

Bakalářská práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra radioelektroniky

Využití sledování očních pohybů pro hodnocení orientace ve virtuálním akustickém prostoru

Vojtěch Mičín

Vedoucí: Ing. František Rund, Ph.D.
Květen 2022

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Mičín** Jméno: **Vojtěch** Osobní číslo: **492038**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávající katedra/ústav: **Katedra radioelektroniky**
Studijní program: **Elektronika a komunikace**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

Využití sledování očních pohybů pro hodnocení orientace ve virtuálním akustickém prostoru

Název bakalářské práce anglicky:

Application of Eye-Tracking for Assessment of Orientation in Virtual Auditory Space

Pokyny pro vypracování:

Brýle pro virtuální realitu HTC Vive Pro Eye obsahují možnost sledování očních pohybů. Navrhněte využití této funkce pro testování orientace ve virtuálním akustickém prostoru. Navržené řešení implementujte. Funkčnost otestujte na skupině uživatelů a výsledky vyhodnoťte.

Seznam doporučené literatury:

- [1] Jarolímek, V: Virtual Acoustic Space Test using HMD, Diplomová práce FEL ČVUT 2019, dostupné z <http://hdl.handle.net/10467/82888>
[2] Eye and Facial Tracking SDK - Developer Resources. HTC Corp. 2022 [cit 2022-01-29] dostupné z <https://developer.vive.com/resources/vive-sense/eye-and-facial-tracking-sdk/>

Jméno a pracoviště vedoucí(ho) bakalářské práce:

Ing. František Rund, Ph.D. katedra radioelektroniky FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) bakalářské práce:

Datum zadání bakalářské práce: **07.02.2022**

Termín odevzdání bakalářské práce: **20.05.2022**

Platnost zadání bakalářské práce: **30.09.2023**

Ing. František Rund, Ph.D.
podpis vedoucí(ho) práce

doc. Ing. Stanislav Vítek, Ph.D.
podpis vedoucí(ho) ústavu/katedry

prof. Mgr. Petr Páta, Ph.D.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Student bere na vědomí, že je povinen vypracovat bakalářskou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v bakalářské práci.

Datum převzetí zadání

Podpis studenta

Poděkování

Rád bych poděkoval Ing. Františku Rundovi, Ph.D. za veškerou pomoc při řešení problémů, rychlou odezvu na mé časté dotazy a vedení celého projektu. Dále chci poděkovat Bc. Martinu Novotnému a Radku Nesnídalovi za spolupráci v oblasti virtuální reality.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

Abstrakt

Tato práce pojednává o návrhu využití funkce sledování očních pohybů pro orientaci ve virtuálním akustickém prostoru (VAS). Popisuje použité zařízení a experimenty, jež měly na návrh vliv. Je uveden proces návrhu a implementace nové funkcionality. Popisuje, jaká rozhodnutí byla při implementaci učiněna a jaké možnosti zváženy. Jsou zhodnoceny výsledky získané od 12 dobrovolníků ze subjektivního testování. Práce udává možné spojitosti mezi zvuky ve VAS a jeho vnímáním skrze zrak. Výsledky potvrzují možnost užití zraku jako média hodnocení orientace ve VAS, ovšem za dodržení určitých podmínek.

Klíčová slova: virtuální realita, VR, virtuální akustický prostor, VAS, sledování očních pohybů, HTC Vive Pro Eye, orientace

Vedoucí: Ing. František Rund, Ph.D.

Abstract

This paper discusses the design of the use of eye tracking for orientation in virtual acoustic space (VAS). It describes the equipment used and the experiments that influenced the design. The process of designing and implementing the new functionality is presented. It describes what decisions were made and what options were considered in the implementation. Results obtained from 12 volunteers from subjective testing are reviewed. The paper indicates possible connections between sounds in the VAS and its perception through vision. The results confirm the possibility of using gaze as a medium to assess orientation in the VAS, but under certain conditions.

Keywords: virtual reality, VR, virtual acoustic space, VAS, eye tracking, HTC Vive Pro Eye, orientation

Title translation: Application of Eye-Tracking for Assessment of Orientation in Virtual Auditory Space

Obsah

Seznam zkratk	I
1 Úvod	1
1.1 Motivace	1
1.2 Návrh využití funkce sledování očních pohybů	2
1.3 Struktura práce	2
2 Teoretický úvod	3
2.1 Lokalizace zdroje zvuku u člověka	3
2.2 Vztah zraku a sluchu	4
3 Eye tracking a použité zařízení	6
3.1 Snímání očních pohybů uvnitř VR	6
3.2 Použité zařízení	7
3.3 Výběr Software Development Kitu (SDK) pro vývoj aplikace	8
4 Návrh úprav aplikace a testů	10
4.1 Původní verze aplikace	10
4.2 Vylepšení aplikace	10
4.2.1 Způsob testů	10
4.2.2 Výběr zdroje zvuku	11
4.2.3 Výběr testů	11
4.2.4 Získávání dat	12
4.2.5 Výstupy aplikace	13
4.3 Implementace	17
4.4 Příprava a průběh testování	18
5 Výsledky a jejich zhodnocení	19
5.1 Zhodnocení výsledků testování	19
5.2 Zhodnocení dotazníků	21
5.3 Možné úpravy do budoucna	22
6 Závěr	24
Literatura	25
Příloha A	28
Příloha B	29

Obrázky

2.1 Závislost průměrné reakční doby na frekvenci a cíli pohledu	5
3.1 Ukázka principu foveated rendering	6
3.2 Obroučky okolo čoček zajišťující sken očních pohybů	7
4.1 Zobrazení jednoho testu včetně colliderů a proloženou rovinou	13
4.2 Metoda 1: Relativní vzdálenost od zdroje	14
4.3 Metoda 2: Výpis pomocí úhlů ..	14
4.4 Ukázka zápisu dat do souboru typu "XZ.txt"	15
4.5 Příklady grafických zobrazení v pořadí: heat map, opacity map, gaze plot	15
4.6 Ukázka grafického výstupu pro jeden test jednoho subjektu	16
4.7 Výběr otázek z dotazníku hodnocení kvality testování	18
5.1 Ukázka principu skriptu "Zavrene_oci.m"	20

Tabulky

2.1 Průměrné reakční doby [ms] pohybů očí směrem ke zdrojovému a tichému reproduktoru s různou frekvencí tónu a různou náběžnou hranou	4
3.1 Specifikace headsetu HTC Vive Pro Eye	7
3.2 Specifikace funkce eye tracking HTC Vive Pro Eye.....	8
5.1 Tabulka výsledků testů	19



Seznam zkratk

FOV	Field Of View	12
G2OM	Gaze To Object Mapping	12
VR	Viruální realita	5, 28
HMD	Head Mounted Display	5, 6, 10-12, 14, 18, 26
HRTF	Head Related Transfer Function	7, 27
VAS	Virtual Acoustic Space	5-7, 15, 27, 28
SDK	Software Development Kits	12, 13, 16, 26

Kapitola 1

Úvod

1.1 Motivace

Možnosti použití se spolu s rostoucí kvalitou headsetů virtuální reality (VR) neustále rozšiřují. Už jen v komerčním odvětví existuje velká řada využití (např. tvorba různých VR únikových místností [1] nebo ozvláštňení atrakcí v zábavních parcích [2]). Kromě zábavy lze virtuální realitu využít k nespočtu užitečných činností: reálnější zobrazení probírané látky ve školách, vojenské či policejní řešení mimořádných situací, praktické činnosti, jež by v realitě byly příliš nákladné, anebo při řešení problémů spjatých s různými nemocemi a dysfunkcemi lidského organismu.

Zajímavým příkladem poslední mnou zmíněné kategorie je systém [3] pro vytváření titulků v reálném čase. Jeho užití je určeno především pro lidi s částečnou nebo úplnou ztrátou sluchu. Funguje na základě rozpoznání řeči pomocí několika kamer. Umělá inteligence pohyby úst dešifruje a zobrazí titulky přímo na Head Mounted Display (HMD), jež má uživatel na hlavě. Prozatím je testován pouze v divadlech, ale jeho využití sahá i za hranice divadla.

Existuje také systém [4], který umožňuje slepým lidem "vidět". Kamery snímají okolí uživatele a získávají tak informaci o struktuře, povrchu a rozmístění objektů v jeho blízkosti. Software následně každý objekt převede na specifický zvuk. Slepý člověk pak s určitou zkušeností s tímto zařízením dokáže vnímat rozmístění prostoru, ve kterém se zrovna nachází.

S tímto posledním výzkumem se pojí termín lokalizace (tj. určení polohy zdroje zvuku na základě parametrů zvuku). Tento pojem je rozvinut v kapitole [2.1]. Ve virtuálním prostoru je lokalizace možná díky metodě zvaná virtuální akustický prostor VAS. Jedná se o metodu zpracování signálu, která zvuku dodá prostorový vjem. Jak bude uvedeno dále, člověk v reálném světě na zvuky mimo jiné reaguje i pohybem očí. HMD HTC Vive Pro Eye, který jsem měl během této práce k dispozici nabízí nejen zlepšení kvality obrazu a dalších parametrů, ale i zabudovaný senzor očních pohybů. Cílem práce bude zjistit, jakým způsobem se člověk očima ve VAS orientuje a jak na zvuky očima reaguje. Dále bych rád zjistil, zda je interakce uživatele a jeho zraku přirozená.

1.2 Návrh využití funkce sledování očních pohybů

Za účelem získání informace o tom, odkud lidé zvuk ve virtuální realitě vnímají, vytvořil Ing. Viktor Jarolímek testovací aplikaci. Ve své diplomové práci „Virtual Acoustic Space (VAS) Test Using HMD“ [5] popisuje proces vytváření této aplikace a možnosti jejího následného využití.

Testování měli v rámci subjektivních testů za úkol vybrat takový objekt, který se jim zdál jako zdroj určitého zvuku. Testy se mezi sebou lišily rozmístěním těchto objektů, jejich vzájemnou vzdáleností a přídatnými rušivými elementy. Z těchto subjektivních testů byla zaznamenávána data o správnosti odpovědí, délce testu a případné odchylce špatně označeného cíle od správného.

Aplikace je dobře navržena, ale má i určité nedostatky. Například neříká nic o chování testované osoby během jejího výběru. Kdyby se tedy subjekt rozhodl mezi dvěma objekty a na poslední chvíli se rozhodl jinak, aplikace to nepozná. Taktéž není možné zjistit, zda testovaný nevybral cíl omylem. Bylo by tedy vhodné tuto aplikaci vylepšit a zjistit i tyto dodatečné informace.

Existující aplikaci bych rád obohatil o možnost snímání očních pohybů a zaznamenávání dodatečných dat v rámci jednotlivých testů. Nabízí se také možnost, aby testování svým zrakem chod aplikace řídili. Tato možnost bude diskutována později. Data bych chtěl zaznamenávat do textových souborů a následně graficky zpracovávat.

1.3 Struktura práce

První část práce je věnována základům teorie, která se zaměřuje na témata lokalizace zdroje zvuku u člověka a propojení zrakové a sluchové soustavy. Také jsou zde popsány experimenty, jež udávají směr dalšímu postupu při úpravě aplikace.

Následující část popisuje funkci sledování očních pohybů, obsahuje specifikace použitého zařízení a nakonec vývojářský kit, který byl použit pro úpravu aplikace resp. získávání dat z očních kamer.

Třetí část se zaměřuje na přetvoření původní aplikace. Je zde popsána původní aplikace a dále volby, jež se při jejím vývoji naskytly a rozhodnutí, která byla učiněna. V závěru kapitoly jsou shrnuty všechny úpravy, jež byly na aplikaci provedeny.

V poslední kapitole jsou zobrazeny a diskutovány výsledky subjektivního testování. Nachází se zde zhodnocení výsledků, poznatky k průběhu testů od testovaných osob a závěry, které z toho vyplynuly. V úplném závěru práce jsou rozepsány další možnosti a doporučení pro případné navazující práce.

Kapitola 2

Teoretický úvod

2.1 Lokalizace zdroje zvuku u člověka

V běžném životě je lokalizace zvuku umožněna díky šíření vln v prostoru. Zvukové vlny se v prostoru odrážejí, rozptylují a tlumí na základě rozměrů, materiálech a rozmístění objektů v prostoru. Méně odrazů lokalizaci napomáhá, více ji naopak komplikuje. V rámci slyšení však zvuk interaguje i společně s tělem posluchače (především pak hlavou a uchem). Head Related Transfer Function (HRTF)¹ je funkce, která charakterizuje, jak ucho přijímá zvuk z bodu v prostoru [6]. Jak již vyplývá z textu, HRTF každého člověka je jedinečná. Tato funkce je frekvenčně závislá, neboť různé povrchy a rozměry předmětů interagují jinak s různými frekvencemi.

Už jen to, že se HRTF definuje pro každé ucho zvlášť, naznačuje, že dvě uši budou mít v lokalizaci určitý význam. Pojem binaurální slyšení [7] se v oblasti akustiky v posledních desetiletích velice zpopularizoval. Nejedná se jen o provázanost s lokalizací zdroje zvuku. Binaurální slyšení člověk (resp. zvíře) využívá například při detekci zvuku v rámci šumu, analýzu hudby či rozpoznání řeči. Pro jeho pochopení existuje celá škála modelů pro specifické případy. Pro tuto práci jsou tyto modely nepodstatné, mají však nastínit komplexnost celé problematiky.

Jak udává Sharon M. Abel ve své práci "The effect of aging on horizontal plane sound localization" [8], proces lokalizace je silně ovlivněn věkem. Výsledky subjektivního testování, které v rámci výzkumu prováděl, byly s rostoucím věkem horší.

Bylo by tedy vhodné subjektivní testy na orientaci ve VAS provádět v různých věkových kategoriích, aby nedocházelo ke zkreslení dat stárnutím.

¹Česky přenosová funkce související s hlavou

2.2 Vztah zraku a sluchu

Otázka, zda je zrak a sluch navzájem provázán, je bezesporná. Obecně je možné říct, že při obdržení odlišných informací těchto smyslů, zrak upřednostňujeme v prostorovém zpracování, sluch v časovém [9]. Tomuto vzájemnému propojení smyslů se říká cross-modalní.

Dominanci zraku cross-modality můžeme pozorovat na efektu zvaném břichomluvecký [10]. Pod tímto názvem se skrývá jev, kdy je zvuk vnímán z jednoho zdroje, i když ve skutečnosti vychází ze zdroje jiného. Efekt je nejintenzivnější při zvucích řeči. Má-li člověk detekovat rychle se měnící vizuální cíle, je možné jeho reakční čas měnit zvukem. Pokud je vizuální cíl synchronně doplněn zvukem, reakční doba člověka se zkrátí [11].

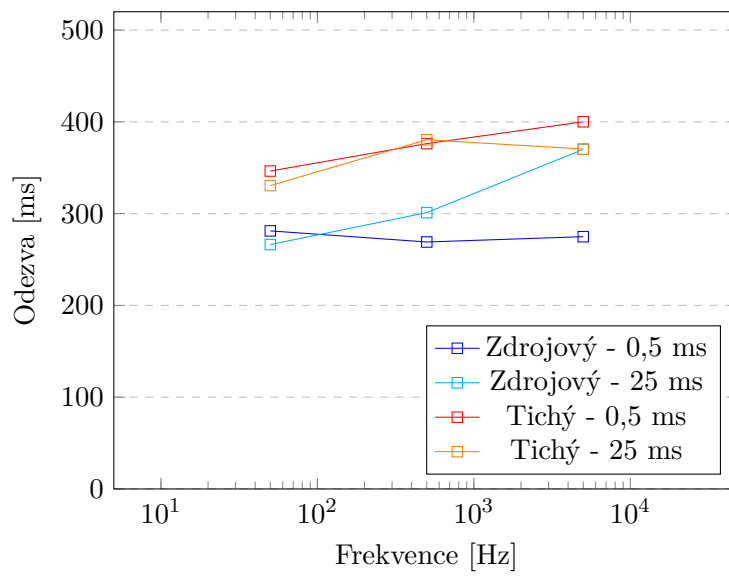
Další ukázkou propojenosti zraku a sluchu je následující studie. V roce 1978 byl proveden výzkum [12], který rozdělil testované osoby do dvou skupin. Každá osoba byla umístěna mezi dva reproduktory. Jedna skupina se měla po zaznění zvuku z jednoho reproduktoru podívat na zdrojový reproduktor, druhá pak na ten, který zvuk nevydával. Zvuky reproduktorů se lišily frekvencí a náběžnými hranami. Z tabulky výsledků [2.1] vyplývá, že lidem je přirozenější podívat se na zdroj zvuku. Pohled na tichý zdroj vyžaduje uvědomění, které je ve výsledcích časově vidět.

Pokud by bylo možné z těchto výsledků udělat závěr, že označení zdroje zvuku je intuitivní, znamenalo by to, že by měl člověk být vnitřně schopen určit správný zdroj. Nejdříve by se totiž podíval na místo, odkud zvuk slyší. Je však nutné zmínit, že studie se dvěma reproduktory nepracovala s blízkými zdroji zvuku. Bylo vždy jasné, odkud zvuk vychází. V testovací aplikaci tato podmínka splněna není. Cíle mohou být libovolně blízko, resp. daleko od sebe.

Z toho plyne otázka, zda testovaná osoba dokáže intuitivně poznat zdroj zvuku tím, kam se podívá nejdříve. Je také možné, že osoba bude očima mezi možnými zdroji zvuku přeskakovat, anebo oči například úplně zavře.

Tabulka 2.1: Průměrné reakční doby [ms] pohybů očí směrem ke zdrojovému a tichému reproduktoru s různou frekvencí tónu a různou náběžnou hranou [12]

Reproduktor	Náběžná hrana [ms]	Frekvence [Hz]		
		50	500	5000
Zdrojový	0,5	281,2	269,1	274,9
	25	266,3	301,1	370,0
Tichý	0,5	346,4	376,3	400,1
	25	330,6	380,5	370,4



Obrázek 2.1: Závislost průměrné reakční doby na frekvenci a cíli pohledu¹

¹Použita jsou data z tabulky [2.1](#).

Kapitola 3

Eye tracking a použité zařízení

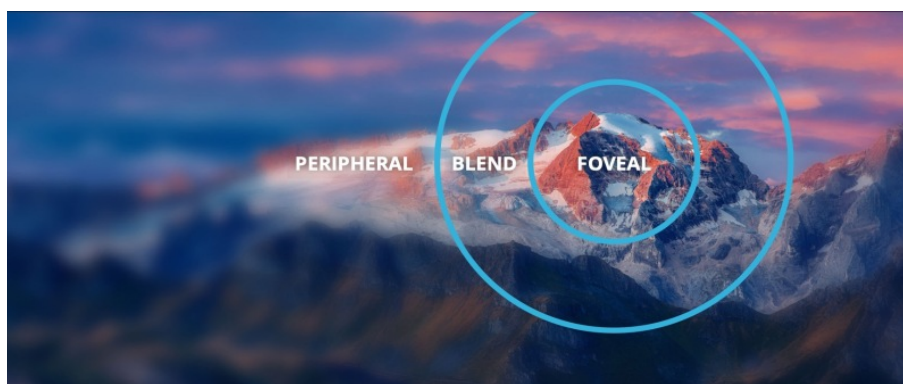
3.1 Snímání očních pohybů uvnitř VR

Zavedení funkce snímání očních pohybů uvnitř HMD je poměrně významný pokrok ve vývoji virtuální reality. Existují dva důvody, proč jsem si dovolil tuto funkcionalitu označit jako velký posun vpřed.

Prvním důvodem je možnost využívání samotné funkce pro mnoho dalších aplikací. Prozatím mnoho takto fungujících aplikací neexistuje. V dnešní době však funkci zkoušejí někteří prodejci. Ve virtuálním prostředí si svůj obchod namodelují a nechají dobrovolníky, aby si tento virtuální obchod procházeli stejně jako ten běžný. Po získání dostatečného množství dat majitel vidí, na které vystavené zboží se dobrovolníci dívali nejvíce a kam by tedy měl umístit jaké zboží.

Obdobnou metodu bych chtěl použít i při přestavbě původní aplikace pana Jarolínka. Pouze bych místo zboží použil možné zdroje zvuku a vyhodnocoval každého dobrovolníka zvlášť.

Mnohem důležitějším důvodem je však tzv. funkce foveated rendering [13]. Ta zajišťuje, že je lokálně možné zvýšit kvalitu obrazu tím, že ji snížíme v periferním poli zraku (viz obr. 3.1). HMD umožňující funkci eye tracking pak může mít v kombinaci se softwarem využívajícím foveated rendering pocitové rozlišení mnohem lepší než ty, které tuto možnost nemají nebo nevyužijí.



Obrázek 3.1: Ukázka principu foveated rendering [13]

3.2 Použité zařízení

Jak jsem již zmiňoval, pro vývoj aplikace a zajištění snímání očních pohybů jsem měl k dispozici HMD HTC Vive Pro Eye [14]. Jedná se o první headset, který má tuto funkci řešenou implicitně. V tomto případě je snímání očních pohybů zajištěno pomocí obrouček okolo čoček (viz obr. 3.2), které sledují pohyb sítnice zabudovaným systémem kamer a malých zdrojů světla. Tyto kamery mimo jiné získávají informace i o pohybech svalů kolem oblasti očí. Další vlastností tohoto HMD jsou zabudovaná sluchátka.



Obrázek 3.2: Obroučky okolo čoček zajišťující sken očních pohybů [15]

V tabulkách 3.1 a 3.2 jsou vypsány parametry HMD Vive Pro Eye. První tabulka obsahuje parametry headsetu, druhá parametry snímání očních pohybů.

Tabulka 3.1: Specifikace headsetu HTC Vive Pro Eye [16]

Screen:	Dual OLED 3.5" diagonal
Resolution:	1440 x 1600 pixels per eye (2880 x 1600 pixels combined)
Refresh rate:	90 Hz
Field of view:	110 degrees
	Hi-Res-certified headset
Audio:	Hi-Res-certificate headphone (removable)
	High-impedance headphone support
	Enhanced headphone ergonomics
Input:	Dual Integrated microphones
Connections:	USB-C 3.0, DP 1.2, Bluetooth
Sensors:	SteamVR Tracking, G-sensor, gyroscope, proximity, IPD sensor, eye tracking
	Eye relief with lens distance adjustment
Ergonomics:	Adjustable IPD
	Adjustable headphones
	Adjustable headstrap

Tabulka 3.2: Specifikace funkce eye tracking HTC Vive Pro Eye [16]

Gaze data output frequency (binocular):	120Hz
Accuracy (Within FOV 20°):	0.5°–1.1°
Calibration:	5-point
Trackable field of view ¹ :	110°
	Timestamp (device and system)
	Gaze origin
Data output (eye information):	Gaze direction
	Pupil position
	Pupil size
	Eye openness
Interface:	HTC SRanipal SDK
SDK engine compatibility:	Unity, Unreal
Sensors:	SteamVR Tracking, G-sensor, gyroscope, proximity, IPD sensor, eye tracking
	Eye relief with lens distance adjustment
Ergonomics:	Adjustable IPD
	Adjustable headphones
	Adjustable headstrap

3.3 Výběr Software Development Kitu (SDK) pro vývoj aplikace

Data ze skenování očních pohybů (viz minulá kapitola 3.1) uvnitř HTC Vive Pro Eye nejsou přímočaře dostupná. Z tohoto důvodu byly vytvořeny tzv. Software Development Kits (SDK) umožňující přístup ke většině datům, které poskytovaný HMD dokáže zaznamenávat. Ke dnešnímu dni se tvorbou SDK zabývají dvě firmy.

SRanipal SDK je software vyvinutý přímo firmou HTC Corporation, jejíž HMD jsem pro vývoj aplikace využil. Tento kit umožňuje přístup k parametrům jako je normalizovaný vektor pohledu jednotlivých očí, průměr zornice, mhouření očí nebo míru zavření očí. Kit je zdarma poskytován všem uživatelům HTC Vive Pro Eye a má sloužit především k developerským a výzkumným účelům.

Druhou vývojářskou firmou je Tobii XR [17]. Její kit umožňuje přístup k vektorům pohledu v rámci HMD (souřadnice v rámci prostoru očí a displeje) i světa (souřadnice v rámci virtuálního prostoru), neposkytuje však tolik funkcí zaměřujících se na mimiku obličeje. Stejně jako SRanipal SDK je zdarma poskytován ke svým produktům, využívá se však především na marketingové účely a vývoj her. Oproti prvnímu SDK se tento kit pyšní funkcí Gaze To Object Mapping (G2OM) [18]. Tato funkce na základě strojového učení a velkého množství uživatelských dat dokáže přesněji rozlišit, zda se uživatel na určitý objekt dívá, či nikoliv. Skenování očních pohybů v dnešní době totiž stále není perfektní, a tak jsou často data zkreslována. Díky této funkci se

¹Operace očí, oční onemocnění, silný make-up a vysoká krátkozrakost mohou ovlivnit výkon sledování očí.

Kapitola 4

Návrh úprav aplikace a testů

4.1 Původní verze aplikace

Aplikace Viktora Jarolímkova psaná v prostředí Unity (verze 2018.2.6f1) [19] je ke dnešnímu dni nespustitelná. Důvodem je plugin OpenVR, na kterém byla aplikace postavena. Tento plugin již v dnešní době není plně podporován, neboť byl zahrnut v pluginu OpenXR (mixed reality). Při této změně ale nebyly zachovány původní funkce, a tak se aplikace častokrát odvolávala na něco, co již neexistuje. Tento problém odstranil kolega Bc. Martin Novotný tím, že některé původní funkce nahradil jejich novými ekvivalenty. Aplikace je po úpravě opět spustitelná ve verzi Unity 2018.2.6f1. Bylo snahou, tuto verzi změnit na jinou dlouho podporovanou verzi. Z důvodu špatné kompatibility však tento krok nebyl učiněn. Rozdíl mezi původní a novou verzí je také v použitém HMD. Ve svém výzkumu pan Jarolímek využíval HTC Vive. Nyní je k dispozici novější HTC Vive Pro Eye. Pokud by však aplikace dále nepoužívala funkci snímání očních pohybů, mohl by být nadále využíván původní headset.

4.2 Vylepšení aplikace

V rámci této kapitoly se budu snažit původní aplikaci přetvořit tak, aby bylo možné získávat informace o testovaných osobách v průběhu testu. Zaznamenávány by byly oční pohyby a chování člověka.

4.2.1 Způsob testů

Ještě před začátkem přetváření bylo nutné rozhodnout, jakým způsobem by se měla upravená verze spouštět. Nabízely se dvě možnosti:

1. vytvořit kopii aplikace a tu upravit tak, aby zajišťovala to, co bych potřeboval
2. přidat do aplikace volbu, zda by si člověk přál spustit původní nebo upravenou verzi

Volba po spuštění aplikace by mohla být zprostředkována přidáním textovým oknem nebo tlačítkem. Experimentátor by vybral tu variantu, kterou by chtěl spustit. Nevýhodou tohoto řešení je nutnost programu "SRanipal.exe". Před začátkem aplikace je nutné ho spustit, neboť umožňuje zaznamenávání veškerých dat získávaných z očních kamer.

Z tohoto jsem vybral možnost spuštění upravené verze aplikace jako samostatného programu. Sice experimentátor nebude mít možnost měnit typ testování uvnitř spuštěné aplikace. Zároveň ale jistě před spuštěním programu ví, jaké testy potřebuje provést. Navíc při přepisu aplikace nedojde ke kolizím, které by mohly při tvorbě jedné kombinované verze vzniknout.

4.2.2 Výběr zdroje zvuku

Další důležité rozhodnutí se týká způsobu vybírání akustického zdroje. Celkem jsem se rozhodoval mezi třemi odlišnými přístupy:

1. vybírání cílů pomocí paprsků vycházejících z ovladačů, potvrzování stisknutím tlačítka, snímání očních pohybů jen jako dodatková informace
2. vybírání cílů pohledem, potvrzování dlouhým mrknutím, snímání očních pohybů i jako dodatková informace
3. vybírání cílů pohledem, potvrzování stisknutím tlačítka, snímání očních pohybů i jako dodatková informace

Metoda 1 je funkční. Paprsky z ovladačů by ale mohly osobu rušit. Za účelem získání přesnější představy o tom, jak se člověk ve VAS chová, nejsou tyto rušivé elementy žádoucí.

Metoda 2 funkční není. Stvrzování výběru by mohlo ovlivňovat výsledky jednotlivých testů. Problém by byl se záznamem času. Dlouhé zavření očí pro stvrzení výběru by vedlo ke zkreslení délky ukončujícího testu a mohlo by vést i k ovlivnění testu následujícího. Závažnějším problémem by bylo zavírání očí testovaných během testu. Je sice možné, že žádný člověk oči za účelem lepší lokalizace nezavře, ale kdyby ano, aplikace by to vyhodnotila jako stvrzení výběru.

Metoda 3 je funkční, protože zrak není ovlivňován ničím jiným než zvuky z testovací aplikace. Potvrzení pomocí stisknutí tlačítka na ovladači by mělo být dostatečně intuitivní a rychlé.

Z těchto důvodů jsem vybral metodu 3. Ta celkově odpovídá požadovaným cílům a zároveň negativně neovlivňuje průběh testování.

4.2.3 Výběr testů

Původní počet 31 testů by byl pro mou analýzu nadměrný, proto jsem se rozhodl vybrat pouze některé z nich. Při výběru konkrétních testů do testové sekvence jsem přihlížel celkem ke třem hlavním faktorům. Prvním faktorem je počet. Při rychlosti zápisu 45 řádků za sekundu (rychlost zápisu je komentována v kapitole 4.2.4) do textového souboru by se při větším počtu

testů dalo bez přídavného softwaru jen těžko orientovat. Bude proto vhodné použít menší počet testů.

Druhý faktor je opět spojen s výstupy testů. Pro přehledné vykreslení jakéhokoliv grafického výstupu by bylo vhodné, aby cíle byly statické (tj. neměnily v průběhu testu svou polohu v rámci prostoru). Kdyby se měly cíle pohybovat, bylo by velice obtížné zaznačit data do grafu tak, aby předávala nějakou smysluplnou informaci. Samotná funkčnost aplikace (vč. výběru cílů pohledem) by však při jejich zahrnutí měla být zachována. Data budou získávána na testech se statickými objekty, ale pro ověření funkčnosti aplikace budou zahrnuty i testy s pohyblivými objekty.

Posledním důležitým faktorem je práce Martina Novotného [20]. Pan Novotný taktéž provádí subjektivní testy uvnitř této aplikace. Z tohoto důvodu by bylo vhodné, aby co nejvíce testů bylo společných pro případné porovnání výsledků či práci, která by tyto dvě práce propojovala.

Z těchto důvodů bylo vybráno celkem 8 statických testů (tj. se statickými objekty) pro získání dat a 3 dodatkové testy pro ověření funkčnosti. Výpis zvolených testů je uveden v tabulce 5.1. U všech testů jsem zároveň upravil zdrojový zvuk tak, aby byl ve všech testech stejný¹.

4.2.4 Získávání dat

Jak jsem již zmiňoval v kapitole 3.3, pomocí SRanipal SDK je možné zaznamenávat velice konkrétní data o mimice a pohybu očí. Během přetváření aplikace jsem však vyzoroval, že mimika je v rámci mého výzkumu irelevantní. Z tohoto důvodu jsem se rozhodl zaznamenávat pomocí SDK pouze kombinovaný vektor očních pohledů (tj. vektor vytvořený jako kombinace vektorů pohledu levého a pravého oka). Aplikaci však není obtížné upravit tak, aby informace o mimice zaznamenávala. Data jsou získávána a zapisována s každým snímkem. Tento cyklus proto proběhne přibližně 45 krát za sekundu (záleží na plynulosti chodu aplikace).

Projekty pro návrh aplikací virtuální reality v prostředí Unity fungují na bázi vektorů, paprsků a 3D objektů. Aby bylo možné zaznamenat data o tom, do kterého místa se uživatel dívá, je nutné převést kombinovaný vektor pohledu a polohu očí na bod v prostoru s určitými souřadnicemi. Za tímto účelem existuje funkce RaycastHit, která vrací souřadnice průniku paprsku a objektu. Pro získání souřadnic mimo možné zdroje zvuku (již vytvořené objekty), je nutné je doplnit o další objekt. Do úvahy připadají dva geometrické objekty: sféra (povrch koule, uprostřed které by se nacházela hlava uživatele) nebo rovinná plocha (proložena všemi možnými zdroji). Z hlediska obecné funkčnosti by vhodnější variantou jistě byla sféra, neboť by zajišťovala kolizi objektu s paprskem ve všech směrech. Jak jsem však zjistil, objekty v prostředí Unity reagují s okolím pouze z jedné strany. V rámci sféry je tato reagující strana stranou vnější. Kombinovaný vektor, který by tedy prošel sférou zevnitř by s ní v podstatě nekolidoval, a tak by nedošlo k vypsání žádných souřadnic. Možnou náhradou tohoto nedostatku je využití tzv. Mesh Collider, který by

¹Rušivé zvuky jsem ponechal původní.

bylo možné ke sféře zevnitř připnout a tím tento celý problém vyřešit. Jeho užití se však nedoporučuje z důvodu různých chyb, které mohou při jeho použití vzniknout. Z funkčního hlediska této práce je nutné zjistit, na jaký cíl (resp. blízké okolí cíle) se testovaný dívá anebo zda se dívá obecně jinam. Za tímto účelem je možné využít i proložení zdrojů rovinou a usnadnit tak navíc interpretaci výstupů (je možné vynechat hodnoty osy kolmé na rovinu).

I přes kalibraci má z důvodu menší přesnosti očních kamer paprsek očního pohledu menší přesnost než paprsky vycházející z ovladačů v původní verzi aplikace. Proto bylo nutné v rámci zachování plné funkčnosti zvětšit collidery (nehmotné prostředí objektu, které reaguje s okolím) možných zdrojů.

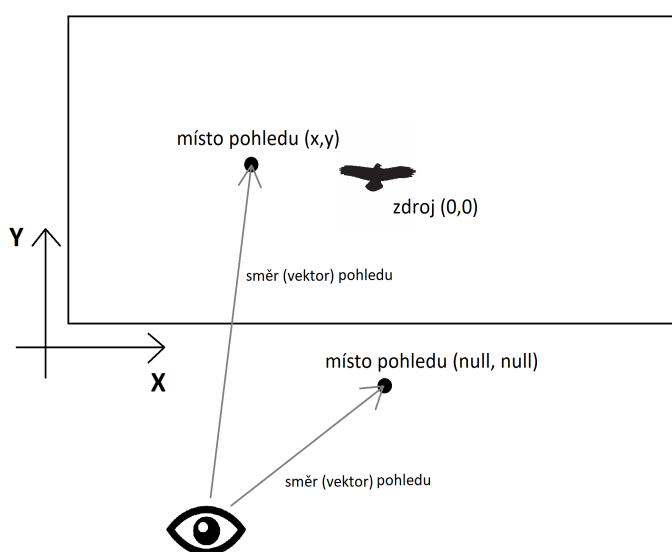


Obrázek 4.1: Zobrazení jednoho testu včetně colliderů a proloženou rovinou

■ 4.2.5 Výstupy aplikace

■ Textové dokumenty

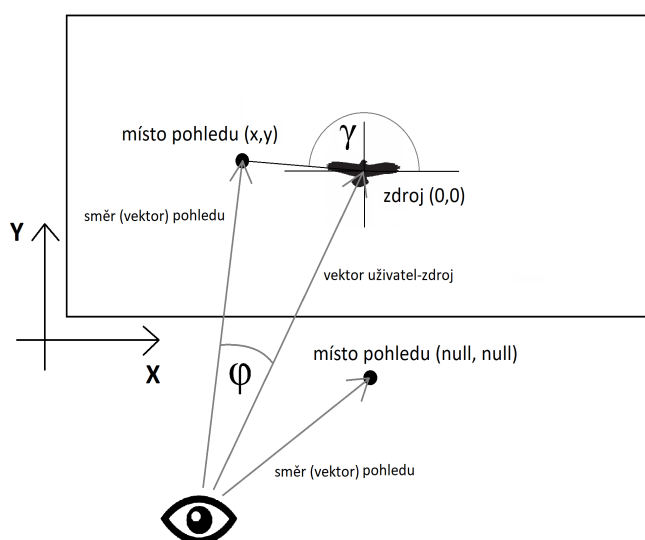
Již jsem určil, jakým způsobem budou data z aplikace zaznamenávána. Tato kapitola se bude zabývat možnostmi, jakými mohou být data do textových dokumentů zapisována. Rozhodoval jsem se mezi dvěma metodami. První metoda je výpis relativní vzdálenosti od zdroje. Zdroj zvuku by se nacházel v počátku kartézské souřadnicové soustavy. Hodnoty jakéhokoliv pohledu mimo rovinu by byly nahrazeny slovy "null". Obrázek 4.2 metodu graficky zobrazuje.



Obrázek 4.2: Metoda 1: Relativní vzdálenost od zdroje

Druhá metoda nevypisuje souřadnice, nýbrž úhly. Počátek v rovině je opět určen polohou zdroje. Úhel γ je svírají vektor směru osy X a vektor od počátku k místu pohledu v rovině. Úhel φ je určen vektorem uživatel-zdroj a směrem (vektorem) pohledu. Stejně jako u první metody by pohled mimo rovinu vypsal hodnoty "null". Grafické znázornění metody se nachází v obrázku [4.3](#).

Problém druhé metody je však v pohybu nositele HMD. Metoda 2 by byla použitelná pouze v případě statického uživatele, což je v rámci tohoto testování nesplnitelná podmínka.



Obrázek 4.3: Metoda 2: Výpis pomocí úhlů

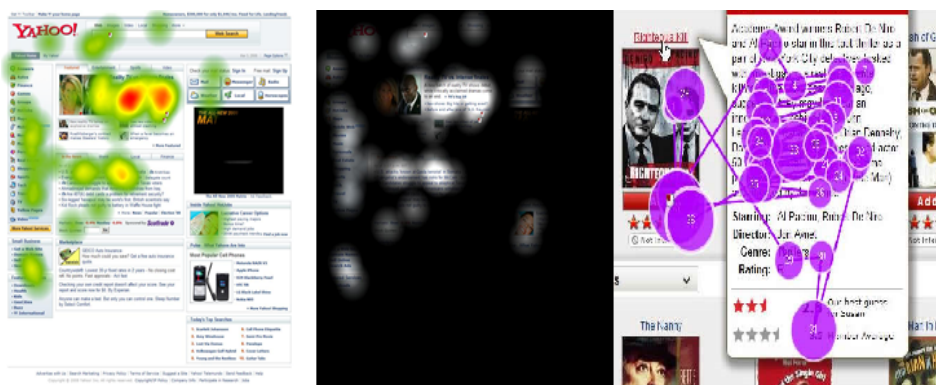
Údaje z první metody společně s časem uplynulým od spuštění aplikace a hlavičkou budou zapisovány do souboru "XZ.txt"¹ do složky s názvem Data. Informace o tom, zda měla testovaná osoba zavřené oči je obsažena v nadměrném čase mezi výpisy dat. Do druhé složky s názvem MATLAB budou vypisována data do dvou dalších souborů "GazeToTable_XZ.txt" a "MultipleTests_XZ.txt" k dalšímu grafickému zpracování.

		63.02384	0.93	0.07
		63.04551	0.91	0.07
Zavření očí na 1,5 s	➔	63.06681	0.87	0.06
		64.57703	0.63	0.15
		64.59992	0.69	0.16
Pohled mimo okolí ptáků	➔	64.71121	null	null
		64.73304	null	null

Obrázek 4.4: Ukázka zápisu dat do souboru typu "XZ.txt"

■ Grafické zobrazení

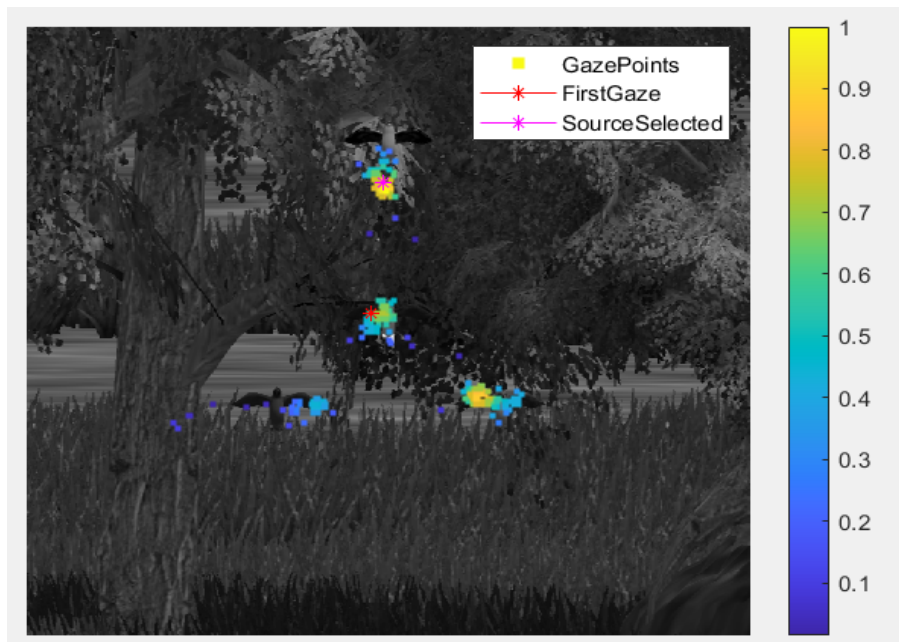
Jak už název složky s daty pro grafické zobrazení napovídá, pro další zpracování jsem vybral program MATLAB [21]. Důvody volby tohoto softwaru jsou jeho rychlé výpočty a relativně snadný programovací jazyk. V rámci programu jsem zvažoval celkem tři možné zobrazovací metody: gaze plot (obr. 4.5), heat map (obr. 4.5), opacity map (obr. 4.5). Protože na pořadí míst, na která se testovaný dívá, nezáleží, vyloučil jsem zobrazení gaze plot. Druhým rozhodovacím faktorem byla informace o prostředí. Opacity map totiž zakrývá různými odstíny černé barvy ta místa, kam se testovaný nedívá. Tím by však došlo k odstranění informace o vzdálenosti jednotlivých objektů. Data jsem se proto rozhodl vykreslovat do grafického zobrazení heat map. Graf jsem proložil obrázkem pozadí daného testu. Obrázek v pozadí jsem zároveň převedl do černobílého podání barev pro vyšší přehlednost.



Obrázek 4.5: Příklady grafických zobrazení v pořadí: heat map, opacity map, gaze plot [22]

¹X nabývá hodnot M (Male), F (Female); Z je celé číslo, jež udává pořadí osoby; Pokud by třetí testovaná osoba byla žena, zápis by byl proveden do souboru F3.txt.

Prostředí MATLAB však nenabízí přímou tvorbu heat mapy z dat souřadnicového typu. Nahradil jsem ji tedy velice podobným grafem hustoty rozptylu (angl. density scatter plot). Pro jeho vykreslení jsem využil funkci `dscatter()` [23], která mým požadavkům vyhovovala nejvíce. Tato funkce body získané z očních kamer vykreslí do grafu a udělí jim barvu podle množství těchto bodů v jejím blízkém okolí (viz Obrázek 4.6). Funkce nabízí možnost upravení velikosti vykreslovaných bodů, jejich podoby (čtverec, kruh, atp.), výplně bodů a způsobu vykreslení dat. Nejvíce mi vyhovovalo základní nastavení, tj. vykreslení plných čtverců se způsobem vykreslení 'scatter'. Tento způsob vykreslení umožňuje, aby bylo pozadí grafu (obrázek) lépe vidět. Z grafu funkce jsem odebral číslování os a přidal barevnou škálu naznačující hustotu bodů, schématickou fotku testu na pozadí, legendu a dva barevné body označující specifická místa. Růžový značí, kam se testovaná osoba dívala v době, kdy stiskla tlačítko ovladače (tj. jaký cíl označila za zdroj). Červený, kam se podívala hned po začátku testu. Červený bod nevzniká hned při prvním snímku každého testu, nicméně zhruba 400 ms po jeho začátku¹. Protože testované osoby nebyly stejně vysoké a nestály na stejném místě, při vykreslování bodů na scénou proložené pozadí dochází k nesrovnalostem. Přeškálování bodů každého testu tak, aby sedělo na pozadí (resp. pozadí na body) je možné. Muselo by se však pro každou osobu a každý test dělat samostatně. Automatické škálování je pak příliš obtížné. Grafické vyobrazení dat proto neslouží k přesnému zobrazení výsledků, pouze jako přibližná dodatková informace.



Obrázek 4.6: Ukázka grafického výstupu pro jeden test jednoho subjektu

¹Tato hodnota byla určena experimentálně v rámci testování a zároveň přibližně odpovídá hodnotě z tabulky 2.1.

4.3 Implementace

V předchozích kapitolách byly teoreticky popsány změny aplikace. Tato kapitola se naopak zaměřuje na praktickou část. Chronologicky zde bude naznačen proces předělání původní (opravené) verze až po verzi použitou pro subjektivní testování.

Nejdříve bylo nutné umožnit Unity, resp. projektu, přístup ke snímání očních pohybů. Proto jsem importoval a nainstaloval package¹ SRanipal, který tuto funkci umožňoval. Do scény, ve které chceme tuto funkci dále využívat bylo nutné přidat objekt "SRanipal Framework". Jak bylo uvedeno v kapitole 4.2.4, pro získání souřadnic místa pohledu byla do všech vybraných testů přidána rovina procházející možnými zdroji zvuku. Poslední úprava v prostředí Unity byla ve změně velikosti colliderů jednotlivých objektů. Jejich novou velikost jsem experimentálně určil během přetváření aplikace.

Po tom, co byl eye tracking umožněn, bylo třeba napsat skript za účelem získání zpřístupněných dat. Jako základ tohoto skriptu jsem použil šablonu [25], poskytovanou na fóru jedním z hlavních vývojářů SRanipal SDK. Tu jsem upravil a dopsal do verze tak, aby bylo možné přistupovat k vektoru kombinovaného pohledu v každém snímku. V skriptu "EyeDataCallback.cs" se dále nachází část, jež vypisuje data o místech pohledu do jednotlivých textových dokumentů. Pro jeho plnou funkčnost bylo třeba upravit i další již vytvořené skripty.

V rámci úprav jsem ostatní skripty upravil tak, aby:

- byla spuštěna kalibrace očních pohybů dříve než bude testovaný vpuštěn do samotné aplikace
- nebyly vidět paprsky vycházející z ovladačů
- cíle byly vybírány s pomocí očního pohledu
- byla vypisovaná data přehlednější (v souboru "XZ.txt" byly přidány řádky oznamující začátek jednotlivých testů) a plně funkční

Jak již bylo zmíněno. Textové výstupy aplikace byly graficky zobrazeny pomocí skriptu "Plotting.m", který je umístěn v příloze B. V rámci získání dodatečné zpětné vazby jsem vytvořil online dotazník pro hodnocení kvality testování. Tento dotazník obsahoval otázky týkající se především výběru cílů očima a celkové náročnosti testování. Na obrázku 4.7 je vidět výstřižek dotazníku. Přehled všech otázek se nachází v příloze B.

¹Package v prostředí Unity je balík dat, který obsahuje různé funkce, nástroje a prostředky, pro zajištění určité funkce [24].

Přišlo Vám, že vybírání cílů zrakem bylo intuitivní? (ANO / NE, protože...)

Text dlouhé odpovědi

Jak moc jist/á jste si byl/a svými odpověďmi?

	1	2	3	4	5	
Čistý tip	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	100%

Bylo označování pomocí zraku přesné? *

- Ano, bez problému
- Ano, jen u vzdálených cílů bylo trochu nepřesné
- Ano, jen u hýbajících se cílů byl trochu problém
- Jiná...

Obrázek 4.7: Výběr otázek z dotazníku hodnocení kvality testování

4.4 Příprava a průběh testování

Za účelem snadnějšího provádění testů byl vytvořen instruktážní PDF dokument. V něm se nacházejí přesné informace o přípravě, průběhu testů a záznamu dat. Dokument je umístěn v příloze A.

Kapitola 5

Výsledky a jejich zhodnocení

Testováno bylo celkem 12 osob. Čtvrtina testovaných (3) byly ženy. Věkové rozmezí všech testovaných bylo přibližně od 20 do 60 let, s mediánem 30 let. První dvě osoby byly podrobeny pre-testu. Byly testovány pouze na základních osmi testech a nevyplňovaly dotazník. Tito první dva testovaní mi poskytli zpětnou vazbu a případné poznámky pro plynulejší průběh následujících testů. Zbýlých 10 testovaných osob bylo podrobena všem 11 vybraným testům. Každý z nich po úspěšném subjektivním testování vyplnil dotazník se zpětnou vazbou. Výsledky všech testů společně s jejich popisy zobrazuje tabulka 5.1.

Tabulka 5.1: Tabulka výsledků testů

NO.	Popis testu	CA	MW	FC	FS
1	8deg angle - 1 source, 1 distractor, side by side	10 (83%)	11 (92%)	8 (67%)	10 (83%)
2	15deg angle - 1 source, 1 distractor, side by side	7 (58%)	9 (75%)	2 (17%)	7 (58%)
3	5deg angle - 1 source, 2 distractors, triangle vertexes	3 (25%)	7 (58%)	5 (42%)	7 (58%)
4	15deg angle - 1 source, 2 distractors, triangle vertexes	7 (58%)	9 (75%)	3 (25%)	4 (33%)
5	8deg angle - 1 source, 3 distractors, triangle vertexes + center	5 (42%)	6 (50%)	9 (75%)	4 (33%)
6	3deg angle - 1 source, 3 distractors, triangle vertexes + center	6 (50%)	6 (50%)	3 (25%)	4 (33%)
7	3deg angle - 1 source, 3 distractors, square	4 (33%)	11 (92%)	0 (0%)	5 (42%)
8	15deg angle - 1 source, 3 distractors, square	4 (33%)	6 (50%)	3 (25%)	2 (17%)
9	1 source, 6 distractors, random omnidirectional movement	4 (40%)			
10	3 sources, 7 distractors, moving distractor audio	3 (30%)			
11	2 sources, 6 distractors, static distractor audio behind the player, with rain	8 (80%)			

NO. = číslo testu

CA (Correct Answers) = počet správných odpovědí

MW (Most Watched) = počet označení cíle, na který se subjekt nejvíce díval

FC (First Correct) = počet prvních pohledů na cíl, které byly správné

FS (First Selected) = počet označení cílů, na které se zároveň podíval jako první

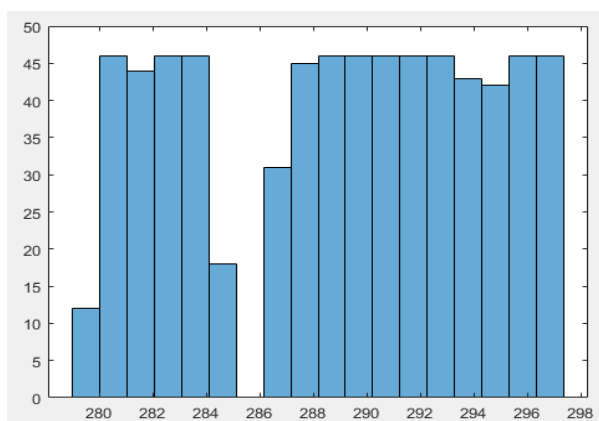
5.1 Zhodnocení výsledků testování

V první řadě je vhodné říct, že aplikace byla funkční i v případech, kdy byla testovaná osoba krátkozraká (celkem 4 z 12 testovaných mělo tuto oční vadu, jež je uváděna v Tabulce 3.2 jako faktor ovlivňující přesnost skenování očních pohybů). Funkčnost aplikace byla zachována i při zahrnutí složitějších testů (čísla 9 - 11). Z těchto poznatků plyne, že je možné užívat zrak jako hodnotící médium. Po kontrole dat jsem navíc zjistil, že byl vždy vybrán ten cíl, na který se testovaná osoba skutečně dívala a který tedy chtěla označit. Přesto existují určité nedostatky tohoto systému, které jsou komentovány v následující kapitole 5.2.

Z popisu jednotlivých testů je jasné, že některé testy budou obtížnější než jiné. Obecně je možné říci, že testy s menším úhlem mezi možnými zdroji jsou obtížnější než testy s úhly většími. Obdobně je to i s množstvím možných zdrojů. Čím více je zdrojů mezi kterými se testovaný rozhoduje, tím obtížnější je určit ten správný. Sloupec správných odpovědí však těmto tvrzením přímo neodpovídá. Červenou barvou jsem označil ty testy, jejichž množství správných odpovědí se blížilo množství testovaných dělený počtem možných zdrojů (tj. číslu, které by odpovídalo situaci, kdyby každý testovaný pouze hádal). Z takto barevně označených polí by bylo možné udělat několik závěrů:

1. Čtvercové rozložení zdrojů dělá lidem při lokalizaci větší problém, než rozložení jiná.
2. Statický rušivý zdroj záporně neovlivní schopnost lokalizace.
3. Je možné, že existují specifické případy testů, které lidem dělají větší problém než jiné.
4. Z důvodu, že nebylo zahrnuto více testů s pohyblivým zdrojem zvuku, není možné z výsledků přímo stanovit, zda je obtížnější lokalizace statického či pohyblivého zdroje.

Tři testovaní si snažili pomoci nakláněním hlavy v různých rovinách. K lepší lokalizaci to však dle výsledků přímo nevedlo. Další metoda, kterou čtvrtina testovaných (3) zkoušela bylo zavírání očí. Se zavřenýma očima člověk soustředí veškerou schopnost lokalizace do sluchu. Pro získání názornější informace o tom, zda a kdy člověk oči zavíral jsem napsal skript "Zavrene_oci.m" (viz Obrázek 5.1). Skript byl psán v prostředí MATLAB a pomocí funkce `histogram()` vykresluje získaná data. Na ose x zobrazuje čas a na ose y množství dat. V místech, kde jsou sloupce nízké, člověk zavřel oči. Ani zavření očí však ze statistického hlediska nepomohlo k lepší lokalizaci.



Obrázek 5.1: Ukázka principu skriptu "Zavrene_oci.m"

Z pohledu na sloupec MW, by se dalo stanovit, že se člověk dívá nejčastěji tam, odkud zvuk vnímá. Je nutné zmínit, že tato data jsou mírně zkreslena.

Člověk se musí dívat na onen cíl, pokud se ho rozhodne označit. Toto zkreslení je ale ve většině případů zanedbatelné, neboť testovaný obvykle mačkal tlačítko ovladače jen chvíli po tom, co se na cíl podíval.

Data týkající se prvních pohledů jsem v tabulce uvedl především v reakci na kapitolu 2.2. V návaznosti na tuto kapitolu by se dalo předpokládat, že první pohledy budou směřovat ke správnému zdroji zvuku. Tento předpoklad data z tabulky nepotvrzují, a tak je možné říci, že lidské sluchové, resp. zrakové ústrojí podvědomě tímto způsobem správný zdroj neurčí. Důvodem je nejspíš příliš malý úhel mezi možnými zdroji. Zajímavé je, že v prvním testu celkem 10 z 12 testovaných označilo cíl, na který se podívalo jako první. S dalšími testy toto množství označení klesá. Tato klesající tendence navozuje dojem, že má tento ukazatel co do činění se zvykem na způsob testování, anebo únavou. Protože ale test trvá maximálně 3 minuty, únavu lze vyloučit. Je tedy možné, že je tento parametr závislý na délce testování. Pokud bych pořadí testů promíchal, tuto možnou domněnku by to vyvrátilo, nebo naopak potvrdilo.

Parametr FC se z tabulky jeví jako zcela náhodný. Je ale jisté, že větší množství testovaných osob by vyloučilo náhodnost, která mohla mít na testování vliv.

Posledním významným faktorem je délka testů, která v tabulce není zobrazena. I když si testovaní obvykle svými odpověďmi nebyli příliš jisti (viz 5.2), délka trvání testů vypovídá o něčem jiném. Testy, jejichž odpověď byla správná trvaly průměrně 11,9 sekund (s průměrnou odchylkou 6 sekund), zatímco testy "nesprávné" trvaly průměrně 14,5 sekund (s průměrnou odchylkou 6,8 sekund). Z toho by mohl plynout závěr, že testovaní u některých testů mohli podvědomě tušit, odkud zvuk vychází. Tento závěr je však diskutabilní, neboť rozdíl 2,6 sekund nemusí být z důvodu vyšší průměrné odchylky dostatečný.

5.2 Zhodnocení dotazníků

Z výsledků dotazníků¹ je jasné, že sledování pohybu očí a jeho následné užití jako média výběru cíle v rámci virtuální reality je intuitivní. Oproti označování cíle ovladačem, resp. paprskem vycházejícím z ovladače je označení velice rychlé (člověk ani nepozná odezvu systému, a tím pádem se stačí na cíl podívat a ihned zmáčknout tlačítko) a snadné (u pohyblivých objektů nemusí člověk obtížně sledovat pohyb objektu ovladačem, ale přirozeně a jednoduše zrakem).

Otázka přesnosti neviditelných paprsků vycházejících z očí je diskutabilní. Jak už bylo uvedeno v závěru kapitoly 4.2.4, pro zachování funkčnosti bylo nutné zvětšit collidery objektů. Již z této úpravy plyne problém překryvu těchto nehmotných okolí. Kdyby se objekty posunuly ještě blíže k sobě, docházelo by ke kolizi a tím pádem i nepřesnosti výběru cíle pohledem. U 70 % testovaných bylo označování přesné, u zbytku docházelo k mírným nepřesnostem způsobeným malou velikostí cílů, anebo jejich pohybem. Zde se střetáváme s dalším problémem. Testovaní byli seznámeni s tím, že nesmějí

¹Přehled všech otázek z dotazníku se nachází v příloze.

hýbat s headsetem po tom, co provedli kalibraci. Při běžném užívání si ale uživatel HMD posouvá tak, aby mu bylo jeho nošení pohodlné. V rámci aplikace fungující na bázi přesných optických vektorů je toto posunutí kritické, neboť prozatímni SDK nenabízí možnost automatické kalibrace během chodu jakékoliv aplikace.

Dalších několik otázek se týkalo náročnosti orientace ve virtuálním akustickém prostoru. Pouze jeden testovaný uvedl, že si byl svým výběrem spíše jist, než že by tipoval. Zbytek tázaných označilo volbu cílů jako úplný tip, převážný tip, anebo volbu 50/50. Po ústní konzultaci s testovanými jsem se dozvěděl, že jako tipy označili i testy, které z tabulky plynou jako většinou úspěšné testy.

V tabulce není uvedena informace o rozdílu v obtížnosti testů se statickými a pohyblivými zdroji. Proto jsem jednu otázku položil i na toto téma. Polovina těch, kteří dělali i dodatkové testy (5) řekla, že označování pohyblivých i statických cílů bylo srovnatelně obtížné, 40 % považuje pohyblivé cíle za snáze lokalizovatelné a jednomu člověku dělalo větší problém lokalizovat pohyblivé cíle.

Z dotazníku lze dále říci, že přítomnost dalších zvuků obvykle (70 %) neztíží výběr zdroje. Tento výsledek však plyne z případů, kdy byly přídavné zvuky srovnatelně hlasité s tím, který měli testovaní označovat. Kdyby byly rušivé zvuky silnější než zvuk označovaných zdrojů, výsledek by byl pravděpodobně jiný.

Poslední otázku jsem směřoval na plynulost aplikace. Dva dotazovaní uvedli, že se aplikace občas na chvíli zasekla. I přes to, že se zbytku aplikace jevila jako plynulá, mohlo by další upravování aplikace vést k častějším a delším přerušením.

Sice se nejedná o softwarový problém samotné aplikace, ale je nutno zmínit nedokonalost zabudovaných sluchátek. I když se jedná o vylepšení oproti předchozí verzi HMD HTC Vive, nejsou tato sluchátka zpracována kvalitně. Parametry sice odpovídají lepším sluchátkům, špatná polohovatelnost však jejich kvalitu silně zhoršuje. Stávalo se tedy, že některým testovaným sluchátka nezakrývala celé uši. Na tuto nedokonalost upozorňovali celkem 4 testovaní (tj. třetina vzorku) bez jakýchkoliv otázek, a tak je nutné zmínit, že zabudovaná nekvalitní sluchátka mohla silně ovlivnit výsledky testování.

5.3 Možné úpravy do budoucna

I přesto, že by aplikace neměla být nikterak náročná, stává se, že v určitých chvílích nedosahuje požadované rychlosti snímků. Tento nedostatek způsobuje, že se během testování aplikace čas od času na krátký časový úsek zasekne. Zaseknutí není kritické, neboť trvá pouze přibližně 30 ms. Aplikaci ovlivňuje pouze takovým způsobem, že testovaná osoba pozná, že obraz nebyl plynulý. Ve vynechaném snímku také nedojde k vypsání žádných dat. Je-li v onom vynechaném snímku volána určitá funkce, daná funkce se neprovede. Pro další postup by tedy bylo vhodné aplikaci optimalizovat, aby tyto situace nenastávaly vůbec. Tato optimalizace by se možná dala vyřešit implementací

funkce foveated rendering, které byla věnována část kapitoly 3.1.

Je možné, že při tvorbě navazující práce již bude k dispozici přesnější snímání očních pohybů. Tato zvýšená přesnost by mohla vést ke zlepšení výsledků testů, především pak v případech, kdy byly cíle pohyblivé anebo byl úhel mezi možnými zdroji malý a kde občas docházelo k nepřesnému označování cílů pohledu.

Posledním návrhem na další vývoj aplikace je propojení této práce s prací M. Novotného, který zkoumá, jaký dopad má na výstupy původní aplikace vložení specifické HRTF daného člověka. Teoreticky by pak měl člověk uvnitř VAS reagovat stejně jako v reálném akustickém prostoru. Reálný výsledek však může být jiný.

Z důvodu snadného přístupu a dalších možných úprav v budoucnu jsem projekt v Unity zanechal ve stavu, kdy je možné provádět úpravy. V opačném případě by se aplikace vyexportovala jako ".exe"spustitelný program.

Kapitola 6

Závěr

Správná orientace ve virtuálním akustickém prostoru (VAS) je jedním ze základních stavebních kamenů věrohodné virtuální reality (VR). Abych zjistil, jak se lidé ve VAS orientují, rozhodl jsem se vylepšit již existující aplikaci V. Jarolímka. Do aplikace jsem přidal možnost snímání očních pohybů a záznamu jejich dat. Zároveň jsem uživatelům umožnil svým pohledem cíle v aplikaci vybírat. Tuto aplikaci jsem otestoval na 12 dobrovolnících z různých věkových kategorií. Data ze subjektivních testů jsem zaznamenal a graficky zobrazil.

Ukázalo se, že zrak je možné využívat za účelem označení statických i pohyblivých cílů. Oproti standardnímu označování cílů ovladačem je tento způsob rychlejší a snazší. Z důvodu nižší přesnosti kamer zajišťujících skenování očí je však třeba dbát, aby objekty volby byly dostatečně velké a daleko od sebe. Při nesplnění těchto podmínek začíná systém selhávat.

V rámci VAS bylo zjištěno, že se člověk nejvíce dívá do místa, odkud zvuk slyší. Mezi místem, kam se člověk při změně zdroje zvuku podívá nejdříve s jeho následným výběrem nevyplývá žádná spojitost.

Dále bylo zjištěno, že testy, na něž bylo odpovězeno správně trvaly průměrně o 2,6 sekund méně, než ty, na něž bylo odpovězeno chybně. Testování ale ve vytvořeném dotazníku označili výběr odpovědi jako převážný tip. Je tedy možné, že lidé dokáží podvědomě určit správný zdroj rychleji, než když odpověď doopravdy hádají.

Ze získaných dat dále plyne, že si někteří testovaní snažili orientaci ve VAS usnadnit natáčením hlavy nebo zavíráním očí. Tyto metody však k lepším výsledkům přímo nevedly. Obecně je také možné říct, že každý člověk reaguje individuálně.

Ze zpětné vazby byl také objeven nedostatek použitého hardwaru. Chyba byla v zabudovaných sluchátkách, resp. jejich omezené nastavitelnosti polohy. Tato nedokonalost mohla vést k silnému ovlivnění získaných dat.

K aplikaci jsem vytvořil přehledný návod pro další experimentátory. V něm se nachází informace o přípravě testování, jeho průběhu a obdržených datech. Společně se stávající verzí aplikace se nachází v příloze A. V příloze B se pak nachází všechna data získaná z testování vč. skriptů pro jejich následné zpracování. V textu byl doporučen návrh na další rozvoj aplikace. Zadáání práce bylo splněno.



Literatura

- [1] PENDIT, U.C. et al. Virtual reality escape room: The last breakout. In: *2017 2nd International Conference on Information Technology (INCIT)* [online]. IEEE, 2017, 1-4 [cit. 2021-11-28]. ISBN 978-1-5386-1431-0. Dostupné z: doi:10.1109/INCIT.2017.8257884
- [2] AQUAPALACE PRAHA [online]. CERYCH-IT David Čerych. [26.4.2022] Dostupné z: <https://www.aquapalace.cz/virtualni-realita-v-toboganech>
- [3] TEÓFILO, Mauro et al. Exploring Virtual Reality to Enable Deaf or Hard of Hearing Accessibility in Live Theaters: A Case Study. ANTONA, Margherita a Constantine STEPHANIDIS, ed. *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments* [online]. Cham: Springer International Publishing, 2018, 132-148 [cit. 2021-11-28]. Lecture Notes in Computer Science. ISBN 978-3-319-92051-1. Dostupné z: doi:10.1007/978-3-319-92052-8_11
- [4] GONZALEZ-MORA, J.L. et al. Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people. In: *2006 2nd International Conference on Information & Communication Technologies* [online]. IEEE, 2006, 837-842 [cit. 2021-11-28]. ISBN 0-7803-9521-2. Dostupné z: doi:10.1109/ICTTA.2006.1684482
- [5] JAROLÍMEK, Viktor. *Virtual Acoustic Space Test using HMD*. Praha, 2019. Diplomová práce. ČVUT v Praze, Fakulta elektrotechnická, Katedra radioelektroniky. 10.1111/1467-9280.00319
- [6] Wikipedia [online]. Wikipedia contributors, 2022. [9.2.2022] Dostupné z: https://en.wikipedia.org/wiki/Head-related_transfer_function
- [7] BRAASCH, J. Modelling of Binaural Hearing. In: BLAUERT, J. *Communication Acoustics* [online]. Springer, Berlin: Heidelberg, 2005. pp. 75-108. ISBN 978-3-540-27437-7
- [8] ABEL, S. M. et al. The effect of aging on horizontal plane sound localization. In: *J Acoust Soc Am* [online]. August 2000, 108(2), pp. 743-752. [6.5.2022]. Dostupné z doi: 10.1121/1.429607.

- [9] GUTTMAN, S. E. et al. Hearing What the Eyes See. In: *Psychological science* [online]. March 2005, 16(3), pp. 228-235. [6.5.2022]. Dostupné z doi: 10.1111/j.0956-7976.2005.00808.x
- [10] CALLAN, A. et al. An fMRI Study of the Ventriloquism Effect. In: *Cerebral cortex* [online]. January 2015, 25(11), pp. 4248-4258. [10.5.2022]. Dostupné z doi: 10.1093/cercor/bhu306
- [11] VROOMEN, J. et al. Sound enhances visual perception: Cross-modal effects of auditory organization on vision. In: *Journal of Experimental Psychology Human Perception & Performance* [online]. November 2000, 26(5), pp. 1583-1590. [9.5.2022]. Dostupné z doi: 10.1037//0096-1523.26.5.1583
- [12] DOWNEY, Paul a Leonard BROSGOLE. Effects of sound localization stimuli on eye-movement reaction time. *Bulletin of the Psychonomic Society* [online]. 1978, 68-70 [cit. 2021-11-28]. ISSN 0090-5054. Dostupné z: doi:10.3758/BF03336767
- [13] Tobii VR [online]. Tobii VR, 2018-2022. [28.11.2021]. Dostupné z: <https://vr.tobii.com/foveated-rendering/>
- [14] HTC Vive [online]. HTC Corporation, 2011-2021. [29.11.2021] Dostupné z: <https://www.vive.com/eu/product/vive-pro-eye/overview/>
- [15] Windows Central [online]. Future US, Inc. [28.11.2021] Dostupné z: <https://www.windowscentral.com/htc-vive-pro-eye-hands-touch-your-eyes-not-your-face>
- [16] HTC Vive [online]. HTC Corporation, 2011-2022. [9.2.2021] Dostupné z: <https://developer.vive.com/resources/hardware-guides/vive-pro-eye-specs-user-guide/>
- [17] Tobii VR [online]. Tobii VR, 2018-2022. [15.3.2022] Dostupné z: <https://vr.tobii.com/developer/>
- [18] Tobii VR [online]. Tobii VR, 2018-2022. [11.4.2022] Dostupné z: <https://vr.tobii.com/sdk/solutions/tobii-g2om/>
- [19] Unity [online]. Unity Technologies, 2022. [26.4.2022] Dostupné z: <https://unity.com/>
- [20] Novotný, M. Vliv HRTF na polohu zdroje zvuku ve VR, Praha 2022, Diplomová práce, předpokládaná obhajoba červen 2022, ČVUT v Praze, Fakulta elektrotechnická, Katedra radioelektroniky
- [21] MATLAB [ONLINE]. The MathWorks, Inc., 1994-2022 . [12.4.2022] Dostupné z: <https://www.mathworks.com/products/matlab.html>
- [22] ROSS, Jim. Eyetracking: Is It Worth It?. In: *UXmatters* [online]. October 19, 2009. [13.4.2022] Dostupné z: <https://www.uxmatters.com/mt/archives/2009/10/eyetracking-is-it-worth-it.php>

- [23] HENSON, Robert. Flow Cytometry Data Reader and Visualization. In: *The MathWorks, Inc.* [online]. April 13, 2022. [13.4.2022] Dostupné z: <https://www.mathworks.com/matlabcentral/fileexchange/8430-flow-cytometry-data-reader-and-visualization?focused=6779476&tab=function>
- [24] Unity [online]. Unity Technologies, 2021. [6.5.2022]. Dostupné z: <https://docs.unity3d.com/Manual/Packages.html>
- [25] CORVUS. VIVE Eye Tracking at 120hz. In: *VIVE Developer Forum* [online]. 12 February 2021. [6.5.2022]. Dostupné z: <https://forum.vive.com/topic/9341-vive-eye-tracking-at-120hz/>



Příloha A

- *UNITY_VAS* je složka obsahující celý Unity projekt, tj. upravenou verzi aplikace.
- *Instrukcni_dokument* je PDF dokument obsahující podrobnou instrukci k obsluze testovací aplikace.



Příloha B

- *Data* je složka obsahující data o pohybech očí a výsledcích testů.
- *MATLAB* je složka s daty a skripty pro následné grafické zpracování.
- *Dotaznik* je PDF dokument s výpisem všech otázek z vyplňovaného dotazníku.