

KAUNO TECHNOLOGIJOS UNIVERSITETAS
VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

JULIUS GELŠVARTAS

**DAUGIAFUNKCĖS VARTOTOJO SAŠAJOS
PARALYŽIUOTIESIEMS KŪRIMAS IR TYRIMAI**

Daktaro disertacijos santrauka

Technologijos mokslai, informatikos inžinerija (T 007)

2019, Kaunas

Disertacija rengta 2014-2018 metais Kauno technologijos universiteto Elektros ir elektronikos fakultete, Automatikos katedroje. Mokslinius tyrimus rėmė Lietuvos mokslo taryba.

Mokslinis vadovas: prof. habil. dr. Rimvydas SIMUTIS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, T 007).

Mokslinis konsultantas: dr. Rytis MASKELIŪNAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, T 007).

Redagavo: Aurelija Gražina Rukšaitė (leidykla „Technologija“)

Informatikos inžinerijos mokslo krypties disertacijos gynimo taryba:

prof. habil. dr. Antanas VERIKAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, T007) – pirmininkas;

dr. Lalit GARG (Maltos universitetas, technologijos mokslai, informatikos inžinerija, T007);

prof. dr. Arnas KAČENIAUSKAS (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija, T007);

prof. dr. Olga KURASOVA (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija, T007);

doc. dr. Evaldas VAIČIUKYNAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, T007).

Disertacija bus ginama 2019 m. rugpjūčio 22 d. 10.00 val. viešame Informatikos inžinerijos mokslo krypties disertacijos gynimo tarybos posėdyje Kauno technologijos universiteto Disertacijų gynimo salėje.

Adresas: K. Donelaičio g. 73-403, 44249 Kaunas, Lietuva.

Tel. (370) 37 30 00 42; faks. (370) 37 32 41 44; el. paštas doktorantura@ktu.lt.

Disertacijos santrauka išsiųsta 2019 m. liepos 22 d.

Su disertacija galima susipažinti Kauno technologijos universiteto (K. Donelaičio g. 20, Kaunas), Vilniaus Gedimino technikos universiteto (Saulėtekio al. 14, Vilnius) bibliotekose bei internete (<http://ktu.edu>).

Padėka

Autorius nori padėkoti savo vadovui Rimvydui Simučiiui. Šis nueitas kelias būtų daug sunkesnis be Rimvydo padrašinimo ir pagalbos. Taip pat noriu padėkoti savo moksliniam konsultantui Ryčiui Maskeliūnui. Esu dėkingas visai savo šeimai ir draugams, kurie palaikė mane šiame kelyje. Ypatingą padėką skiriu savo žmonai Eglei ir vaikams Ramunei ir Liudvikui už jų kantrybę ir paramą. Galiausiai noriu padėkoti kolegoms iš Rubedo sistemų už galimybę derinti mokslą su darbu.

Šią disertaciją skiriu savo a.a. seneliui Algirdui Karaliui. Pokalbiai su juo paskatino domėtis ir pamilti mokslą.

”...no-one is actually dead until the ripples they cause in the world die away... The span of someone’s life, they say, is only the core of their actual existence.”

— Terry Pratchett, Reaper Man

Turinys

1	Ižanga	7
1.1	Problemos pristatymas ir aktualumas	7
1.2	Tyrimo objektas	8
1.3	Tikslas	8
1.4	Disertacijos uždaviniai	8
1.5	Tyrimų metodologija	9
1.6	Mokslinis naujumas	9
1.7	Praktinė reikšmė	10
1.8	Ginamieji teiginiai	11
1.9	Mokslinis pripažinimas	12
1.10	Santraukos struktūra	13
2	Esamų technologijų neįgaliesiems analizė	14
2.1	Daugiafunkcės vartotojo sąsajos struktūra	14
2.2	Pagalbiniai įrenginiai	14
2.2.1	Jungiklio įrenginiai	14
2.2.2	Liežuviu valdomi įrenginiai	15
2.2.3	Smegenų kompiuterio sąsaja	15
2.2.4	Akių judesio sekimas	15
2.2.5	Kalbos atpažinimas	15
2.2.6	Raumenų judesiai	16
2.2.7	Pagalbinių įrenginių palyginimas	16
2.3	Informacijos apdorojimas	16
2.3.1	Signalų jungimas ir klaidų aptikimas	16
2.3.2	Teksto nuspėjimas	16
2.4	Informacijos pristatymas	18
2.4.1	Virtuali realybė	18
2.4.2	Papildyta realybė	18
2.4.3	Projektuojami vaizdai	18
2.4.4	Informacijos pristatymo metodų palyginimas	18
2.5	Aplinkos suvokimas	19
2.5.1	Kameros valdymas	20
2.5.2	Stereoskopinių vaizdų suliejimas	20

2.5.3	Objektų aptikimas 3D vaizduose	21
2.6	Sąveika su aplinka	21
2.7	Analizės apibendrinimas	21
3	Daugiafunkcės vartotojo sąsajos architektūra	22
3.1	Daugiafunkcės vartotojo sąsajos reikalavimai	22
3.2	Daugiafunkcės vartotojo sąsajos panaudos atvejai	22
3.3	Sistemos diegimo aprašas	24
3.4	Daugiafunkcės vartotojo sąsajos architektūros aprašas	24
3.5	Detali komponentų įgyvendinimo informacija	24
3.6	Komponentų informacijos tėkmė	24
3.7	Architektūros apibendrinimas	24
4	Pasiūlyti daugiafunkcės vartotojo sąsajos algoritmai	25
4.1	Projektuojamų vaizdų vartotojo sąsaja	25
4.1.1	Objektų aptikimo algoritmas	28
4.1.2	Objekto paryškimas naudojant projektorius	29
4.2	Kameros valdymas	29
4.3	Stereoskopinių vaizdų apdorojimas maksimaliai stabilių ekstremalių regionų metodu	32
4.4	Specializuotas teksto nuspėjimo algoritmas	32
4.5	Pasiūlytų algoritmų apibendrinimas	33
5	Eksperimentų apibūdinimas ir rezultatai	34
5.1	Eksperimentų metodologija	34
5.2	Daugiafunkcės vartotojo sąsajos eksperimentai	34
5.3	Kameros valdymo eksperimentai	35
5.4	Stereosuliejimo eksperimentai	35
5.5	Teksto nuspėjimo algoritmo eksperimentai	37
5.6	Eksperimentų apibendrinimas	38
6	Išvados	39
7	CV	47

8	Summary	48
8.1	Introduction	48
8.1.1	Problem statement and relevance	48
8.1.2	Thesis object	48
8.1.3	Thesis goal	48
8.1.4	Tasks of the thesis	48
8.1.5	Scientific novelty	49
8.1.6	The defended statements	49
8.1.7	Scientific approval	50
8.1.8	Thesis organization	51
8.2	Analysis of existing assistive technology solutions	51
8.3	Architecture of multifunctional user interface	51
8.4	Proposed algorithms and methods for multifunctional user interface	52
8.5	Experimental setup and results	53
8.6	Conclusions	53

1. IŽANGA

1.1. Problemos pristatymas ir aktualumas

Bendravimas yra vienas svarbiausių įgūdžių, kuriuos išmokstame pirmaisiais gyvenimo metais ir naudojame visą likusį gyvenimą. Bendravimas reikalingas išsakant savo mintis, bendradarbiaujant ir prašant pagalbos. Bendravimo sutrikimai gali vesti prie socialinės atskirties, streso ir nuolatinės baimės. Didelę negalią turintiems žmonėms dažnai pasireiškia kalbos sutrikimai. Kalbos sutrikimai gali trukdyti lengvai bendrauti. Ši problema dažnai paveikia neįgaliųjų bei jų artimųjų gyvenimo kokybę [Westgren and Levi, 1998]. Žmonėms, turintiems tokias negalias, kaip tetraplegija, šios problemos pasireiškia labai dažnai.

Tetraplegija - tai negalia, pasireiškianti visų keturių galūnių paralyžiumi. Tetraplegija atsiranda dėl galvos arba stuburo smegenų pažeidimų. Šių pažeidimų priežastys gali būti traumos, ligos arba įgimti sutrikimai. Tetraplegiją sukeltantys pažeidimai dažniausiai būna tarp C1 (aukščiausio stuburo slankstelio, esančio prie kaukolės) ir C7. Tetraplegija gali būti dviejų tipų: visiška ir nevisiška. Esant nevisiškai tetraplegijai, žmonės gali dalinai valingai judinti raumenis. Kai yra visiška tetraplegija, žmonės visiškai nevaldo raumenų ir neturi pojūčio jausmo. Ilgalaikės tokių negalių prognozės dažniausiai yra labai individualios ir sunkiai nuspėjamos.

Sunkias negalias turintiems žmonėms reikalinga nuolatinė priežiūra. Tokia priežiūra reikalauja daug laiko ir yra brangi. Tetraplegija labai dažnai yra stuburo traumų (*Spinal cord injury (SCI)*) pasekmė. SCI slauga brangiausiai kainuoja pirmaisiais metais po traumos. Taip yra dėl didelių medicininių išlaidų. Yra apskaičiuota, kad SCI kainuoja iki 800 000USD pirmaisiais metais po traumos ir iki 200 000USD kiekvienais kitais metais. Tik apie 35 proc. SCI patyrusių žmonių gali grįžti į darbą. Tokių žmonių gyvenimo trukmė dažnai yra trumpesnė dėl padidėjusių rizikos susirgti kvėpavimo takų infekcija ir inkstų problemų. Neįgaliųjų, kuriems suteikiama galimybė gyventi nepriklausomai, gyvenimo kokybė yra daug geresnė. Tai taip pat leidžia sumažinti slaugymo išlaidas.

Neįgaliesiems skirtos technologijos (*Assistive technology (AT)*) leidžia jiems atlikti naujus veiksmus. AT susideda iš dviejų pagrindinių dalių, t.y. pagalbinių įrenginių ir žmogaus kompiuterio sąveikos (*Human-Computer*

interaction (HCI)) sąsajos. Kai kurie pagalbinaie įrenginiai neturi HCI dalies ir yra naudojami tiesiogiai paveikti aplinką. Pavyzdžiui, siurbimo ir pūtimo įrenginys gali būti prijungiamas prie neįgaliųjų vežimėlio, kad jį valdytų. Tačiau tokia sistema nėra labai lanksti. Pagalbinaie įrenginiai, prijungti prie kompiuterio, gali būti naudojami daugeliui įvairių veiksmų atlikti.

Kiekvienas stuburo pažeidimas yra skirtingas, todėl geriausių rezultatų pasiekama taikant individualų reabilitacijos procesą ir individualiai parenkant AT. Šiuo metu naudojami pagalbinaie įrenginiai veikia neproduktyviai. Kiekvienas įrenginys turi savo vartotojo sąsają (*User interface (UI)*) ir programinę įrangą. Dėl šios priežasties yra labai sunku arba neįmanoma naudoti keleto įrenginių vienu metu. Taip pat, keičiant įrenginį, kiekvieną kartą reikia išmokti naudotis nauja UI. Tai sunkina AT individualaus pritaikymo galimybes. Dėl to dažnai yra naudojamas tik vienas įrenginys. Keičiant įrenginius, kai keičiasi vartotojo galimybės, galima užtikrinti efektyvesnę sistemos naudojimą. UI, veikianti su keliais skirtingais įrenginiais, išplėstų individualaus pritaikymo galimybes. Tokia sistema taip pat pagerintų sistemos naudojimo našumą ir vartotojų patitenkinimą.

Adaptyvios AT sistemos yra svarbios ne tik tetraplegijos negalią turintiems žmonėms. Tokios sistemos gali būti naudojamos, kai žmonės serga tokiomis ligomis, kaip Hantingtono liga arba cerebrinis paralyžius.

1.2. Tyrimo objektas

Šios disertacijos tyrimo objektas yra neįgaliems vartotojams skirti įrenginiai ir modernūs HCI metodai.

1.3. Tikslas

Šios disertacijos tikslas yra padėti žmonėms, turintiems dideles motorines negalias, bendrauti efektyviai. Pasiūlytame sprendime didelis dėmesys yra skiriamas lengvam sistemos pritaikymui kiekvieno vartotojo poreikiams. Nuolatinis sistemos koregavimas užtikrina efektyvų žmogaus kompiuterio bendravimą (HCI), nepriklausomai nuo vartotojo motorinės negalios lygio.

1.4. Disertacijos uždaviniai

Darbo tikslas pasiektas išsprendus šias užduotis:

1. Išanalizuoti ir palyginti neįgaliems vartotojams skirtas technologijas ir modernius HCI metodus;
2. Suprojektuoti daugiafunkčę UI sistemos architektūrą;
3. Pasiūlyti realaus laiko aplinkos suvokimo ir papildytos realybės (*Augmented reality (AR)*) vartotojo sąsajos algoritmus;
4. Pasiūlyti nuspėjimo algoritmus, sutrumpinančius sistema atliekamų veiksmų laiką;
5. Sukurti adaptyvią AT vartotojo sąsają, veikiančią su įvairiais pagalbinais įrenginiais. Vaizdų apdorojimo ir AR algoritmai buvo integruoti į sukurtą sistemą;
6. Atlikti eksperimentus ir įvertinti sukurtą UI bei pasiūlytus algoritmus.

1.5. Tyrimų metodologija

Du pagrindiniai metodai buvo naudojami siekiant gauti disertacijoje pristatomus rezultatus. Pirmiausia buvo prašoma žmonių grupės naudotis sukurta sistema. Eksperimentų metu buvo įrašinėjami sistemos naudojimo duomenys. Šie duomenys leido nustatyti sistemos naudojimo efektyvumą. Antra - pasiūlyti algoritmai buvo įvertinami naudojant egzistuojančius arba rankiniu būdu sužymėtus duomenų rinkinius. Detalesnę metodologijos aprašymą žr. 5.1.

1.6. Mokslinis naujumas

Tyrimų metu buvo pasiekti šie nauji rezultatai:

1. Sukurta daiktų parinkimo vartotojo sąsaja UI. Novatoriškas automatinis projektuojamų vaizdų metodas buvo gautas sujungus kameros-projektoriaus sistemos kalibravimą su objektų aptikimo (spalvos ir gylis kameros (*Color and depth camera (RGB-D)*) vaizduose) algoritmu. Ši sistema automatiškai aptinka objektus ir juos paryškina realioje aplinkoje. Naudodamas tokią sąsają, vartotojas gali pasirinkti norimą daiktą.
2. Pasiūlytas regionais paremtas stereoskopinių vaizdų apdorojimo algoritmas. Šis algoritmas gali realiuoju laiku apdoroti vaizdus įterptiniame kompiuteryje. Pagrindinis pasiūlyto stereoskopinių vaizdų

sujungimo algoritmo naujumas - regionais paremtų taškų naudojimas, siekiant sumažinti sutapdinimo algoritmo paieškos erdvę. Stereoskopinė kamera su įterptiniu kompiuteriu yra naudojama kaip aplinkos suvokimo įrenginys.

3. Kameros, veikiančios kintančių apšvietimo sąlygų aplinkose, turi turėti sparčius kameros parametrų valdymo algoritmus. Buvo pasiūlytas naujas kameros parametrų valdymo algoritmas. Algoritmo naujumas yra keleto išvesties parametrų valdymas vienu metu. Dar vienas algoritmo naujumas yra automatinis stiprinimo koeficientų keitimas, užtikrinantis valdiklio greitaveiką ir stabilumą. Šis algoritmas leidžia užtikrinti patikimą vaizdų apdorojimo algoritmo veikimą.

1.7. Praktinė reikšmė

JAV kasmet yra patiriama iki 12 000 SCI. SCI traumas išgyvenusiems žmonėms yra reikalinga nuolatinė slauga. Tokia slauga užima daug laiko ir brangiai kainuoja. Šioje disertacijoje yra pristatoma adaptyvi pagalbinių technologijų sistema. Tokia sistema turi keletą privalumų. Pirmas - vartotojas gali pradėti naudoti sistemą ankstyvoje reabilitacijos stadijoje. Antras - keičiantis vartotojo būklei, sistemos adaptyvumas užtikrina, kad nesumažės sistemos naudojimo efektyvumas. Šie veiksniai leidžia sumažinti slaugos išlaidas. Išlaidos taupomos dėl potencialiai sutrumpėjusio reabilitacijos laiko ir dėl to, kad vartotojas gali nepriklausomai atlikti naujas užduotis. Sistemos adaptyvumas taip pat yra labai svarbus progresuojančiomis motorinėmis ligomis sergantiems žmonėms. Čia sistema gali būti iš naujo pritaikoma, kai padidėja naudotojo negalios lygis.

Kelių įrenginių integravimas į sistemą padidina jos adaptyvumą. Tokia sistema taip pat turi ir kitų praktinių pranašumų. Nebrangių pagalbinių įrenginių integravimas padaro sistemą prieinamą žmonėms, gaunantiems mažas pajamas. Tai tampa vis svarbiau, nes rinkoje daugėja nebrangių pagalbinių įrenginių. Nebrangūs įrenginiai dažnai neturi specializuotos programinės įrangos. Sistema, pristatoma šioje disertacijoje, siekia išspręsti šią problemą.

Disertacijoje taip pat analizuojama, kaip naujos technologijos, tokios kaip papildyta realybė (AR), gali būti pritaikomos pagalbinėse sistemose. Darbe pristatoma, kaip projektuojamų vaizdų metodas gali būti naudoja-

mas kuriant intuityvų objektų pasirinkimo HCI metodą. Dabartinės AR technologijos vis dar susiduria su problemomis realiose taikymo srityse. Disertacijoje pristatomi vaizdų apdorojimo algoritmai gali būti naudojami siekiant išspręsti dalį šių problemų. Tai leistų sukurti patikimesnes AR sistemas. Papildytos realybės sąsajos gali būti naudojamos kuriant pagalbinės technologijas ir kitose srityse.

Pasiūlyta sistema buvo kuriama Mokslo, inovacijų ir technologijų agentūros (MITA) EUREKA projekto QUADRIBOT metu. Projekto metu buvo parengtos sistemos adaptacijos ir konfigūracijos rekomendacijos. Pasiūlytas stereovaizdų apdorojimo algoritmas buvo kurtas Lietuvos mokslo tarybos projekto REP-15114 metu. Projektuojamų vaizdų vartotojo sąsaja buvo kuriama Europos bendradarbiavimo moksle ir technologijose COST veiklos CA15122 metu.

1.8. Ginamieji teiginiai

1. Neįgaliesiems skirtos vartotojo sąsajos, naudojančios keletą įvesties įrenginių vienu metu, yra adaptyvesnės. Sistemos adaptyvumas užtikrina, kad sistema bus tinkama didesnei vartotojų grupei. Įvesties įrenginiai yra naudojami daugiafunkcei vartotojo sąsajai valdyti.
2. Kameros išlaikymo ir jutiklio stiprinimo parametrai gali būti efektyviai valdomi vienu metu naudojant pakopinį valdiklį. Valdiklio stabilumas gali būti užtikrinamas naudojant stiprinimo koeficientų automatinį adaptavimą. Taip yra sumažinama proceso netiesiškumo įtaka valdiklio darbui. Šis algoritmas yra naudojamas daugiafunkcės vartotojo sąsajos vaizdų apdorojimo algoritmo patikimumui užtikrinti.
3. Stereovaizdų apdorojimo algoritmas, paremtas maksimaliai stabiliais ekstremaliais regionais, gali sumažinti algoritmo paieškos erdvę. Sumažinus paieškos erdvę, sutrumpėja algoritmo skaičiavimo laikas ir sumažėja naudojamas atminties kiekis. Šis algoritmas gali būti naudojamas įterptinėse sistemose, tokiose kaip Jetson TK1. Tokia sistema yra naudojama kaip patikimas aplinkos suvokimo įrenginys.
4. Specializuotas teksto nuspėjimo algoritmas gali padėti įvesti tekstą greičiau negu standartiniai teksto nuspėjimo algoritmai. Toks nuspėjimo algoritmas yra naudojamas efektyviai įvesti tekstą teminių

pokalbių metu. Efektyvus teksto įvedimas yra viena iš pasiūlytos sistemos funkcijų.

1.9. Mokslinis pripažinimas

Doktorantūros metu buvo parengtos ir išspausdintos 9 publikacijos. Trys iš jų „ISI Web of Science“ moksliniuose žurnaluose su citavimo indeksu. Viena publikacija kituose moksliniuose žurnaluose. Likusios publikacijos išspausdintos konferencijų medžiagoje.

Tyrimų rezultatai pristatyti šiose tarptautinėse konferencijose:

1. The 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE) 2018, Funšalis, Portugalija;
2. The 20th International Conference on Methods and Models in Automation and Robotics (MMAR) 2015, Międzyzdroje, Lenkija;
3. The 20th International Conference Biomedical Engineering 2016, Kaunas, Lietuva;
4. The 21st International Conference Information and Software Technologies (ICIST) 2015, Druskininkai, Lietuva;
5. The 9th International Conference Electrical and Control Technologies (ETC) 2014, Kaunas, Lietuva;

Taip pat 2015 m. dalyvauta VISion Understanding and Machine intelligence (VISUM) vasaros mokykloje, rengiamoje Universidade do Porto, Portugalijoje. 2016 m. ir 2017 m. dalyvauta Europos robotikos forume Liublijanoje, Slovėnijoje, ir Edinburge, Jungtinėje Karalystėje. Dalyvauta tarptautinėse konferencijose ROSCon 2016 Seule, Pietų Korėjoje, ir ROSCon 2017 Vankuveryje, Kanadoje. Studijų metu gautos dvi Lietuvos mokslo tarybos stipendijos už akademinis pasiekimus (2016 ir 2018) ir dvi Kauno technologijos universiteto aktyviausių doktorantų konkurso stipendijos. Taip pat su komanda buvo laimėta pirma vieta Hacker Games 2016 Smart Energy konkurse.

1.10. Santraukos struktūra

Ši santrauka yra suskirstyta į 6 skyrius. Antras skyrius apžvelgia naujausias neįgaliesiems skirtas technologijas. Šiame skyriuje taip pat apžvelgiamos kitos technologijos, kurios yra naudojamos kuriant modernias neįgaliesiems skirtas technologijas. Trečiame skyriuje pristatoma kurtos daugiafunkcės vartotojo sąsajos architektūra. Sukurti nauji algoritmai aprašyti ketvirtajame skyriuje. Penktame skyriuje aprašomi atlikti eksperimentai. Šeštajame skyriuje pateikiamos išvados bei rekomendacijos.

2. ESAMŲ TECHNOLOGIJŲ NEĮGALIESIEMS ANALIZĖ

Šiame skyriuje apžvelgiamos naujausios AT, aktualios šiai disertacijai. AT yra labai platus terminas, apimantis pagalbinius, adaptyvius ir reabilitacinius įrenginius bei specializuotas HCI vartotojo sąsajas. Modernios AT yra sudėtingos sistemos, jungiančios nesudėtingus ir patikimus pagalbinius įrenginius su išmaniomis UI programomis. Vartotojai gali atlikti sudėtingas užduotis naudodamiesi tokiomis sistemomis. Šiame skyriuje nagrinėjami penki pagrindiniai AT HCI sistemų aspektai, t.y. įvesties įrenginiai, informacijos apdorojimas, informacijos pateikimas, aplinkos suvokimas ir sąveika su aplinka.

AT padeda neįgaliesiems būti savarankiškesniems, taip smarkiai pagerinant jų gyvenimo kokybę [Manns and Chad, 2001]. Neįgalieji naudoja daug skirtingų technologijų, padedančių jiems atlikti įvairias užduotis. Naujojamų AT spektro detalių aprašymą galima rasti [Cook and Polgar, 2014].

2.1. Daugiafunkcės vartotojo sąsajos struktūra

Daugiafunkcė vartotojo sąsajos sistema gali būti suskirstyta į keletą komponentų. Informacijos įvesties įrenginiai apžvelgiami 2.2 skyriuje. Gautos informacijos apdorojimo algoritmai aptariami 2.3 skyriuje. Vartotojui informacija dažniausiai pateikiama grafinėje vartotojo sąsajoje Graphical user interface (GUI) (žr. 2.4 skyrių). Sudėtingesnės sistemos taip pat gali būti naudojamos su aplinka sąveikauti. Tam dažnai reikalingi aplinkos suvokimo metodai, aptariami 2.5 skyriuje. 2.6 skyriuje aprašomi sąveikavimo su aplinka metodai.

2.2. Pagalbiniai įrenginiai

Šiame skyriuje pristatomi įvairūs pagalbiniai įrenginiai, leidžiantys neįgaliesiems siųsti valdymo komandas kompiuteriui.

2.2.1. Jungiklio įrenginiai

Jungiklio įrenginiai - tai nesudėtingi ir plačiausiai naudojami pagalbiniai įrenginiai. Tokie įrenginiai gali būti pritaikomi įvairioms užduotims atlikti. Šių įrenginių panaudojimo galimybės plačiai aprašomos [Angelo, 2000].

2.2.2. Liežuviu valdomi įrenginiai

Šio įrenginio veikimas paremtas ant liežuvio pritvirtinto magneto padėties nustatymu [Krishnamurthy and Ghovanloo, 2006]. Eksperimentiniuose tyrimuose šie įrenginiai buvo panaudoti sėkmingai įvairioms užduotims atlikti. Nepaisant šių pasiekimų rinkoje, vis dar nėra vartotojams skirtų liežuviu valdomų įrenginių.

2.2.3. Smegenų kompiuterio sąsaja

Smegenų kompiuterio sąsajos Brain computer interface (BCI) yra labai plačiai tyrinėjamos. Pagrindinė to priežastis yra tai, kad BCI leidžia valdyti kompiuterį mintimis. Kompiuterio valdymas mintimis yra vienas iš intuityviausių valdymo būdų. Neseniai mokslininkams pavyko pademonstruoti, kad neįgalusis gali valdyti roboto ranką mintimis [Hochberg ir kt., 2012]. Tačiau tokios sistemos reikalauja labai brangių invazinių BCI įrenginių. O vartotojams skirtų BCI galimybės vis dar labai ribotos [Maskeliunas ir kt., 2016].

2.2.4. Akių judesio sekimas

Akių judesio sekimo technologija taip pat yra sparčiai besivystanti. Net ir vartotojams skirti akies judesio sekimo įrenginiai pasiekia gerų rezultatų [Dalmaijer, 2014] ir gali būti naudojami įvairioms užduotims atlikti. Pagrindinė šių įrenginių problema yra vadinamasis Mido prisilietimas, pirmą kartą aprašytas [Jacob, 1991]. Ši problema sako, kad be papildomų įrenginių yra labai sunku atskirti, kada vartotojas akis judina norėdamas atlikti valdymo operaciją, o kada tik norėdamas pamatyti naują informaciją.

2.2.5. Kalbos atpažinimas

Kalbos atpažinimas nereikalauja papildomos įrangos, bet gali būti naudojamas kaip menamas vartotojo sąsajos valdymo įrenginys. Kalbos atpažinimo technologijų vystymosi istoriją galima rasti [Huang ir kt., 2014]. Pagrindinės kalbos atpažinimo naudojimo problemos yra triukšmingos aplinkos. Šia technologija taip pat negali naudotis kalbos sutrikimų turintys asmenys.

2.2.6. Raumenų judesiai

Šie įrenginiai yra panašūs į BCI, tačiau vietoje smegenų signalų yra fiksuojami raumenų susitraukimai. Kadangi raumenų susitraukimų generuojami signalai yra daug stipresni, juos yra lengviau aptikti ir išskirti. Nereikia tam pakanka nedidelių skaičiavimo pajėgumus turinčio mikroprocesoriaus. Norint gauti gerus rezultatus, reikia tinkamai parinkti raumenis, kurių generuojami signalai bus stebimi.

2.2.7. Pagalbinių įrenginių palyginimas

Kiekvienas pagalbinis įrenginys turi privalumų ir trūkumų. Taip pat gali skirtis ir rezultatai, gaunami skirtingų to paties tipo įrenginių. 2.1 lent. yra apibendrinami kiekvieno įrenginio trūkumai bei privalumai.

2.3. Informacijos apdorojimas

Šiame skyriuje aptariami informacijos apdorojimo algoritmai, kurie pagerina sistemos naudojimo efektyvumą. Informacijos apdorojimo algoritmai gali būti naudojami signalams sujungti bei vartotojo veiksmams nuspėti.

2.3.1. Signalų jungimas ir klaidų aptikimas

Signalų jungimas naudojamas, kai sistemoje yra prijungiamas daugiau negu vienas pagalbinis įrenginys. Pakoreguoti signalų sujungimo algoritmai taip pat gali būti naudojami aptinkant vieno iš pagalbinių įrenginių klaidas [Leeb ir kt., 2010].

2.3.2. Teksto nuspėjimas

Teksto nuspėjimo algoritmai pagreitina teksto įvedimo laiką. Tokio algoritmo efektyvumas labai priklauso nuo naudojamo kalbos modelio. Įvairūs teksto nuspėjimo algoritmai detaliam aprašomi [Garay-Vitoria and Abascal, 2006]. Teksto nuspėjimo algoritmai skirstomi į tris pagrindines grupes - tai sintaksiniai, semantiniai ir statistiniai. Statistiniai algoritmai yra lengvai įgyvendinami ir duoda gerus rezultatus. Sintaksinių ir semantinių algoritmų rezultatai gali būti geresni, tačiau jiems yra reikalingi dideli sužymėtų duomenų kiekiai. Dar vienas būdas, leidžiantis pagerinti teksto nuspėjimo rezultatus, yra nuspėjimo algoritmo adaptavimas veikimo metu. Toks algoritmas gali prisitaikyti prie

2.1 lentelė Pagalbinių įrenginių palyginimas

Pagalbinis įrenginys	Privalumai	Trūkumai	Pagrindinis panaudos atvejis
Jungiklis	Nesudėtingas, nebrangus, nereikia signalų apdorojimo	Nedidelis generuojamų signalų skaičius	GUI navigacija, teksto įvedimas
Liežuviu valdomi įrenginiai	Tikslus pozicijos nustatymas	Jutikliai, dedami į burną	Neįgaliųjų vežimėlio valdymas
BCI	Natūralus HCI metodas	Vartotojams skirti įrenginiai netikslūs, įterpiami įrenginiai brangūs	Roboto rankos valdymas
Akies judesių	Tiksliai nustato 2D poziciją ekrane	Mido prisilietimo problema	UI navigacija
Kalbos	Gali atpažinti daug komandų	Neveikia triukšmingose aplinkose	Naudoti atliekant daug komandų
Raumenų judesių	Gali būti naudojamas kaip jungiklis ir pozicijos nustatymo įrenginys	Reikalingi elektrodai ant stebimų raumenų	Įrenginių valdymui bei UI navigacijai

konkreto vartotojo kalbėjimo stiliaus.

2.4. Informacijos pristatymas

Įvairūs informacijos pristatymo būdai yra naudojami pateikiant UI sistemos naudotojui. Dažniausiai naudojamas informacijos pristatymo būdas yra kompiuterio ekranas. Toks būdas yra tinkamas atliekant individualias nesudėtingas užduotis. Naujos technologijos, tokios kaip virtualios realybės akiniai arba ant galvos montuojami ekranai Head-mounted display (HMD), leidžia kurti novatoriškas vartotojo sąsajas.

2.4.1. Virtuali realybė

Virtuali realybė - tai kompiuteryje generuojama aplinka, panaši į tikrą pasaulį. Tokios aplinkos dažnai naudojamos reabilitacijoje arba mokant vartotoją naudotis tam tikrais įrenginiais [Perdikis ir kt., 2018]. Kita vertus, virtuali realybė nėra tinkama atliekant kasdienes užduotis, nes neįgalusis yra dar labiau atskiriamas nuo jį supančios aplinkos.

2.4.2. Papildyta realybė

Papildyta realybė sulieja žmogaus regimus vaizdus su kompiuterio generuojamais vaizdais. Dažniausiai tam yra naudojami HMD arba mobilieji telefonai. Detali šios technologijos naujovių apžvalga pateikiama [Billinghurst ir kt., 2015]. Papildytos realybės kūrimas yra sudėtingesnis negu virtualios realybės, nes tokia sistema turi ne tik generuoti kompiuterio vaizdus, bet ir suprasti vartotoją supančią aplinką. Dėl šios priežasties papildytos realybės įranga vis dar yra labai brangi.

2.4.3. Projektuojami vaizdai

Projektuojami vaizdai yra tam tikra papildytos realybės forma, naudojami projektorius informacijai atvaizduoti aplinkoje. Pagrindinė tokios sistemos nauda yra ta, kad projektuojamą vaizdą mato visi aplinką stebintys žmonės. Pagrindiniai šio metodo trūkumai yra projektoriaus generuojamo vaizdo šviesumas ir iškraipymas žiūrint iš skirtingų taškų.

2.4.4. Informacijos pristatymo metodų palyginimas

Informacijos pristatymo metodai parenkami atsižvelgiant į vykdomos užduoties tipą. Kitas svarbus faktorius yra naudojamų įrenginių kaina. Pagrindiniai informacijos pristatymo metodų privalumai ir trūkumai pristato-

2.2 lentelė Informacijos pristatymo metodų palyginimas

Informacijos pristatymo metodas	Privalumai	Trūkumai	Pagrindinis panaudos atvejis
Monitorius	Nebrangus, didelės rezoliucijos	GUI atskirta nuo aplinkos	Naudojama daugelyje sistemų dėl didelio prieinamumo
Virtuali realybė (VR)	Įtraukiantis potyris, vartotojas patiria iki galo kompiuterio sugeneruotą aplinką	Vartotojas atskiriamas nuo realios aplinkos, reikalinga brangi įranga	Reabilitacija ir mokymasis
AR	GUI atvaizduojama realioje aplinkoje	Brangi įranga, reikalingi 3D aplinkos jutikliai	Sąveikaujant su aplinka
Projektuojami vaizdai	Panašus į AR pojūtis, bet nebrangi įranga	Prastas matomumas šviesioje aplinkoje, vaizdas priklauso nuo žiūrėjimo kampo	Sąveikaujant su aplinka

mi 2.2 lent.

2.5. Aplinkos suvokimas

Išmaniosioms AT reikalingi jutikliai, gebantys suvokti aplinką. Labiausiai tam tinkami yra 3D jutikliai. 3D jutikliai skirstomi į tris pagrindines kategorijas: struktūrizuotos šviesos, šviesos aptikimo ir atstumo nustatymo (*Light Detection and Ranging (LIDAR)*), keliavimo laiko (Time of flight (TOF)) ir stereoskopinių vaizdų. Kiekvienas šių metodų turi privalumų ir trūkumų. Struktūrizuotos šviesos jutikliai turi nedidelį veikimo atstumą bei prastai veikia lauke. TOF kameros dažnai būna labai mažos rezoliucijos, o LIDAR jutikliai vis dar yra labai brangūs.

Stereoskopinių vaizdų apdorojimas yra vienintelė pasyvi iš aukščiau išvardintų technologijų. Tai reiškia, kad šis jutiklis nespinduliuoja į aplinką šviesos spindulių. Toks jutiklis naudoja mažiau energijos ir gali lengvai prisitaikyti prie kintančio aplinkos šviesumo. Tolimesniuose skyriuose aptariami kameros parametrų valdymo algoritmai 2.5.1 ir stereoskopinių vaizdų apdorojimo algoritmai 2.5.2.

2.5.1. Kameros valdymas

Kameros valdymas - tai procesas, kurio metu yra parenkami kameros parametrai, leidžiantys gauti gerai eksponuotus vaizdus. Šie algoritmai leidžia kamerali prisitaikyti prie kintančių apšvietimo sąlygų. Kameros valdymo algoritmai yra sudėtingi, nes dažniausiai bent du kameros parametrai yra valdomi vienu metu. Daugelis komercinių kamerų turi kameroje veikiančius valdymo algoritmus, tačiau tokie algoritmai nėra tinkami, kai kelios kameros turi būti valdomos vienu metu (pvz., stereokamera). Taip pat ne visos kameros turi galimybę nustatyti regioną vaizde, naudojamą parametrų parinkti. Kameros vaizdo ekspozicijai nustatyti gali būti naudojami skirtingi algoritmai. Keletas tokių algoritmų yra aptariami [Nourani-Vatani and Roberts, 2007].

2.5.2. Stereoskopinių vaizdų suliejimas

Stereoskopinių vaizdų suliejimas yra daug skaičiavimo resursų reikalaujantis procesas. Dėl šios priežasties stereovaizdams apdoroti dažnai naudojami galingi kompiuteriai. Tačiau tokie kompiuteriai yra dideli ir naudoja daug elektros energijos. Norint sukurti stereokamerą gebančią atlikti skaičiavimus kameros viduje, reikalingi efektyvūs stereovaizdų apdorojimo algoritmai. Pastaraisiais metais atsiradusios įterptinės platformos, tokios kaip NVIDIA's Jetson, turi pakankamai skaičiavimo resursų stereovaizdams apdoroti.

Stereovaizdų algoritmai skirstomi į dvi grupes - lokalius ir globalius. Globalūs algoritmai leidžia gauti labai tikslius rezultatus, tačiau jie yra netinkami realaus laiko sistemoms. Lokalūs algoritmai yra mažiau tikslūs, bet gali atlikti skaičiavimus realiu laiku. Stereovaizdų algoritmams įvertinti yra naudojamos paruoštos duomenų bazės, turinčios tikslius apdorotų vaizdų pavyzdžius ir jų palyginimo metodus. Middlebury version 3 yra viena populiariausių stereovaizdų duomenų bazių [Scharstein ir kt., 2014].

Pusiau globalus stereovaizdų apdorojimo algoritmas (aprašomas [Hirschmuller, 2008]) yra vienas iš populiariausių ir tiksliausių realaus laiko algoritmų. Šioje disertacijoje pristatomas algoritmas yra panašus į [Geiger ir kt., 2010] algoritmą. Pagrindinis šio algoritmo pranašumas yra skaičiuojamas pirminis netikslus sutapdintas vaizdas, kuris vėliau yra tikslinamas.

Šiame darbe pristatomas metodas naudoja maksimaliai stabilių ekstremalių regionų (*Maximally stable extremal region (MSER)*) algoritmą skaičiuojant pradinį sutapdintą stereovaizdą. MSER buvo pasirinkti dėl greito jų skaičiavimo laiko ir stabilumo savybių.

2.5.3. Objektų aptikimas 3D vaizduose

Objektų aptikimo metu yra nustatomas objekto tipas ir jo 3D pozicija erdvėje. Objekto tipas gali būti lyginamas su kompiuteriu sukurtu modeliu arba iš nufotografuotų 3D objekto vaizdų. Šis procesas yra vykdomas apskaičiuojant vaizdo charakteringus taškus ir juos sutapdinant su objekto modelio charakteringais taškais. Neretai po šio veiksmo dar yra reikalingas papildomas 3D objekto pozicijos nustatymo algoritmas.

2.6. Sąveika su aplinka

Didžioji dauguma iki šiol aptartų AT yra naudojami bendrauti su kitu žmogumi arba kompiuteriu. AT taip pat gali būti naudojama sąveikaujant su aplinka. Pasyvios sąveikos sistemos leidžia vartotojui kontroliuoti jo aplinkoje esančius daiktus, tokius kaip lempos ar televizoriai. Aktyvios sistemos leidžia vartotojui pačiam paveikti aplinką. Dažniausiai sutinkamos tokios sistemos yra neįgaliųjų vežimėliai ir roboto rankos. Tobulėjant smegenų vartotojo sąsajoms, tokios sistemos tampa vis lengviau valdomos. Ateityje tokios technologijos gali leisti neįgaliesiems būti visiškai savarankiškiems.

2.7. Analizės apibendrinimas

Šiame skyriuje yra apžvelgiamos naujausios AT. Apžvelgtos temos buvo suskirstytos į penkias kategorijas: pagalbinių įrenginių, informacijos pristatymo, informacijos apdorojimo, aplinkos suvokimo ir sąveikos su aplinka. Kiekviena iš šių sričių atspindi skirtingą daugiafunkcės vartotojo sąsajos sistemos sluoksnį.

3. DAUGIAFUNKCĖS VARTOTOJO SĄSAJOS ARCHITEKTŪRA

Šiame skyriuje aptariama kuriamos daugiafunkcės vartotojo sąsajos architektūra. Sistemos kūrimo metu buvo atsižvelgiama į du svarbius aspektus:

1. **Galimybė perkonfigūruoti** sistemą užtikrina, kad sistemos elgesys bus lengvai keičiant sistemos parametrus. Taip sistema lengvai pritaikoma prie vartotojo poreikių.
2. **Galimybė išplėsti** leidžia nesudėtingai pridėti naujus įrenginius ir funkcionalumą. Pasiūlytos sistemos išplėtimas nereikalauja didelių pastangų ir gali būti atliekamas be programuotojo pagalbos.

3.1. Daugiafunkcės vartotojo sąsajos reikalavimai

Sistema buvo kurta pagal reikalavimus, surinktus Quadribot projekto metu. Šiame skyriuje yra aprašyti sistemos kūrimui iškelti reikalavimai. Kaip jau buvo paminėta, didelis dėmesys buvo skiriamas sistemos išplėtimo ir adaptyvumo reikalavimams. Sistemos reikalavimai taip pat darė įtaką sukurtos sistemos architektūrai.

3.2. Daugiafunkcės vartotojo sąsajos panaudos atvejai

Sistemoje išskiriami tokie vartotojai:

1. **Įprastas vartotojas** yra neįgalus žmogus, naudojantis sistemą kaip pagalbines priemones.
2. **Slaugytojas** yra žmogus, bendraujantis ir padedantis įprastam sistemos vartotojui.
3. **Sistemos administratorius** yra atsakingas už sistemos parengimą darbui.

Kiekvienas panaudos atvejis yra detalizuotas žemiau:

1. **Valdyti sąsają naudojant pagalbinius įrenginius.** Šis panaudos atvejis apibrėžia, kad sukurta sistema yra valdoma prie sistemos prijungtais pagalbiniais įrenginiais.

2. **Naudoti keletą įrenginių vienu metu** yra prieš tai buvusio panaudos atvejo plėtinys. Šiuo atveju keletas pagalbinių įrenginių yra naudojami vienu metu.
3. **Įvesti tekstą** panaudos atvejis leidžia vartotojui įvesti tekstą ir taip bendrauti su aplinkiniais.
4. **Pasirinkti iš sąrašo** leidžia vartotojui pasirinkti vieną iš norimų dalykų sąrašė. Sąrašai yra kuriami sistemos derinimo metu arba generuojami dinamiškai iš aplinkoje matomų objektų. Vienas sąrašas taip pat gali būti susietas su keletu kitų sąrašų hierarchiniais ryšiais.
5. **Valdyti aplinką** panaudos atvejis yra **Pasirinkti iš sąrašo** plėtinys. Čia kiekvienas sąrašo elementas yra susiejamas su aplinką valdančiu įrenginiu.
6. **Pasirinkti taip arba ne** Leidžia vartotojui atsakinėti į taip arba ne atsakymo reikalaujančius klausimus.
7. **Pasirinkti objektą** leidžia vartotojui pasirinkti vieną iš automatiškai aptiktų objektų. Tokiems objektams aptikti yra naudojami aplinkos jutikliai.
8. **Projektuoti aptiktų objektų vaizdą** yra **Pasirinkti objektą** plėtinys. Aptikti aplinkos objektai yra paryškunami naudojant projektuojamų vaizdų metodą.
9. **Peržiūrėti sistemos istoriją** leidžia sistemos vartotojui peržiūrėti anksčiau sistemoje atliktus veiksmus. Šis funkcionalumas taip pat naudojamas atliekant eksperimentus.
10. **Konfigūruoti sistemą** leidžia pakeisti nustatymus ir taip pritaikyti sistemą kiekvienam vartotojui individualiai.
11. **Sukurti naują sąrašą** leidžia lengvai išplėsti sistemos veikimą. Šią operaciją gali atlikti prižiūrintis personalas, ir tam nereikia kreiptis į sistemos kūrėją.

3.3. Sistemos diegimo aprašas

Šiame skyriuje yra detaliam aprašomas sistemos diegimas. Sistema yra diegiama viename kompiuteryje. Visi naudojami pagalbiniai įrenginiai yra prijungiami prie šio kompiuterio. Disertacijoje yra detalizuojami visi diegiant sistemą naudojami artefaktai.

3.4. Daugiafunkcės vartotojo sąsajos architektūros aprašas

Šiame skyriuje aprašoma daugiafunkcės vartotojo sąsajos architektūra. Sistema yra suskirstyta į keletą atskirų komponentų. Grafinės vartotojo sąsajos komponentai projektuojami remiantis Modelio-Vaizdo programavimo paradigma [Vlissides ir kt., 1995]. Disertacijoje plačiau yra aprašomas kiekvienas komponentas. Komponentai tarpusavyje bendrauja tiksliai aprašytomis sąsajomis. Disertacijoje yra detalizuojama kiekvieno komponento sąsaja ir jame naudojamos duomenų struktūros.

3.5. Detali komponentų įgyvendinimo informacija

Kiekvienas sistemos komponentas yra įgyvendinamas viena arba keliomis klasėmis. Sukurtos klasės griežtai atitinka komponentu apibrėžtus interfeisus. Kai kurie komponentai turi daugiau negu vieną įgyvendinimo variantą. Informacija apie kiekvienos klasės įgyvendinimo detales yra pateikiama disertacijoje.

3.6. Komponentų informacijos tėkmė

Šiame skyriuje yra aprašoma, kaip informacija keliauja sistemos komponentais. Valdymo informacija yra kuriama pagalbiniuose įrenginiuose. Ši informacija yra apdorojama sistemoje ir iš jos yra kuriamos vartotojo sąsajos valdymo komandos. Vartotojo sąsajai taip pat yra pateikiama teksto nuspėjimo informacija. Sugeneruotas sąsajos vaizdas yra perduodamas informacijos pateikimo įrenginiui ir parodomas vartotojui.

3.7. Architektūros apibendrinimas

Šiame skyriuje buvo pristatoma sukurta daugiafunkcės vartotojo sąsajos architektūra. Skyriuje detaliam aprašomi sistemos panaudos atvejai, sistemos diegimas ir jos komponentai. Taip pat yra pateikiama detali sukurtų sistemos komponentų informacija.

4. PASIŪLYTI DAUGIAFUNKCĖS VARTOTOJO SĄSAJOS ALGORITMAI

Šis skyrius aprašo disertacijoje pristatomus naujus algoritmus. Sukurti algoritmai yra naudojami daugiafunkcės vartotojo sąsajos veikimui pagerinti. Pasiūlytų algoritmų apžvalga ir jų naudojimas yra iliustruojamas 4.1 pav.

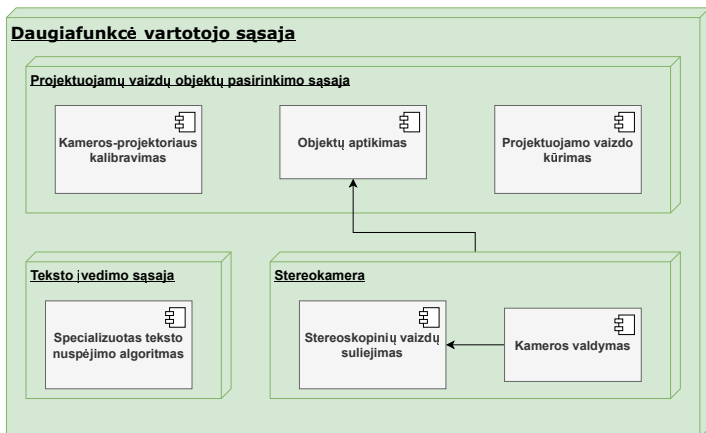
Viena iš sukurtos sistemos dalių yra projektuojamais vaizdais paremta vartotojo sąsaja, skirta objektams pasirinkti. Ši sistema leidžia vartotojui matyti aptiktus daiktus realioje aplinkoje. Projektuojamų vaizdų sąsaja susideda iš trijų dalių: kameros-projektoriaus kalibravimo, objektų aptikimo ir aptiktų objektų paryškimo. Šie algoritmai detaliai aprašyti 4.1 skyriuje. Pagrindinis šio metodo naujumas yra objektų aptikimo algoritmo integravimas. Tai leidžia objektus aplinkoje paryškinti automatiškai, netgi kintant jų padėčiai.

Objektų aptikimo algoritmas veikia su RGB-D tipo duomenimis. Tokie duomenys dažnai gaunami naudojant struktūrizuotos šviesos jutiklius. Tačiau tokie jutikliai prastai veikia labai šviesiose aplinkose. Stereokameros taip gali būti naudojamos RGB-D vaizdams kurti. Pagrindinis stereokamerų privalumas yra tai, kad jos yra daug mažiau jautrios apšvietimo pokyčiams. Pagrindinis stereokamerų trūkumas yra stereovaizdų suliejimo algoritmams reikalingas didelis skaičiavimų skaičius. Sukurtas realaus laiko stereovaizdų suliejimo algoritmas yra pristatomas 4.3 skyriuje. Pagrindinis šio algoritmo naujumas yra vaizdo regionų sutapdinimas tam, kad būtų sumažinta algoritmo paieškos erdvė.

Stereokameros patikimumui užtikrinti buvo pasiūlytas greitai veikiantis kameros parametrų valdymo algoritmas. Šis algoritmas geba prisitaikyti prie staiga pakitusių apšvietimo sąlygų. Kameros valdymo algoritmo aprašymas pateikiamas 4.2. Paskutiniame šios disertacijos dalies skyriuje yra pristatomas teksto nuspėjimo algoritmas.

4.1. Projektuojamų vaizdų vartotojo sąsaja

Šioje disertacijoje yra pristatoma novatoriška vartotojo sąsaja, skirta objektams pasirinkti. Sistema geba aptikti ant lygaus paviršiaus padėtus objektus. Tam yra naudojama RGB-D kamera. Aptikti objektai gali būti

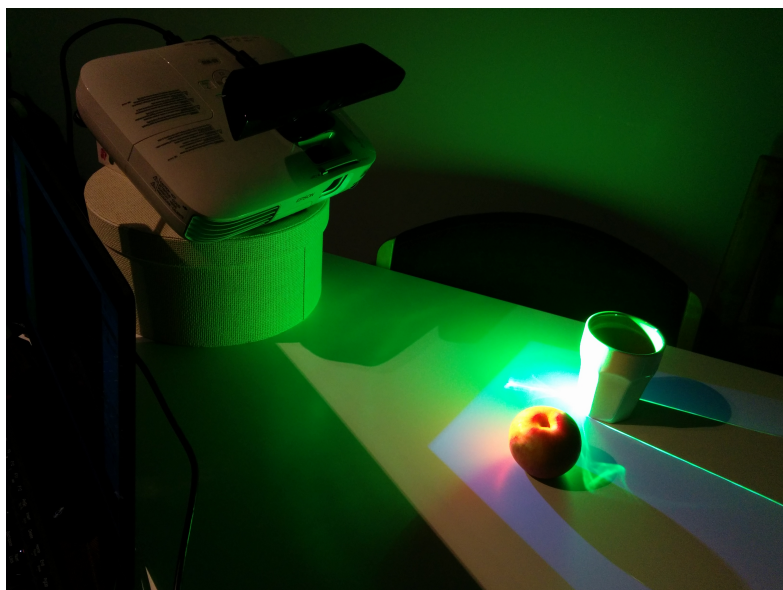


4.1 pav. Pasiūlytų algoritmų apžvalga ir jų naudojimas daugiafunkcėje vartotojo sąsajoje

pažymimi scenoje naudojant projektorius. Objektų pažymėjimo scenoje pavyzdys pateiktas 4.2 pav.

Šiai vartotojo sąsajai yra naudojamas automatinis projektuojamų vaizdų metodas. Šis metodas susideda iš trijų etapų. Pirmiausia yra kalibruojama kameros-projektoriaus sistema. Antrame etape yra automatiškai aptinkami scenos objektai naudojant RGB-D kamerą. Trečiajame etape objektų 3D modeliai yra naudojami projektuojamam vaizdai sukurti. Pirmas etapas atliekamas sistemos paleidimo metu, o antras ir trečias etapas vykdomi nuolat sistemos veikimo metu. Projektuojamų vaizdų metodas plačiai naudojamas vaizduojant 3D videovaizdus ant pastatų fasadų. Tam dažniausiai yra naudojamas rankinis projektoriaus kalibravimo metodas. Pristatyta sąsaja jungia kameros-projektoriaus kalibravimo procedūrą, aprašytą [Yang ir kt., 2016], su realiu laiku veikiančiu objektų atpažinimo algoritmu, įgyvendintu naudojant Point Cloud Library. Taip yra gaunama novatoriška sąsaja, gebanti automatiškai projektuoti aptiktų objektų vaizdus realioje aplinkoje.

Tokia vartotojo sąsaja naudojama, kai vartotojas, prižiūrinčiam žmogui nori parodyti objektą, kurio šiuo metu jis nori. Tokia sistema taip pat gali



4.2 pav. Pavyzdinė scena, kurioje matomi pažymėti objektai bei kameros ir projektoriaus sistema

būti naudojama valdant roboto ranką. Šiuo atveju vartotojas roboto rankos valdikliui parodo, kokio objekto jis nori, ir roboto ranka šį objektą paima. Pagrindinė tokios sąsajos nauda yra tai, kad vartotojas gali matyti aptinkamus objektus realioje aplinkoje ir jam nereikia keisti žvilgsnio nuo kompiuterio ekrano į sceną.

Kalibravimo metu projektoriui ir kamerai aprašyti yra naudojamas *pinhole* kameros modelis. Nors realybėje projektorius nėra kamera, jam aprašyti gali būti naudojamas tas pats kameros modelis. Tiksliau - kalibravimo metu projektorius yra laikomas atvirkštine kamera. *Pinhole* kameros modelis aprašo matematinį ryšį tarp 3D taško koordinatės ir jos projekcijos idealiame *pinhole* kameros jutiklyje. Tam yra naudojama perspektyvinė transformacija:

$$c = K [R|T] W, \quad (4.1)$$

čia c yra taškas, suprojektuotas kameros jutiklyje, W aprašo 3D taško koordinatės pasaulio koordinačių sistemoje, o K yra kameros matrica. R ir T aprašo kameros judėjimą (posūkį ir transformaciją) apie statinę aplinką. Visas kameros modelio aprašymas yra pateikiamas disertacijoje.

4.1.1. Objektų aptikimo algoritmas

Objektų aptikimo algoritmas buvo įgyvendintas sujungus kelis 3D taškų debesies apdorojimo algoritmus, įgyvendintus Point Cloud Library [Rusu and Cousins, 2011] bibliotekoje. Įgyvendintas algoritmas susideda iš tokių žingsnių:

1. Duomenų surinkimo ir paruošimo etapas. Šiame etape iš kameros yra gaunamas RGB-D vaizdas. Šis vaizdas turi būti transformuojamas į 3D taškų debesį. Transformacijos proceso metu yra naudojami kalibruojant kamerą gauti kameros modelio parametrai.
2. Fono išskyrimo etapas. Šio etapo metu 3D taškų debesis yra suskirstomas į du taškų debesis. Pirmo debesies turinys yra 3D taškai, priklausantys scenos fonui, antro - likę taškai, priklausantys scenoje esamiems daiktams. Algoritmo metu yra randama dominuojanti plokštuma ir visi taškai, esantys nedideliame atstume nuo šios plokštumos, yra priskiriami fonui.

3. Scenos objektams priklausančių taškų grupavimas. Šio etapo metu taškai yra sugrupuojami į individualius objektus. Taško atstumas nuo visų jo kaimynystėje esančių taškų nulemia kuriai grupei (objektui) šis taškas bus priskirtas.
4. Grupių klasifikavimo etapas. Šiame etape kiekvienai taškų grupei yra priskiriama klasė. Grupės klasė nusako aptikto objekto tipą. Ši informacija yra naudojama tinkamam 3D modeliui pasirinkti, kuriant projektuojamą vaizdą.
5. Likusioms taškų grupėms yra paskiriama viena iš bazinių formų. Palaikomos bazinės formos yra: sfera, cilindras ir kubas. Bazinės formos parametrams nustatyti yra naudojamas RANSAC metodas.

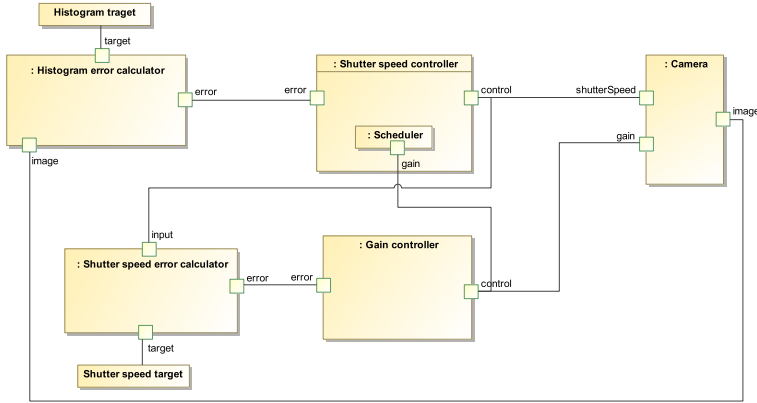
4.1.2. Objekto paryškinimas naudojant projektorius

Nustacius objekto tipą, galima sukurti jo atvaizdą, kurį parodžius projektoriumi yra paryškinamas konkretus objektas. Šiam procesui yra reikalingas objekto 3D modelis. Pirma, šis objekto modelis yra patalpinaamas virtualioje 3D erdvėje. Šis procesas atliekamas kiekvienam aptiktam objektui. Objekto 3D koordinatės yra nustatomos objekto aptikimo metu. Tuomet sukurta virtuali 3D aplinka yra projektuojama į menamą kamerą. Šios kameros parametrai yra parenkami pagal kalibravimo metu gautus projektoriaus parametrus. Menamoje kameroje gaunamas vaizdas gali būti rodomas projektoriumi, ir aplinkoje bus apšviečiami visi aptikti objektai.

4.2. Kameros valdymas

Kameros valdymo algoritmo tikslas yra keisti kameros jutiklio jautrumo ir išlaikymo parametrus taip, kad būtų gaunamas gerai eksponuotas vaizdas. Gerai eksponuoti vaizdai pagerina vaizdo apdorojimo algoritmo veikimą. Kameros vaizdo ekspozicija yra priklausoma nuo keleto vidinių ir išorinių veiksnių. Šie veiksniai yra: jutiklio jautrumas, išlaikymas, diafragma ir scenos šviesos lygis. Visi šie veiksniai nulemia į kamerą patenkančią šviesos kiekį.

Kai į kamerą patenkančios šviesos kiekis yra didelis, kamera tampa jautresnė mažiems parametrų pokyčiams. Kita vertus, jeigu patenkančios šviesos kiekis yra mažas, dideli parametrų pokyčiai mažai paveikia vaizdo ekspoziciją. Šis ryšys buvo nustatytas eksperimentų būdu. Disertacijoje



4.3 pav. Kameros parametrų valdymo algoritmo schema

yra detaliai aprašomi šie eksperimentai. Į šį pastebėjimą buvo atsižvelgta kuriant kameros valdymo algoritmą.

Algoritme buvo naudojamas histogramos masės centro nustatymo algoritmas apskaičiuojant dabartinę vaizdo ekspozicijos vertę. Histogramos masės centras yra apskaičiuojamas pagal tokią formulę:

$$H_{ct} = \frac{\sum_{i=0}^{255} i \times H(i)}{\sum_{i=0}^{255} H(i)}, \quad (4.2)$$

čia H_{ct} yra apskaičiuotas histogramos masės centras, o funkcija $H(i)$ gražina i -jį histogramos stulpelį. Algoritme yra laikoma, kad H_{ct} reikšmė, artima 128, yra laikoma kaip gerai eksponuoto vaizdo indikatorius. Eksperimentų metu buvo nustatyta, kad kameros valdymo procesas yra tiesinis beveik visame parametrų diapazone (kai vienas iš valdomų parametrų nekinta). Dėl šios priežasties kamrai valdyti buvo naudojamas Proportional Integral Derivative (PID) reguliatorius. Kadangi du kameros parametrai buvo valdomi vienu metu, buvo pritaikytas kaskadomis paremtas PID reguliatorius. Algoritmo schema pateikiama 4.3 pav.

Pirmasis valdiklio lygis valdo kameros užlaikymą. Šio valdymo ciklo klaidos reikšmė $e_{ss}(t)$ yra apskaičiuojama taip:

$$e_{ss}(t) = H_{ct} - H_{target}, \quad (4.3)$$

čia H_{ct} yra esamam vaizdui apskaičiuotas histogramos centras, o H_{target} yra histogramos centro tikslo reikšmė. Dažniausiai ši reikšmė yra nustatoma į $H_{target} = 128$. Tuomet kameros valdiklio valdymo funkcija gali būti išreiškiama taip:

$$u_{ss}(t) = K_p^{ss} e_{ss}(t) + K_i^{ss} \int_0^t e_{ss}(t') dt' + K_d^{ss} \frac{de_{ss}(t)}{dt}, \quad (4.4)$$

čia $u_{ss}(t)$ yra nauja kameros užlaikymo reikšmė, nustatoma kameroje. K_p^{ss} , K_i^{ss} ir K_d^{ss} yra proporcinio, integralinio ir diferencialinio valdiklio koeficientai, naudojami kameros užlaikymo valdiklyje.

Antrame valdiklio lygyje yra valdomas kameros jautrumas. Šis valdiklio ciklas yra skirtas prisitaikyti prie lėtai kintančių apšvietimo sąlygų. Valdiklis yra valdomas atsižvelgiant į dabartinę kameros užlaikymo reikšmę. Matematiškai jautrumo valdiklio klaida yra apskaičiuojama:

$$e_g(t) = u_{ss}(t) - T_{ss}, \quad (4.5)$$

čia T_{ss} yra norima gauti kameros užlaikymo reikšmė. Tiksli T_{ss} reikšmė yra pasirenkama eksperimentiniu būdu, taip, kad būtų kiek galima labiau sumažinamas vaizdo suliejimas dėl judančių objektų. Tuomet visas valdiklio aprašymas yra:

$$u_g(t) = K_p^g e_g(t) + K_i^g \int_0^t e_g(t') dt' + K_d^g \frac{de_g(t)}{dt}, \quad (4.6)$$

čia $u_g(t)$ yra kameros jautrumo reikšmė, nustatoma kameroje. O valdiklio proporcinis, integralinis ir diferencialinis koeficientai yra K_p^g , K_i^g ir K_d^g .

Kaip jau buvo minėta, kameros procesas yra tiesinis tik tuomet, kai vienas iš valdomų parametrų yra fiksuotas. Norint efektyviai valdyti abu kameros parametrus vienu metu, yra naudojama kameros užlaikymo valdymo ciklo stiprinimo koeficientų adaptavimo strategija. Ši strategija leidžia kompensuoti valdomo proceso netiesiškumus. Formaliai kameros užlaikymo ciklo proporcinis koeficientas K_p^{ss} yra sumažinamas, kai kameros jautrumas didėja:

$$K_p^{ss} = K_{gs} \frac{1}{u_g(t)}, \quad (4.7)$$

šioje formulėje K_{gs} yra konstanta, kuri gali būti paskaičiuojama eksperimentiniu būdu. Šis procesas yra detalai aprašytas disertacijoje.

4.3. Stereoskopinių vaizdų apdorojimas maksimaliai stabilių ekstremalių regionų metodu

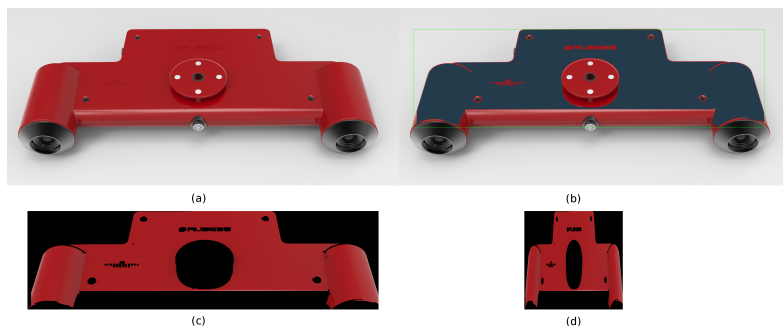
Šiame skyriuje pristatomas stereovaizdų apdorojimo algoritmas *Cyclops*, paremtas MSER metodu. Stereovaizdų suliejimo algoritmui yra perduodami du rektifikuoti vaizdai iš kameros. Algoritmo rezultatas yra skirtuminis vaizdas, kuris yra naudojamas nustatant atstumus iki scenoje esančių objektų. Algoritmas yra suskirstytas į šiuos etapus:

1. MSER apskaičiavimas kiekvienam iš stereovaizdų.
2. MSER normalizavimas. Šis žingsnis reikalingas tam, kad kiekvieno vaizdo regionai galėtų būti lengvai sutapdinami. Normalizavimo procesas yra iliustruojamas 4.4 pav.
3. Sutapdinti normalizuotus kairiojo ir dešiniojo vaizdo MSER regionus. Sutapdinti regionai yra naudojami pirminiam grubiam skirtuminiam vaizdai apskaičiuoti.
4. Skirtuminio vaizdo apskaičiavimas. Šiame žingsnyje yra naudojamas pusiau globalaus sutapdinimo algoritmas, aptartas apžvalgos skyriuje.

Pagrindinė pristatomo algoritmo nauda yra tai, kad pirminis skirtuminis vaizdas yra apskaičiuojamas labai greitai. Taip atsiranda galimybė naudoti tikslesnį algoritmą paskutiniame algoritmo etape. Šis algoritmas yra naudojamas vaizdams apdoroti realiuoju laiku.

4.4. Specializuotas teksto nuspėjimo algoritmas

Praktikoje dažnai naudojami bendriniai teksto nuspėjimo algoritmai. Tokie algoritmai apmokomi dideliu duomenų kiekiu ir duoda gerus rezultatus daugeliu atvejų. Tačiau tokie algoritmai ne visada yra tinkami neįgaliesiems. Visų pirma yra sunku surinkti didelį kiekį duomenų, kurio reikia



4.4 pav. MSER regionų normalizavimo procedūra. (a) pavyzdinis vaizdas I , (b) aptiktas MSER regionas (mėlyna) ir jo ribos b (žalia), (c) laikinas MSER vaizdas I_l , (d) normalizuotas MSER vaizdas I_f

nuspėjimo algoritmui išmokyti. Specializuotas teksto nuspėjimo algoritmas, pristatomas šioje disertacijoje, naudoja keletą nedidelių duomenų rinkinių algoritmui apmokymui. Kiekvienas duomenų rinkinys yra parenkamas taip, kad jame būtų tik duomenys, skirti vienai pokalbio temai. Tuomet skirtingi specializuoti teksto nuspėjimo algoritmai yra keičiami priklausomai nuo dabartinės pokalbio temos.

4.5. Pasiūlytų algoritmų apibendrinimas

Šiame skyriuje yra detalai aprašomi keturi disertacijoje pristatyti algoritmai. Pirmieji trys algoritmai yra panaudoti kuriant ir optimizuojant novatorišką projektuojamų vaizdų vartotojo sąsają. Ketvirtasis algoritmas yra naudojamas teksto įvedimui pagerinti. Pagreitinimas yra pasiekiamas išmokant keletą konkrečiai pokalbio temai pritaikytų teksto nuspėjimo kalbos modelių.

5. EKSPERIMENTŲ APIBŪDINIMAS IR REZULTATAI

Pasiūlyti algoritmai buvo įvertinti atliekant šiame skyriuje aprašytus eksperimentus. Taip pat buvo atliekami eksperimentai, skirti sukurtos daugiafunkcės vartotojo sąsajos efektyvumui įvertinti.

5.1. Eksperimentų metodologija

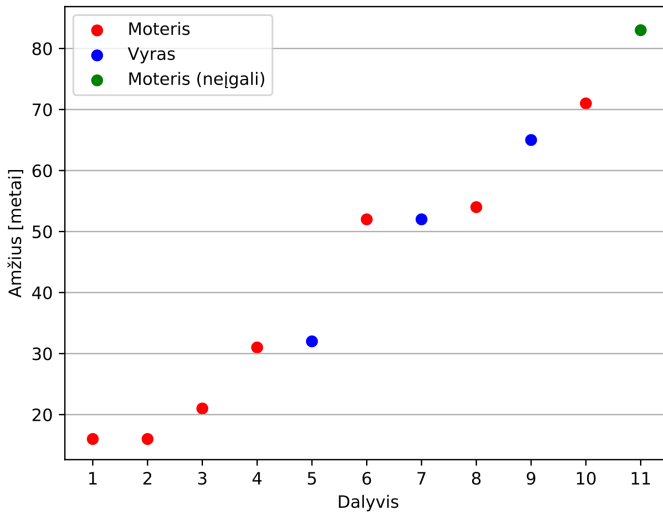
Sukurtai sistemai įvertinti yra naudojami trijų tipų eksperimentai: panaudojimo efektyvumo, vartotojų išpūdžių ir algoritmų įvertinimo. Panaudojimo efektyvumo ir vartotojų išpūdžių eksperimentai yra atliekami su tyrimo dalyviais. Algoritmai yra vertinami naudojant duomenų rinkinius ir įrenginius.

Efektyvumo eksperimentai yra naudojami vertinant daugiafunkcę vartotojo sąsają. Projektuojamų vaizdų sąsaja yra vertinama naudojant vartotojų išpūdžių klausimyną. Kamos valdymo algoritmas yra vertinamas naudojant Universal Serial Bus (USB) kamerą. Stereoskopinių vaizdų apdorojimo algoritmas naudoja Middlebury duomenų rinkinį, o teksto nuspėjimo algoritmo eksperimentai buvo atlikti naudojant rankiniu būdu sukurtą duomenų rinkinį.

Tyrimuose dalyvavo 11 žmonių. Iš jų 8 moterys ir 3 vyrai. Dalyvių amžius buvo nuo 16 iki 83 metų. Kiekvieno dalyvio amžius ir lytis yra pavaizduota 5.1 pav. Dešimt iš tyrimo dalyvavusių vienuolikos žmonių buvo sveiki. Vienas tyrimo dalyvis turėjo nustatytą didelį specialiųjų poreikių lygį.

5.2. Daugiafunkcės vartotojo sąsajos eksperimentai

Šių eksperimentų metu buvo tikrinamas sukurtos daugiafunkcės vartotojo sąsajos efektyvumas naudojant skirtingus įrenginius. Eksperimentų metu buvo naudojama pasirinkimo iš sąrašo užduotis. Eksperimento metu buvo nustatyta, kad skirtingiems žmonėms efektyvesni skirtingi įrenginiai. Kai kurie dalyviai iš viso negalėjo naudoti tam tikrų įrenginių, bet visi dalyviai galėjo efektyviai naudoti bent vieną įrenginį. Atlikto eksperimento įvedimo laiko diagrama pateikiama 5.2 pav. Detalesni rezultatai, parodantys kiekvieno dalyvio rezultatus, yra pateikiami disertacijoje.



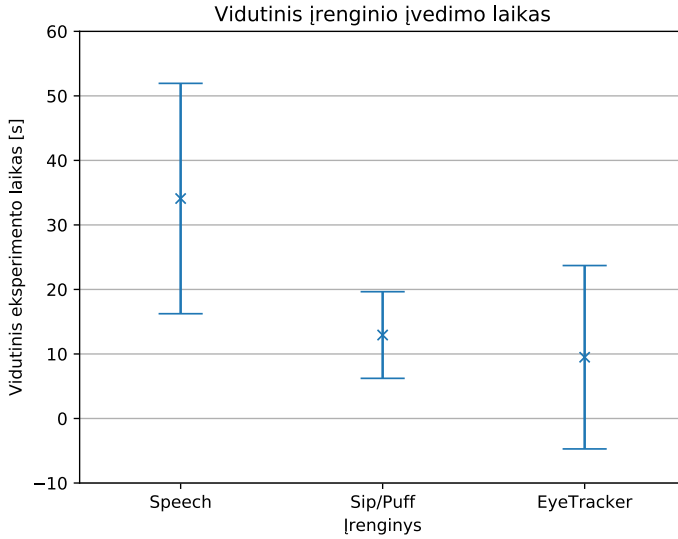
5.1 pav. Eksperimentų dalyvių amžius ir lytis

5.3. Kameros valdymo eksperimentai

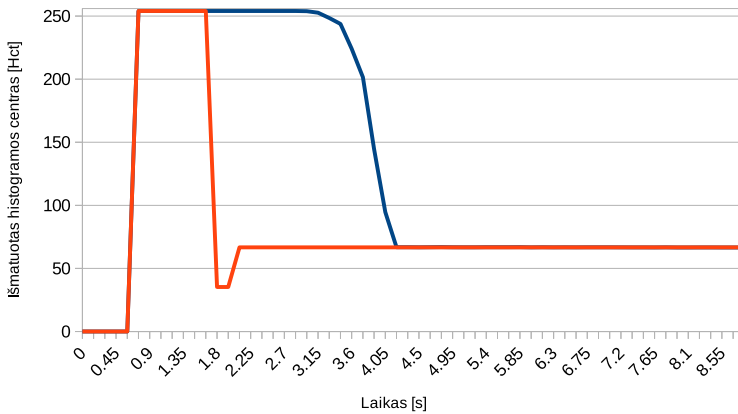
Kameros valdymo eksperimentų metu buvo lyginamas automatinis PointGrey Flea3 kameros valdymo algoritmas su šioje disertacijoje pristatomu algoritmu. Eksperimentų metu buvo tiriama, kaip algoritmai sugeba prisitaikyti prie staiga pasikeitusių apšvietimo sąlygų. Išmatuoto prisitaikymo laiko grafikas pateikiamas 5.3 pav. Eksperimentu buvo pademonstruota, kad pasiūlyto algoritmo prisitaikymo laikas greitesnis, bet yra šiek tiek viršijamas tikslo taškas.

5.4. Stereosuliejimo eksperimentai

Disertacijoje pristatomas algoritmas *Cyclops* buvo lyginamas su dviem realaus laiko algoritmais Semi-Global block matching (SGBM) ir Efficient large-scale stereo matching (ELAS). Algoritmams įvertinti buvo naudojama Middlebury Stereo Evaluation duomenų rinkinio trečioji versija. Eksperimentų metu buvo vertinamas algoritmų skaičiavimo laikas, naudojama atmintis, blogai apskaičiuotų ir neapskaičiuotų taškų kiekis ir blogai apskai-



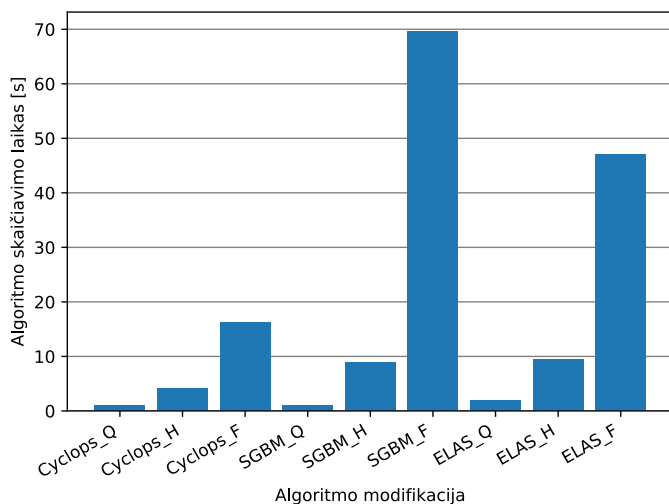
5.2 pav. Skirtingų įrenginių įvesties laiko diagrama



5.3 pav. Kameros parametrų valdymo algoritmo pritaikymo laikas. Gaminio algoritmas (mėlynas) ir pasiūlytas algoritmas (oranžinė)

čiuotų taškų vidutinė paklaida.

Eksperimentais buvo nustatyta, kad pasiūlyto algoritmo skaičiavimo laikas yra trumpesnis už kitų algoritmų. Skaičiavimo laikai yra pavaizduoti 5.4 pav. Taip pat buvo nustatyta, kad pristatomo algoritmo skaičiavimo laikas auga lėčiau, didėjant pateikiamų vaizdų rezoliucijai. Tai parodo, kad šis algoritmas yra tinkamas didelės rezoliucijos vaizdams apdoroti. Algoritmas taip pat turėjo mažesni naudojamos atminties kiekį. Deja, sumažintas atminties ir skaičiavimo laiko kiekis turėjo neigiamos įtakos pristatomo algoritmo tikslumo rodikliams. Disertacijoje yra pateikiamas detalus visų atliktų eksperimentų aprašymas.

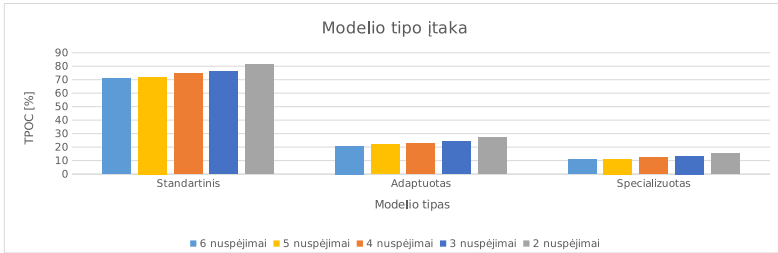


5.4 pav. Vidutinis algoritmo skaičiavimo laikas vienam vaizdui sekundėmis

5.5. Teksto nuspėjimo algoritmo eksperimentai

Šiame skyriuje buvo tirtas specializuoto teksto nuspėjimo algoritmo efektyvumas. Specializuotas teksto nuspėjimo algoritmas buvo lyginamas su bendro pobūdžio teksto nuspėjimo algoritmu ir adaptuotu algoritmu. Adaptuotas algoritmas naudoja standartinį algoritmą, papildomai išmokytą specializuoto algoritmo duomenimis. Algoritmų gerumas buvo vertina-

mas naudojant bendrą reikiamų įvesti skaitmenų procentą (*Total percentage of characters (TPOC)*). Šio eksperimento rezultatai pateikiami 5.5 pav. Grafike aiškiai matoma, kad net adaptuotasis algoritmas nesugeba pasiekti tokių gerų rezultatų, kaip specializuotas teksto nuspėjimo algoritmas.



5.5 pav. TPOC pateikiamas standartiniam, adaptuotam ir specializuotam teksto nuspėjimo algoritmui. Kiekvienas algoritmas buvo sukongfigūruotas grąžinti nuo 6 iki 2 žodžių pasiūlymų

5.6. Eksperimentų apibendrinimas

Šiame skyriuje aprašomi disertacijos metu atlikti eksperimentai. Atlikti eksperimentai gali būti suskirstyti į dvi grupes: eksperimentai su tyrimo dalyviais ir eksperimentai, skirti disertacijoje pristatomiems algoritmams įvertinti. Tyrimo dalyviai įvertino sukurtą daugiafunkcę vartotojo sąsają ir projektuojamų vaizdų vartotojo sąsają. Visi likę eksperimentai vertino pasiūlytų algoritmų efektyvumą. Tyrimo dalyviai šiuose eksperimentuose nedalyvavo.

6. IŠVADOS

1. Šioje disertacijoje yra pateikiama detali vartotojams pritaikytų pagalbinių technologijų ir HCI metodų analizė. Dauguma egzistuojančių sprendimų yra pritaikyti vienam panaudos atvejui. Dėl to vartotojas, atlikdamas skirtingas užduotis, turi išmokti naudotis skirtingomis sistemomis. Vartotojas taip pat privalo keisti sistemą pasikeitus jo fizinei būklei.
2. Egzistuojančių sistemų trūkumai buvo sprendžiami projektuojant daugiafunkcės vartotojo sąsajos architektūrą, skirtą sistemos adaptyvumui užtikrinti. Pasiūlyta architektūra buvo panaudota kurianti daugiafunkcę vartotojo sąsaja, kurioje buvo integruoti keli skirtingi pagalbiniai įrenginiai. Sukurta sistema buvo išbandyta 11 eksperimento dalyvių. Kiekvienas dalyvis naudojo sistemą su visais sistemoje integruotais įrenginiais. Eksperimentų metu nustatyta, kad įrenginių tinkamumas buvo labai skirtingas tarp eksperimento dalyvių. Kai kurie dalyviai negalėjo naudoti tam tikrų įrenginių, bet visi dalyviai galėjo naudoti bent vieną įrenginį efektyviai. Šis eksperimentas patvirtina adaptyvios, pagalbinės keleto įrenginių sistemos pranašumą. Dalyviai galėjo naudoti sistemą iki 60 proc. efektyviau, naudojami kelis įrenginius vienu metu.
3. Automatinis projektuojamų vaizdų generavimas leido sukurti novatorišką bei nebrangią vartotojo sąsają, skirtą objektams parinkiti. Ši sąsaja sujungia RGB-D kamerą su projektoriumi, taip vartotojui sukurdamą interaktyvią aplinką. Pagrindinis sistemos trūkumas yra projektuojamo vaizdo neryškumas šviesiose aplinkose. Ši problema gali būti sprendžiama naudojant HMD įrenginius, kai jie taps plačiai prieinami. Pasiūlytas algoritmas yra tinkamas panaudoti HMD įrenginiuose. Projektuojamų vaizdų sąsaja buvo įgyvendinta ir integruota į daugiafunkcę vartotojo sąsają. Tyrimo dalyviai įvertino dabartinę sistemą, atsakydami į vartotojo patirties klausimyną. Sukurta sąsaja vidutiniškai buvo įvertinta 2,84 balo skalėje nuo 1 iki 7 (čia 1 yra geriausias įvertinimas).
4. Kameros valdymo algoritmas realiu laiku reguliuoja jutiklio jautrumo

ir kadro užlaikymo parametrus. Norint pasiekti patikimą objektų atpažinimo algoritmo veikimą, yra reikalingi gera ekspoziciją turintys kameros vaizdai. Šioje disertacijoje yra pademonstruota, kad kameros parametrų sistemos procesas yra beveik tiesinis, jeigu vienas iš valdomų parametrų yra nekeičiamas. Kaskadomis paremtas PID valdiklis buvo panaudotas dviem parametrams valdyti vienu metu. Procesu netiesiškumas, atsirandantis dėl parametrų tarpusavio ryšių, buvo panaikintas panaudojus stiprinimo koeficientų automatinį koregavimą. Šis valdiklis sugebėjo prisitaikyti prie staiga pakitusių apšvietimo sąlygų per 50 proc. laiko (du kartus greičiau), kurį užtruko gamintojo pateikiamas algoritmas. Valdiklio veikimo laikas gali būti pagerintas panaudojant automatinius PID valdiklio derinimo metodus.

5. Stereokamera yra naudojama aplinkai suvokti ir objektams aptikti. Tokia kamera yra nejautri tiesioginiams saulės spinduliams ir gali apskaičiuoti aukštos rezoliucijos trimatį aplinkos vaizdą. Šioje disertacijoje yra pristatomas efektyvus stereovaizdų apdorojimo algoritmas, tinkamas įterptinėms sistemoms. Pasiūlytas algoritmas geba apskaičiuoti atstumų vaizdą iki dviejų kartų greičiau negu lyginamieji efektyvūs stereovaizdų apdorojimo algoritmai. Tuo pat metu algoritmo naudojamas atminties kiekis buvo sumažintas net iki 32 proc. Ateityje turėtų būti tobulinama antra algoritmo stadija, kad galima būtų sumažinti dabartinio algoritmo skirtuminio vaizdo klaidas.
6. Teksto nuspėjimo algoritmai yra naudojami siekiant sutrumpinti teksto įvedimo laiką. Šioje disertacijoje buvo pademonstruota, kad specializuotas teksto nuspėjimo algoritmas yra efektyvesnis negu bendro pobūdžio teksto nuspėjimo algoritmas. Specializuotas teksto nuspėjimo algoritmas buvo išmokytas naudojant duomenų rinkinį, sukurtą iš medicininių klausimynų. Šis algoritmas veikia iki 60 proc. efektyviau negu bendro pobūdžio algoritmas. Toks būdas gali būti panaudotas kuriant keletą specifiniams pokalbiams išmokytų algoritmų. Konkretus nuspėjimo algoritmas yra aktyvuojamas sistemos veikimo metu.
7. Pasiūlyti algoritmai buvo integruoti į daugiafunkcę vartotojo sąsają. Ši sąsaja leidžia vartotojams atlikti naujus veiksmus. Pastarai-

siais metais prie interneto prijungtų įrenginių kiekis smarkiai išaugo. Tokių įrenginių integravimas į pasiūlytą vartotojo sąsają leistų vartotojams atlikti dar daugiau užduočių savarankiškai. Vartotojui tinkami pagalbiniai įrenginiai turėtų būti parenkami individualiai. Įrenginių parinkimą reikėtų kartoti kiekvieną kartą pasikeitus vartotojo galimybėms.

Literatūros sąrašas

- [Angelo, 2000] Angelo, J. (2000). Factors affecting the use of a single switch with assistive technology devices. *Journal of rehabilitation research and development*, 37(5):591.
- [Billingham ir kt., 2015] Billingham, M., Clark, A., Lee, G., ir kt. (2015). A survey of augmented reality. *Foundations and Trends® in Human-Computer Interaction*, 8(2-3):73–272.
- [Cook and Polgar, 2014] Cook, A. M. and Polgar, J. M. (2014). *Assistive Technologies-E-Book: Principles and Practice*. Elsevier Health Sciences.
- [Dalmaijer, 2014] Dalmaijer, E. (2014). Is the low-cost eyetribe eye tracker any good for research? Technical report, PeerJ PrePrints.
- [Garay-Vitoria and Abascal, 2006] Garay-Vitoria, N. and Abascal, J. (2006). Text prediction systems: a survey. *Universal Access in the Information Society*, 4(3):188–203.
- [Geiger ir kt., 2010] Geiger, A., Roser, M., and Urtasun, R. (2010). Efficient large-scale stereo matching. In *Asian conference on computer vision*, pages 25–38. Springer.
- [Hirschmuller, 2008] Hirschmuller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2):328–341.
- [Hochberg ir kt., 2012] Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., Haddadin, S., Liu, J., Cash, S. S., van der Smagt, P., ir kt. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375.
- [Huang ir kt., 2014] Huang, X., Baker, J., and Reddy, R. (2014). A historical perspective of speech recognition. *Communications of the ACM*, 57(1):94–103.

- [Jacob, 1991] Jacob, R. J. (1991). The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS)*, 9(2):152–169.
- [Krishnamurthy and Ghovanloo, 2006] Krishnamurthy, G. and Ghovanloo, M. (2006). Tongue drive: A tongue operated magnetic sensor based wireless assistive technology for people with severe disabilities. In *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*, pages 4–pp. IEEE.
- [Leeb ir kt., 2010] Leeb, R., Sagha, H., Chavarriaga, R., ir kt. (2010). Multimodal fusion of muscle and brain signals for a hybrid-bci. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4343–4346. IEEE.
- [Manns and Chad, 2001] Manns, P. J. and Chad, K. E. (2001). Components of quality of life for persons with a quadriplegic and paraplegic spinal cord injury. *Qualitative Health Research*, 11(6):795–811.
- [Maskeliunas ir kt., 2016] Maskeliunas, R., Damasevicius, R., Martisius, I., and Vasiljevas, M. (2016). Consumer-grade eeg devices: are they usable for control tasks? *PeerJ*, 4:e1746.
- [Nourani-Vatani and Roberts, 2007] Nourani-Vatani, N. and Roberts, J. M. (2007). Automatic camera exposure control. In *Proceedings of the Australasian Conference on Robotics and Automation 2007*, pages 1–6. Australian Robotics & Automation Association ARAA.
- [Perdikis ir kt., 2018] Perdikis, S., Tonin, L., Saeedi, S., Schneider, C., and Millán, J. d. R. (2018). The cybathlon bci race: Successful longitudinal mutual learning with two tetraplegic users. *PLoS biology*, 16(5):e2003787.
- [Rusu and Cousins, 2011] Rusu, R. B. and Cousins, S. (2011). 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–4, Shanghai, China.
- [Scharstein ir kt., 2014] Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., and Westling, P. (2014). High-

resolution stereo datasets with subpixel-accurate ground truth. In *German Conference on Pattern Recognition*, pages 31–42. Springer.

[Vlissides et al., 1995] Vlissides, J., Helm, R., Johnson, R., and Gamma, E. (1995). *Design patterns: Elements of reusable object-oriented software*. Reading: Addison-Wesley, 49(120):11.

[Westgren and Levi, 1998] Westgren, N. and Levi, R. (1998). Quality of life and traumatic spinal cord injury. *Archives of physical medicine and rehabilitation*, 79(11):1433–1439.

[Yang et al., 2016] Yang, L., Normand, J.-M., and Moreau, G. (2016). Practical and precise projector-camera calibration. In *Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on*, pages 63–70. IEEE.

AUTORIAUS PUBLIKACIJOS DISERTACIJOS TEMA

Moksliniai straipsniai tarptautiniuose žurnaluose, turinčiuose citavimo indeksą

1. Ivanavičius, A., Simonavičius, H., Gelšvartas, J., Lauraitis, A., Maskeliūnas, R., Cimmerman, P., & Serafinavičius, P. (2018). Real-time CUDA-based stereo matching using Cyclops2 algorithm. *EURASIP Journal on Image and Video Processing*, 2018(1), 12.
2. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2018). Projection Mapping User Interface for Disabled People. *Journal of healthcare engineering*, 2018.
3. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2016). User adaptive text predictor for mentally disabled Huntington's patients. *Computational intelligence and neuroscience*, 2016, 2.

Moksliniai straipsniai kituose moksliniuose žurnaluose

1. Gelšvartas, J., Simonavičius, H., Lauraitis, A., Maskeliūnas, R., Cimmerman, P., & Serafinavičius, P. (2016). Realtime Stereo Matching Using Maximally Stable Extremal Regions. *Transylvanian Review*, (1).

Moksliniai straipsniai tarptautinėse konferencijose

1. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2018). Augmented reality object selection user interface for people with severe disabilities. 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health (pp. 156-160).
2. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2015, August). Multifunctional user interface implementation details and evaluation. In *Methods and Models in Automation and Robotics (MMAR)*, 2015 20th International Conference on (pp. 501-504). IEEE.
3. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2015, October). Text Predictor for Lithuanian Language. In *International Conference on*

Information and Software Technologies (pp. 460-468). Springer, Cham.

4. Gelšvartas, J., Simutis, R., & Maskeliūnas, R. (2014). Multifunctional user interface to control robotic arm for paralyzed people. *Electrical and Control Technologies*, 22.
5. Gelšvartas, J., Lauraitis, A., Simutis, R., & Maskeliūnas, R. Review of assistive technologies for disabled people. *BIOMEDICAL ENGINEERING* 2016, 20(1).

7. CV

Julius Gelšvartas

El.-paštas: julius.gelsvartas@gmail.com, Tel.: +37061551059

Įgudžiai ir patirtis:

Robotika, ROS	C++/C, Python
Vaizdų apdorojimas/Kompiuterių rega	OpenCV, PCL
Procesų valymas, padėties nustatymas	PID, Kalman filtrai

Išsilavinimas:

2014-2018

Kauno technologijos universitetas Informatikos inžinerijos doktorantūra
Daugiafunkcės vartotojo sąsajos paralyžiuotiesiems kūrimas ir tyrimai

2009-2010

The University of Edinburgh Dirbtinio intelekto magistras
Pagrindinė mokslo šaka Robotų intelektas

2005-2009

Kauno technologijos universitetas Informatikos bakalauras.

Darbo patirtis:

2019-03 iki Dabar Hexagon Technology Center (<https://hexagon.com/>)

Vaizdų apdorojimo inžinierius. Vaizdais paremtų pozicionavimo sistemų kūrimas.

2013-08 iki 2019-03 Rubedo sistemos (<http://www.rubedo.lt/>)

Vyr. Programuotojas. Autonominių mobilių robotų architektūros kūrimas. Robotų vaizdų apdorojimo ir valdymo algoritmų kūrimas.

2012-05 iki 2013-07 Cognitec (<http://www.cognitec-systems.de/>)

Programinės įrangos inžinierius. Veidų atpažinimo sistemų architektūra ir kūrimas.

2010-09 iki 2012-04 SeeByte (<http://www.seebyte.com/>)

Programinės įrangos inžinierius. Autonominių povandeninių laivų valdymo sistemų kūrimas.

2007-03 iki 2009-06 InfoEra (<http://www.infoera.lt/>)

Programinės įrangos kūrėjas. GIS sistemų architektūra ir kūrimas.

8. SUMMARY

8.1. Introduction

8.1.1. Problem statement and relevance

Existing assistive technologies are not adaptable. They are usually only developed for a single device. This makes it difficult to select an assistive technology for each individual user. This thesis presents a multifunctional user interface for people suffering from severe motor disabilities such as quadriplegia. We propose a system that is easily adaptable to the needs of each individual user.

8.1.2. Thesis object

The object of this thesis is to study consumer grade assistive technology devices and modern HCI methods.

8.1.3. Thesis goal

The goal of this thesis is to help people, suffering from severe motor disabilities, to communicate efficiently. The proposed solution concentrates on easy system adaptability for each individual user. Constant system customization ensures efficient HCI independently of users motor disability level.

8.1.4. Tasks of the thesis

The goal of this thesis has been achieved by solving the following tasks:

1. Analyze and compare consumer grade AT devices and modern HCI methods;
2. Design multifunctional UI system architecture;
3. Propose real time scene recognition and augmented reality (AR) user interface presentation algorithms;
4. Propose predictive algorithms that improve system interaction time;
5. Develop an adaptive AT user interface application that integrates multiple input devices. The created application also contains an im-

age processing pipeline and AR information presentation component;

6. Perform experiments and evaluate the developed UI and proposed algorithms.

8.1.5. Scientific novelty

The research achieved these novel results:

1. Development of projection mapping based object selection UI. The novel automatic projection mapping is achieved by combining camera-projector system calibration with RGB-D based object detection pipeline. The system automatically detects objects and uses projector to highlight them in real world. User can select desired object using this UI.
2. Proposed region based stereo matching algorithm that can process images in real-time on embedded computer. Stereo matching algorithm novelty is using region based features to reduce stereo matching search space. Stereo camera with onboard embedded computer is used as an environment sensing device.
3. Camera systems operating in changing lighting conditions require fast and stable camera parameter control algorithms. Cascade camera parameter control algorithm has been proposed. Main algorithm novelty is using cascade controller to control multiple camera parameters simultaneously. Another novelty is using controller parameter gain scheduling to increase controller efficiency and stability. The algorithm is used to ensure image processing pipeline reliability.

8.1.6. The defended statements

1. Assistive user interfaces that integrate multiple assistive devices are more adaptive. Adaptability ensures that the system is suitable for a larger user group. Assistive devices are used to control the proposed multifunctional user interface.

2. Cascade controller can be used to efficiently control camera shutter speed and sensor gain parameters simultaneously. Controller stability can be ensured by performing gain scheduling to address process non-linearity. This controller is used to improve multifunctional user interface image processing pipeline reliability.
3. Maximally stable extremal region based stereo matching algorithm can be used to reduce the search space. Reducing search space can significantly improve algorithm computation time and reduce memory consumption. This algorithms can be used to perform computations on embedded Jetson TK1 computer. Such system is used as a reliable environment sensing device of multifunctional user interface.
4. Specialized text predictor can help to input words more efficiently than standard text predictor. Such predictor is used for efficient text input during topic specific conversations. Efficient text input is one of the capabilities offered by the proposed system.

8.1.7. Scientific approval

Results presented in this thesis are original and correspond to a total of 9 publications. 3 publications have been published in "ISI Web of Science" scientific journals with Impact Factor and 1 publications in "ISI Web of Science" journals. The remaining five publications are published in conference proceedings.

Research results have been presented in the following international conferences:

1. The 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE) 2018, Funchal, Portugal;
2. The 20th International Conference on Methods and Models in Automation and Robotics (MMAR) 2015, Międzyzdroje, Poland;
3. The 20th International Conference Biomedical Engineering 2016, Kaunas, Lithuania;

4. The 21st International Conference Information and Software Technologies (ICIST) 2015, Druskininkai, Lithuania;
5. The 9th International Conference Electrical and Control Technologies (ETC) 2014, Kaunas, Lithuania;

8.1.8. Thesis organization

The thesis is organized into 5 chapters. Chapter 2 presents and discusses state of the art assistive technologies. This chapter also discusses other technologies that can be used to create modern assistive technologies. The detailed architecture of presented multifunctional user interface is presented in chapter 3. Chapter 4 describes novel algorithms proposed in this thesis. The experiments performed to evaluate the multifunctional user interface and proposed algorithms are presented and described in details in chapter 5. Chapter 6 is the conclusions.

8.2. Analysis of existing assistive technology solutions

This chapter reviews the state-of-the-art AT related to this thesis. AT is a very broad term covering a range of assistive, adaptive and rehabilitative devices as well as specialized HCI user interfaces. Modern AT are complex systems that integrate simple but reliable assistive devices with advanced UI programs. Patients are able to perform complex tasks using such systems. We will mainly focus on five key aspects of AT HCI systems, namely input modalities, information processing, information presentation, environment sensing and interaction.

Majority of this chapter is dedicated to studying HCI methods that are used in AT applications. This chapter reviews 6 different assistive devices, namely Sip/Puff, BCI, Tongue, Electromyography (EMG), Eye tracking and Speech recognition. We present a detailed device comparison and their advantages and disadvantages. Other sections in this chapter concentrate on environment sensing and information presentation methods.

8.3. Architecture of multifunctional user interface

In this chapter we discuss the architecture of multifunctional UI presented in this thesis. The UI was designed taking system reconfigurability and extendability into account.

The sections in this chapter discuss the detailed architecture of system components. This chapter first details the developed system requirements, use cases and describes system deployment. Next we define the system components and interfaces that are used for information exchange. All system components have well defined interfaces. Each system component is dynamically loaded ensuring system extensibility. This makes it possible to add new device and UI screens without the need to change the remaining system code. Lastly, the detailed information of classes used to implement the components is described and system component information flow diagram is presented.

8.4. Proposed algorithms and methods for multifunctional user interface

This chapter describes the novel algorithms presented in this thesis. The presented algorithms are used to improve multifunctional user interface for disabled people. The developed system contains a projection mapping based UI used for object selection. This interface offers a natural HCI method because the user sees detected objects directly in the scene. The projection mapping user interface consists of three parts, namely camera-projector calibration, object detection and detected object highlight rendering. The main novelty of the proposed projection mapping user interface is that it integrates object detector and is therefore able to highlight the objects in the scene automatically.

The first three algorithms presented in sections are used in novel Projection mapping based UI. These three algorithms are camera-projector calibration process, image processing pipeline and object highlight rendering.

Camera control algorithm is also described in details. Many cameras have built-in camera control algorithms but their quality varies significantly. Stable and fast camera control ensures that results obtained using image processing are accurate and reliable. The proposed algorithm is therefore important in ensuring reliability of multifunctional UI.

Stereo matching is another algorithm that is used to improve the Projection mapping UI. Automatic projection mapping used in this thesis requires a RGB-D camera to perform object detection. The proposed stereo matching algorithm can be used to create a stereo RGB-D camera capable of processing images in real-time. When combined with proposed camera

parameter controller such stereo camera would be able to work outdoors and improve projection mapping UI reliability.

Final chapter in this section describes the specialized text predictor used in multifunctional user interface. This text predictor can significantly reduce the text input time and improve UX.

8.5. Experimental setup and results

This chapter describes the experiments that were performed to evaluate the proposed multifunctional user interface architecture and algorithms. The described experiments measure the performance of individual algorithms as well as the usability of the whole system. The following sections describe the experimental setup and the results of the performed experiments.

Participant experiments were used to evaluate developed multifunctional UI and projection mapping UI. All remaining experiments were performed to evaluate the proposed algorithms and had no participants. Camera control experiments used PointGrey camera to evaluate the algorithm. Middlebury Stereo Evaluation data-set was used for stereo algorithm evaluation. Finally, a data-set generated from medical questionnaires was used for text predictor evaluation.

8.6. Conclusions

1. This thesis contains a comprehensive analysis of existing consumer grade AT solutions and HCI methods. Majority of existing solutions are intended for single use case. As a result, the user has to learn to use multiple systems to perform different tasks. Moreover, the user would have to change a system as their abilities change.
2. The limitations of existing solutions have been addressed by designing a multifunctional UI architecture optimized for system adaptability. The proposed architecture was used to implement a multifunctional user interface that integrates multiple assistive devices. The developed system has been tested by 11 participants. Each participant used the system with a number of assistive devices. The experiments revealed that device suitability varied significantly among the participants. Some participants were not able to use some devices, but every participant managed to use at least one device ef-

ficiently. This confirms the advantage of multiple device adaptive assistive system. Moreover, the participants were able to use the system up to 60% faster with two devices (used simultaneously).

3. Automatic projection mapping is a novel, low-cost user interface for object selection. This interface combines RGB-D camera with projector to create real time interactive UX. Main proposed method limitation is projected view brightness in bright environments. These limitations can be addressed with novel HMD devices as they become more widespread. The presented algorithms are suitable for HMD devices. The projection mapping UI has been implemented and integrated into the multifunctional user interface. Experiment participants evaluated current system using PSSUQ questionnaire. The UI received average score of 2.84 points on the scale of 1 to 7 (where 1 is highest score).
4. Camera control algorithm adjusts sensor gain and shutter speed parameters in real time. This algorithm is used to ensure reliability of image processing algorithms. This thesis demonstrated that camera image exposure process is mostly linear when one of the controlled parameters is fixed. Cascaded PID controller is used to control two output signals simultaneously. The non-linearity resulting from camera parameter interdependence was eliminated by applying PID gain scheduling techniques. We demonstrate that proposed algorithm can adapt to sudden lighting change in 50% of time taken by the manufacturer provided camera control algorithm. Further algorithm performance improvements could be achieved by using automatic PID tuning techniques.
5. Stereo camera is used for environment sensing and object detection. Such camera is not sensitive to direct sunlight and can produce high resolution depth images of the environment. This thesis presented an efficient stereo matching algorithm suitable for embedded computers. The proposed algorithm computes depth images up to two times faster than state of the art efficient stereo matching algorithms. At the same time this algorithm used at least 32% less memory compared to other evaluated algorithms. Future algorithm developments

will be concentrated on reducing disparity errors produced by second algorithm stage.

6. Text predictors are used to reduce text input time. This thesis demonstrated that conversation specific text predictor is more efficient than general purpose text predictor. Proposed specialized predictor has been trained using data-set created from medical questionnaire questions. This predictor works up to 60% more efficiently than standard text predictor. This approach can be used to create several topic specific predictors and switching them during system operation.
7. The proposed algorithms have been integrated into a multifunctional UI that enables users to perform novel tasks. In recent years the amount of internet connected devices has grown significantly. Such devices can be integrated into the proposed system and enable users to perform more tasks independently. It is recommended to choose the combination of devices for each user individually. This combination should be constantly updated as the users abilities change.

UDK 004.5-056.29 + 004.932](043.3)

SL344. 2019-05-31; 3,5 leidyb. apsk. 1. Tiražas 50 egz.

Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas