VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

TELEPRESENCE V MICROSOFT HOLOLENS

TELEPRESENCE IN MICROSOFT HOLOLENS

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE **ONDŘEJ ŠEBESTA**
AUTHOR

VEDOUCÍ PRÁCE **Ing. DANIEL BAMBUŠEK**
SUPERVISOR

BRNO 2022

Vysoké učení technické v Brně
Fakulta informačních technologií

Ústav počítačové grafiky a multimédií (UPGM) Akademický rok 2021/2022

Zadání bakalářské práce

Student: **Šebesta Ondřej**
Program: Informační technologie
Název: **Telepresence v Microsoft HoloLens**
Telepresence in Microsoft HoloLens
Kategorie: Uživatelská rozhraní

Zadání:

1. Prostudujte koncept rozšířené reality a seznamte se s brýlemi Microsoft HoloLens a hloubkovou kamerou Azure Kinect.
2. Prozkoumejte možnosti realizace telepresence s využitím hloubkových a obrazových dat z Kinectu a brýlí HoloLens.
3. Vyberte vhodné metody a nástroje a navrhněte aplikaci, která umožní vzdáleně přenést snímanou osobu zařízením Kinect a vizualizovat v brýlích Microsoft HoloLens za účelem virtuální telepresence.
4. Navrženou aplikaci implementujte.
5. Proveďte experimenty a vyhodnoťte vlastnosti výsledného řešení.
6. Vytvořte video prezentující klíčové vlastnosti výsledného řešení.

Literatura:

- SCHMALSTIEG Dieter, HÖLLERER Tobias. *Augmented Reality: Principles and Practice.* Addison-Wesley, 2016. ISBN 978-0321883575.
- BRADSKI Gary R., KAEHLER Adrian. *Learning OpenCV: Computer Vision with the OpenCV Library*, ISBN 978-0-596-51613-4, September 2008.
- Dále dle pokynů vedoucího.

Pro udělení zápočtu za první semestr je požadováno:

- Body 1, 2, 3 a rozpracovaný bod 4.

Podrobné závazné pokyny pro vypracování práce viz https://www.fit.vut.cz/study/theses/

Vedoucí práce: **Bambušek Daniel, Ing.**
Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.
Datum zadání: 1. listopadu 2021
Datum odevzdání: 11. května 2022
Datum schválení: 1. listopadu 2021

Abstrakt

Digitalizace světa umožnila mnoha lidem pracovat pohodlně z domova. Nepříznivý vývoj epidemie koronaviru a s ním spojené restrikce zapříčinily nárůst počtu takto pracujících lidí, což způsobilo zvýšený zájem o nástroje umožňující vzdálenou komunikaci. Tato práce se zabývá dalším potencionálním stupněm vzdálené komunikace – telepresencí. Cílem této práce je navrhnout a vytvořit systém telepresence v rozšířené realitě. Osoba, účastnící se vzdálené komunikace, je snímána zařízením Azure Kinect, který obrazová data převádí kabelem do počítače. V počítači jsou data zpracována, vykreslena a pomocí funkcionality Holographic Remoting přenesena přes lokální síť do brýlí Microsoft Hololens 2.

Abstract

Digitizing the world allows many people to work comfortably from home. The unfavorable development of the coronavirus epidemic and the associated restrictions caused an increase in the number of people working this way, which has also meant an increased interest in tools enabling remote communication. This work deals with another potential stage of the remote communication - telepresence. The aim of this work is to design and create the telepresence system in augmented reality. The person involved in the remote communication is captured by the Azure Kinect device, which transfers the image data to the computer through a cable. The data is processed, plotted on the computer and transferred into the Microsoft Hololens 2 glasses via the local network using the Holographic Remoting functionality.

Klíčová slova

Rozšířená realita, Telepresence, Azure Kinect DK, Microsoft Hololens

Keywords

Augmented reality, Telepresence, Azure Kinect DK, Microsoft Hololens

Citace

ŠEBESTA, Ondřej. *Telepresence v Microsoft HoloLens*. Brno, 2022. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Daniel Bambušek

Telepresence v Microsoft HoloLens

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana Ing. Daniela Bambuška. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....
Ondřej Šebesta
11. května 2022

Poděkování

Rád bych touto cestou poděkoval mému vedoucímu Ing. Danielovi Bambuškovi za jeho odborné rady, trpělivý přístup a mimořádnou ochotu při vedení této práce. Také bych chtěl poděkovat ústavu počítačové grafiky a multimédií za zapůjčení potřebného vybavení k vývoji a otestování výsledného systému. V neposlední řadě bych chtěl vyjádřit poděkování mé rodině, která mi byla oporou po celou dobu studia.

Obsah

1 Úvod **3**

2 Rozšířená realita a telepresence **4**
2.1 Metody zobrazování rozšířené reality 4
2.2 Zobrazovací zařízení pro rozšířenou realitu 8
2.3 Uživatelská rozhraní pro rozšířenou realitu 10
2.4 Telepresence 13

3 Microsoft HoloLens 2 **14**
3.1 Hardware 14
3.2 Ovládání a interakce 15
3.3 Prostorové mapování 18
3.4 Srovnání s Microsoft HoloLens 1 19

4 Zpracování obrazových dat **21**
4.1 Azure Kinect DK 21
4.2 Srovnání RGBD zařízení 22
4.3 3D reprezentace obrazových dat 23

5 Návrh řešení **24**
5.1 Analýza existujících řešení 24
5.2 Architektura systému telepresence 27
5.3 Návrh uživatelského rozhraní 28
5.4 Zpracování a vykreslování obrazových dat 28
5.5 Bezdrátový přenos dat 29

6 Implementace **32**
6.1 Použité technologie 32
6.2 Zpracování obrazů a vykreslení mračna bodů 34
6.3 Bezdrátový přenos do brýlí HoloLens 36
6.4 Implementace uživatelského rozhraní 36

7 Testování a vyhodnocení **38**
7.1 Měření 38
7.2 Uživatelské testování 40
7.3 Možná rozšíření 40

8 Závěr **43**

Literatura **44**

Kapitola 1

Úvod

Globalizace spolu s rozvojem informačních technologií umožnily lidem pracovat vzdáleně. Důležitou součástí pracovního života se staly nástroje pro vzdálenou komunikaci, jakými jsou například Zoom[1], FaceTime[2] nebo Skype[3]. Epidemie koronaviru COVID-19 a s ní spojená protiepidemiologická opatření způsobila nebývalý nárůst lidí pracujících z domova a poptávka po těchto nástrojích tak vzrostla ještě více.

Kombinací zařízení umožňujícího snímat hloubková a barevná obrazová data se zařízením, které využívá trojrozměrný prostor pro vizualizaci volumetrických obrazů, lze posunout hranice telekomunikace o stupeň dále. Výsledkem této práce je vzdálený přenos obrazových dat, snímaných zařízením Azure Kinect, skrze počítačovou síť a jejich následné zobrazení v brýlích Microsoft HoloLens při zachování snímkovací frekvence blížící se třiceti snímkům za vteřinu, což je maximální snímkovací frekvence zařízení Azure Kinect. Vytvořený systém telepresence v rozšířené realitě lze využít jako demonstrátor na prezentačních akcích a výsledky experimentů mohou posloužit jako odrazový bod při implementaci podobného řešení.

Kapitola 2 popisuje problematiku rozšířené reality, metody zobrazování rozšířené reality, zobrazovací zařízení a uživatelská rozhraní pro rozšířenou realitu. Kapitola 3 představuje brýle pro rozšířenou realitu Microsoft HoloLens 2. Poslední kapitolou v teoretické části je kapitola 4, která se věnuje zařízení Azure Kinect, jeho alternativám a možnostem při zpracování obrazových dat.

První kapitolou v praktické části je kapitola 5, která popisuje existující řešení, návrh systému telepresence a návrh uživatelského rozhraní. Další kapitolou praktické části je kapitola 6 popisující programové řešení a kapitola 7, ve které jsou zhodnoceny výsledky této práce.

[1] `https://zoom.us/`

[2] `https://support.apple.com/cs-cz/HT204380`

[3] `https://www.skype.com/cs/`

Kapitola 2

Rozšířená realita a telepresence

Rozšířená realita, z anglického *Augmented Reality (AR)*, je technologie, která umožňuje zobrazit počítačem vytvořenou informaci do reálného prostředí [39]. Podle Kippera a Rampoly [25] lze kromě vizuálních informací využít také jiné formy informací a zacílit tak na další smysly uživatele, jakými jsou čich, chuť, hmat nebo sluch. Azuma [2] publikoval ve svém výzkumu tři základní vlastnosti, které by měla rozšířená realita splňovat:

1. AR by měla kombinovat reálné a virtuální.
2. AR by měla být interaktivní v reálném čase.
3. AR by měla být registrována ve 3D.

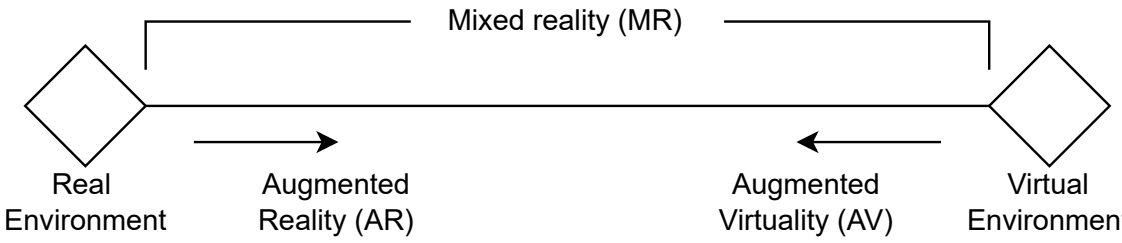
První zmíněná vlastnost vymezuje rozšířenou realitu vůči virtuální realitě (VR). Zatímco cílem AR je umožnit koexistenci reálného a virtuální světa, cílem VR je vytvořit virtuální svět, do kterého je uživatel ponořen a ve kterém je od reálného světa izolován. Zhigeng a spol. [36] definují virtuální realitu jako využití systémů založených na počítačové grafice v kombinaci s různými zobrazovacími zařízeními a rozhraními k poskytnutí efektu ponoření v interaktivním trojrozměrném, počítačem vytvořeném, prostředí.

Milgram a Kishino [32] považují virtuální a rozšířenou realitu za opačné konce téhož kontinua. V roce 1994 představili reálně-virtuální kontinuum, které je zobrazeno na obrázku 2.1. Levá strana kontinua reprezentuje reálné prostředí s reálnými objekty. Pravá strana kontinua naopak reprezentuje virtuální prostředí s virtuálními, počítačem vytvořenými, objekty.

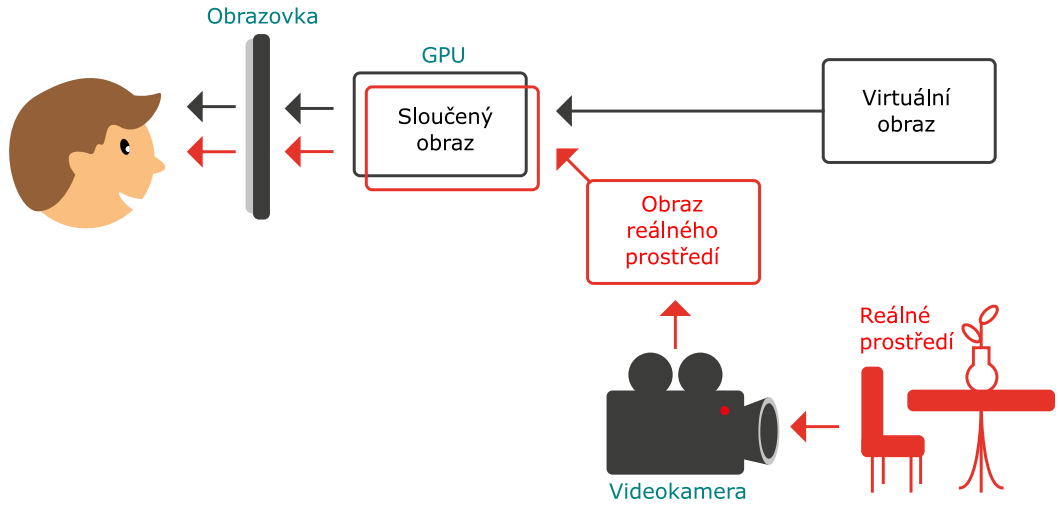
Na základě reálně-virtuálního kontinua Milgram a Kishino [32] definují smíšenou realitu. Smíšená realita se podle nich nachází kdekoli mezi opačnými konci kontinua a je prostředím, ve kterém jsou prvky virtuálního a reálného světa prezentovány společně v rámci jednoho zobrazovacího zařízení.

2.1 Metody zobrazování rozšířené reality

Pro zobrazení rozšířené reality existují tři základní metody – *video see-through*, *optical see-through* a prostorová projekce. Každá z těchto metod přistupuje k zobrazení rozšířené reality jinak a liší se také v použitém hardwarovém vybavení. V následujících odstavcích jsou tyto metody definovány včetně jejich silných a slabých stránek.



Obrázek 2.1: Obrázek reálně-virtuálního kontinua [32].



Obrázek 2.2: U metody *video see-through* je k pořízení obrazu reálného prostředí využita videokamera. Obraz z videokamery je pomocí grafické výpočetní jednotky upraven a rozšířen o počítačem vytvořený obraz. Takto upravený obraz je následně prezentován uživateli na obrazovce [39].

Video see-through (VST)

Metoda *video see-through* vytváří rozšířenou realitu zcela elektronicky. Obraz reálného prostředí je zaznamenán videokamerou a následně přenesen do grafického procesoru. Grafický procesor spojí obraz reálného prostředí s počítačem vytvořeným obrazem. Obraz reálného prostředí je obvykle použit jako pozadí, které je překryto počítačem vytvořeným obrazem. Takto sloučený obraz je poté zobrazen uživateli na obrazovce [39]. Schéma fungování metody je zobrazeno na obrázku 2.2. Praktická aplikace metody je zobrazena na obrázcích 2.4 a 2.3.

Synchronizace a registrace obrazu reálného prostředí s obrazem počítačem vytvořeným je u této metody jednodušší než u metody *optical see through* [22]. U metody *video see-through* lze garantovat registraci obrazu reálného prostředí s obrazem počítačem vytvořeným, avšak na úkor nesouladu mezi pohledem uživatele a jeho propriocepcí [38]. Výhodou metody je jistá míra flexibility, jelikož výměnou videokamery nebo čočky lze vylepšit zorné pole a rozlišení obrazu reálného prostředí. Komerční výhodou této metody je dostupnost hardwaru, která je způsobena dostupností mobilních telefonů, tabletů a počítačů.

Kompozice reálného a počítačem vytvořeného obrazu může vést k nesprávnému vyhodnocení okluze mezi reálnými a virtuálními objekty. Tento problém vzniká překrytím počítačem vytvořeného obrazu nad obrazem reálného prostředí bez znalosti hloubkových

Obrázek 2.3: Ruční zobrazovací zařízení pro rozšířenou realitu – mobilní telefon. V levé části obrázku je zobrazena virtuální raketa, která je umístěna v reálném prostředí. V pravé části obrázku je virtuální lampa umístěna na stůl v reálném prostředí.
Zdroj: `https://www.apple.com/cz/augmented-reality/`

Obrázek 2.4: Náhled barvy pokoje v rozšířené realitě [46].

informací z reálného prostředí. Problém lze vyřešit dodatečným přidáním hloubkového senzoru a získaná hloubková data z reálného prostředí následně testovat oproti počítačem vytvořené scéně, která již hloubková data zpravidla obsahuje [5].

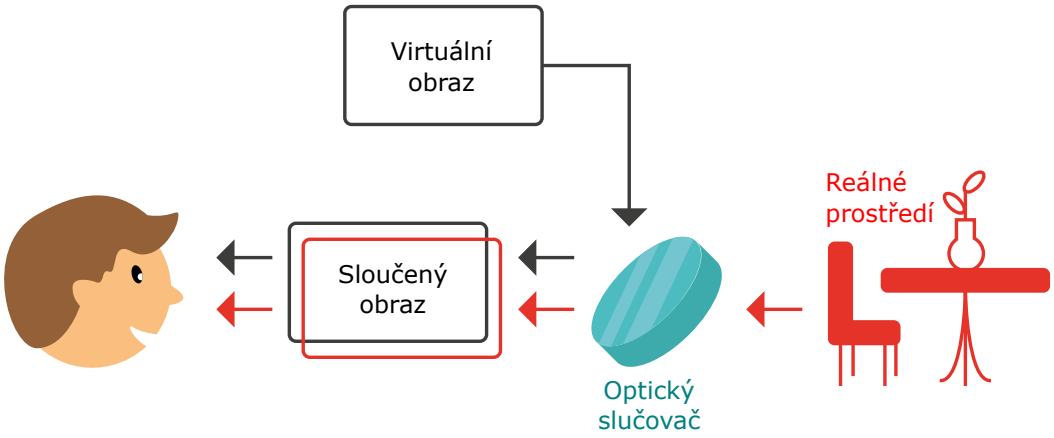
Obraz reálného prostředí je zaznamenáván pomocí videokamery, která má omezené rozlišení a zorné pole, což může vést ke zkreslení pohledu uživatele na reálné prostředí. Dalším problémem u metody *video see-through* je její závislost na elektrické energii, což má při výpadku elektrické energie za následek ztrátu vizuálního kontaktu uživatele s okolím [26]. Dalším problémem je zpoždění obrazu reálného prostředí a lišící se umístění videokamery vůči pozici očí uživatele [5].

Optical see-through (OST)

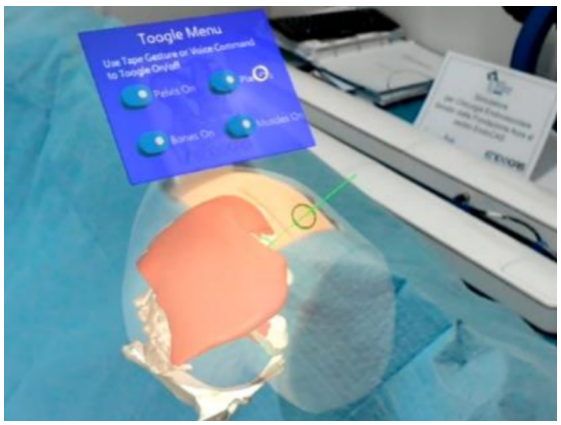
Narozdíl od metody *video see-through*, tato metoda zachovává přímý, nezprostředkovaný pohled na reálné prostředí. K zobrazení rozšířené reality jsou u této metody využívány poloprůhledné povrchy, které propouštějí dostatečné množství světla k tomu, aby uživatel viděl reálné prostředí kolem sebe a současně odrážejí dostatečné množství světla k zobrazení počítačem vytvořeného obrazu [53]. Schéma fungování metody *optical see-through* je zobrazeno na obrázku 2.5. Praktická aplikace metody je zobrazena na obrázcích 2.6 a 2.7.

Nezprostředkovaný pohled na reálné prostředí má mnoho výhod. Obraz reálného prostředí je narozdíl od metody *video see-through* bez časové prodlevy, v plném rozlišení a bez posunutí pohledu uživatele. Tyto vlastnosti mohou být důležité v odvětvích náročných na bezpečnost, jakými jsou zdravotnictví nebo armáda [5]. Pro zobrazení obrazu reálného prostředí není potřeba elektrická energie a při jejím výpadku uživatel neztrácí vizuální kontakt se svým okolím.

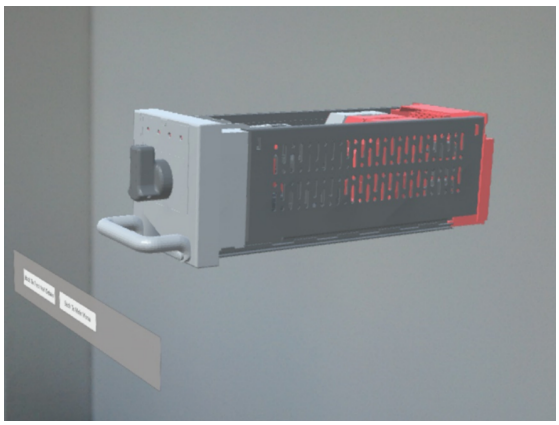
Oproti metodě *video see-through* je u této metody méně přesná registrace reálného a počítačem vytvořeného obrazu. V mnoha případech vyžadují zobrazovací zařízení využívající metodu *optical see-through* manuální kalibraci, která vede k horší kvalitě registrace obrazů



Obrázek 2.5: U metody *optical see-through* je využit optický prvek ke spojení uživatelova pohledu na reálné prostředí s počítačem vytvořeným obrazem [39].



Obrázek 2.6: Simulátor artroplastiky kyčelního kloubu. Snímek zobrazuje virtuální anatomii kyčelního kloubu v brýlích Microsoft HoloLens [45].



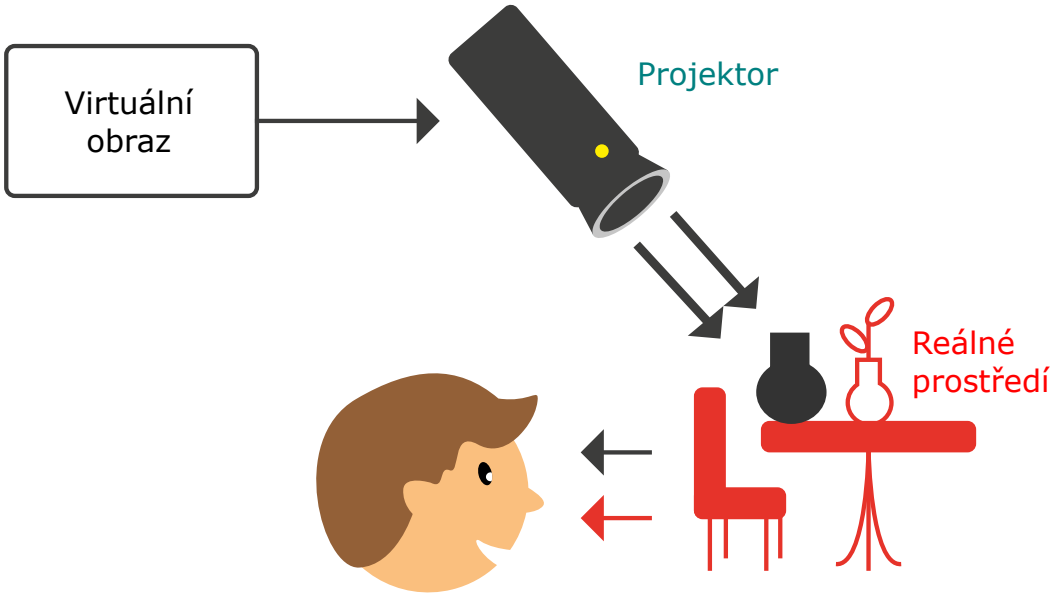
Obrázek 2.7: Snímek zobrazuje 3D model průmyslového zařízení v brýlích Microsoft HoloLens [1].

v kontrastu s automatickou kalibrací metody *video see-through*, která je založena na algoritmech počítačového vidění. Dalším problémem metody je časová prodleva mezi okamžitým obrazem reálného prostředí a zpožděným virtuálním obrazem [5].

Prostorová projekce

Prostorová projekce odděluje zobrazovací technologii od uživatele a integruje ji do prostředí [8]. Systémy založené na prostorové projekci využívají světelné projektory, které promítají virtuální obraz přímo na objekty reálného prostředí. Promítání virtuálního obrazu na fyzické objekty je formou optické kombinace, která, narozdíl od předchozích metod, nevyžaduje použití optického slučovače nebo elektronického displeje [39]. Schéma fungování metody je zobrazeno na obrázku 2.8. Praktická aplikace metody je zobrazena na obrázcích 2.9 a 2.10.

Světelné projektory jsou obvykle připevněny na stěně nebo na stropu místnosti, což omezuje jejich využití na aplikace, které nevyžadují mobilitu. Dalším omezením je nutnost



Obrázek 2.8: U metody prostorové projekce je počítačem vytvořený obraz promítán přímo na objekty reálného prostředí. Narozdíl od předchozích metod není u této metody vyžadován žádný slučovací prvek [39].

mít objekty, na které je promítán obraz, v blízkosti projektoru. Z tohoto důvodu je prostorová projekce nepoužitelná ve venkovním prostředí, kde mohou být objekty od projektoru vzdálené desítky metrů. Další nevýhodou je citlivost projektoru na světelné podmínky – například stín uživatele může promítaný obraz zcela znehodnotit. Výhodou této metody je, že uživatel nemusí pro zobrazení rozšířené reality nic nosit [5].

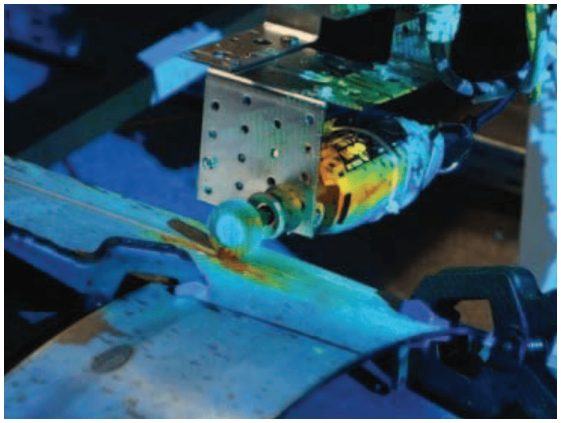
Problémem prostorové projekce je její závislost na vlastnostech povrchu, na který je promítán počítačem vytvořený obraz – pro prostorovou projekci je ideální světlý difúzní povrch s hladkou geometrií. Dalším faktorem, ovlivňujícím kvalitu promítaného obrazu, je ambientní světlo, které může ovlivňovat kontrast obrazu, což limituje použití této metody do prostředí s kontrolovaným osvětlením [37]. Heinrich a spol. [17] porovnávali metody používané k zobrazování rozšířené reality pro podporu zavádění lékařské jehly. Zjistili, že oproti ostatním zobrazovacím metodám byla projekce zkreslena z důvodu povrchových deformací a měla horší kontrast a ostrost vizualizace. Při experimentování také zjistili, že stín ruky uživatele často způsobuje okluzi s promítaným obrazem.

2.2 Zobrazovací zařízení pro rozšířenou realitu

Zařízení pro zobrazení rozšířené reality jsou rozdělena do třech hlavních kategorií – ruční zobrazovací zařízení, obrazovky připevněné k hlavě a prostorové projektory. V následujících odstavcích jsou tyto kategorie popsány podrobněji.

Ruční zobrazovací zařízení

Ruční zobrazovací zařízení, z anglického *handheld displays*, jsou komerčně nejdostupnější skupinou zařízení pro zobrazování rozšířené reality. Do této skupiny zařízení patří mobilní



Obrázek 2.9: Promítání tepelných dat na fyzické objekty během obrábění [9].

Obrázek 2.10: Projekce rozšířené reality na pracovní stůl za účelem programování kolaborativního robota [4].

telefony, tablety a dnes již ne tak běžní osobní digitální asistenti (PDA[1]). Tato skupina zařízení využívá pro zobrazování rozšířené reality výhradně metodu *video see-through*. Podle Billinghursta [5] mají dnešní mobilní telefony a tablety dostatečný výkon a obsahují všechny potřebné senzory k běhu aplikací využívajících rozšířenou realitu. Současně uvádí, že tato zařízení jsou již plně integrována do běžného života uživatelů.

Rozšířená realita v ručních zařízeních již našla své uplatnění v mnoha oblastech. Typickou ukázkou využití rozšířené reality v herním průmyslu je hra *Pokémon GO*[2] od amerického vývojáře Niantic. V této hře může uživatel umístit model Pokémona do reálného prostředí a následně jej s pomocí *pokéballu*[3] chytat. Dalším zajímavým využitím je aplikace pro náhled barvy na stěně [46] (viz obrázek 2.4), která umožňuje v reálném čase překrýt reálnou barvu pokoje barvou virtuální. Uživatel si tak může vytvořit představu o budoucím vzhledu pokoje i bez nutnosti jej fyzicky vymalovat.

Náhlavní displeje

Náhlavní displeje, z anglického *head-mounted displays*, zobrazují virtuální obraz přímo před očima uživatele, což vede k tomu, že se žádný objekt z reálného prostředí nedostane mezi oči uživatele a virtuální obraz. Tím lze garantovat, že virtuální obraz nebude zakryt objekty fyzického prostředí. Ačkoli je pro náhlavní displeje využívána obvykle metoda *video see-through* a metoda *optical see-through*, existují výzkumy, které se zabývají také využitím prostorové projekce připevněné k hlavě [5].

Zařízení, která spojují rozšířenou realitu s náhlavními displeji, jsou v dnešní době již komerčně dostupná. Metoda *optical see-through* je využita například u brýlí Microsoft HoloLens první a druhé generace.

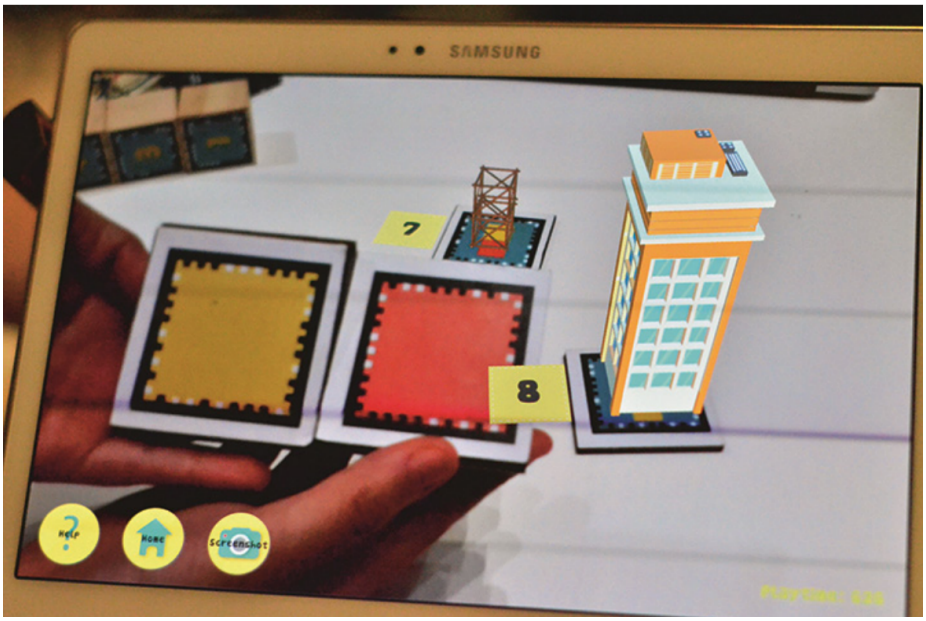
Projektory

Ve srovnání s předchozími zobrazovacími zařízeními jsou prostorové projektory obvykle určeny pro nemobilní aplikace a jsou obvykle instalovány na pevno na jedno místo. Prostorové

[1] PDA je zkratka pro anglické slovní spojení Personal Digital Assistant.

[2] `https://pokemongolive.com/`

[3] Pokébal je předmět ve hře Pokémon GO, využívaný pro chytání pokémonů.



Obrázek 2.11: Hmatatelné uživatelské rozhraní. Manipulací s reálnými kartičkami mění uživatel vlastnosti virtuálního objektu – barvu budovy [54].

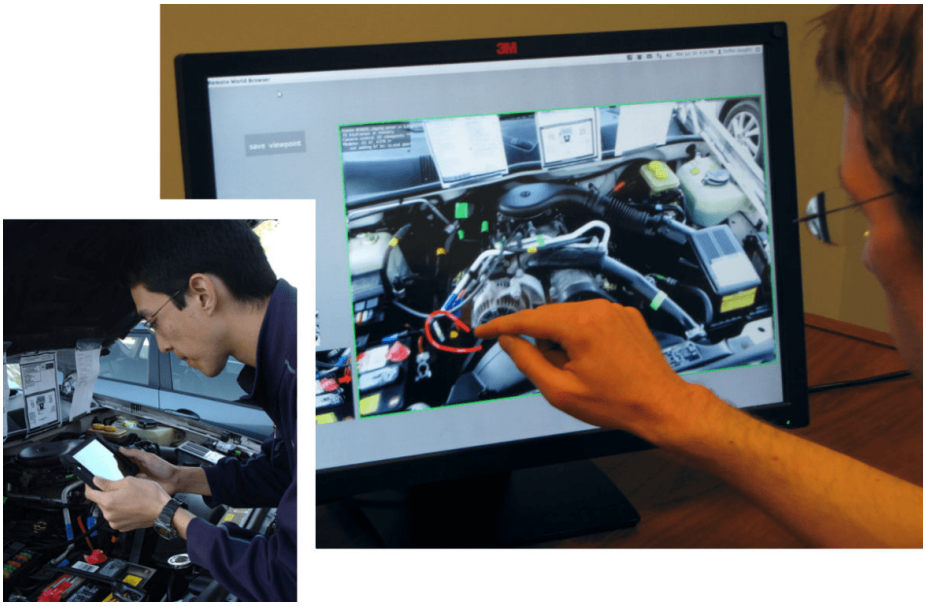
projektory obvykle poskytují větší obraz a jsou tak využitelné jako veřejné displeje, které sdílí více uživatelů [5]. Narozdíl od ručních a náhlavních zobrazovacích zařízení, u prostorových projektorů uživatel nepotřebuje mentálně přenést informaci z vizualizačního zařízení do reality [9].

2.3 Uživatelská rozhraní pro rozšířenou realitu

Uživatelská rozhraní jsou neméně důležitým odvětvím rozšířené reality. Cílem rozhraní je poskytnutí bezproblémové interakce mezi člověkem a počítačem. Uživatelská rozhraní jsou v této práci rozdělena do pěti hlavních kategorií – hmatatelné uživatelské rozhraní, kolaborativní uživatelské rozhraní, virtuální uživatelské rozhraní, multimodální uživatelské rozhraní a rozhraní *augmented paper*.

Hmatatelné uživatelské rozhraní

Hmatatelné rozhraní, z anglického *Tangible interface*, je rozhraní pro rozšířenou realitu, které využívá hmatatelných fyzických objektů k interakci s virtuálními objekty. Každý virtuální objekt musí být registrován na fyzický objekt. Uživatel manipulací s fyzickými objekty manipuluje s virtuálními objekty [6]. Fitzmaurice a Buxton [12] považují hmatatelné rozhraní pro rozšířenou realitu za velmi intuitivní, protože fyzické objekty jsou namapovány jedna ku jedné k operacím s virtuálními objekty. Na obrázku 2.11 je zobrazena aplikace využívající hmatatelné uživatelské rozhraní ke změně barvy virtuálního objektu. Hmatatelné uživatelské rozhraní je využíváno ve vzdělávání [35] nebo k navigaci [33].



Obrázek 2.12: Kolaborativní uživatelské rozhraní. Vzdálený uživatel (vpravo) označí prvek v prostředí tím, že kolem něj nakreslí obrys. Anotace je registrována v prostoru a zobrazena na tabletu v rozšířené realitě lokálního uživatele (vlevo) [13].

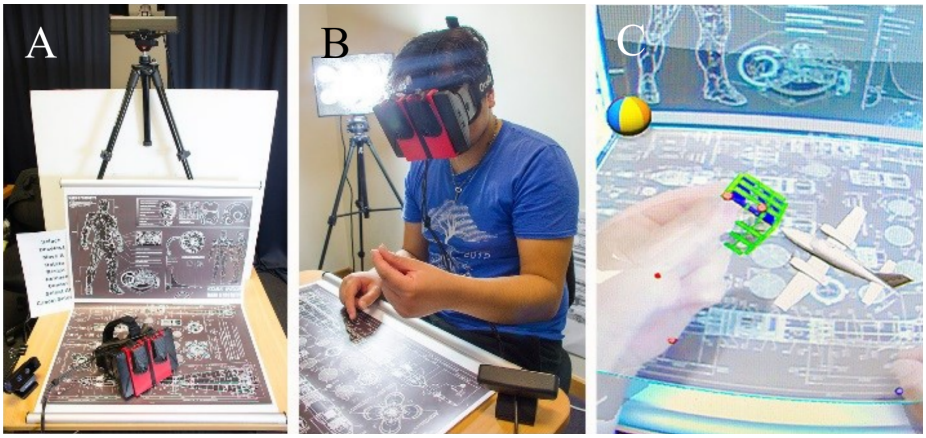
Kolaborativní uživatelské rozhraní

Kolaborativní uživatelské rozhraní využívá několik zobrazovacích zařízení a umožňuje uživatelům přistupovat ke sdílenému trojrozměrnému prostoru, rozšířenému o virtuální objekty, k dosažení počítačem podporované spolupráce [7]. Ačkoli se kolaborativní rozšířená realita využívá především pro spolupráci uživatelů nacházejících se ve stejném fyzickém prostoru [49, 52], existují výzkumy zabývající se vzdálenou spoluprací [13]. Na obrázku 2.12 je zobrazena vzdálená spolupráce v rozšířené realitě.

Wagner a spol. [49] porovnal efektivitu učení při využití počítačové obrazovky, papíru a ručního zařízení pro kolaborativní rozšířenou realitu. Tématem, na kterém byla studie provedena, byla historie umění prezentovaná v muzeu. Pro porovnání byla implementována aplikace Virtuoso pro každé médium zvlášť. Dle výsledků studie, uživatelům připadala rozšířená realita zábavnější a nejvíce přínosná pro aplikaci v muzeích. Na druhou stranu, nejjednodušší ovládání poskytovala verze pro počítač.

Virtuální uživatelské rozhraní

Jedním z přímých přístupů pro podporu interakce s virtuálními objekty je využití tradičních 3D technik uživatelského rozhraní [5]. Řešení známá z počítačových nebo mobilních uživatelských rozhraní mohou být použita také v rozšířené realitě. Výhodou je, že většina uživatelů umí s těmito rozhraními pracovat. Umístění virtuálních dotykových widgetů na povrchy reálného prostředí je pohodlným způsobem, jak přidat komplexní rozhraní k zážitku rozšířené reality. Prvky virtuálního uživatelského rozhraní jsou umístěny na povrchy fyzických objektů – například zeď, stůl nebo na ruku uživatele [39].



Obrázek 2.13: Multimodální uživatelské rozhraní. Uživatel využívá kombinaci gest a hlasových příkazů k manipulaci s virtuální Rubikovou kostkou [54].

Rozhraní Augmented paper

Papír je důležitou součástí každodenního života. Zatímco při práci na počítači je papír simulován a oddělen od reálného papíru, rozhraní pro rozšířenou realitu *augmented paper* papír doplňuje [39]. Wellner [50] argumentoval o interakci mezi fyzickým a virtuálním dokumentem. Ve své práci představil systém DigitalDesk, který umožňuje snímání papíru na stole pomocí kamery, zatímco pomocí projektoru jsou na papír promítány informace a nástroje, které usnadňují práci s dokumentem.

Grubert a spol. [14] prozkoumali využití rozšířené reality na mobilních telefonech v kombinaci s tištěnými plakáty. Bylo zjištěno, že uživatelé jsou obvykle ve spěchu, když přistupují k plakátu, a proto by měl obsah rozšířené reality setrvat i v případě následného vzdálení od papírového média. Na principu rozšíření papírového média funguje také aplikace Applaydu[4], kterou představila společnost Kinder. Uživatel si může prohlížet příbalový leták k hračce z Kinder vajíčka i bez použití digitálních technologií. Jakmile však v mobilní aplikaci *Applaydu* naskenuje QR kód z letáčku, objeví se 3D postavička, se kterou může uživatel dále manipulovat.

Multimodální uživatelské rozhraní

Multimodální uživatelské rozhraní je takové, které kombinuje dvě a více modalit. Modalitou je nazýván každý nezávislý komunikační kanál, který zprostředkovává komunikaci mezi člověkem a počítačem. Vizuálními modalitami jsou například rozpoznávání gest nebo rozpoznávání výrazu ve tváři. Zvukovými modalitami jsou rozpoznávání hlasových příkazů nebo rozpoznávání řečníka na základě jeho hlasu. Senzorickými modalitami jsou počítačová myš, klávesnice, pera, haptické senzory nebo *joystick*. Dotykovou modalitou je rozpoznávání dotyků – jednoho nebo více v jednom okamžiku [47, 34].

Podle Martina a spol. [31] lze modality dále rozdělit dle pozornosti uživatele. Vědomé používání modality z ní činí aktivní modalitu, naopak při nevědomém používání modality je modalita označena za pasivní. Pokud uživatel využívá gesta k ovládání virtuálního objektu, pak tak činí vědomě, a proto je modalita aktivní. Jestliže je během tohoto ovládání snímána

[4] `https://www.kinder.com/cz/cs/applaydu`

uživatelova poloha pomocí GPS, je toto snímání považováno za pasivní modalitu, jelikož nevyžaduje uživatelovu pozornost.

2.4 Telepresence

Pojem telepresence vychází z definice pojmu *presence* neboli přítomnosti. Přítomnost je definována jako prožitek vlastní zkušenosti ve fyzickém prostředí. Přítomnost se nevztahuje přímo na okolí člověka, ve kterém se fyzicky nachází, ale je subjektivním vnímáním tohoto okolí. Subjektivní vnímání okolí je zprostředkováno automatickými a kontrolovanými mentálními procesy [43].

V nezprostředkovaném vnímání je přítomnost považována za samozřejmost, jelikož člověk vnímá pouze své bezprostřední fyzické okolí. Pokud je vnímání přítomnosti zprostředkováno pomocí komunikačních technologií, pak je člověk nucen vnímat dvě oddělená prostředí – fyzické prostředí, ve kterém se skutečně nachází a prostředí prezentované prostřednictvím komunikačního média. Telepresence je míra, do jaké se člověk cítí přítomen ve zprostředkovaném prostředí, spíše než v bezprostředním fyzickém prostředí [43]. V kontextu této práce je telepresence uvažována jako pocit přítomnosti druhé osoby prostřednictvím komunikačního média.

Často skloňovaným tématem spolu s telepresencí je nevolnost z pohybu. Nevolnost z pohybu, z anglického *motion sickness*, je běžně vyvolaná v situacích zahrnujících neobvyklé vzorce zrychlení (například mořská nemoc) nebo smyslové stimulace (například optokinetické bubny). Příznaky nevolnosti z pohybu jsou často vyvolané také v leteckých a automobilových simulátorech – tam, kde zažitý pohyb a setrvačný pohyb nekoresponduje. Nevolnost z pohybu je diskutovaným tématem ve virtuální realitě, zejména v aplikacích, které vyvolávají zdánlivý pohyb uživatele. Příznaky nevolnosti z pohybu se netýkají pouze pocitů nevolnosti, ale také ospalosti, únavy očí nebo ztráty iniciativy [28].

Kapitola 3

Microsoft HoloLens 2

Brýle Microsoft HoloLens 2 jsou plně samostatným holografickým počítačem od společnosti Microsoft[1]. Brýle patří do kategorie náhlavních displejů a využívají výhradně metodu *optical see-through*. Operačním systémem brýlí Microsoft HoloLens je *Windows Holographic OS*, který je založen na systému Windows 10 a poskytuje robustní, výkonnou a bezpečnou platformu pro uživatele [20]. Brýle jsou tvořeny nastavitelnou čelenkou, na které jsou připevněna plastová ramena, hledí a zadní kryt. Součástí hledí jsou senzory a displeje. Na levé straně hledí jsou umístěna tlačítka pro ovládání jasu. Na pravé straně hledí jsou umístěna tlačítka pro ovládání hlasitosti. Na zadním krytu je umístěn USB-C port pro nabíjení brýlí, tlačítko napájení a kolečko pro seřízení čelenky. Brýle jsou zobrazeny na obrázku 3.1.

Kapitola popisuje brýle pro rozšířenou realitu Microsoft HoloLens 2, jejich hardwarové vybavení, možnosti ovládání a prostorové mapování. V závěru kapitoly jsou popsány rozdíly mezi brýlemi Microsoft HoloLens 2 a brýlemi Microsoft HoloLens 1.

3.1 Hardware

Brýle obsahují 8MP RGB kameru, která dokáže snímat video v rozlišení 1080p se snímkovací frekvencí 30ti snímků za sekundu. Brýle Microsoft HoloLens obsahují čtyři černobílé kamery pro sledování okolního prostředí a dvě infračervené kamery pro sledování očí uživatele. Brýle obsahují také inerciální měřící jednotku (IMU), která obsahuje akcelerometr, gyroskop a magnetometr [20]. Dále brýle obsahují hloubkovou kameru, která funguje ve dvou režimech. První režim, nazvaný *AHAT*, má snímkovací frekvenci 45ti snímků za sekundu a je využíván pro sledování rukou uživatele. Druhý režim, nazvaný *Long-throw*, má snímkovací frekvenci pohybující se mezi jedním a pěti snímky za sekundu. Hloubková kamera v režimu *Long-throw* je využívána pro prostorové mapování [21]. Umístění senzorů v hledí je znázorněno na obrázku 3.2.

Displej brýlí Microsoft HoloLens je binokulární neboli určen pro pozorování dvěma očima. Displej brýlí HoloLens je tvořen kombinací vlnovodů a světelných projektorů. Uživatelé se při nošení náhlavní soupravy dívají skrz vlnovody — čočky uvnitř hledí. Světelné projektory jsou umístěny uvnitř krytu nad obočím a jsou využívány k osvětlení displeje [18]. Rozlišení obrazu je 2K pro každé oko [20].

Pro obsluhu displejů a senzorů obsahují brýle tři výpočetní jednotky – centrální výpočetní jednotku (CPU), grafickou výpočetní jednotku (GPU) a holografickou výpočetní

[1] `https://docs.microsoft.com/en-us/hololens/`

Obrázek 3.1: Brýle Microsoft HoloLens druhé generace.
Zdroj: `https://www.amazon.com/Microsoft-Hololens2-64GB-Headset-Device/dp/B08GP98FG7`

Systém na čipu	Výpočetní platforma Qualcomm Snapdragon 850
Holografická výpočetní jednotka	2. generace holografické výpočetní jednotky
Operační paměť	4 GB LPDDR4x system DRAM
Úložiště	64 GB UFS 2.1
Konektivita	WiFi 802.11ac 2x2, Bluetooth 5.0, USB Typ-C

Tabulka 3.1: Hardwarová specifikace brýlí Microsoft HoloLens 2.
Zdroj: `https://docs.microsoft.com/en-us/hololens/hololens2-hardware`

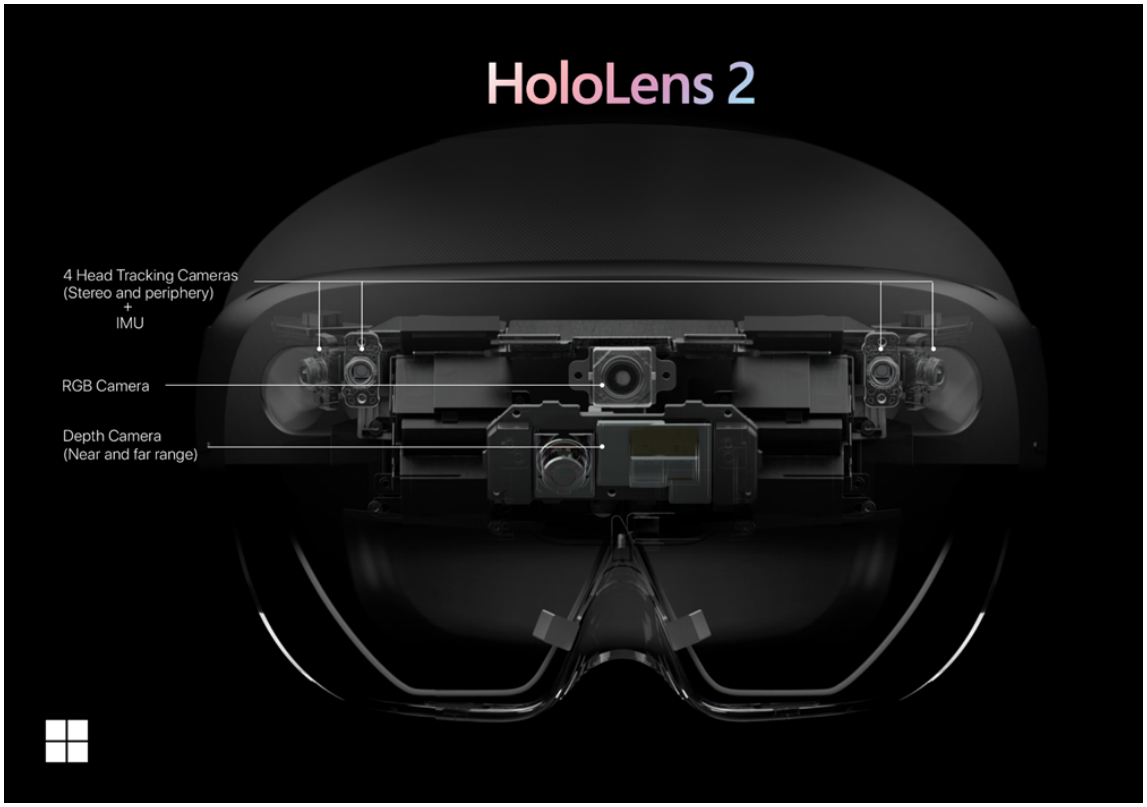
jednotku (HPU). Holografická výpočetní jednotka slouží k zajištění provozu prostorového mapování. Další detaily hardwarové specifikace jsou uvedeny v tabulce 3.1.

3.2 Ovládání a interakce

Pro komunikaci uživatele s brýlemi HoloLens jsou využity pokročilé vstupní systémy. Uživatel je schopen ovládat obsah, zobrazený v brýlích HoloLens, pomocí svých rukou, pomocí pohybů hlavy, pohybů očí nebo hlasovými příkazy. Pro tyto způsoby interakce není potřeba žádné dodatečné vstupní zařízení, protože veškerý potřebný hardware je součástí brýlí.

Hlasové příkazy

Vývoj umělé inteligence a pokroky v rozpoznávání hlasu umožnily Microsoftu integrovat do brýlí možnost interakce založené na hlasových příkazech. Hlasové příkazy lze zadávat v jazyce, který je pro operační systém brýlí nastaven jako výchozí. Pomocí hlasu lze interagovat s uživatelským rozhraním bez nutnosti pohnout končetinami, hlavou nebo očima. V brýlích je, stejně jako v ostatních univerzálních aplikacích pro systémy Windows (UWP), využita technologie rozpoznávání hlasu *Speech Recognition*. Tato technologie využívá integrovaný mikrofon [48, 42].



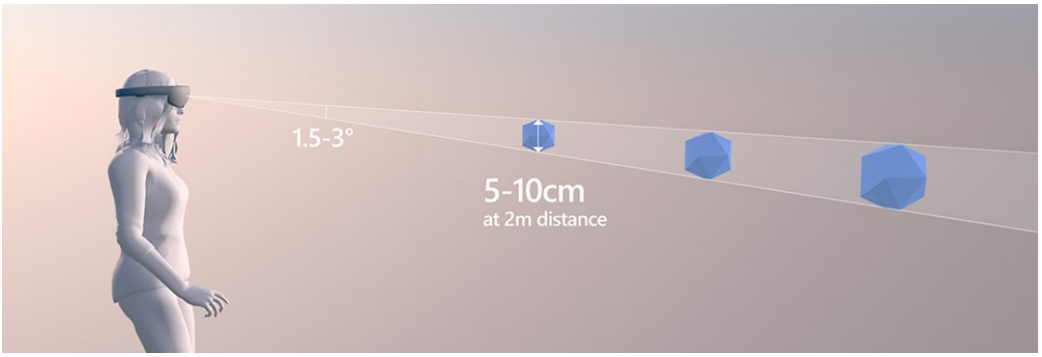
Obrázek 3.2: Umístění senzorů v hledí brýlí Microsoft HoloLens 2.
Zdroj: `https://docs.microsoft.com/en-us/hololens/hololens2-hardware`

Aby zařízení rozumělo příkazům uživatele, musí mít definovanou gramatiku. Pro základní příkazy jsou využívány předdefinované gramatiky, které jsou obsluhovány vzdálenou webovou službou. Pro předdefinované gramatiky jsou využívány výsledky vyhledávání na webu a těch je mnoho. Současně jsou dostupné pouze online – nachází-li se uživatel v místě se špatným připojením, může nastat zpoždění mezi vysloveným příkazem, jeho rozeznáním a jeho vykonáním. Z tohoto důvodu Microsoft nabízí vývojářům možnost definice vlastních *offline* gramatik, které mohou integrovat do svých aplikací [42]. Další možností hlasové interakce je využití hlasové asistentky Cortany, která využívá sadu přednastavených hlasových příkazů.

Head-gaze

Head-gaze je způsob interakce, který využívá sledování pohybu a orientace hlavy uživatele v prostoru. Vstupní model *Head-gaze* je rozdělen na vstupní model *Head-gaze and commit* a vstupní model *Head-gaze and dwell*. Pro vstupní model *Head-gaze and commit* je akce uživatele potvrzena pomocí gesta nebo hlasového příkazu [15].

U přístupu *Head-gaze and dwell* je využito tlačítko setrvání, které se aktivuje, pokud je kurzor umístěn na virtuálním objektu po určitou dobu. Užití tohoto přístupu je vhodné v situacích, kdy je uživatel v hlučném prostředí a má v rukou nástroje nebo součástky, což mu znemožňuje využít běžná gesta nebo hlasové příkazy [16].



Obrázek 3.3: Optimální vzdálenost hologramu je dva metry od pozorovatele.
Zdroj: `https://docs.microsoft.com/en-us/windows/mixed-reality/design/eye-tracking`

Eye tracking

Brýle Microsoft HoloLens 2 využívají technologii *Eye tracking* pro sledování pohybu očí uživatele v reálném čase. Pro přesné sledování pohybu očí je nutné, aby uživatel prošel vstupní kalibrací, během které postupně zaměřuje zrak na virtuální objekty před sebou. Optimální vzdálenost hologramu[2] je 2 metry, jak je zobrazeno na obrázku 3.3. Technologii lze použít pro interakci uživatele s virtuálním obsahem nebo pro kontrolu uživatelovy pozornosti [11].

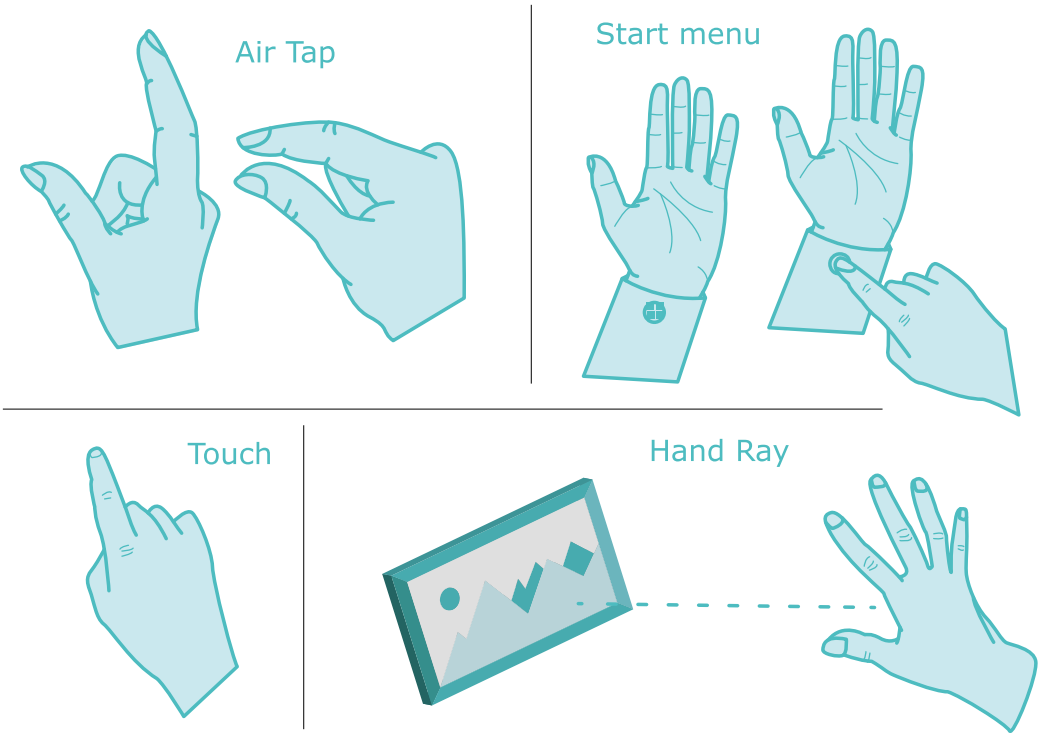
Gesta

Sledování rukou uživatele poskytuje přirozený způsob interakce s virtuálním obsahem pomocí gest v brýlích Microsoft HoloLens 2. Senzory, které sledují pohyb rukou, mají omezené zorné pole a vzdálenost, na kterou dokáží detekovat gesta. Jestliže má uživatel ruce mimo dosah senzorů, pak nejsou gesta detekována a na prvcích uživatelského rozhraní se zobrazí alternativní způsob interakce – kurzor ovládaný pohybem hlavy [19].

Brýle Microsoft HoloLens 2 mají několik základních gest – *Touch*, *Air tap*, *Air tap and hold*, *Hand Ray* a gesto pro vyvolání nabídky Start. Gesta jsou zobrazena na obrázku 3.4. První gesto, *Touch*, reprezentuje prostý dotek ukazováčkem. Druhé gesto, *Hand Ray*, využívá nataženou ruku s dlaní odvrácenou od uživatele pro označení nebo manipulaci s hologramy. Akce je u gesta *Hand Ray* potvrzena tím, že uživatel zachová pozici ruky a spojí prsty k sobě [19].

Dalším gestem je gesto *Air tap* a jeho varianta *Air tap and hold*. Gesto *Air tap* je využíváno pro potvrzení akce uživatele a funguje na principu levého tlačítka počítačové myši. Uživatel využívá pouze ukazováček a palec, jak je zobrazeno v levém horním rohu na obrázku 3.4. Pokud po sevření prstů k sobě uživatel prsty znovu oddálí, je aktivováno gesto *Air tap*. Pokud uživatel setrvá se sevřenými prsty, je aktivováno gesto *Air tap and hold*, které slouží pro manipulaci s hologramy. Posledním gestem je gesto pro vyvolání nabídky Start. Pro aktivaci gesta uživatel otočí dlaň jedné ruky směrek k sobě, čímž se mu na zápěstí zobrazí ikona nabídky Start, kterou aktivuje pomocí gesta *Touch* [19].

[2]Pojem hologram vyjadřuje virtuální obsah v brýlích Microsoft HoloLens.



Obrázek 3.4: Základní gesta pro brýle Microsoft HoloLens 2.

3.3 Prostorové mapování

Brýle Microsoft HoloLens 2 jsou schopny porozumět okolnímu prostředí. Prostorové mapování, z anglického *spatial mapping*, je funkcionalita, která využívá hloubkovou kameru a informace o pozici brýlí v prostoru k vytvoření 3D modelu okolního prostředí [41]. Pomocí prostorového mapování dokáží brýle HoloLens identifikovat povrchy reálného prostředí. Toho lze využít při umístění hologramů na objekty reálného prostředí.

Model reálného prostředí je neustále aktualizován. Prostorové mapování využívá dva hlavní komponenty – *Spatial Surface Observer* a *Spatial Surface*. Pomocí *Spatial Surface Observer* jsou vytvořeny ohraničené oblasti v prostoru, ve kterých jsou požadovány prostorové informace. Každé ohraničené oblasti jsou přiděleny komponenty *Spatial Surface*. Komponenty *Spatial Surface* popisují povrchy reálného prostředí na malém prostoru, který je reprezentován 3D trojúhelníkovou sítí. S novými informacemi z reálného prostředí se mění, vytvářejí a zanikají komponenty *Spatial Surface* [41]. Ohraničené oblasti spolu s popisem povrchů utvářejí trojúhelníkovou síť, jak je zobrazeno na obrázku 3.5. Možným využitím prostorového mapování je vyhodnocení okluzí. Okluze je významná vlastnost prostorového mapování, jelikož má obrovský dopad na realističnost hologramů a pomáhá vytvořit pocit, že virtuální hologramy koexistují s reálným prostředím. Prostorové mapování lze využít také při simulaci fyziky, kdy virtuální objekty, které mají fyzikální vlastnosti, nepropadnou podlahou nebo neprojdou zdí [41].

Obrázek 3.5: Vizualizace prostorového mapování.

3.4 Srovnání s Microsoft HoloLens 1

První generace brýlí HoloLens byla představena v roce 2016[3] společností Microsoft. Druhá generace brýlí HoloLens byla představena o tři roky později[4]. V této sekci jsou popsány nejdůležitější rozdíly mezi těmito generacemi. Prvním mezigeneračním rozdílem je zorný úhel. Zorný úhel brýlí Microsoft HoloLens 2 má 52 stupňů, což je téměř dvojnásobek zorného úhlu brýlí Microsoft HoloLens 1, jejichž zorný úhel je 30 stupňů. Porovnání zorných úhlů je zobrazeno na obrázku 3.6. Současně se zorným úhlem bylo vylepšeno také rozlišení obrazu ze 720p na 2K pro každé oko[5].

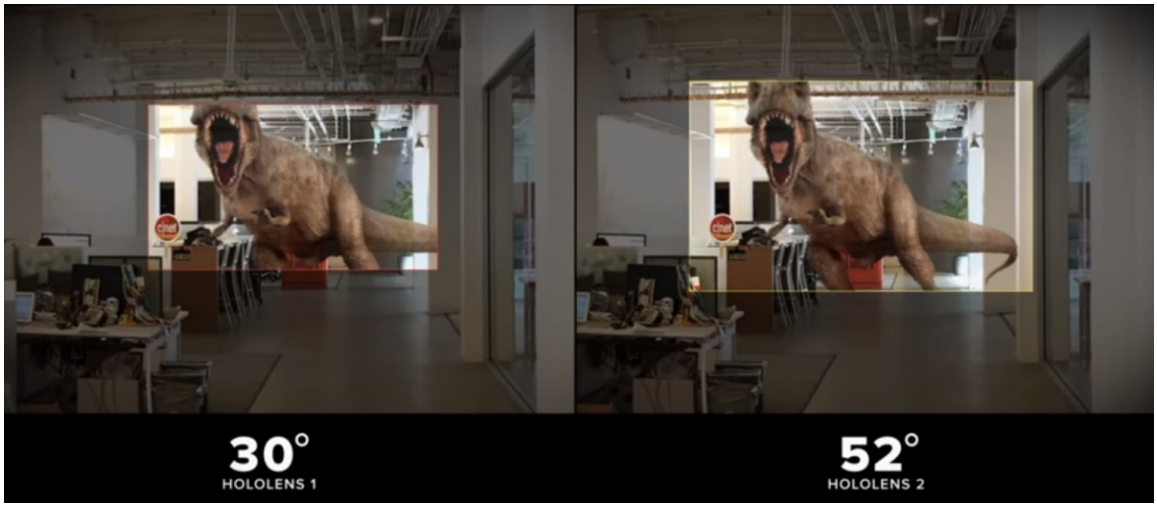
U brýlí HoloLens 2 je využita výpočetní jednotka založená na ARM architektuře narozdíl od HoloLens 1, kde je využita výpočetní jednotka založená na x86 architektuře. Nová generace brýlí má větší RAM paměť, vyšší výkon výpočetních jednotek a podporuje vyšší standardy pro konektivitu (viz tabulka 3.2). Oproti předchozí generaci brýlí přibylo sledování pohybů očí. Gesto pro vyvolání nabídky Start *Bloom* bylo nahrazeno gestem využívajícím ikonu na zápěstí.

Vylepšena byla také ergonomie zařízení. Baterie a výpočetní jednotky byly přesunuty do zadní části brýlí, což vedlo k lepšímu vyvážení náhlavní soupravy. Byla snížena také hmotnost nové generace brýlí HoloLens. Tyto designové změny vedly k menší námaze při nošení brýlí.

[3] `https://docs.microsoft.com/en-us/archive/msdn-magazine/2016/november/hololens-introduction-to-the-hololens`

[4] `https://news.microsoft.com/microsoft-at-mwc19/`

[5] `https://4experience.co/hololens-2-vs-hololens-1-whats-new/`



Obrázek 3.6: Vizualizace zorného úhlu brýlí Microsoft HoloLens 1 (vlevo) a brýlí Microsoft HoloLens 2 (vpravo).

Zdroj: `https://mspoweruser.com/controversy-around-hololens-2-field-of-view-continues/`

	Microsoft HoloLens 1	Microsoft HoloLens 2
Procesor	Intel Atom x5-Z8100P, 4 jádra	QC Snapdragon 850, 8 jader
Operační paměť	1 GB	4 GB
Úložiště	64 GB	64 GB
WiFi	802.11ac	802.11ac 2x2
Bluetooth	4.1 LE	5.0
HPU	1. generace HPU	2. generace HPU

Tabulka 3.2: Srovnání hardwarové specifikace brýlí Microsoft HoloLens 1 a brýlí Microsoft HoloLens 2.

Zdroj: `https://docs.microsoft.com/en-us/hololens/hololens2-hardware`, `https://docs.microsoft.com/en-us/hololens/hololens1-hardware`

Kapitola 4

Zpracování obrazových dat

V úvodu této kapitoly je představeno zařízení Azure Kinect DK[1] a senzory, které obsahuje. Dále jsou představeny a srovnány komerčně dostupné RGBD zařízení. V závěru kapitoly jsou popsány dva možné způsoby reprezentace 3D dat.

4.1 Azure Kinect DK

Azure Kinect DK je elektronické zařízení od společnosti Microsoft. Zařízení obsahuje barevnou kameru, hloubkový senzor, inerciální měřící jednotku a pole sedmi mikrofonů. Zařízení je cíleno na vývojáře a komerční firmy, čímž se odlišuje od svých předchůdců, zařízení Xbox Kinect a Kinect pro Windows [3]. Tölgyessy a spol. [44] označují Azure Kinect DK za malé univerzální zařízení s mnoha případy užití. Zařízení lze využít například k rozpoznávání objektů, navigaci, rozpoznávání gest nebo telepresenci.

Hloubkový senzor

Hloubková data jsou snímána 1MP senzorem *Continuous-Wave (CW) Time-of-Flight (ToF)*. U tohoto senzoru je světlo vysíláno z amplitudou modulovaného světelného zdroje – vysílače. V prostoru před senzorem je zpětně rozptylováno objekty v zorném poli kamery. Zpětně rozptýlené světlo následně zachytí přijímač senzoru. Senzor měří časový rozdíl mezi vysílaným a přijímaným světlem pomocí fázového posunu. Tento posun je následně převeden na hodnotu hloubky pro každý pixel v zobrazovacím poli. U zařízení Azure Kinect je hodnota hloubky měřena v milimetrech[2]. Hloubkový senzor poskytuje obraz ve formátu 16ti bitové hloubkové mapy. Provozní režimy hloubkové kamery jsou zobrazeny v tabulce 4.1.

Barevná kamera

Barevný obraz je snímán pomocí RGB snímače *OV12A10 12MP CMOS* s pohyblivou závěrkou. Kamera má, podobně jako hloubkový senzor, několik nastavitelných provozních režimů. Provozní režimy barevné kamery jsou zobrazeny v tabulce 4.2. Programově lze upravovat rozlišení kamery a tím i poměr stran obrazu. Nejmenší možné rozlišení je 1280×720 px a největší možné rozlišení je 4096x3072 px. Dále lze programově upravovat snímky za vteřinu a formát obrazových dat. Formát obrazových dat *BGRA* je podporován knihovnou obsluhující zařízení Azure Kinect, ale není podporován přímo hardwarem, a proto je při využití

[1] DK je zkratka pro anglický výraz *Developer Kit* neboli vývojářskou sadu

[2] `https://docs.microsoft.com/cs-cz/azure/kinect-dk/depth-camera`

Obrázek 4.1: Zařízení Azure Kinect DK.

Mód	Rozlišení	Snímkovací frekvence	Operační rozsah	Čas expozice
NFOV unbinned	640x576 px	0, 5, 15, 30	0.5 m – 3.86 m	12.8 ms
NFOV 2x2 binned	320x288 px	0, 5, 15, 30	0.5 m – 5.46 m	12.8 ms
WFOV 2x2 binned	512x512 px	0, 5, 15, 30	0.25 m – 2.88 m	12.8 ms
WFOV unbinned	1024x1024 px	0, 5, 15	0.25 m – 2.21 m	20.3 ms
Passive IR	1024x1024 px	0, 5, 15, 30	N/A	1.6 ms

Tabulka 4.1: Provozní režimy hloubkové kamery.
Zdroj: `https://docs.microsoft.com/cs-cz/azure/kinect-dk/hardware-specification`

tohoto formátu spotřebován procesorový čas navíc. Formát *BGRA* je důležitý, protože je vyžadován metodami při registraci barevného obrazu s hloubkovým obrazem [24].

Ostatní senzory

Zařízení Azure Kinect DK dále obsahuje kruhové pole sedmi, vysoce kvalitních, mikrofonů. Mikrofony jsou uspořádány do kruhu pro záznam 360ti stupňového prostorového zvuku. Data o okolním prostoru a orientaci zařízení v něm poskytuje inerciální měřící jednotka, která kombinuje senzor gyroskopu se senzorem akcelerometru [24].

4.2 Srovnání RGBD zařízení

Azure Kinect je následovníkem zařízení Kinect pro Windows neboli zařízení Kinect v2. Ačkoli obě zmíněná zařízení Kinect sdílejí podobné vlastnosti, Tölgyessy a spol. [44] zjistili, že z hlediska přesnosti je výkon zařízení Azure Kinect lepší než výkon zařízení Kinect pro Windows. V závěru svého výzkumu vyzdvihují poloviční hmotnost (440g) nového zařízení Azure Kinect oproti zařízení Kinect pro Windows a nově přidané provozní režimy, které zajišťují větší míru variability při vývoji aplikací.

Alternativním zařízením k zařízení Azure Kinect je zařízení *Lidar Camera L515*[3] od společnosti Intel. *Lidar Camera L515* obsahuje sadu senzorů určených pro počítačové vidění ve vnitřních prostorách. Hloubková kamera pracuje v operačním rozsahu 0.25 m až 9 m při rozlišení 1024x768 px. Barevná kamera má maximální rozlišení 1920×1080 px. Obě kamery poskytují snímkovací frekvenci třiceti snímků za vteřinu.

[3] `https://www.intelrealsense.com/lidar-camera-l515/`

Rozlišení	Poměr stran	Formát	Snímkovací frekvence
3840 × 2160	16:9	MJPEG	0, 5, 15, 30
2560 × 1440	16:9	MJPEG	0, 5, 15, 30
1920 × 1080	16:9	MJPEG	0, 5, 15, 30
1280 × 720	16:9	MJPEG, YUY2, NV12	0, 5, 15, 30
4096 × 3072	4:3	MJPEG	0, 5, 15
2048 × 1536	4:3	MJPEG	0, 5, 15, 30

Tabulka 4.2: Provozní režimy barevné kamery.
Zdroj: `https://docs.microsoft.com/cs-cz/azure/kinect-dk/hardware-specification`

Dalšími zařízeními od společnosti Intel jsou modely SR305, D455, D415 a D435. Lourenço a Araujo [30] zkoumali přesnost hloubkových kamer z rodiny RealSense SR305, D415 a L515 a experimentálně zjistili, že model L515 poráží ostatní kamery precizností a přesností ve vnitřních, kontrolovaných podmínkách.

4.3 3D reprezentace obrazových dat

V počítačové grafice jsou plochy trojrozměrných objektů reprezentovány obvykle trojúhelníkovou sítí. Postupně však přibývá aplikací v průmyslu a zdravotnictví, které vyžadují skenování složitých třírozměrných objektů s následnou úpravou v počítači. Z tohoto důvodu byly vyvinuty měřící techniky a zařízení, které produkují velké množství bodů ležících na povrchu 3D objektů reálného prostředí.

Mračno bodů

Množina bodů, která představuje hranice trojrozměrného objektu, je nazývána mračnem bodů. Podle Linsena [29] je vytváření trojúhelníkové sítě časově náročné, a proto lze využívat přímo reprezentaci mračna bodů bez nutnosti jej převádět na trojúhelníkovou síť. V počítači je mračno bodů uloženo jako seznam souřadnic jednotlivých vrcholů.

Trojúhelníková síť

Trojúhelníková síť je kolekce vrcholů, hran a stěn. Proces vytvoření trojúhelníkové sítě se nazývá triangulace. V počítači je trojúhelníková síť uložena jako seznam souřadnic vrcholů trojúhelníků, seznam hran trojúhelníků a seznam trojúhelníků [10].

Jednou z nejznámějších technik triangulace je *Delaunay triangluation*. Technika byla vynalezena za účelem spojování bodů v prostoru do trojúhelníkových skupin, takových, že minimální úhel všech trojúhelníků v triangulaci je také maximálním úhlem. To znamená, že technika *Delaunay* se snaží vyhnout protáhlým trojúhelníkům během triangulace [10].

Kapitola 5

Návrh řešení

V této práci bude vytvořena aplikace, která poběží na počítači a bude využívat možnosti funkcionality Holographic Remoting. Uživatel bude v brýlích HoloLens vidět obrazová data z Azure Kinectu a bude moct s hologramem dále manipulovat. Veškeré výpočty a vykreslování bude probíhat na Remote počítači.

V této kapitole jsou analyzovány existující řešení související s telepresencí a využitím zařízení Azure Kinect. Dále je v kapitole popsán návrh architektury systému telepresence, návrh zpracování obrazových dat, bezdrátového přenosu a návrh uživatelského rozhraní.

5.1 Analýza existujících řešení

LiveScan3D

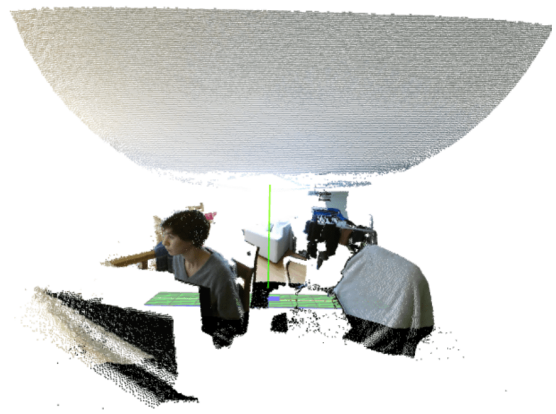
Marek Kowalski a spol. [27] představili systém navržený pro 3D rekonstrukci dat v reálném čase s využitím několika zařízení Azure Kinect nebo zařízení Microsoft Kinect. 3D rekonstrukce je ve formátu barevného mračna bodů, přičemž všechny body ze všech připojených zařízení Kinect jsou umístěny ve stejném souřadnicovém systému.

Systém LiveScan3D umožnuje kromě simultánního zachytávání 3D struktury z několika zorných úhlů také zvýšit hustotu bodů jedné snímané scény pomocí využití více zařízení Kinect. Pro každé zařízení Kinect, připojené do systému, je vytvořena instance klientské aplikace. V případě využití zařízení Azure Kinect lze mít více klientských instancí na jednom počítači, naopak v případě využití zařízení Microsoft Kinect musí být pro každou klientskou aplikaci samostatný počítač. Klientské aplikace jsou připojeny k serveru, kde probíhá kalibrace, filtrace, synchronizace zachytávání obrazů a vykreslení mračna bodů.

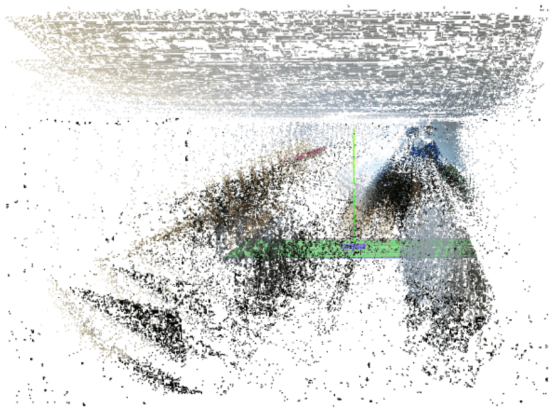
Streaming Volumetric Images Online

V práci od Lisy Jamhoury [23] je popsán systém přenosu hloubkového a barevného obrazu ze zařízení Microsoft Kinect do internetového prohlížeče. Přenos obrazů do prohlížeče je realizován pomocí datových kanálů aplikačního rozhraní WebRTC. V práci jsou popsány dva způsoby reprezentace hloubkového a barevného obrazu a jejich vliv na rychlost přenosu a kvalitu obrazu.

V prvním přístupu je barevný i hloubkový obraz uložen samostatně ve formátu WebP. Barevný obraz je uložen do WebP formátu se ztrátovou kompresí. 13ti bitový hloubkový obraz je rozdělen a uložen do WebP formátu na pozici kanálů R a G. Kanály B a A nejsou využity. Hloubkový obraz je uložen do WebP formátu s bezztrátovou kompresí. Oba



Obrázek 5.1: Mračno bodů zrekonstruované z hloubkové mapy, na kterou byla aplikována bezztrátová komprese [23].



Obrázek 5.2: Mračno bodů zrekonstruované z hloubkové mapy, na kterou byla aplikována ztrátová komprese [23].

obrazy jsou odeslány přes datové kanály WebRTC ke klientovi, kde jsou sloučeny a je z nich zrekonstruováno mračno bodů, které je následně vykresleno. Tento přístup ukládání obrazů je efektivní pro přenos barevného obrazu, u kterého je aplikována ztrátová komprese. Přístup je neefektivní při přenosu hloubkového obrazu, ale zachovává jeho maximální rozlišení.

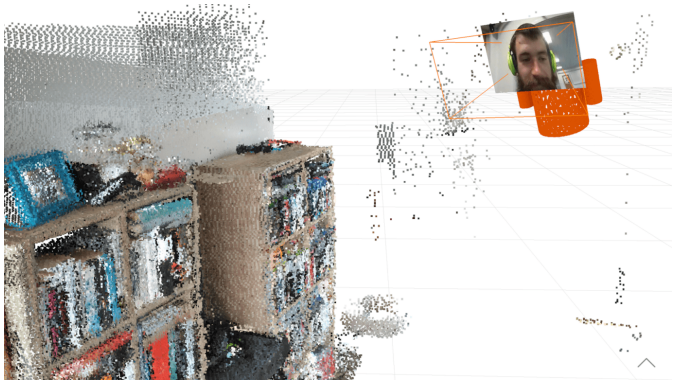
V druhém přístupu jsou oba obrazy sloučeny do jednoho WebP formátu před jejich odesláním. Pro barevný obraz jsou využity RGB kanály a pro hloubkový obraz je využit A kanál. WebP formát má nastavenu bezztrátovou kompresi. Sloučený obraz je odeslán přes datový kanál WebRTC do prohlížeče, kde je převeden na mračno bodů a vykreslen. Tento přístup je neefektivní pro přenos barevného obrazu, na který nelze aplikovat ztrátovou kompresi. Pro přenos hloubkového obrazu je přístup výhodný, jelikož není potřeba samostatného obrazu pro hloubková data. Na druhou stranu je rozlišení hloubkových dat sníženo, což vede k nepřesné hloubkové mapě a zkreslení výsledného mračna bodů.

Za přínos práce považuji demonstraci využití WebP formátu pro ukládání obrazových dat z Microsoft Kinectu a představení dvou způsobů ukládání, včetně jejich kladů a záporů. Práce také demonstruje, jakým způsobem jsou hloubková data znehodnocena při použití ztrátové komprese (viz obrázek 5.2) oproti bezztrátové kompresi (viz obrázek 5.1).

Mobileportation

Young a spol. [51] představili ve své práci prototyp telepresenčního systému *Mobileportation*. V systému je použit mobilní telefon se zabudovaným RGBD senzorem, pomocí kterého je inkrementálně vytvářena 3D rekonstrukce prostředí *lokálního uživatele*. Vzdálený uživatel může prozkoumávat zrekonstruované prostředí v 6DoF z exocentrického pohledu. V systému je použita také 360ti stupňová kamera, která doplňuje data z RGBD senzoru o barevná data ve vyšším rozlišení.

Pro zachytávání a zpracování hloubkových dat je využito aplikační rozhraní Google Tango, které zpracovává data z hloubkového senzoru a rekonstruuje z nich mračno bodů. Mračno bodů je generováno s frekvencí pěti snímků za sekundu. Mračno je transformováno na souřadnicový systém scény vynásobením vektoru souřadnic každého bodu s aktuální poziční maticí *lokálního uživatele*. Tato transformace je provedena v *OpenGL Compute shaderu*. Stejný *shader* provádí projekci každého bodu do nejnovějšího barevného obrazu pořízeného barevnou kamerou s využitím kalibračních matic barevné a hloubkové kamery.



Obrázek 5.3: Aplikace Mobileportation. Pohled na rekonstruované mračno bodů [51].

Obrázek 5.4: Aplikace Mobileportation. Osoba na obrázku snímá pomocí RGBD kamery své prostředí [51].

Tím je vytvořeno barevné mračno bodů v aktuálním zorném poli hloubkové kamery. Barevné mračno bodů je předáno renderovacímu vláknu a uloženo v grafické kartě. Každé další mračno bodů je přidáno na konec *vertex bufferu*. Ve své práci Young a spol. zkoušeli také složitější datové struktury (*octree*), ale zjistili, že nejsou vhodné, jelikož jejich využití vede k duplicitám.

Pro zobrazení panoramatického obrazu z 360ti stupňové kamery je vytvořena kruhová *mesh* okolo *lokálního uživatele*. Vytvořený kruh má poloměr dva metry a hloubková data, která jsou blíže než dva metry od pozorovatele jsou tak stále viditelná.

Pro síťový přenos je v řešení využito aplikační rozhraní WebRTC[1]. Pro ustavení spojení mezi lokálním a vzdáleným klientem je použit server. Po ustavení spojení probíhá komunikace napřímo mezi klienty. V řešení je implementována podpora pro *STUN* a *TURN* servery a byl otestován také mezinárodní přenos. Každá strana komunikace má svůj video kanál pro přenos videa a audio kanál pro přenos zvuku. Pro přenos uživatelovy pozice v prostoru a přenos 3D mračna bodů jsou využity vlastní datové kanály. Jelikož je mračno obarveno již před přenosem, není vyžadována synchronizace mezi datovými kanály.

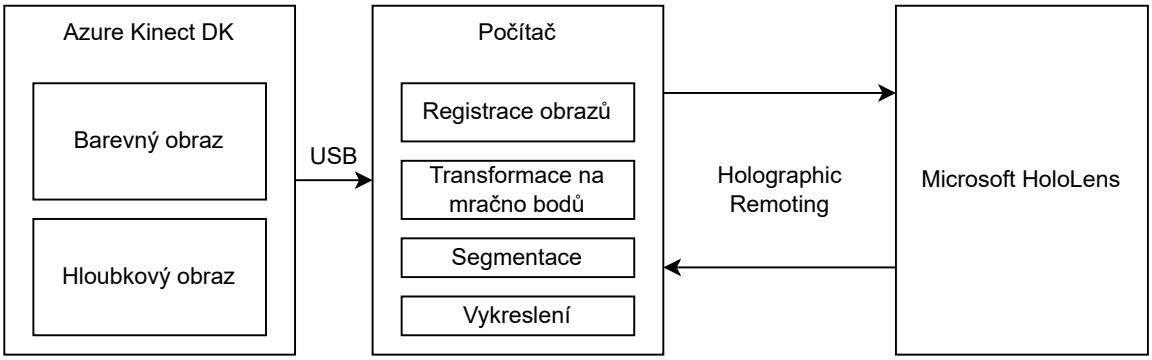
Systém *Mobileportation* je dle autorů schopen vykreslit 5 milionů bodů s frekvencí 15ti snímků za sekundu. Autoři očekávají, že při obvyklých podmínkách bude systém dosahovat frekvence pohybující se mezi 20ti až 30ti snímky za sekundu. Každý rámec má průměrně 5000 bodů a po zachycení mračna hloubkovou kamerou je potřeba 31 ms na jeho transformaci, obarvení a předání grafické kartě. Řešení je limitováno hloubkovou kamerou, která poskytuje mračno bodů s maximální frekvencí 5ti snímků za sekundu. V práci není uvedeno, v jakém datovém formátu je obarvené mračno odesíláno.

Azure Kinect Unity Sample

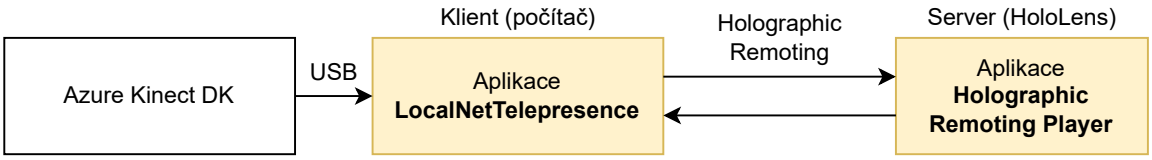
Práce Takashiho Yoshinagy[2] se zabývá vykreslením obrazových dat, snímaných zařízením Azure Kinect, v Unity. Práce využívá Unity ve verzi *2019.4.25*. Pro práci se zařízením Azure Kinect je využita oficiální knihovna od společnosti Microsoft – *Sensor SDK*. Obrazová data jsou po jejich záchytu registrována a hloubkový obraz je převeden na mračno bodů. Následně

[1] `https://webrtc.org/`

[2] `https://github.com/TakashiYoshinaga/Azure-Kinect-Sample-for-Unity`



Obrázek 5.5: Diagram architektury systému telepresence.



Obrázek 5.6: Diagram zobrazující komunikaci a aplikace systému.

je barevný obraz převeden na pole ve formátu *BGRA* a hloubkový obraz je převeden na pole ve formátu *Short3*. Tato pole jsou využita pro naplnění dalších pomocných polí ve formátu *Vector3* a *Color32*, což jsou formáty používané v nativní třídě *Mesh*. V cyklu jsou pole postupně překopírována a u hloubkových dat je proveden převod z milimetrů na metry. Následně jsou pole přiřazena instanci třídy Mesh. *Mesh* je pomocí komponent *MeshRenderer* a *MeshFilter* vykreslena v Unity. Ve výchozím nastavení využívá komponenta *MeshRenderer* materiál, který je jí přiřazen, což by znamenalo, že ačkoli mají vrcholy v instanci třídy *Mesh* přiřazeny barvy, nebyly by viditelné. Takashi Yoshinaga proto vytvořil materiál, který má přiřazen vlastní *shader*, který zapříčiní, že místo přiřazeného materiálu budou použity barvy přiřazené každému jednotlivému vrcholu. Na této práci oceňuji její jednoduchost a přímočarost, a proto z ní ve své práci částečně vycházím.

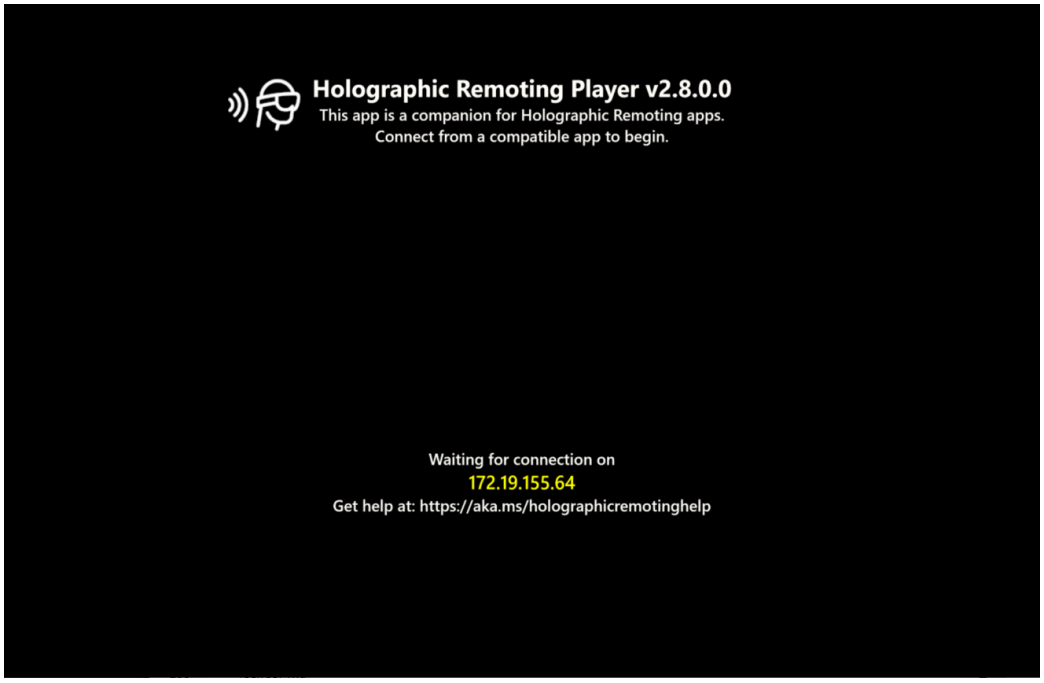
5.2 Architektura systému telepresence

V rámci řešení bylo nutné navrhnout části systému telepresence tak, aby spolu spolupracovaly. Při navrhování architektury systému telepresence bylo potřeba zohlednit fakt, že zařízení Azure Kinect není samonosné, a proto bylo do systému přidáno zařízení, které dokáže se zařízením Azure Kinect pracovat. Navržená architektura je zobrazena na obrázku 5.5.

Návrh architektury systému telepresence vychází z principu fungování síťové funkcionality *Holographic Remoting*. Navržený systém zahrnuje dvě aplikace, které spolu komunikují po síti (viz obrázek 5.6). První aplikací je *Holographic Remoting Player*[3], která je dílem společnosti Microsoft a slouží jako podpůrná aplikace pro systémy využívající funkcionalitu *Holographic Remoting*. Aplikace musí být v minimální verzi 2.4.0[4], aby fungovala funkciona-

[3] `https://docs.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-player`

[4] `https://docs.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-player`



Obrázek 5.7: Pohled na aplikaci *Holographic Remoting Player* v brýlích Microsoft HoloLens.

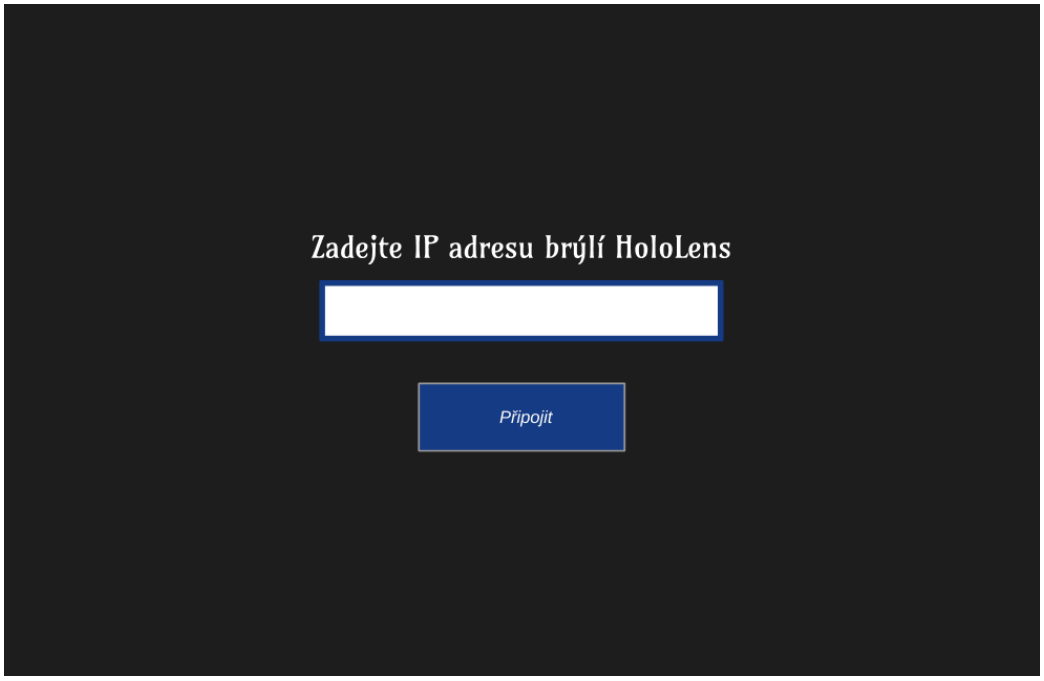
lita *Holographic Remoting* poskytovaný aplikačním rozhraním OpenXR. Pohled na aplikaci *Holographic Remoting Player*, před připojením vzdáleného uživatele, je zobrazen na obrázku 5.7. Druhou aplikací, tou, která je v rámci této práce navržena a implementována, je aplikace *LocalNetTelepresence*, která běží na počítači, zpracovává obrazová data ze zařízení Azure Kinect, vykresluje je a odesílá do aplikace *Holographic Remoting Player*, kde jsou data prezentována uživateli.

5.3 Návrh uživatelského rozhraní

Při návrhu uživatelského rozhraní je potřeba znát cílovou platformu, pro kterou je uživatelské rozhraní vytvářeno. V kontextu této práce jsou cílové platformy dvě. První platformou je počítač, který zpracovává data ze zařízení Azure Kinect, vykresluje je a odesílá je na druhou platformu, do brýlí Microsoft HoloLens 2. Obě zmíněné platformy využívají jiná vstupní zařízení. Počítač využívá klávesnici a myš, brýle HoloLens využívají rozpoznávání gest a sledování pohybu hlavy uživatele v prostoru. Počítač využívá 2D rozhraní, zatímco v brýlích HoloLens ovládá uživatel aplikace ve pomocí 3D rozhraní.

5.4 Zpracování a vykreslování obrazových dat

Pro vytvoření telepresence je důležité, aby výsledný 3D objekt odpovídal co nejlépe realitě. Takovým objektem, v kontextu této práce, je výsledek vhodné kombinace hloubkových dat s daty barevnými. Zařízení Azure Kinect považuji, pro poskytování obrazových dat, za jedno z nejvhodnějších. Pro snímání bude využita knihovna *Sensor SDK*, kterou Microsoft vyvinul pro přístup k jednotlivým senzorům zařízení Azure Kinect. Na obrázku 5.10 je zobrazení



Obrázek 5.8: Návrh uživatelského rozhraní pro počítač. Obrazovka před připojením k brýlím HoloLens.

tvarů obrazů. Barevný obraz s poměrem stran 4:3 má větší plochu překrytí s hloubkovým obrazem než barevný obraz s poměrem stran 16:9.

Obrazy budou vzájemně registrovány a hloubkový obraz bude transformován na mračno bodů. Vizualizace výsledku těchto kroků bude řešena napojením na práci Takashiho Yoshinagy, která byla představena v sekci 5.1.

Výpočetní výkon brýlí HoloLens není dostatečný pro vykreslování velkých a detailních 3D objektů. Povazuji tedy za výhodné přenést výpočetní zátěž na zařízení, které potřebný výkon má – například počítač s dedikovanou grafickou kartou NVIDIA. Přenesení výpočetní zátěže bude docíleno pomocí funkcionality *Holographic Remoting*.

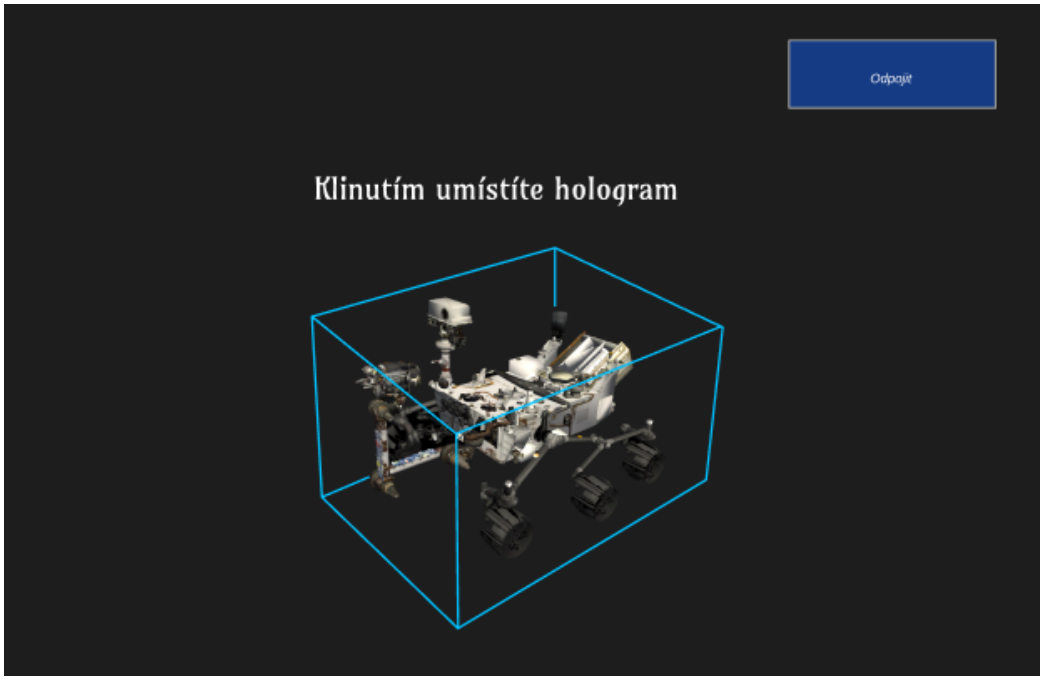
5.5 Bezdrátový přenos dat

Bezdrátový přenos bude realizován pomocí funkcionality *Holographic Remoting*. *Holographic Remoting* využívá pro komunikací UDP[5] protokol na portech 8265 nebo 8266. Z brýlí HoloLens jsou odesílána data o pozici uživatele v prostoru, o vstupních událostech a o prostorovém mapování. Na počítači jsou tato data zpracována, je vytvořen virtuální rámec, do kterého je vykreslen obsah[6]. U rámce je využita komprese H.264 nebo H.265. Rámec je odeslán do brýlí, kde je dekódován a umístěn do prostoru podle informací, které byly přijaty počítačem od brýlí HoloLens. Komunikace je tedy duplexní.

Holographic Remoting má dva případy užití. Prvním případem užití je náhled a ladění aplikací, určených pro brýle HoloLens. Aplikaci lze spustit v Unity editoru na počítači a streamovat obsah do brýlí HoloLens. Lze tak ladit aplikace bez nutnosti jejich sestavení

[5] `https://www.ietf.org/rfc/rfc768.txt`

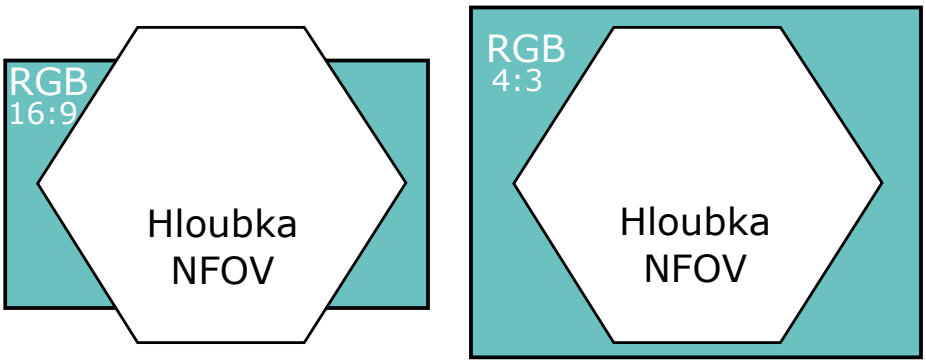
[6] `https://docs.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-create-remote-openxr`



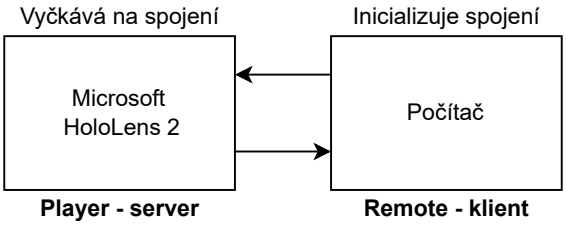
Obrázek 5.9: Návrh uživatelského rozhraní pro brýle HoloLens. Obrazovka po úspěšném připojení k aplikaci na počítači.

a instalace v brýlích HoloLens. Druhým případem užití, který je v této práci relevantní, je delegace výpočetní zátěže z brýlí na výkonnější počítač. Aplikace v brýlích HoloLens je označena jako *Player* a aplikace běžící na počítači je označena jako *Remote*.

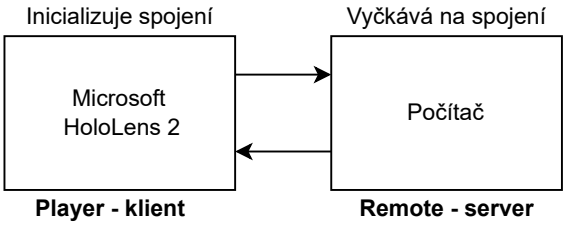
Funkcionalitu Holographic Remoting lze využít ve dvou režimech. První režim využívá *Player* jako server, který vyčkává na příchozí spojení a *Remote* jako klient který se připojuje (viz obrázek 5.11). Druhý režim funguje naopak, využívá *Remote* jako server a *Player* se k němu připojuje (viz obrázek 5.12). V kontextu této práce je využit první režim, a to z důvodu, že uživatel může využít oficiální aplikace *Holographic Remoting Player*.



Obrázek 5.10: Tvar obrazu kamer, jak je vidět zepředu ze vzdálenosti 2 m
Zdroj: `https://docs.microsoft.com/cs-cz/azure/kinect-dk/hardware-specification`



Obrázek 5.11: Režim holographic remoting. Player v brýlích HoloLens funguje jako server a vyčkává na příchozí spojení od klienta Remote, běžícího na počítači.



Obrázek 5.12: Režim holographic remoting. Remote na počítači funguje jako server a vyčkává na příchozí spojení od klienta Player, běžícího v brýlích HoloLens.

Kapitola 6

Implementace

Jak již bylo zmíněno v návrhu, navržený systém se skládá ze dvou aplikací. Aplikace, která běží v brýlích Microsoft HoloLens je již implementována společností Microsoft. Z tohoto důvodu je tato kapitola zaměřena na popis vývoje druhé aplikace, určené pro počítač – *LocalNetTelepresence*. V kapitole jsou popsány použité technologie, zpracování obrazových dat a jejich vykreslení, bezdrátový přenos do brýlí HoloLens a implementace uživatelského rozhraní.

Původně jsem se pokoušel implementovat systém telepresence pro brýle Microsoft HoloLens první generace, ale narazil jsem na spoustu problémů, které jsem nedokázal vyřešit. Jedním z problémů byla nekompatibilita knihoven. Pro první generaci brýlí byla určena knihovna HoloToolkit, která již není udržována. Navíc některé části moderních knihoven na hardwaru brýlí HoloLens 1 nefungovaly a dokumentace často neposkytovala náhradní řešení. Z tohoto důvodu jsem se po domluvě s vedoucím práce rozhodl vyvíjet na zařízení Microsoft HoloLens druhé generace.

6.1 Použité technologie

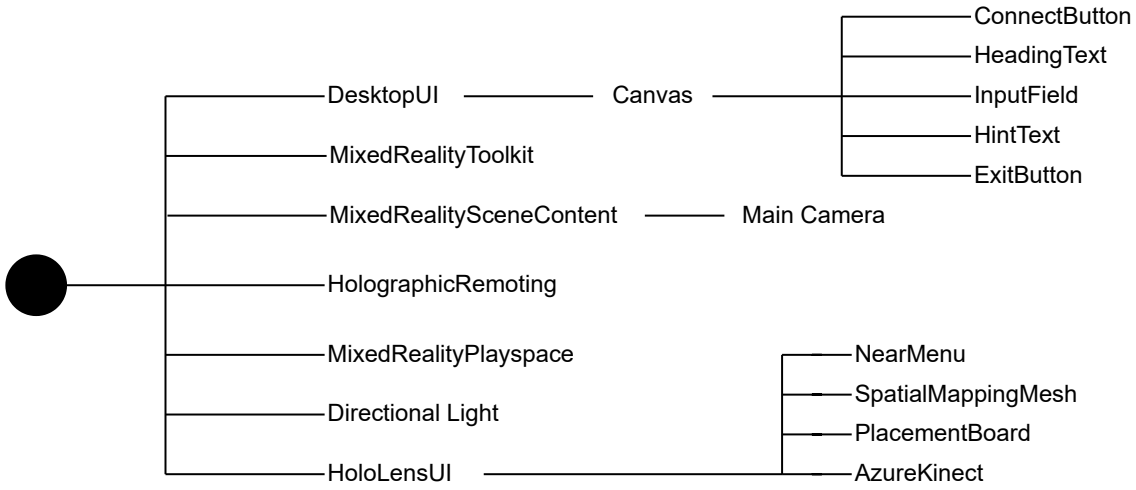
Hlavním programovacím jazykem této práce je jazyk C#. Tento jazyk je výchozím jazykem pro vývoj aplikací v Unity. Při vývoji jsem využíval paradigma objektově orientovaného programování. V této sekci budou popsány nejdůležitější nástroje a knihovny, využité při vývoji aplikace *LocalNetTelepresence*.

Unity

Unity je herní engine a integrované vývojové prostředí (IDE[1]) spravované společností Unity Technologies. Typickým případem užití Unity je vývoj her. Unity je multiplatformní a podporuje většinu operačních systémů. Kromě klasických operačních systémů, jakými jsou Microsoft Windows, macOS, Linux, Android nebo iOS, lze využít Unity také k programování aplikací pro zařízení využívající operační systémy pro rozšířenou realitu[2]. Důležitou součástí Unity je skriptování. Skriptování umožňuje vývojářům vytvářet skripty tak, aby reagovaly na vstup od uživatele nebo aby řídily události ve scéně. Skriptování lze využít také pro vytváření grafických efektů, ovládání fyzického chování objektů nebo pro implementaci vlastních systémů umělé inteligence. Výchozím skriptovacím jazykem Unity je C#.

[1] IDE je zkratka pro anglické spojení slov *Integrated Development Environment*.

[2] `https://unity.com/solutions/multiplatform`



Obrázek 6.1: Schéma hierarchie objektů v aplikaci LocalNetTelepresence.

Pro skriptování lze využít také další .NET jazyky, pokud tyto jazyky umožňují kompilaci kompatibilních DLL knihoven. V této práci je Unity použito ve verzi *2020.3.13f1*.

Mixed Reality Toolkit a OpenXR

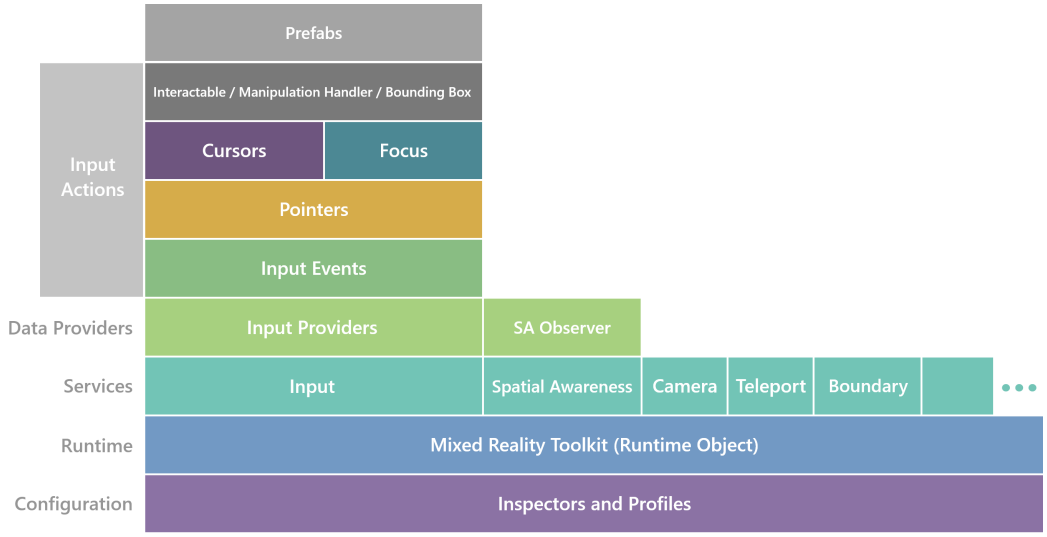
Mixed Reality ToolKit (MRTK) je rozsáhlá knihovna spravovaná společností Microsoft. Knihovna podporuje herní engine Unity a lze ji využít pro vývoj aplikací určených na zařízení Microsoft HoloLens, Oculus Quest a další zařízení využívající rozšířenou nebo virtuální realitu[3]. Architektura knihovny je zobrazena na obrázku 6.2. Knihovna poskytuje veškeré nástroje potřebné pro vývoj aplikací využívajících rozšířenou realitu. V kontextu této bakalářské práce je tato knihovna využita jako hlavní knihovna pro interakci a uživatelský zážitek v rozšířené realitě.

OpenXR[4] je rozhraní pro programování aplikací vyvíjené společností Khronos. Rozhraní poskytuje jednotné aplikační rozhraní pro vývoj softwaru na platformy využívající rozšířenou nebo virtuální realitu. Rozhraní spolupracuje s knihovnou *MRTK*. Vývoj tohoto rozhraní je podporován mnoha velkými společnostmi, jakými jsou například Google, Intel, nVidia nebo Microsoft. Rozhraní OpenXR je v této práci využito pro implementaci funkcionality Holographic Remoting.

Sensor SDK

Sensor SDK je multiplatformní sada vývojových nástrojů od společnosti Microsoft. Sada poskytuje programový přístup k jednotlivým senzorům zařízení Azure Kinect DK. S pomocí sady lze programově konfigurovat každý ze senzorů a zachytávat jejich výstup. Kromě programového přístupu sada obsahuje také nástroje pro záznam a zobrazení dat ze zařízení Azure Kinect, které lze použít pro ověření správné funkce senzorů [40].

[3] `https://github.com/microsoft/MixedRealityToolkit-Unity`
[4] `https://www.khronos.org/openxr/`



Obrázek 6.2: Architektura knihovny Mixed Reality ToolKit.

Zdroj: `https://docs.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/architecture/overview?view=mrtkunity-2021-05`

Knihovna je do řešení přidána ve formě NuGet[5] balíčku *Microsoft.Azure.Sensor.SDK*[6]. Pro integraci knihovny do Unity jsou DLL soubory z balíčku překopírovány do složky *Plugins* v adresáři Unity projektu. *NuGet* balíček neobsahuje všechny funkce, které obsahuje výchozí knihovna pro jazyk C++. Poskytované funkce však pro tuto bakalářskou práci zcela postačují.

6.2 Zpracování obrazů a vykreslení mračna bodů

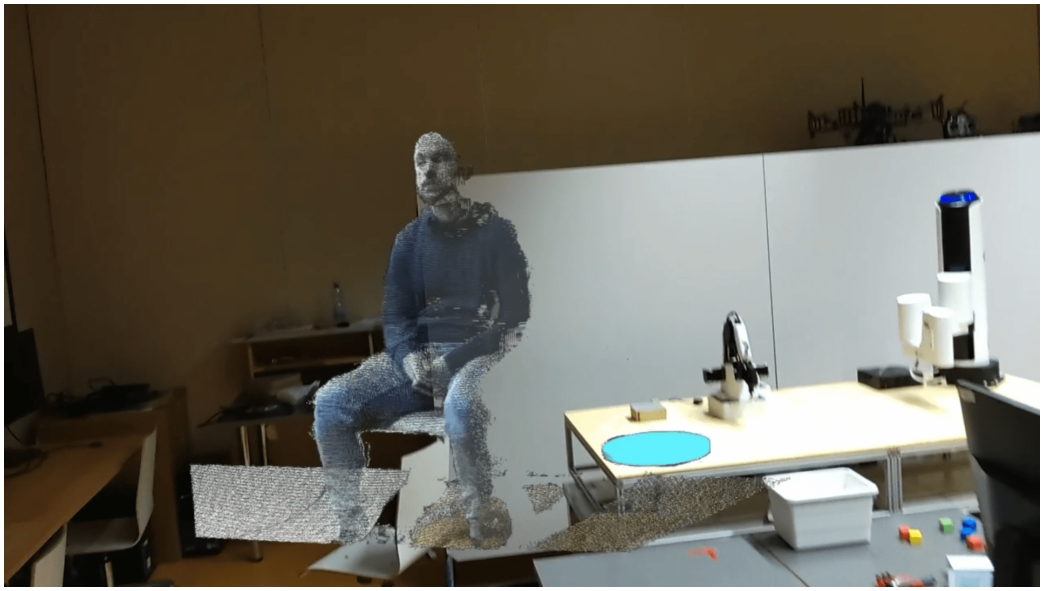
Pro snímání barevného a hloubkového obrazu je využíváno zařízení Azure Kinect, které je obsluhováno knihovnou *Sensor SDK*. Pro pohodlnější práci s touto knihovnou jsem implementoval třídu *AzureKinect*. Třída *AzureKinect* usnadňuje inicializaci a konfiguraci zařízení. Obsahuje také metody, které byly použity při ladění aplikace.

Před zahájením bezdrátového přenosu je programově ověřeno, zda je zařízení Azure Kinect připojeno k počítači. Pokud není nalezeno žádné připojené zařízení Azure Kinect, je uživateli na obrazovce vypsáno upozornění a přenos není zahájen. V opačném případě je proveden pokus o konfiguraci zařízení a pokus o spuštění senzorů. Zařízení je nakonfigurováno na rozlišení barevného obrazu 720p se snímkovací frekvencí 30ti snímků za sekundu. Hloubková kamera je nakonfigurována na využití módu *NFOV unbinned* se snímkovací frekvencí 30ti snímků za sekundu. Až po úspěšném ukončení těchto operací je zařízení připraveno zachytávat obrazy.

Po úspěšné inicializaci je spuštěn záchyt a zpracování obrazů. Záchyt a zpracování obrazů jsou implementovány v třídě *AzureKinectRenderer*. Tato třída přejímá část kódu z práce Takashiho Yoshinagy, jehož řešení je zmíněno v sekci 5.1. Konkrétně přejímá asynchronní zápis záchytu obrazů ze zařízení Azure Kinect a vykreslování mračna bodů. Zachytávání obrazů probíhá v nekonečné smyčce za pomocí asynchronního volání funkce *Get-*

[5]NuGet balíček označuje soubor s koncovkou *.nupkg*, který obsahuje obvykle kompilované knihovny a soubory související s těmito knihovnami.

[6]`https://www.nuget.org/packages/Microsoft.Azure.Kinect.Sensor/`



Obrázek 6.3: Pohled z první osoby z brýlí Microsoft HoloLens 2 po umístění hologramu.

Capture, která je součástí knihovny *Sensor SDK* a jejímž výstupem je struktura *Capture*, která obsahuje barevný i hloubkový obraz.

Obrazy ve struktuře *Capture* mají různá rozlišení a různá zorná pole. Z tohoto důvodu bylo potřeba je registrovat. V práci registruji barevný obraz na rozlišení obrazu hloubkového. Je vytvořena instance třídy *Transformation* z knihovny *Sensor SDK*, která obsahuje informace o kalibraci zařízení Azure Kinect a je potřebná pro volání funkce *ColorImageToDepthCamera*, která je taktéž součástí knihovny *Sensor SDK*. Po vykonání funkce je barevný obraz registrován s hloubkovým obrazem. Dále je provedena transformace hloubkového obrazu na mračno bodů, která taktéž využívá instanci třídy *Transformation*.

Po úspěšné registraci barevného obrazu s hloubkovým obrazem a úspěšné transformaci hloubkového obrazu na mračno bodů přichází na řadu práce Takashiho Yoshinagy. Barevný obraz a transformovaný hloubkový obraz jsou převedeny na pole. Barevný obraz je převeden na *BGRA* pole pomocí kombinace funkcí *GetPixels<BGRA>()* a *ToArray()*. Hloubkový obraz je převeden na *Short3 XYZ* pole kombinací funkce *GetPixels<Short3>()* a *ToArray()*.

Je vytvořena instance třídy *Mesh* a inicializována na počet bodů, daný rozlišením obrazu. Do instance třídy *Mesh* nelze vkládat vrcholy a barvy napřímo, musejí být vloženy do dočasného pole zvlášť a až toto pole je poté překopírováno do vrcholů a barev instance třídy *Mesh*. Při překopírovávání polí je provedena hloubková segmentace, kdy jsou souřadnice bodu, který má hloubku před prahem 1200 mm nebo za prahem 2200 mm, nastaveny na nulu. To způsobí, že při vykreslování nebudou tyto body vykresleny. Každý vrchol instance Mesh má přidělenu barvu a hloubku z polí vytvořených dříve. Po překopírování vrcholů a barev z polí do instance třídy Mesh je za pomoci Unity objektů *MeshRenderer* a *MeshFilter* mračno vykresleno. Jelikož objekt *MeshRenderer* využívá materiál pro vykreslení *meshe*, je mu přiřazen materiál, který využívá shader, který způsobí, že je místo materiálu využita barva přiřazená každému bodu při překopírování vrcholů a barev.

6.3 Bezdrátový přenos do brýlí HoloLens

Bezdrátový přenos do brýlí Microsoft HoloLens je implementován v třídě *HolographicRemoting*. Třída *HolographicRemoting* využívá k ovládání funkcionality *Holographic Remoting* metody třídy *AppRemoting*, která je součástí aplikačního rozhraní OpenXR[7]. Implementovaná třída *HolographicRemoting* obsahuje metody *Awake*, *Start*, *StartConnection* a *StopConnection*. Metoda *Awake* je volána při aktivaci skriptu. Metoda vytvoří instanci konfigurace *RemotingConfiguration* pro *Holographic Remoting* a aktivuje 2D uživatelské rozhraní pro počítač. Vykonání metody Start následuje po vykonání metody *Awake*. Metoda *Start* přiřadí tlačítka uživatelského rozhraní k příslušným akcím. Metoda *Start Connection* ověří uživatelem zadanou IP adresu a nastaví ji do objektu *RemotingConfiguration*, který byl vytvořen v metodě *Awake*. Poté je zahájen pokus o připojení ke vzdálenému uživateli. V případě úspěšného pokusu o připojení je navázáno spojení a brýle HoloLens odesílají do aplikace *LocalNetTelepresence* data o uživatelově pozici v prostoru a uživatelský vstup. Tyto informace jsou přijaty počítačem a na jejich základě je vytvořen rámec, do kterého je vykreslen obsah, který je následně odeslán zpět do brýlí.

6.4 Implementace uživatelského rozhraní

Uživatelské rozhraní jsem dle návrhu rozdělil na uživatelské rozhraní pro počítač a uživatelské rozhraní pro brýle HoloLens. V Unity jsou tato uživatelská rozhraní reprezentována jako oddělené objekty GameObject. *GameObject DesktopUI* obsahuje uživatelské rozhraní pro počítač a *GameObject HoloLensUI* obsahuje uživatelské rozhraní pro brýle HoloLens. Grafické uživatelské rozhraní pro desktop aplikaci využívá *GameObject Canvas*. *GameObject Canvas* vytváří 2D plátno, které je mateřským prvekm pro ostatní prvky 2D uživatelského rozhraní. Tlačítka ve scéně jsou vytvořena pomocí *GameObject Button*, text pomocí *GameObject Text*. Pozadí je vytvořeno jako obrázek s barvou. Vstupní pole je vytvořeno pomocí GameObject InputField. Po úspěšném zadání IP adresy a připojení k brýlím HoloLens je uživatelské rozhraní pro desktop deaktivováno a proběhne aktivace uživatelského rozhraní pro Microsoft HoloLens.

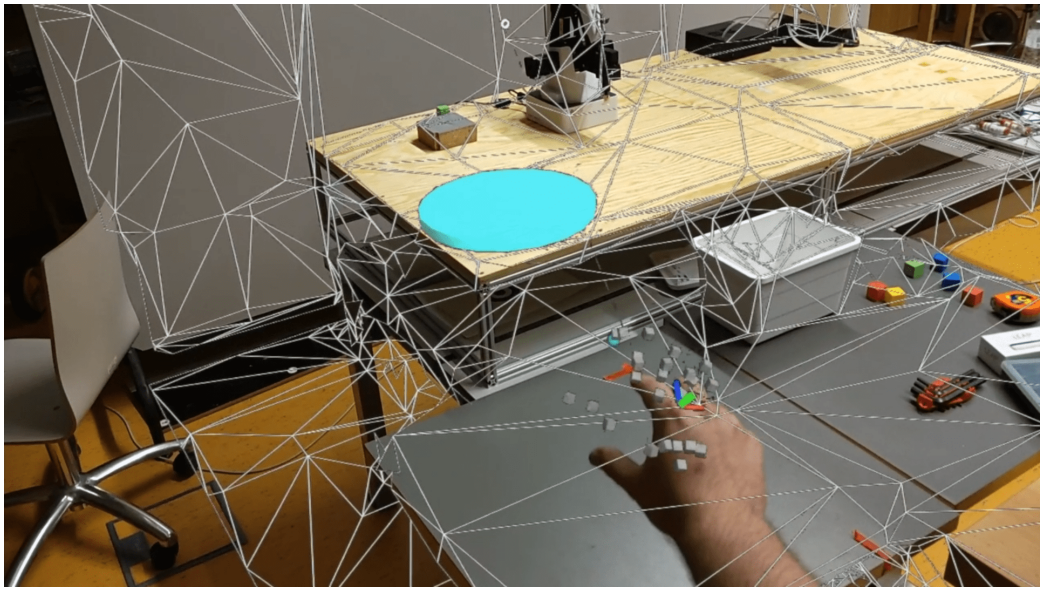
Umístění hologramu

Umístění hologramu je započato ihned po navázání spojení s brýlemi HoloLens. Při umisťování podložky je aktivní interakční komponenta z knihovny MRTK – *Tap to place*[8]. Tato komponenta využívá sledování pohybu hlavy v kombinaci s gestem *Air tap*. Uživatel pohybem hlavy pozicuje položku v prostoru. Tato komponenta je doplněna druhou komponentou – *Surface magnetism*[9]. *Surface Magnetism* zajišťuje, že se podložka přichytává k podloží, v případě této implementace k síti vytvořené prostorovým mapováním. Po umístění podložky je nad ní zobrazen hologram, přijatý pomocí funkcionality *Holographic Remoting*.

[7] `https://docs.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-create-remote-openxr`

[8] `https://docs.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/features/ux-building-blocks/solvers/tap-to-place?view=mrtkunity-2021-05`

[9] `https://docs.microsoft.com/en-us/windows/mixed-reality/design/surface-magnetism`



Obrázek 6.4: Implementovaná aplikce při umisťování hologramu s viditelnou reprezentací prostorového mapování. Pohled z první osoby z brýlí Microsoft HoloLens 2.

Prostorové mapování

Pro ovládání prostorového mapování byla implementována třída *SpatialMesh*. Třída obsahuje soukromou metodu *SetAllObservers*, veřejnou metodu *Show* a veřejnou metodu *Hide*. Metoda *SetAllObservers* je pomocnou metodou. V této metodě jsou v cyklu procházeni všichni aktivní *observers* a podle argumentu funkce jsou nastaveni na viditelné nebo neviditelné. Metoda *Show* aktivuje vizualizaci trojúhelníkové sítě a metoda *Hide* ji naopak ukončí. Trojúhelníková síť má přiřazen základní materiál pro prostorové mapování – *MRTK_Wireframe*. Volání metod *Show* a *Hide* je závislé na umisťování hologramu, pokud je hologram v režimu umisťování, je zavolána metoda *Show*. Po potvrzení pozice hologramu v prostoru je zavolána metoda *Hide* a trojúhelníková síť je skryta. Obrázek 6.4 zobrazuje podložku při umisťování hologramu s aktivní vizualizací trojúhelníkové sítě.

Kapitola 7

Testování a vyhodnocení

Výslednou aplikaci jsem testoval nejprve na simulátoru Microsoft HoloLens 2, který běžel na stejném počítači jako aplikace pro získávání dat z Azure Kinectu a zpracování, potažmo odesílání hologramu. Při nesegmentovaných datech z Azure Kinectu jsem na simulátoru dosahoval průměrně 25-30 snímků za vteřinu s odezvou nepřekračující 50 ms. Byl jsem si vědom, že pokud je spuštěn simulátor na stejném počítači, ze kterého jsou data vysílána, je rychlost přenosu vysoká. Bylo tedy potřeba testovat i na reálném hardwaru.

7.1 Měření

Veškeré měření probíhalo na počítači s procesorem Intel(R) Core(TM) i5-10300H CPU 2.50 GHz, operační pamětí 16.0 GB, systémem Windows 10 PRO 64bit a grafickou kartou nVidia GeForce RTX 2060.

Přenosová rychlost systému

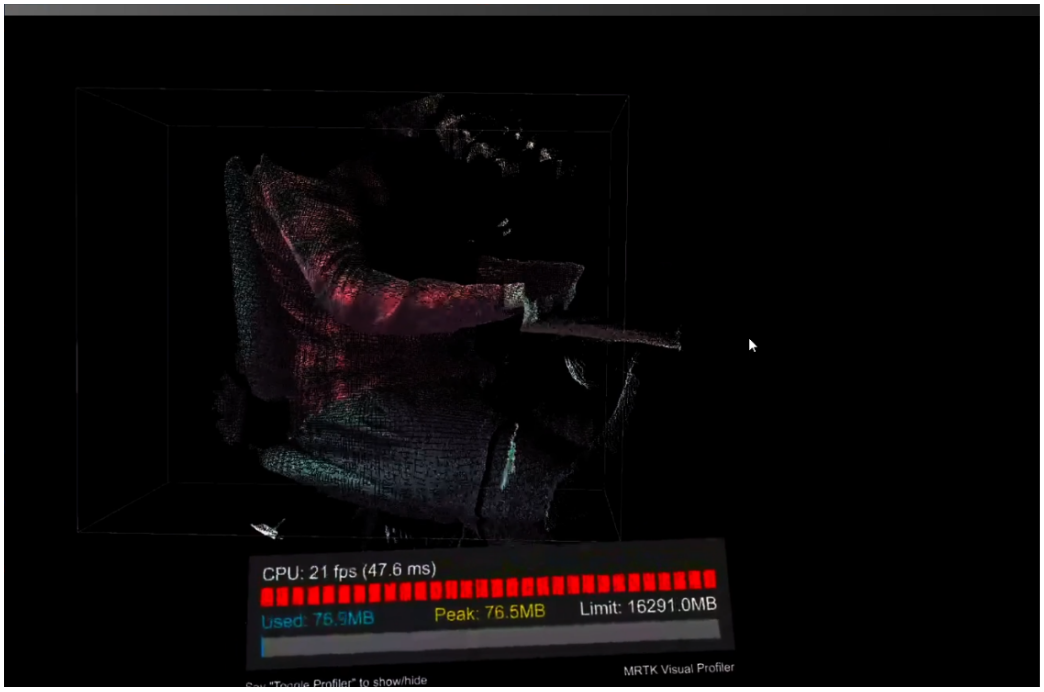
Přenosová rychlost mezi počítačem a brýlemi Microsoft HoloLens byla měřena nástrojem WireShark[1]. Komunikace pomocí *Holographic Remoting* probíhá na portu 8266 nebo na portu 8265. Z tohoto důvodu byl při měření aplikován filtr *(udp.port eq 8265) or (udp.port eq 8266)* Brýle HoloLens i počítač byly při měření připojeny k mobilnímu hotspotu zařízení Samsung Galaxy A40. Přenosová rychlost obrazu mračna limitovaného na 100 tisíc bodů je zobrazena v grafu 7.2. Přenosová rychlost obrazu mračna limitovaného na 200 tisíc bodů je zobrazena v grafu 7.3. Přenosová rychlost při přenášení obrazu mračna bez limitu je zobrazena v grafu 7.4. Přenosová rychlost obrazu mračna, hloubkově segmentovaného, je zobrazena v grafu 7.5.

Latence systému a jednotlivých operací

Latence je metrika vyjadřující dobu, která uplyne mezi akcí a reakcí v pozorovaném systému. Pro měření latence jsem využil dva nástroje. Prvním nástrojem byl *VisualProfiler*, který je součástí knihovny MRTK. Tímto nástrojem byla měřena latence přenosu přes WiFi síť. Druhým použitým nástrojem je *.NET* třída *Stopwatch*[2], která umožnila měření latence jednotlivých operací systému telepresence.

[1] `https://www.wireshark.org/`

[2] `https://docs.microsoft.com/en-us/dotnet/api/system.diagnostics.stopwatch?view=net-6.0`



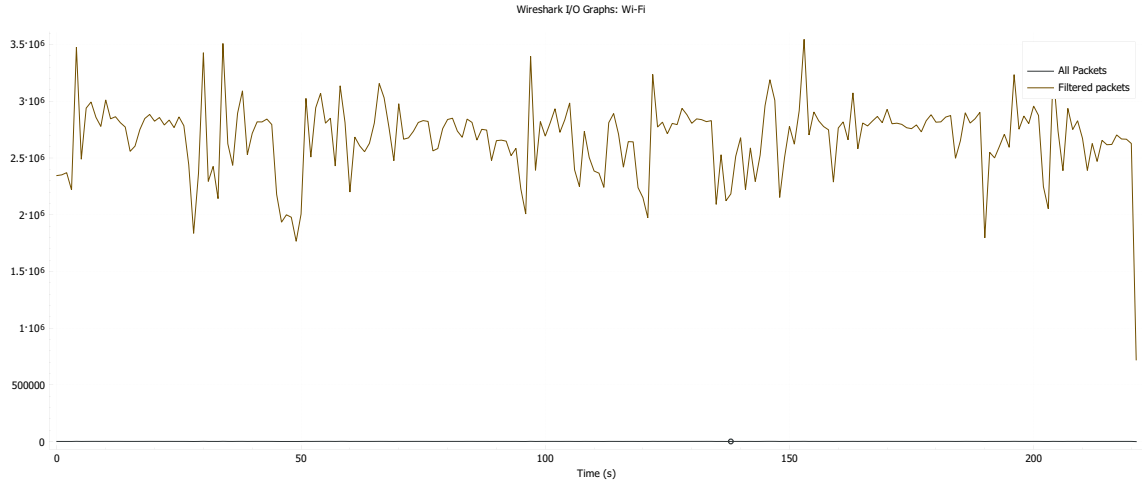
Obrázek 7.1: Vizualizace mračna bodů v simulátoru Microsoft HoloLens 2 s aktivovaným měřením *VisualProfiler*.

Rozlišení barevného obrazu	Naměřená latence v ms
720p	1.81
1080p	3.27
1536p	4.36

Tabulka 7.1: Naměřené latence operace registrace barevného obrazu s hloubkovým obrazem při změně rozlišení barevné kamery. Výsledek měření je průměrem 1000 naměřených údajů.

Naměřené latence operace registrace barevného obrazu na hloubkový je zobrazeno v tabulce 7.1. Hloubkový obraz má konstantní rozlišení, proto bylo provedeno měření transformace hloubkového obrazu na mračno bodů pouze jednou. Měření ukázalo průměrnou latenci transformace 0.63 ms. Výsledek měření je průměrem 2400 naměřených hodnot. Dále byla změřena celková latence mezi záchytem obrazů a jejich vykreslením. Naměřená latence se pohybovala v rozmezí 29 ms až 42 ms. Což přibližně odpovídá maximální snímkovací frekvenci zařízení Azure Kinect.

V tomto odstavci jsou představeny výsledky měření latence síťového přenosu pomocí Holographic Remoting. Měření zahrnuje dobu, mezi kterou jsou odeslána data z brýlí HoloLens na počítač, vykreslení rámce a přijetí obrazových dat brýlemi. Latence přenosu obrazu mračna bodů limitovaného na 100 tisíc bodů se pohybovala v rozmezí 16.6 ms a 18.5 ms. Latence přenosu obrazu limitovaného na 200 tisíc bodů se pohybovala mezi 18.5 ms a 19.2 ms. Latence přenosu obrazu v maximálním rozlišení se pohybovala mezi 17.24 ms a 31.25 ms. Latence přenosu obrazu hloubkově segmentovaného mračna bodů se pohybovala mezi 16.6 ms a 17.8 ms. Latence přenosu na vytíženější WiFi síti se pohybovala mezi 30 ms až 55 ms.



Obrázek 7.2: Přenosová rychlost při přenášení obrazu mračna limitovaného na 100 tisíc bodů. Jednotkou osy y jsou *bytes/s*.

7.2 Uživatelské testování

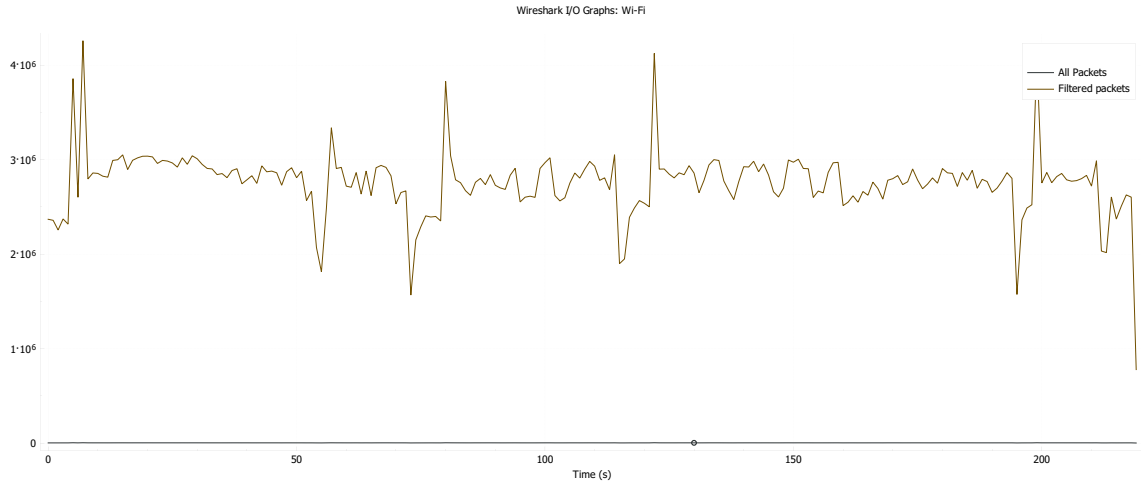
Uživatelské testování proběhlo během řešení práce dvakrát. První uživatelské testování probíhalo během vývoje a odhalené nedostatky byly zapracovány do řešení. Jedním z těchto nedostatků bylo přivrácení hologramu vůči uživateli, což způsobovalo, že hologram byl vůči podlaze nakloněn a uživatel musel manuálně hologramem rotovat, aby jej narovnal. Tato situace je zobrazena na obrázcích 7.6 a 7.7.

Druhé uživatelské testování proběhlo na studentské konferenci ExcelFIT 2022. Na konferenci jsem pozoroval, jakým problémům uživatelé při používání aplikace a brýlí čelí a co lze zlepšit. Uživatelé oceňovali trojrozměrnou reprezentaci obrazových dat a plynulost řešení. Uživatelé, kteří měli na sobě brýle poprvé měli problém s ovládáním uživatelského rozhraní. Brýle HoloLens u některých uživatelů vyvolávaly funkci kalibrace očí, což je na jednu stranu důležité pro správné zarovnání virtuálního obsahu s reálným prostředím, na druhou stranu to uživatele zdržovalo a způsobovalo přerušení aplikace a nutnost jejího opětovného spuštění.

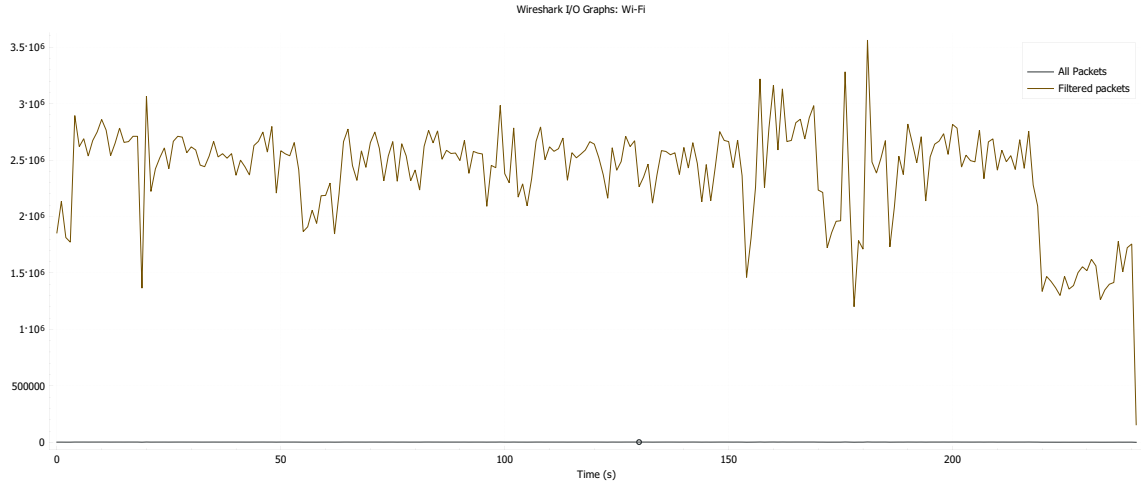
Na obrázku 6.4 je zobrazen pohled z první osoby z brýlí HoloLens při pokládání hologramu. Na obrázku 6.3 je zobrazen pohled z brýlí HoloLens po umístění hologramu.

7.3 Možná rozšíření

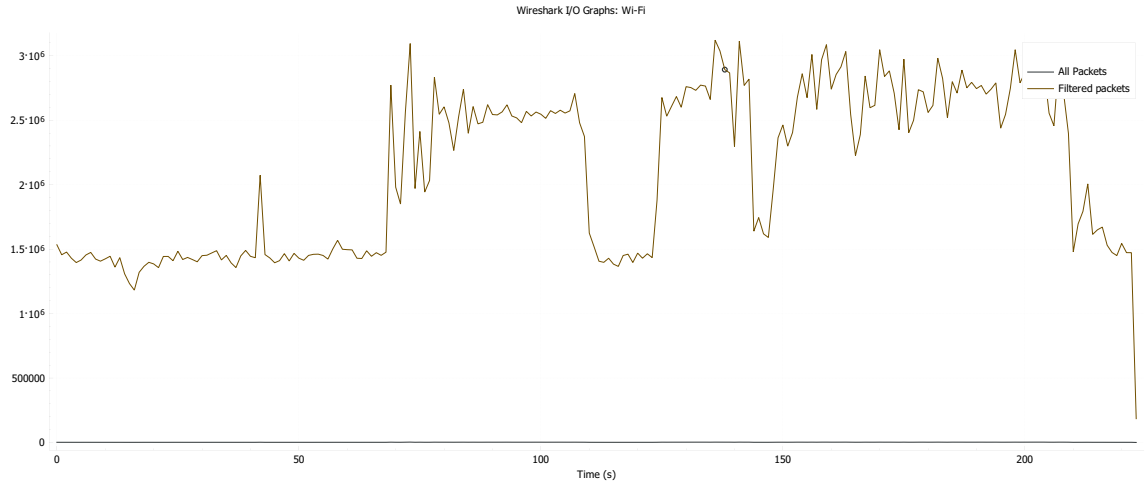
Jedním z potenciálních rozšíření této práce je přidání druhého zařízení Azure Kinect do systému. Zařízení by snímalo osobu z jiného úhlu. Mračno získané z druhého zařízení by bylo za pomoci vhodného algoritmu registrováno s mračnem získaným z prvního zařízení. Dalším možným rozšířením této práce je přidání komunikace přes internet. Toto rozšíření by spočívalo v přenosu dat mezi dvěma počítači, využívajícími implementovaný systém telepresence z této práce.



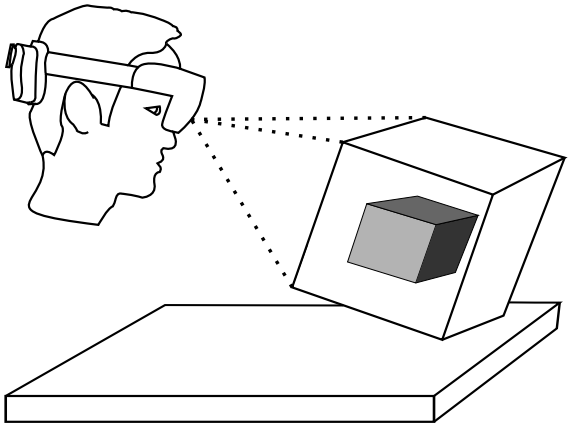
Obrázek 7.3: Přenosová rychlost při přenášení obrazu mračna limitovaného na 200 tisíc bodů. Jednotkou osy y jsou *bytes/s*.



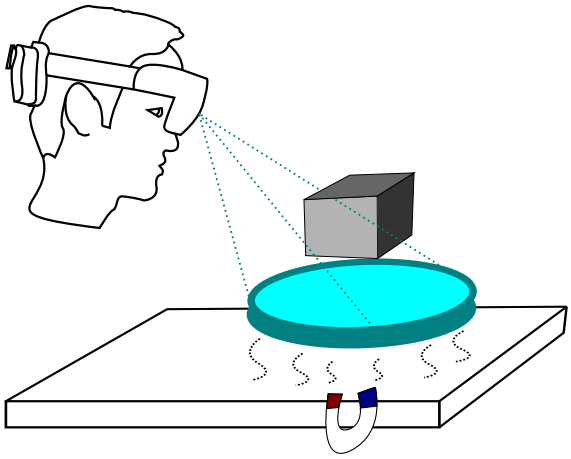
Obrázek 7.4: Přenosová rychlost při přenášení obrazu mračna v plném rozlišení. Jednotkou osy y jsou *bytes/s*.



Obrázek 7.5: Přenosová rychlost při přenášení obrazu hloubkově segmentovaného mračna bodů. Jednotkou osy y jsou *bytes/s*.



Obrázek 7.6: Vlevo je vyobrazen hologram, který se při umisťování řídí pozicí a náklonem hlavy uživatele a s reálným povrchem neinteraguje.



Obrázek 7.7: Řešení tohoto problému, kde uživatel místo hologramu pokládá virtuální podložku, která je k detekovanému povrchu přichytávána.

Kapitola 8

Závěr

Teoretickým cílem této práce bylo prozkoumat koncept rozšířené reality a možnosti zařízení Azure Kinect a Microsoft HoloLens k vytvoření telepresence. Praktickým cílem této práce bylo navrhnout a implementovat systém telepresence, který by umožnil bezdrátový přenos obrazových dat ze zařízení Azure Kinect do zařízení Microsoft HoloLens.

K analýze teoretických informací jsem využil dostupné zdroje, převážně články, související s cílem této práce. Na základě získaných informací jsem navrhl a implementoval systém telepresence, který umožňuje vytvořit z obrazových dat snímaných zařízením Azure Kinect barevné mračno bodů. Toto barevné mračno bodů je vykresleno v aplikaci v Unity a pomocí funkcionality *Holographic Remoting* je bezdrátově přeneseno do aplikace *Holographic Remoting Player* v brýlích Microsoft HoloLens, kde je mračno vizualizováno uživateli.

Implementovaný systém dosahuje na vytížené WiFi síti snímkovací frekvence až dvaceti osmi snímků za vteřinu a latence pohybující se mezi 30 ms a 55 ms. Systém je závislý na rychlosti WiFi připojení a na výkonu počítače. Řešení funguje pouze přes lokální síť. Práce byla prezentována jako demonstrátor na studentské konferenci Excel@FIT 2022.

Literatura

[1] Aleksy, M., Troost, M., Scheinhardt, F. a Zank, G. T. Utilizing hololens to support industrial service processes. In: IEEE. *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*. 2018, s. 143–148.

[2] Azuma, R. T. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*. 1997.

[3] *Azure Kinect DK — Develop AI Models | Microsoft Azure* [online]. 2022 [cit. 2022-05-05]. Dostupné z: `https://azure.microsoft.com/en-us/services/kinect-dk/`.

[4] Bambuŝek, D., Materna, Z., Kapinus, M., Beran, V. a Smrž, P. Combining interactive spatial augmented reality with head-mounted display for end-user collaborative robot programming. In: IEEE. *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 2019, s. 1–8.

[5] Billinghurst, M., Clark, A. a Lee, G. A survey of augmented reality. *Foundations and Trends in Human-Computer Interaction*. 2015, sv. 8, 2-3. ISSN 1551-3963.

[6] Billinghurst, M., Grasset, R. a Looser, J. Designing augmented reality interfaces. *ACM Siggraph Computer Graphics*. ACM New York, NY, USA. 2005, sv. 39, č. 1, s. 17–22.

[7] Billinghurst, M. a Kato, H. Collaborative augmented reality. *Communications of the ACM*. ACM New York, NY, USA. 2002, sv. 45, č. 7, s. 64–70.

[8] Bimber, O. a Raskar, R. *Spatial augmented reality: merging real and virtual worlds*. CRC press, 2005.

[9] Borrmann, D., Leutert, F., Schilling, K. a Nüchter, A. Spatial projection of thermal data for visual inspection. In: IEEE. *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 2016, s. 1–6.

[10] Bradski, G. a Kaehler, A. *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.", 2008.

[11] *Eye tracking overview* [online]. 2022 [cit. 2022-05-02]. Dostupné z: `https://docs.microsoft.com/en-us/windows/mixed-reality/design/eye-tracking`.

[12] Fitzmaurice, G. W. a Buxton, W. An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In: *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. 1997, s. 43–50.

[13] Gauglitz, S., Nuernberger, B., Turk, M. a Höllerer, T. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. In: *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*. 2014, s. 197–205.

[14] Grubert, J., Grasset, R. a Reitmayr, G. Exploring the design of hybrid interfaces for augmented posters in public spaces. In:. říjen 2012, s. 238–246. DOI: 10.1145/2399016.2399053. ISBN 9781450314824.

[15] *Head-gaze and commit – Mixed Reality* [online]. 2022 [cit. 2022-04-16]. Dostupné z: https://docs.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-commit-head.

[16] *Head-gaze and dwell – Mixed Reality* [online]. 2022 [cit. 2022-04-16]. Dostupné z: https://docs.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-dwell-head.

[17] Heinrich, F., Schwenderling, L., Joeres, F., Lawonn, K. a Hansen, C. Comparison of Augmented Reality Display Techniques to Support Medical Needle Insertion. *IEEE Transactions on Visualization and Computer Graphics*. IEEE. 2020, sv. 26, č. 12, s. 3568–3575.

[18] *HoloLens 2 Display Troubleshooting* [online]. 2021 [cit. 2022-04-26]. Dostupné z: https://docs.microsoft.com/en-us/hololens/hololens2-display.

[19] *HoloLens 2 gestures for authoring and navigating in Dynamics 365 Guides* [online]. 2021 [cit. 2022-05-02]. Dostupné z: https://docs.microsoft.com/en-us/dynamics365/mixed-reality/guides/authoring-gestures-hl2.

[20] *HoloLens 2 hardware* [online]. 2021 [cit. 2022-04-26]. Dostupné z: https://docs.microsoft.com/en-us/hololens/hololens2-hardware.

[21] *HoloLens Research Mode* [online]. 2022 [cit. 2022-05-02]. Dostupné z: https://docs.microsoft.com/en-us/windows/mixed-reality/develop/advanced-concepts/research-mode.

[22] Hu, L., Wang, M. a Song, Z. A Convenient Method of Video See-Through Augmented Reality Based on Image-Guided Surgery System. In: *2013 Seventh International Conference on Internet Computing for Engineering and Science*. 2013, s. 100–103. DOI: 10.1109/ICICSE.2013.27. ISBN 978-0-7695-5118-0.

[23] Jamhoury, L. *Streaming Volumetric Images Online* [online]. 2018 [cit. 2022-04-16]. Dostupné z: https://lisajamhoury.medium.com/streaming-volumetric-images-online-a04d68703973.

[24] *Azure Kinect DK hardware specifications* [online]. 2021 [cit. 2021-11-26]. Dostupné z: https://docs.microsoft.com/en-us/azure/Kinect-dk/hardware-specification.

[25] KIPPER, G. a. J. R. *Augmented reality: an emerging technologies guide to AR.* Waltham: Syngress, 2013. ISBN 978-1-59749-733-6.

[26] Kiyokawa, K., Kurata, Y. a Ohno, H. An optical see-through display for mutual occlusion of real and virtual environments. In: IEEE. *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000).* 2000, s. 60–67.

[27] Kowalski, M., Naruniec, J. a Daniluk, M. Livescan3d: A fast and inexpensive 3d data acquisition system for multiple kinect v2 sensors. In: IEEE. *2015 international conference on 3D vision.* 2015, s. 318–325.

[28] Lackner, J. Simulator sickness. *The Journal of the Acoustical Society of America.* Říjen 1992, sv. 92, s. 2458. DOI: 10.1121/1.404501.

[29] Linsen, L. *Point cloud representation.* 2001.

[30] Lourenço, F. a Araujo, H. Intel RealSense SR305, D415 and L515: Experimental Evaluation and Comparison of Depth Estimation. In: *VISIGRAPP (4: VISAPP).* 2021, s. 362–369. ISBN 978-989-758-488-6.

[31] Martin, A., Cheriyan, J., Ganesh, J., Sebastian, J. et al. Indoor Navigation using Augmented Reality. *EAI Endorsed Transactions on Creative Technologies.* 2021, sv. 8, č. 26, s. e1.

[32] Milgram, P. a Kishino, F. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems.* The Institute of Electronics, Information and Communication Engineers. 1994, sv. 77, č. 12, s. 1321–1329.

[33] Moore, A. a Regenbrecht, H. The tangible augmented street map. In: *Proceedings of the 2005 international conference on Augmented tele-existence.* Association for Computing Machinery, 2005, s. 249–250. ISBN 0473106574.

[34] Nizam, S. M., Abidin, R. Z., Hashim, N. C., Lam, M. C., Arshad, H. et al. A review of multimodal interaction technique in augmented reality environment. *Int. J. Adv. Sci. Eng. Inf. Technol.* 2018, sv. 8, 4-2, s. 8–4.

[35] Oberdörfer, S., Elsässer, A., Schraudt, D., Grafe, S. a Latoschik, M. E. Horst - The Teaching Frog: Learning the Anatomy of a Frog Using Tangible AR. In: *Proceedings of the Conference on Mensch Und Computer.* New York, NY, USA: Association for Computing Machinery, 2020, s. 303–307. MuC '20. DOI: 10.1145/3404983.3410007. ISBN 9781450375405. Dostupné z: `https://doi.org/10.1145/3404983.3410007`.

[36] Pan, Z., Cheok, A. D., Yang, H., Zhu, J. a Shi, J. Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics.* 2006, sv. 30, č. 1, s. 20–28. DOI: https://doi.org/10.1016/j.cag.2005.10.004. ISSN 0097-8493. Dostupné z: `https://www.sciencedirect.com/science/article/pii/S0097849305002025`.

[37] Raskar, R., Welch, G. a Fuchs, H. Spatially augmented reality. *Augmented Reality: Placing Artificial Objects in Real Scenes.* AK Peters/CRC Press. 1999, s. 64–71.

[38] Rolland, J. P., Holloway, R. L. a Fuchs, H. Comparison of optical and video see-through, head-mounted displays. In: SPIE. *Telemanipulator and Telepresence Technologies*. 1995, sv. 2351, s. 293–307.

[39] SCHMALSTIEG, D. a. T. H. *Augmented reality: principles and practice*. 1. vyd. Addison Wesley: Boston, 2016. ISBN 978-0-321-88357-5.

[40] *Azure Kinect Sensor SDK download | Microsoft Docs* [online]. 2021 [cit. 2022-04-28]. Dostupné z: https://docs.microsoft.com/en-us/azure/kinect-dk/sensor-sdk-download.

[41] *Spatial mapping* [online]. 2022 [cit. 2022-05-02]. Dostupné z: https://docs.microsoft.com/en-us/windows/mixed-reality/design/spatial-mapping.

[42] *Speech recognition – Windows apps* [online]. 2021 [cit. 2022-04-16]. Dostupné z: https://docs.microsoft.com/en-us/windows/apps/design/input/speech-recognition.

[43] Steuer, J. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*. Wiley Online Library. 1992, sv. 42, č. 4, s. 73–93.

[44] Tölgyessy, M., Dekan, M., Chovanec, L. a Hubinskỳ, P. Evaluation of the azure Kinect and its comparison to Kinect V1 and Kinect V2. *Sensors*. Multidisciplinary Digital Publishing Institute. 2021, sv. 21, č. 2, s. 413.

[45] Turini, G., Condino, S., Parchi, P. D., Viglialoro, R. M., Piolanti, N. et al. A microsoft hololens mixed reality surgical simulator for patient-specific hip arthroplasty training. In: Springer. *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. 2018, s. 201–210.

[46] Vagala, D. *Náhled barvy na stěně pomocí rozšířené reality v mobilu*. Brno, CZ, 2020. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: https://www.fit.vut.cz/study/thesis/23181/.

[47] Vidakis, N., Konstantinos, K. a Triantafyllidis, G. A Multimodal Interaction Framework for Blended Learning. In: *Interactivity, Game Creation, Design, Learning, and Innovation*. Springer, 2016, s. 205–211.

[48] *Voice input – Mixed Reality* [online]. 2022 [cit. 2022-04-16]. Dostupné z: https://docs.microsoft.com/en-us/windows/mixed-reality/design/voice-input.

[49] Wagner, D., Schmalstieg, D. a Billinghurst, M. Handheld AR for collaborative edutainment. In: Springer. *International Conference on Artificial Reality and Telexistence*. 2006, s. 85–96.

[50] Wellner, P. Interacting with paper on the DigitalDesk. *Communications of the ACM*. ACM New York, NY, USA. 1993, sv. 36, č. 7, s. 87–96.

[51] Young, J., Langlotz, T., Mills, S. a Regenbrecht, H. Mobileportation: Nomadic Telepresence for Mobile Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. ACM New York, NY, USA. 2020, sv. 4, č. 2, s. 1–16.

[52] Zhang, W., Han, B., Hui, P., Gopalakrishnan, V., Zavesky, E. et al. CARS: Collaborative augmented reality for socialization. In: *Proceedings of the 19th International Workshop on Mobile computing Systems & Applications*. 2018, s. 25–30.

[53] Zhou, F., Duh, H. B.-L. a Billinghurst, M. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In: IEEE. *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. 2008, s. 193–202.

[54] Zhu, Y. a Wang, S. J. A tangible augmented reality toy kit: Interactive solution for early childhood education. In: *Interactivity, game creation, design, learning, and innovation*. Springer, 2016, s. 12–19.