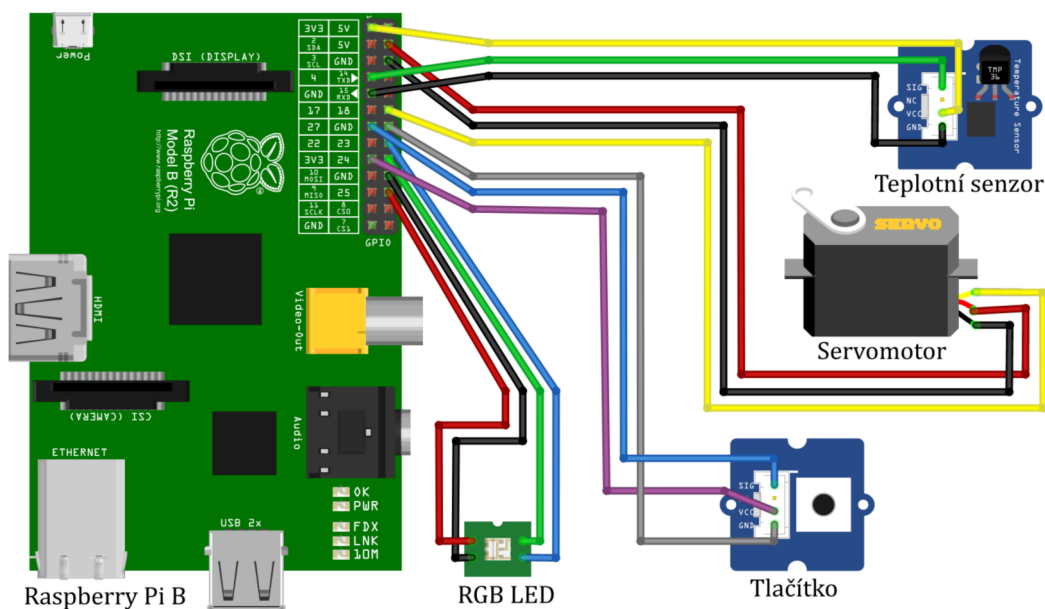
### 3.1.1 Propojení komponent

Raspberry Pi je napájeno z elektrické sítě micro USB adaptérem, který poskytuje stejnosměrný proud 1200 mA s napětím 5 V. Dle výrobce Model B vyžaduje proud mezi 700 mA a 1000 mA v závislosti na odběru zapojených externích komponent.

Do USB portu Raspberry Pi je připojena zvuková karta Behringer UCA222, která je tímto propojením také napájena. Do jejích vstupních konektorů typu cinch je redukcí připojen předzesilovač Audio Buddy 20cm kabely typu 6,3 mm mono jack. Mikrofonní XLR kabely dlouhé 50 cm propojují mikrofony Behringer XM1800S se zadními XLR vstupy předzesilovače.

### 3.1.2 Zapojení vstupně výstupních pinů Raspberry Pi

Externí moduly, které má Raspberry Pi kontrolovat či využívat se připojují na vstupně výstupní konektory vyvedené na desce. Podle charakteristiky modulu se napájecí kabel připojuje na výstup s napětím 3,3 V (senzory) nebo 5 V (motory). Uzemňující kabely se připojují k pinům s označením GND. Protože jsou některé konektory vyhrazené pro speciální použití, je vhodné k obsluze modulů využívat nevyhrazené (pro případ, že by bylo jejich speciálních vlastností potřeba například při komunikaci se složitějším modulem). Přesné schéma zapojení zobrazuje obrázek 5.



**Obrázek 5:** Schéma zapojení modulů do vstupně výstupních pinů desky Raspberry Pi

Komunikaci s teplotním senzorem zprostředkovává konektor číslo 4. Přerušení od tlačítka je čteno z pinu číslo 27. Modulovaný signál řídící servomotor je vysílán pinem

číslo 18. Taktéž modulované signály pro řízení barvy tříbarevné LED jsou vysílány z konektorů číslo 23 (modrá složka), 24 (zelená složka) a 25 (červená složka).

## 3.2 Algoritmus

Kromě skriptů pro automatické spuštění je celý program implementován v jazyce Python verze 2.7. Důvodem této volby je široká podpora Pythonu v operačním systému Raspbian, především existence modulů pro řízení vstupně výstupních pinů, získání signálu z USB zařízení a modulů pro matematické operace.

Pro lepší přehlednost a využitelnost kódu jsou funkce rozčleněny do 7 modulů podle jejich účelu. Hlavní modul *main.py* je řídicí – importuje dílčí moduly, provádí hlavní cyklus a obsahuje konstanty. Modul *tools\_package.py* obsahuje definice funkcí podpůrných výpočtů a definice funkcí potřebných k analýze dat. Dalším nezbytným modulem je *chunk\_recorder.py*. Ten vytváří instanci zvukového zařízení a poskytuje jak binární, tak konvertovanou podobu zaznamenaných dat. Zbývající moduly *temperature\_controller.py*, *servo\_controller.py* a *rgb\_led\_controller.py* definují funkce s výpočty a příkazy k obsluze hardwarových modulů připojených do vstupně výstupních pinů. Modul *turn\_off\_button.py* není importován do hlavního modulu, ale je spouštěn samostatně během automatického spuštění.

### 3.2.1 Automatické spuštění

Aby kromě spuštění celého zařízení nebyla nutná další obsluha, je hlavní program zaveden automaticky po spuštění operačního systému. K zajištění případné diagnostiky procesu s připojeným monitorem je nutné nespouštět program „na pozadí“, ale v konzoli, kde je viditelný výpis standardního výstupu programu.

K nastavení automatického provedení skriptu v konzoli po spuštění operačního systému slouží zápis příkazu v podobě `@lterminal --command [požadovaný příkaz]` do souboru `/etc/xdg/lxsession/LXDE/autostart`. V tomto případě je požadovaným příkazem spuštění souboru na adrese `/home/pi/BPcode/autostart.sh`, jenž obsahuje skript psaný v jazyce Bash.

Úkolem tohoto skriptu je zavedení ovladače ServoBlaster, který umožňuje vysílat PWM signály až na osm výstupních pinů. Dále následuje spuštění modulu *turn\_off\_button.py* „na pozadí“, jenž slouží k obsluze tlačítka vypnutí. Signál z tohoto

tlačítka je přiveden na konektor číslo 27, na kterém je očekávána vzestupná hrana. Je-li tato změna signálu zaznamenána, je odeslán příkaz k zahájení vypínací sekvence. Po aktivaci tlačítka vypnutí je rozsvícením RGB LED zeleně signalizováno následující zavedení souboru *main.py* s hlavním programem, čímž je procedura startovacího skriptu ukončena.

### 3.2.2 Hlavní program

Soubor *main.py* obsahuje hlavní program, ve kterém jsou definovány konstanty, inicializovány proměnné a volány funkce importovaných modulů. Z důvodu snížení počtu výpočetních operací uvnitř hlavního cyklu programu je většina potřebných dat předpočítána v rámci inicializace proměnných mimo cyklus.

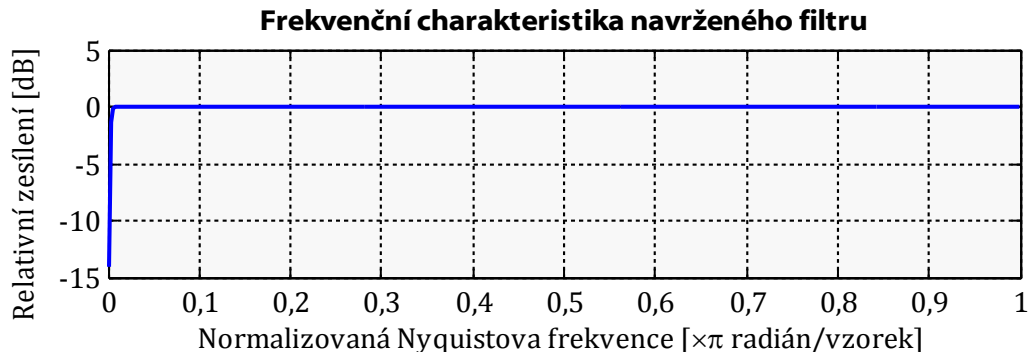
Po spuštění programu jsou importovány podpůrné moduly. Primárně je inicializován teplotní senzor zavedením ovladačů, jež umožňují přístup k datům senzoru podle jeho adresy. Na důkaz bezchybného spuštění hlavního modulu je vytvořeno vlákno *SERVO*, které zajistí vyslání příkazu k natočení servomotoru na úhel  $0^\circ$  (střední pozice).

Na řadu přichází definice konstant a inicializace proměnných, jejichž hodnoty jsou použity při záznamu a analýze signálů. V případě, že by selhal teplotní senzor, je výchozí hodnota rychlosti zvuku nastavena na hodnotu rychlosti zvuku při teplotě  $25^\circ\text{C}$ , tedy  $347.05\text{ ms}^{-1}$ . Vzdálenost mikrofونů, použita při výpočtu rozsahu fázového posunu, je dána konstrukcí zařízení na 11 cm. Vzorkovací frekvence zvukové karty je nastavena na 48 kHz z důvodu získání maximálního počtu zaznamenaných vzorků – při vzdálenosti mikrofونů 11 cm, teplotě  $25^\circ\text{C}$  a vzorkovací frekvenci 48 kHz je maximální hodnota posunu přibližně 15 vzorků. Při nastavení nejnižší možné vzorkovací frekvence použité karty (32 kHz) je nejvyšší možný rozdíl na signálech pouhých 10 vzorků. Hodnota prahu výkonu signálu, podle kterého je rozhodováno o provedení analýzy, je nastavena staticky na hodnotu 50 tisíc (viz kapitolu 4.1 Kalibrace na str. 35).

V hlavním cyklu budou filtrovány signály filtrem typu horní propust, který je navržen podle Butterworthovy aproximace. Parametry filtru jsou navrženy funkcí modulu SciPy podle zadaného řádu filtru a poměru mezní frekvence vůči Nyquistově frekvenci. Frekvenční charakteristika použitého filtru 3. řádu při mezní frekvenci 80 Hz vůči polovině vzorkovací frekvence je vidět na grafu 6 na straně 30.

Inicializaci objektu vstupního streamu dat umožňuje modul *alsaudio*. Mód přejímání dat z bufferu karty je nastaven na NORMAL – data jsou dostupná po blocích

o délce nastavené periody bufferu. Počet kanálů je roven dvěma, vzorkovací frekvence je 48 kHz a délka jedné periody bufferu je rovna 100 vzorků. Datový typ vstupních dat je nastaven na 16bitový znaménkovaný little-endian integer (nejméně významný bajt je uložen na adrese jako první).



*Graf 6: Navržený filtr podle Butterworthovy aproximace*

Hlavnímu cyklu předchází také definice časové značky, která slouží k řízení občasného čtení hodnoty teplotního senzoru, a definice proměnné nesoucí aktuální (případně výchozí) hodnotu rychlosti zvuku. Pro prvotní průběh cyklu jsou také definovány hodnoty Kalmanova filtru – výchozí hodnota natočení je rovna 0. K zapamatování posledního směru natočení je také definována vlastní proměnná.

### 3.2.3 Hlavní cyklus

Po provedení úvodní inicializace zařízení a předpočítání konstant je spuštěn nekonečný *while* cyklus, který lze přerušit pouze vypnutím počítače – přidaným tlačítkem. Tento cyklus obstarává celý proces od záznamu zvuku po rotaci servomotoru.

První operací cyklu je *if* podmínka, která porovnává rozdíl aktuální časové značky a uložené časové značky v proměnné *TEMPERATURE\_TIMESTAMP*. Je-li rozdíl časů větší než 5 minut, podmínka platí a bude proveden pokus o změření teploty. Komunikace s teplotním senzorem je navázána maximálně 5 krát (v případě neúspěšného čtení). Zjištěná teplota je použita k výpočtu rychlosti zvuku v proměnné *SPEED\_OF\_SOUND* a hodnota časové značky posledního čtení je aktualizována. Pokud žádný z pěti pokusů měření nebyl úspěšný, je použita výchozí hodnota rychlosti zvuku (rychlost při teplotě 25 °C) a další čtení je odloženo pouze o 30 sekund.

Po zjištění rychlosti zvuku je vypočítána hodnota maximálního možného fázového posunu, ze které je odvozena šířka pásma *BANDWIDTH*, podle kterého bude počítána křížová korelace a konverzní poměr fázového posunu a úhlu natočení.

Následně je zahájeno čekání na ukončení vlákna *SERVO*, aby záznam zvuku nebyl ovlivněn ruchem pohybu servomotoru. Poté je rozsvícena indikační RGB LED dioda červeně, což značí zahájení záznamu zvuku. Definovaná proměnná *DATA* bude obsahovat dvě pole dat – jedno pro každý kanál. Během nahrávání jsou z bufferu ukládána binární data do proměnné *stream\_data* v podobě řetězce, dokud není počet uložených vzorků větší nebo roven požadované délce dat. Binární data jsou konvertována na 16 bitový znaménkový integer a získaná data jsou rozdělena podle příslušnosti ke kanálům.

Podle hodnot v proměnných *LAST\_ANGLE* a *KALMAN\_ESTIMATED\_VALUE* je rozhodnuto, zda je servomotor natočen na poslední vypočítaný úhel. Pokud se hodnoty výše uvedených proměnných liší, znamená to, že motor není v aktuálně vyhodnoceném stavu a musí být vytvořeno nové vlákno *SERVO*, které zajistí pozvolnou rotaci servomotoru na požadovaný úhel. Do proměnné *LAST\_ANGLE* je také zkopírována hodnota proměnné *KALMAN\_ESTIMATED\_VALUE*, což v případě nedostatečného výkonu signálů (neprovádění další analýzy) zajistí nevytváření nového vlákna k ovládní servomotoru. K řízení PWM signálů ovládajících servomotor je použito vlastní vlákno z důvodu dosažení plynulého pohybu motoru. Podle rozdílu aktuální a požadované pozice je vyslán odpovídající počet příkazů k natočení o vypočítanou část rozdílu s definovanou časovou prodlevou.

Na zaznamenané hodnoty obou kanálů je aplikován výpočet výkonu signálu. Výkony levého i pravého kanálu jsou sečteny uloženy do proměnné *POWER*. Tato hodnota je porovnána s prahovou hodnotou uloženou ve statické proměnné *POWER\_THRESHOLD*. Je-li výkon signálů vyšší nebo roven než prahová hodnota, je splněna podmínka a analýza dat bude pokračovat. Pokud podmínka výkonu splněna není, hlavní cyklus je zahájen znovu.

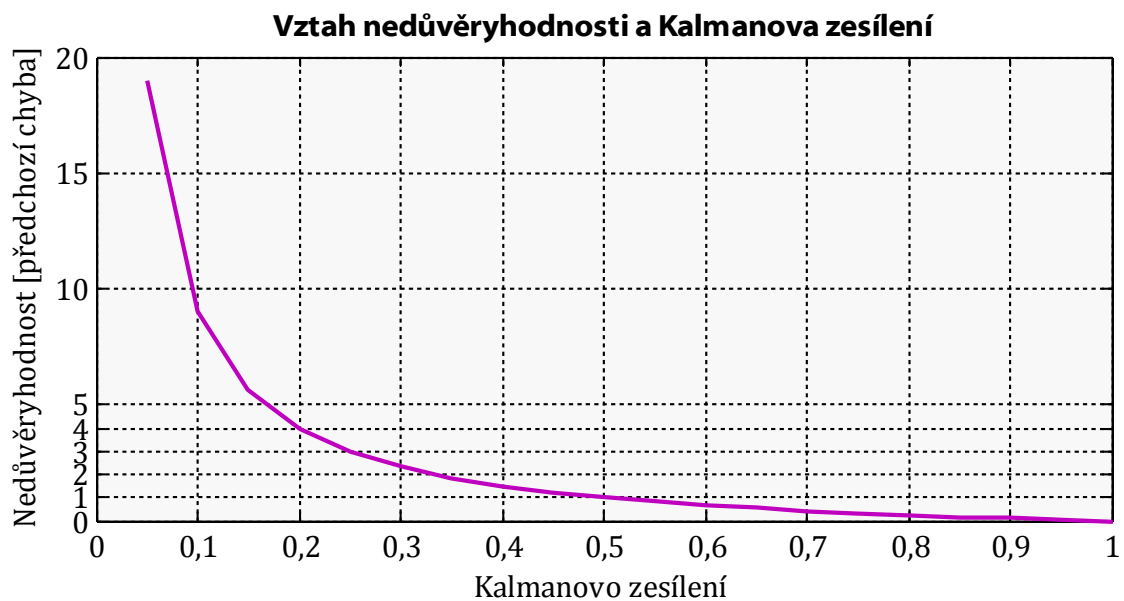
Pokud mají zaznamenané signály dostatečný výkon, na důkaz pokračování analýzy je zhasnuta indikační LED. Jelikož by mohl ruch způsobený rotací servomotoru rušit záznam zvuku, je příkaz k rotaci vyslán až v této fázi, kdy je jisté, že je do dalšího nahrávání dostatek času k dokončení rotace a vypnutí napájení motoru – pravděpodobně z konstrukčních důvodů motor pod napětím v různých pozicích bručí. Funkce ovládající motor přijímá hodnotu parametru jako úhel ve stupních, následně ji převede na údaj v procentech, jenž určuje, do jaké pozice má být motor natočen. Převod těchto hodnot byl odvozen z manuálního nastavování pozic. Jako parametr příkazu k přepočtu a natočení je tedy vložena hodnota proměnné *KALMAN\_ESTIMATED\_VALUE*, jež obsahuje výslednou hodnotu poslední provedené analýzy. Z toho vyplývá, že implementovaný Kalmanův filtr pracuje s hodnotami určujícími výsledné úhly.



Data jsou filtrována navrženým Butterworthovým filtrem typu horní propust. Tyto filtrované signály jsou uloženy do proměnné *FILTERED* se stejnou strukturou jako proměnná *DATA*.

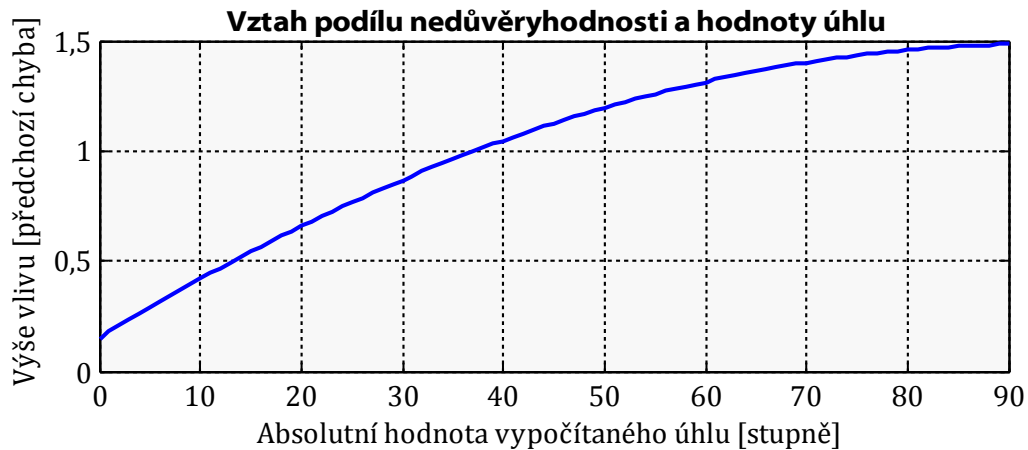
Filtrované signály jsou použity k výpočtu křížové korelace s posunutími v intervalu proměnné *BANDWIDTH*. Výpočet probíhá ve *for* cyklu, kde jsou data signálů doplňována nulami a korelována. Výsledky korelace při těchto posunutích jsou ukládány do vlastního pole, ve kterém je později nalezen nejvyšší výsledek a uložena hodnota jeho indexu. Tento maximální výsledek a jeho sousední prvky pole jsou následně použity k výpočtu parametrů kvadratické funkce a k dosazení do první derivace této funkce pro získání hodnoty, kde je funkční hodnota této křivky maximální. Tato reálná hodnota je uložena do proměnné *PEAK\_INTERPOL* a je považována za hodnotu fázového posunu signálů.

Hodnota fázového posunu je pomocí známé šířky pásma *BANDWIDTH* přepočítána do intervalu od -90 do +90 stupňů a spolu s posledními výsledky Kalmanova filtru *KALMAN\_ESTIMATED\_VALUE* a *KALMAN\_ESTIMATED\_ERROR* jsou odeslány jako parametry funkce, jež vrací nové hodnoty po aplikaci Kalmanova filtru. Hodnoty vypočítané Kalmanovým filtrem naposledy jsou použity ve stavu predikce jako odhadovaná hodnota a odhadovaná chyba. Spolu s naměřenou hodnotou jsou ve stavu korekce použity k výpočtu nedůvěryhodnosti pro určení výše Kalmanova zesílení. Vliv nedůvěryhodnosti na výši Kalmanova zesílení lze vidět na grafu 7.

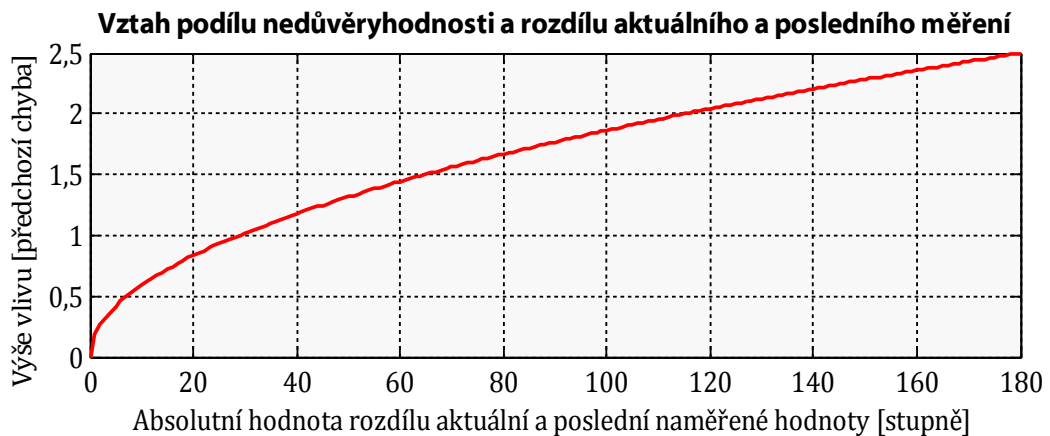


**Graf 7:** Znárodnění vlivu nedůvěryhodnosti na Kalmanovo zesílení

Nedůvěryhodnost je vypočítána jako součet podílu odhadované chyby vypočítaného úhlu (předpočítané hodnoty podle provedených experimentů - viz graf 8) a podílu vlivu rozdílu aktuálního a posledního úhlu (viz graf 9). Hodnoty podílů se pohybují v intervalech 0 do 1,5 a od 0 do 2,5 násobku předchozí chyby (odhad předpokládané chyby z proměnné  $KALMAN\_ESTIMATED\_ERROR$ ) (viz rovnici (13)). Nedůvěryhodnost tak nabývá reálné hodnoty od 0 až do 4 násobku hodnoty odhadované chyby, což vymezuje hodnotu Kalmanova zesílení do intervalu od 0,2 do 1 včetně (viz graf 7 na str. 32).



**Graf 8:** Výše podílu nedůvěryhodnosti dle chyby měření podle úhlu příchodu zvuku



**Graf 9:** Výše podílu nedůvěryhodnosti dle dvou úhlů příchodu zvuku

$$R = P_k^- \cdot \left\{ 1,5 \cdot \left( \text{variance}[\text{floor}(|z_k|)] \right) + 2,5 \cdot \left( \sqrt{\frac{|x_k^- - z_k|}{180}} \right) \right\} \quad (13)$$

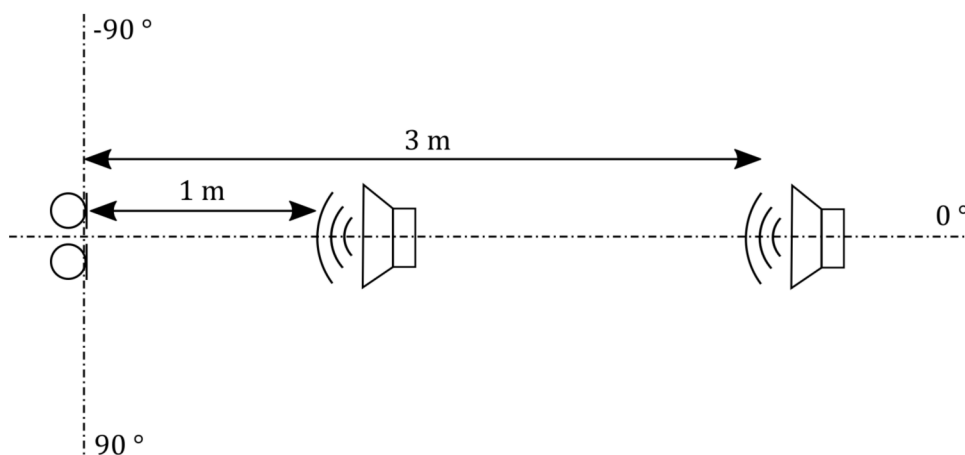
Získáním hodnot vypočítaných pomocí Kalmanova filtru je analýza signálů dokončena a provedení jedné procedury hlavního cyklu je ukončeno.

## 4 Analýza experimentů

Nedílnou součástí vývoje každého zařízení je ověření správnosti fungování. K tomu slouží provedení rozborů uskutečněných experimentů. Aby bylo možné porovnávat výsledky jednotlivých rozborů, musí být všechny pokusy prováděny za použití neměnného akustického zdroje.

Při provádění experimentů uvedených níže byl jako neměnný akustický zdroj použit mobilní telefon vybavený monofonním reproduktorem o průměru 1 cm, jenž opakovaně reprodukoval záznam řeči a zpěvu s hudebním doprovodem, celý o délce 3 minut a 38 sekund. Nastavení hlasitosti přehrávače bylo konstantní, avšak hlasitost záznamu po celou dobu trvání stálá nebyla. Snížení vlivu měnícího se charakteru přehrávaného záznamu na výsledky analýz bylo dosaženo opakováním záznamu ve smyčce po nezbytnou dobu.

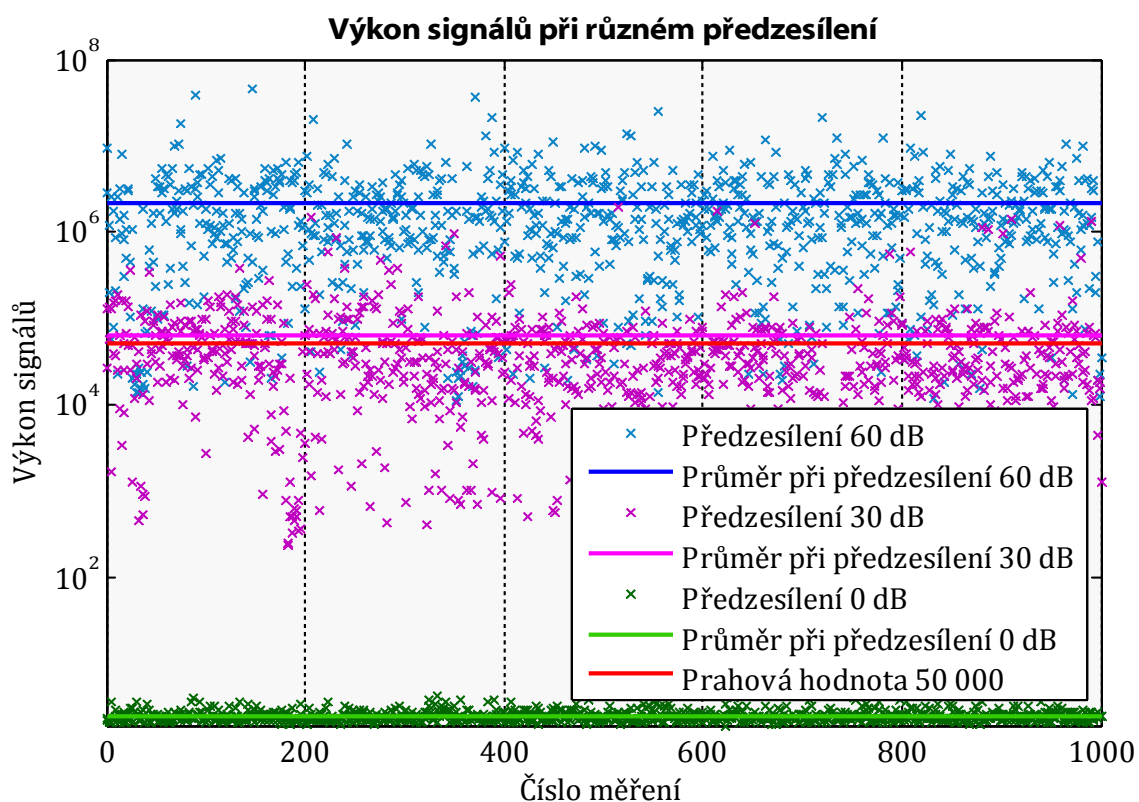
Pokusy probíhaly v místnosti o rozměrech přibližně 2,7 metru na výšku, 2,6 metru na šířku a 4,6 metru na délku. Reprodukční byl umístěn před zařízením, orientován na mikrofony. Vzdálenost reproduktoru od inteligentního ovladače kamery byla různá (jak je uvedeno u jednotlivých pokusů), ale vždy ležel na ose zařízení – tedy v úhlu 0 stupňů, pokud není uvedeno jinak (viz obrázek 10).



**Obrázek 10:** Náčrt rozmístění prvků v prostředí během experimentování

## 4.1 Kalibrace

Jelikož se jedná o zařízení pracující na základě zaznamenaného zvuku, je nutné jej kalibrovat při změně hlasitosti šumu okolního prostředí či při změně hlasitosti zvuku sledovaného zdroje. Ke zjištění, zda zaznamenané signály nesou hodnotné informace, je vypočítáván výkon signálů a podle porovnání s předdefinovanou prahovou hodnotou algoritmus rozhoduje, zda pokračovat v analýze dat. Ona prahová hodnota byla nastavena podle provedených experimentů na hodnotu 50 000, což (jak lze vidět na grafu 11) je hodnota blízká průměrnému výkonu signálu při předzesílení 30 dB. Výkon signálů k překročení této hranice lze regulovat manuální změnou předzesílení od 0 dB až po 60 dB. Jak určit správnou hodnotu předzesílení je popsáno v následujícím odstavci.



**Graf 11:** Hodnoty výkonu signálů včetně středních hodnot a prahové hodnoty

Po spuštění inteligentního ovladače se nastaví otočné voliče předzesilovače na počáteční nulovou pozici. Pokud zařízení reaguje na ruch prostředí, je nutné tento hluk snížit. V ideálním prostředí při nulovém zesílení však není tento ruch dostatečně hlasitý. Pokud tedy zařízení nereaguje na ruchy prostředí, postupným zvyšováním předzesílení je dosaženo hodnoty, kdy na ruchy prostředí zařízení reaguje. V tu chvíli je nutné předzesí-

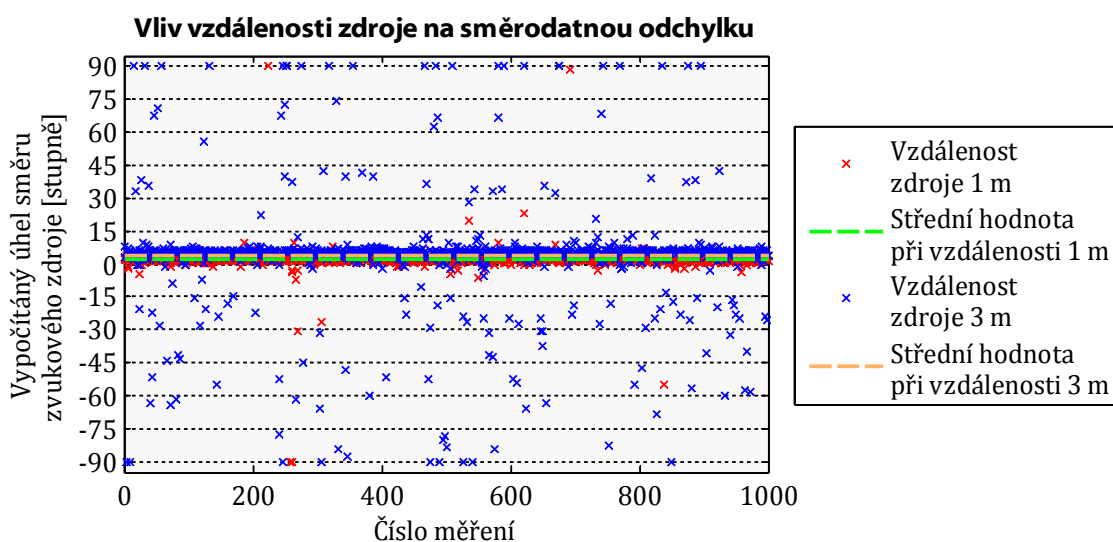
lení naopak snížit o několik decibelů a následně otestovat, zda zařízení nereaguje na ruchy prostředí, ale reaguje na zvuk vydávaný sledovaným zdrojem.

Optimální hodnota předzesílení se pohybuje mezi 20 a 40 decibely. Při nižší hodnotě musí být zvukový zdroj hlasitý a velice blízko mikrofonům. Při vyšší hodnotě reaguje zařízení i na ruchy prostředí.

## 4.2 Vliv vzdálenosti zdroje

Na správnost vyhodnocení směru akustického zdroje má vliv také vzdálenost zdroje od inteligentního ovladače. Čím je zdroj zvuku od mikrofonů dále, tím je nižší hlasitost zaznamenaných signálů. Při kalibraci zařízení je tedy také nutné zohlednit vzdálenost zdroje – s rostoucí vzdáleností zvyšovat hodnotu předzesílení.

Na grafu 12 lze vidět, že se zvyšující se vzdáleností roste rozptýl dat vyhodnocených z naměřených signálů. Analýza pokusů při vzdálenostech 1 a 3 metry prokázala, že směrodatná odchylka byla více než třináásobně vyšší při trojnásobné vzdálenosti. Konkrétně při vzdálenosti 1 metr byla vypočítána směrodatná odchylka přibližně 6,3 a při vzdálenosti 3 metry byla hodnota směrodatné odchylky přibližně 22,7.



*Graf 12: Rozptýly hodnot směrů ovlivněné vzdáleností akustického zdroje*

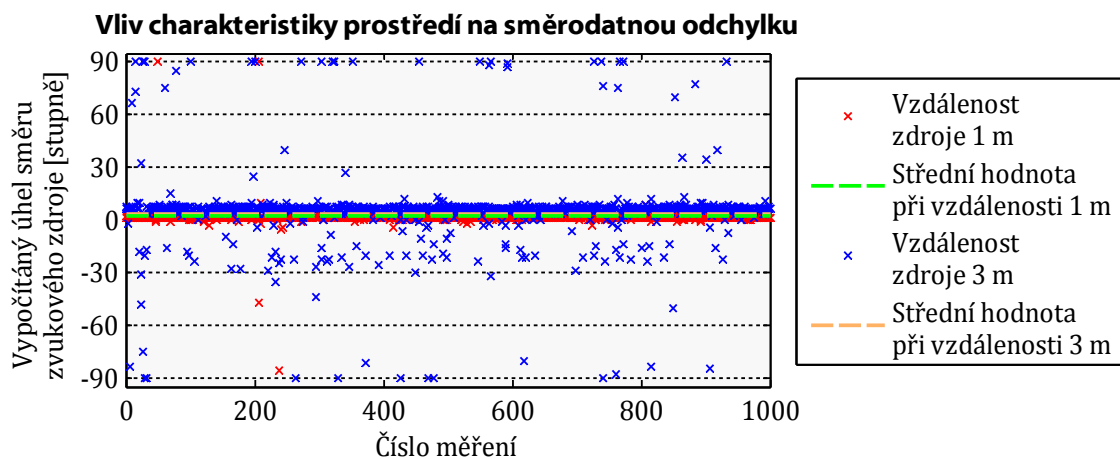
Podle předpokládané vzdálenosti zdroje (několik metrů) byla také zvolena směrová charakteristika mikrofonů. Superkardioidní charakteristika zajišťuje lepší záznam zvuku přicházejícího v širokém úhlu směru zepředu a potlačení příchozího vlnění z ostatních směrů. Pokud by se měl sledovaný zdroj pohybovat ve větší vzdálenosti (například více než

10 metrů), bylo by vhodné vybavit zařízení mikrofony s hyperkardioidní směrovou charakteristikou, která zaznamenává zvukové vlny přicházející v úzkém úhlu zředu a zvuky z ostatních směrů téměř eliminuje. V případě velké vzdálenosti nesměrových mikrofonů od zařízení by totiž mohlo být nutné zvýšení předzesílení signálů natolik, že by byly zaznamenávány nežádoucí ruchy prostředí. Pak by bylo nutné použít mikrofony s úzce směrovou charakteristikou.

### 4.3 Vliv prostředí

V ideálním případě by byl zvuk šířen z bodového zdroje v kulových vlnách, které lze při záznamu mikrofony považovat za vzájemně rovnoběžné, nedochází k jejich odrazům, interferenci vlnění a veškeré předměty v prostředí vlnění absorbují. V reálném prostředí se však zvukové vlny šíří všemi směry, a jelikož překážky nemají ideální absorpční charakteristiku, část energie vlnění je předměty pohlcena a přeměněna na teplo, část je propuštěna předmětem a část zvukového vlnění je odražena zpět do prostředí, které má samo o sobě vliv na pronikání zvuku (rychlost).

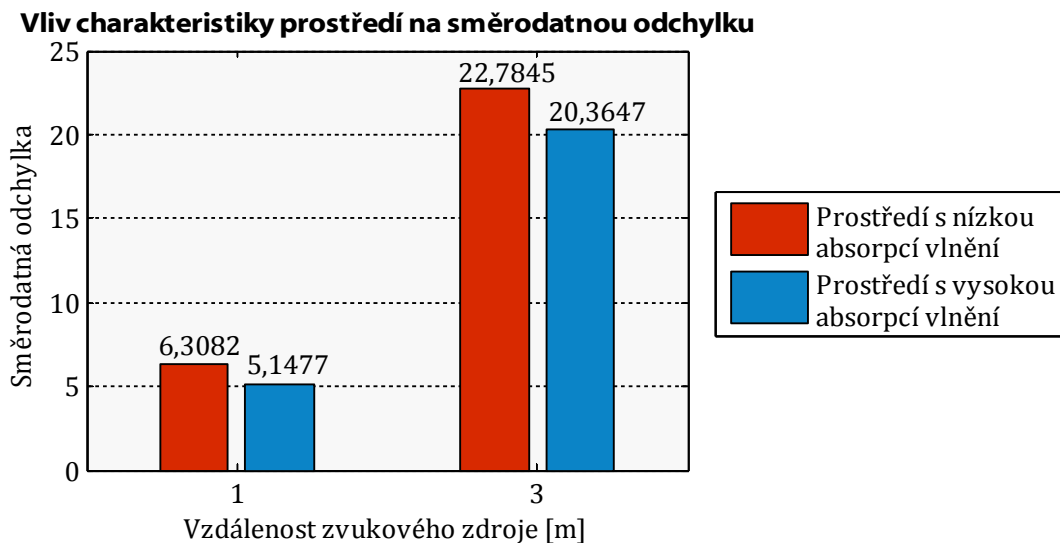
Vliv prostředí lze vyvodit z vypočítaných hodnot rozptylů hodnot naměřených v prostředích s vysokou a nízkou absorpcí zvukových vln. Graf 13 zobrazuje rozptyl vyhodnocených dat v prostředí ohraničeném stěnami, jejichž materiál dobře absorbuje zvukové vlny, a tedy nezpůsobuje příliš odrazů vlnění.



*Graf 13: Rozptyly hodnot směrů ovlivněné charakteristikou prostředí*

Směrodatné odchylky jsou rovny při vzdálenosti od zdroje 1 metru přibližně 5,1 a při vzdálenosti 3 metry přibližně 20,3. Oproti obdobnému experimentu bez polstrovaných stěn (viz graf 12 na str. 36) je směrodatná odchylka naměřených hodnot nižší o 20 %, respektive 10 %.

Srovnání směrodatných odchylek ovlivněných charakteristikou prostředí a vzdáleností zdroje (dat zobrazených v grafech 12 a 13 na stranách 36 a 37) lze vidět na grafu 14.



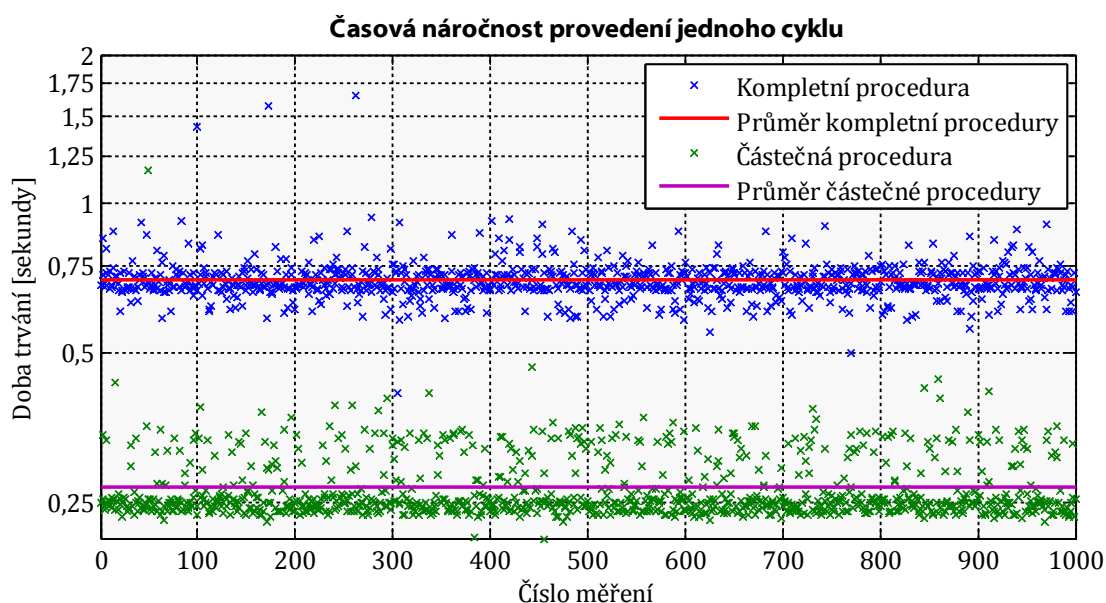
**Graf 14:** Vliv prostředí a vzdálenosti na směrodatnou odchylku naměřených hodnot

Správnost vyhodnocování směru příchodu je tedy lépe vyhodnocována v rozlehlém prostředí, v němž se tvoří minimální dozvuk způsobený odrazem od předmětů v okolí záznamového zařízení. Ke snížení odrazivosti předmětů (například stěn) slouží čalounění či obložení akustickým molitanem. Například v malé místnosti s holými stěnami může docházet k vysokému počtu odrazů, a tak může zařízení nevhodně vyhodnotit podobnost odražených vln jako hledaný směr akustického zdroje a tím selhat.

## 4.4 Časová náročnost

Jelikož má zařízení sledovat akustický zdroj – člověka, pro splnění podmínky aplikace reálného času musí být analýza signálů dokončena dříve, než se poloha zdroje změní.

Jak vyplývá z provedeného měření časové náročnosti zařízení (viz graf 15 na str. 39), tak průměrná doba trvání jednoho průběhu hlavního cyklu při částečné proceduře (nedostatečném výkonu signálů) – je proveden pouze záznam signálů, výpočet výkonu, porovnání s prahovou hodnotou a případně je natočen servomotor – je rovna 0,2671 sekundy. Nejvyšší naměřená doba trvání jednoho takového cyklu je 1,713 sekundy. Je-li výkon signálů dostatečný (je provedena kompletní analýza zaznamenaných dat), je průměrná doba trvání jedné kompletní procedury hlavního cyklu 0,7005 sekundy a nejvyšší naměřená doba trvání je rovna 1,6493 sekundy. Důvodem různého trvání procedur může být potřeba obsluhy operačního systému či provádění výpočetních operací s různě vysokými čísly.



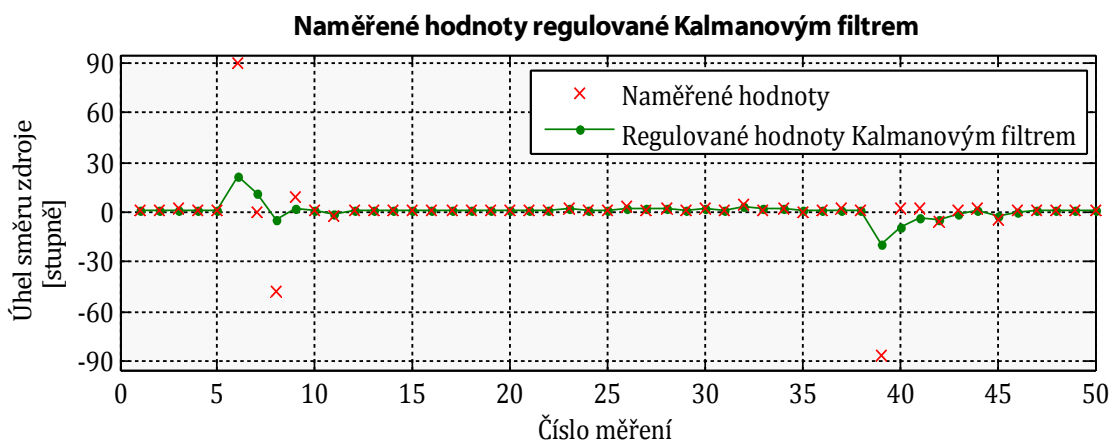
**Graf 15:** Hodnoty dob trvání procedur včetně jejich středních hodnot

K dosažení maximálního výkonu bylo Raspberry Pi přetaktováno na dynamický plný výkon (v případě zvýšeného výkonu se automaticky zvyšuje takt procesoru). Tím je zkrácena doba výpočetních operací. Dobu pro zaznamenání zvuku však lze zkrátit pouze snížením počtu požadovaných vzorků. Průměrné vytížení procesoru je při kompletní proceduře cyklu přibližně 97 %, při částečné proceduře je průměrné vytížení procesoru 93 %.

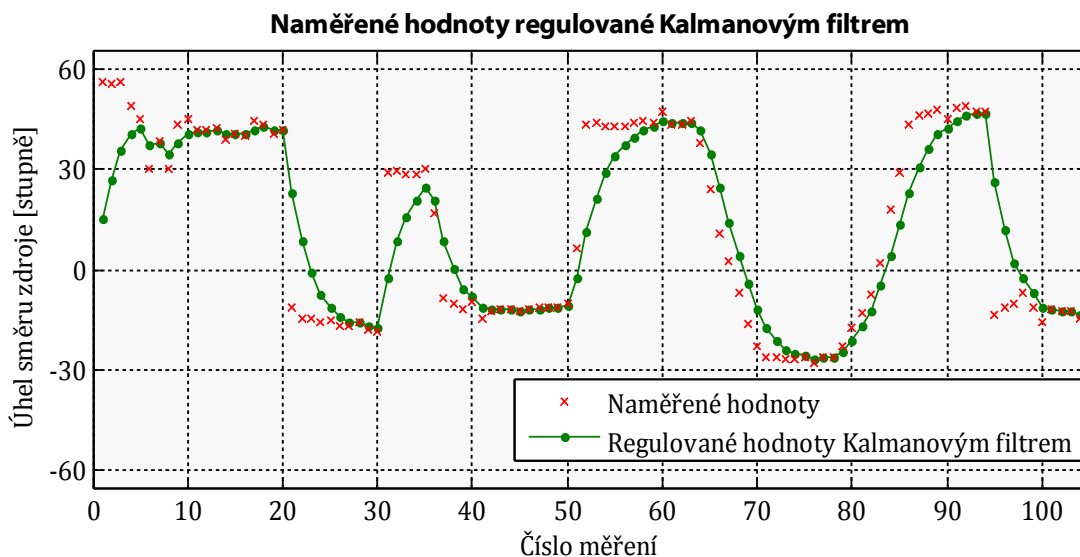


## 4.5 Plynulost změny směru

Pro dosažení plynulého pohybu kamery při změně polohy zdroje je příkaz k rotaci rozdělen do několika kroků. Při chybném vyhodnocení směru zdroje (viz graf 16) nebo při záznamu střídajících se mluvčích (viz graf 17) by mohlo docházet ke skokovým změnám natočení i při využití členěné rotace. Proto je na výslednou vypočítanou hodnotu směru aplikován navržený Kalmanův filtr, který zamezuje nežádoucím rychlým změnám hodnot, a tak koriguje finální směr natočení kamery.



*Graf 16: Ukázka regulace naměřených hodnot implementovaným Kalmanovým filtrem*



*Graf 17: Ukázka regulace hodnot Kalmanovým filtrem – případ střídajících se mluvčích*

## 5 Závěr

Smyslem této bakalářské práce bylo vytvoření funkčního prototypu samostatného mobilního zařízení, jež podle zaznamenání zvuku dvěma mikrofony lokalizuje akustický zdroj a následně provádí rotaci kamery jeho směrem v reálném čase.

Kódy modulů programu byly implementovány v programovacím jazyce Python 2.7 pro minipočítač Raspberry Pi model B revize 2 s operačním systémem Raspbian. Skript zajišťující obsluhu automatické spouštění programu po spuštění operačního systému byl napsán ve skriptovacím jazyce Bash.

Zařízení pracuje na základě metody určování fázového posunu zaznamenaných zvukových signálů. Po provedení záznamu dvou signálů o délce 1000 vzorků je vypočítáván součet jejich výkonů, na jehož základě je rozhodováno, zda signály obsahují informace dostačující k další analýze. Jsou-li signály dostatečně hlasité, jsou filtrovány nízké frekvence navrženým filtrem typu horní propust podle Butterworthovy aproximace. Na filtrovaných datech je prováděn výpočet křížové korelace, jejíž výsledky jsou interpolovány kvadratickou funkcí k získání polohy hledaného nejvyššího výsledku. Takto získávaný fázový posun signálů je přepočítáván na odpovídající úhel natočení, který je z důvodu optimalizace regulován navrženým Kalmanovým filtrem. Regulované hodnoty jsou předávány ke konverzi vláknu, jež obsluhuje servomotor, pro převedení na příslušný PWM signál, kterým je řízena rotace motoru – kamery.

Analýza provedených experimentů prokázala, že inteligentní ovladač kamery nejlépe pracuje v tichém prostředí, ve kterém dochází k odrazům zvukových vln minimálně. Podle měření časové náročnosti procesu byla maximální doba odezvy zařízení pod 2 sekundy. Průměrná doba odezvy se však pohybovala pod hranicí 1 sekundy.

Při vývoji obdobného zařízení s požadavkem nižší odezvy by bylo nutné využít výpočetní jednotku s vyšším výpočetním výkonem než Raspberry Pi, jelikož snížením počtu zaznamenávaných vzorků je kromě počtu prováděných matematických operací snížena také přesnost vyhodnocování směru příchodu zvuku.

Získávání kvalitních dat zvukového vlnění je podmíněno především charakteristikou použitých mikrofonů. Dynamické mikrofony (použité při konstrukci tohoto zařízení) by bylo vhodné nahradit mikrofony kondenzátorovými, jež disponují vyšší citlivostí.

## Seznam použité literatury

- [1] RANJAN, Abhishek, et al. Automatic camera control using unobtrusive vision and audio tracking. In: *Proceedings of Graphics Interface 2010*. Canadian Information Processing Society, 2010. s. 47-54. ISBN: 978-1-56881-712-5. [vid. 25. 4. 2015].  
Dostupné z: [http://www.aranjan.com/Docs/CameraControl\\_GI2010.pdf](http://www.aranjan.com/Docs/CameraControl_GI2010.pdf)
- [2] GAZDA, Benjamin, et al. *ECE 480 - Spring 2010 - Wireless Automated Stage Tracking System* [online]. NORTH CAROLINA STATE UNIVERSITY. 2010 [vid. 25. 4. 2015].  
Dostupné z: [http://www4.ncsu.edu/~bpgazda/Stage\\_Tracking\\_System.html](http://www4.ncsu.edu/~bpgazda/Stage_Tracking_System.html)
- [3] JELEN, Josef a Jindřich LEGO. *Fyzika II*. Dot. 1. vyd. Praha: Vydavatelství ČVUT, 1993, 169 s. ISBN 80-01-01001-5.
- [4] KREIDL, Marcel a Radislav ŠMÍD. *Technická diagnostika*. 1. vyd. Praha: BEN - technická literatura, 2006, 408 s. Senzory neelektrických veličin. ISBN 80-7300-158-6.
- [5] NEVŘIVA, Pavel. *Analýza signálů a soustav*. 1. vyd. Praha: BEN - technická literatura, 2000, 671 s. ISBN 80-7300-004-0.
- [6] HÁJEK, Karel a Jiří SEDLÁČEK. *Kmitočtové filtry*. 1. vyd. Praha: BEN - technická literatura, 2002, 536 s. ISBN 80-7300-023-7.
- [7] NORRIS, Donald. *Raspberry Pi: projekty*. 1. vyd. Brno: Computer Press, 2015, 264 s. ISBN 978-80-251-4346-9.
- [8] DS18B20: Programmable Resolution 1-Wire Digital Thermometer. *Analog, linear, and mixed-signal devices from Maxim* [online]. MAXIM INTEGRATED. 2008 [vid. 18. 3. 2015]. Dostupné z: <http://datasheets.maximintegrated.com/en/ds/DS18B20.pdf>
- [9] TowerPro MG90: Micro Servo. *RC Servo Specifications and Reviews* [online]. SERVO DATABASE. 2015 [vid. 18. 3. 2015].  
Dostupné z: <http://www.servodatabase.com/servo/towerpro/mg90>
- [10] Behringer: ULTRAVOICE XM1800S: 3 Dynamic Cardioid Vocal and Instrument Microphones (Set of 3) [online]. BEHRINGER. 2015 [vid. 18. 3. 2015].  
Dostupné z: <http://www.behringer.com/EN/Products/XM1800S.aspx>
- [11] Audio Buddy. *M-AUDIO* [online]. INMUSIC BRANDS, INC. 2015 [vid. 18. 3. 2015].  
Dostupné z: <http://www.m-audio.com/products/view/audio-buddy>

- [12] Behringer: U-CONTROL UCA222: Ultra-Low Latency 2 In/2 Out USB Audio Interface with Digital Output [online]. BEHRINGER. 2015 [vid. 18. 3. 2015].  
Dostupné z: <http://www.behringer.com/EN/Products/UCA222.aspx>
- [13] *Raspbian* [online]. [vid. 20. 3. 2015].  
Dostupné z: <http://www.raspbian.org>
- [14] *Python 2.7.10rc0 documentation* [online]. [vid. 20. 3. 2015].  
Dostupné z: <https://docs.python.org/2.7>
- [15] *alsaaudio documentation* [online]. [vid. 20. 3. 2015].  
Dostupné z: <http://pyalsaaudio.sourceforge.net>
- [16] *SciPy* [online]. SCIPY. 2015 [vid. 20. 3. 2015].  
Dostupné z: <http://www.scipy.org>
- [17] HIRST Richard. *richardghirst/PiBits/ServoBlaster* [online]. GITHUB, INC. 2015.  
[vid. 20. 3. 2015].  
Dostupné z: <https://github.com/richardghirst/PiBits/tree/master/ServoBlaster>

## Přílohy

### A Obsah přiloženého CD

- Text bakalářské práce
  - Bakalarska\_prace-2015-Vojtech\_Stary.pdf
  - Bakalarska\_prace-2015-Vojtech\_Stary.odt
  - Kopie\_zadani\_bakalarske\_prace-2015-Vojtech\_Stary.pdf
  
- Zdrojový kód programu
  - autostart.sh
  - chunk\_recorder.py
  - main.py
  - rgb\_led\_controller.py
  - servo\_controller.py
  - temperature\_controller.py
  - tools\_package.py
  - turn\_off\_button.py
  - kompilované verze výše uvedených souborů formátu Python – přípona .pyc

## B Fotografie zařízení



*Obrázek 18: Fotografie inteligentního ovladače kamery s připevněnou kamerou*

## C Zobrazení zařízení v praktickém využití



*Obrázek 19: Použití inteligentního ovladače k záznamu divadelního představení*