



**UNIVERSITY
OF OULU**

TIETO- JA SÄHKÖTEKNIIKAN TIEDEKUNTA

**Alexi Lahti
Antti Luukkonen
Jouni Saari**

**IHMISEN JA ROBOTIN SANATTOMAN
KANSSAKÄYNNIN HAASTEET SULAUTETUSSA
JÄRJESTELMÄSSÄ**

Kandidaatintyö
Tietotekniikan tutkinto-ohjelma
Lokakuu 2020

Lahti A., Luukkonen A., Saari J. (2020) Ihmisen ja robotin sanattoman kanssakäynnin haasteet sulautetussa järjestelmässä. Oulun yliopisto, Tietotekniikan tutkinto-ohjelma, 31 s.

TIIVISTELMÄ

Robottien kehityksen myötä niitä voidaan hyödyntää tehtävissä, jotka vaativat robotilta jatkuvaa kanssakäymistä ihmisten kanssa. Ihmisten keskuudessa toimivien robottien pitää pystyä vuorovaikuttamaan sulavasti ja luonnollisesti käyttäjiensä kanssa. Sulava kanssakäynti vaatii robotilta kykyä lukea ihmisen sanatonta ja sanallista viestintää sekä kykyä tuottaa ihmisten ymmärtämää viestintää.

Tässä työssä kehitettiin toimiva katsekontakti robotille. Työssä käytettiin pohjana InMoov-robotin päätä, jonka silmiä ohjattiin Raspberry PI:lle ja Arduino UNO:lle kehitetyillä ohjelmistoilla. Katsekontakti toteutettiin seuraavalla tavalla: Robotin silmään sijoitetulta kameralta tulevasta kuvasta tunnistetaan kasvot, jonka jälkeen silmät ohjataan kohti kuvassa olevia kasvoja.

Avainsanat: Ihmisen ja robotin vuorovaikutus, robotiikka, tekoäly, katsekontakti

Lahti A., Luukkonen A., Saari J. (2020) Challenges of Nonverbal Human-Robot Interaction within an Embedded System. University of Oulu, Degree Programme in Computer Science and Engineering, 31 p.

ABSTRACT

Due to technological improvement of robots they can be utilized in tasks that require constant interaction with humans. Robots that work among humans need to be able to interact naturally and fluently with its users. Smooth interaction requires ability to read both non-verbal and verbal communication of humans and an ability to communicate in a way that can be easily understood.

In this project a working eye contact application was built for a robot. The robot used is built on an InMoov robot head, with eyes controlled by software made for Raspberry Pi and Arduino UNO. The application was designed in the following way: A camera mounted in the eye of the robot takes pictures. A face recognition algorithm then finds the face and the eyes are turned towards it.

Keywords: Human robot interaction, robotics, artificial intelligence, eye contact

SISÄLLYSLUETTELO

TIIVISTELMÄ	
ABSTRACT	
SISÄLLYSLUETTELO	
ALKULAUSE	
LYHENTEIDEN JA MERKKIEN SELITYKSET	
1. JOHDANTO	7
2. TAUSTA.....	8
2.1. Ihmisen näköjärjestelmä.....	8
2.1.1. Näkö	9
2.1.2. Silmän liikkeet	9
2.2. Tekoäly ja konenäkö	10
2.3. Katsekontakti	11
2.4. Ihmisen ja robotin vuorovaikutus	11
3. TOTEUTUS	13
3.1. Laitteisto	13
3.1.1. Kamera.....	14
3.1.2. Raspberry Pi 3 Model B.....	14
3.1.3. Arduino UNO	15
3.1.4. Servot.....	16
3.1.5. InMoov-robotti.....	16
3.2. Kasvontunnistus	17
3.2.1. Kameran asetusten valinta.....	18
3.2.2. Luokittelijan testaus	19
3.3. Sarjakommunikaatio ja ohjelmasilmukka	21
3.4. Servojen ohjaus	22
4. POHDINTA.....	24
4.1. Tulokset	24
4.2. Toiminnalliset parannukset	24
4.3. Jatkokehitys.....	25
5. AJANKÄYTTÖ	26
6. YHTEENVETO.....	27
7. VIITTEET	28

ALKULAUSE

Haluamme kiittää Teemu Tokolaa työn ohjauksesta, sekä Juha Röningiä työn tarkastamisesta. Lisäksi haluamme kiittää kaikkia työssä avustaneita henkilöitä. Kandidaatintyö tehtiin osana Oulun Yliopiston Sulautettujen järjestelmien projekti -kurssia 521275A.

Oulussa 1. lokakuuta 2020

Aleksi Lahti
Antti Luukkonen
Jouni Saari

LYHENTEIDEN JA MERKKIEN SELITYKSET

USB	Universal Serial Bus
GHz	Gigahertsi
GPIO	General Purpose IO
kB	kilotavu
I/O	Input/Output
MHz	Megahertsi
PWM	Pulssinleveysmodulaatio
IDE	Integrated Development Environment
TTL	Time To Live
V	Voltti
Nm	Newtonmetri
fps	kuvataajuus
LBP	Local Binary Patterns
LBPi	Local Binary Patterns improved
s	sekunti
o	aste
h	tunti

1. JOHDANTO

Tekniikan ja tekoälyn kehityksen myötä robottien määrä ihmisten keskuudessa on kasvamassa nopeasti. Valtaosa niistä on edelleen teollisuuskäytössä, missä niillä on ennalta määrätyt tehtävät ja vuorovaikutus ihmisten kanssa on vähäistä. Kehitys on kuitenkin menossa suuntaan, jossa roboteista on tulossa osa ihmisten jokapäiväistä elämää ja vuorovaikutustilanteet niiden kanssa ovat lisääntymässä.

Suurimpana esteenä robottien laajamittaiselle käyttöönotolle kotiympäristöihin on pidetty niiden liikkuvuuden rajoittuneisuutta ja kykyä suorittaa tehtäviä itsenäisesti. Viime vuosina robottien kehityksessä on kuitenkin edetty kohti sellaisia robotteja, jotka osaavat toimia itsenäisesti ja myös ihmisten kanssa vuorovaikutuksessa. Lisäksi robotit kykenevät soveltaviin fyysisiin suorituksiin, joita vain muutamia vuosia sitten pidettiin hyvin vaikeina, kuten itsenäiseen ovien avaamiseen [1], maastossa navigointiin ja akrobatiaan [2]. Kun robottien kyky suorittaa itsenäisiä tehtäviä ihmisille rakennetuissa ympäristöissä paranee, robottien kanssakäynnin kehittämisen merkitys korostuu.

Yhtenä robotin ja ihmisen välisen kanssakäynnin esteenä on niin sanottu uncanny valley -ilmiö. Tässä ilmiössä ihmisen suhtautuminen esimerkiksi robottiin on sitä empaattisempi, mitä ihmismäisempi robotti on, kunnes robotti näyttää lähes ihmiseltä. Tällöin jokin osa robotin ulkoasusta tai käytöksestä jää puutteelliseksi ja siitä tulee sen seurauksena ihmiselle häiritsevän tai uhkaavan näköinen. Uncanny valley -ilmiön aiheuttamaan ongelmaan on olemassa kaksi ratkaisua. Robotista pitää joko saada täysin ihmismäinen niin liikkeiltään, kuin käytökseltään, tai siitä pitää tehdä selvästi robottimainen [3]. Luonnollisen ihmismäisen käytöksen saavuttaminen on kuitenkin vaikeaa, sillä se vaatii täydellistä onnistumista non-verbaalisessa ja verbaalisessa viestinnässä.

Automatisoidun verbaalisen kommunikoinnin parhaimmista edustavista eri yritysten virtuaaliset assistentit, jotka ovat monilta osin jo hyvin lähellä ihmisen tasoa. Kuitenkin niiltä puuttuu kyky sanattomaan viestintään, kuten katsekontaktiin. Katsekontaktia pidetään merkittävänä osana sosiaalista kommunikaatiota [4, 5]. Se on myös nykyteknologialla robotille saavutettavissa oleva toiminto [6], joka tässä työssä toteutettiin inMoov-robotille. Kehityksen kohteena oli silmärakenteen ohjausjärjestelmä, joka kykenee luonnolliseen katsekontaktiin yhden sen näkökentässä olevan henkilön kanssa. Työssä keskitytään sen tekniseen toteutukseen.

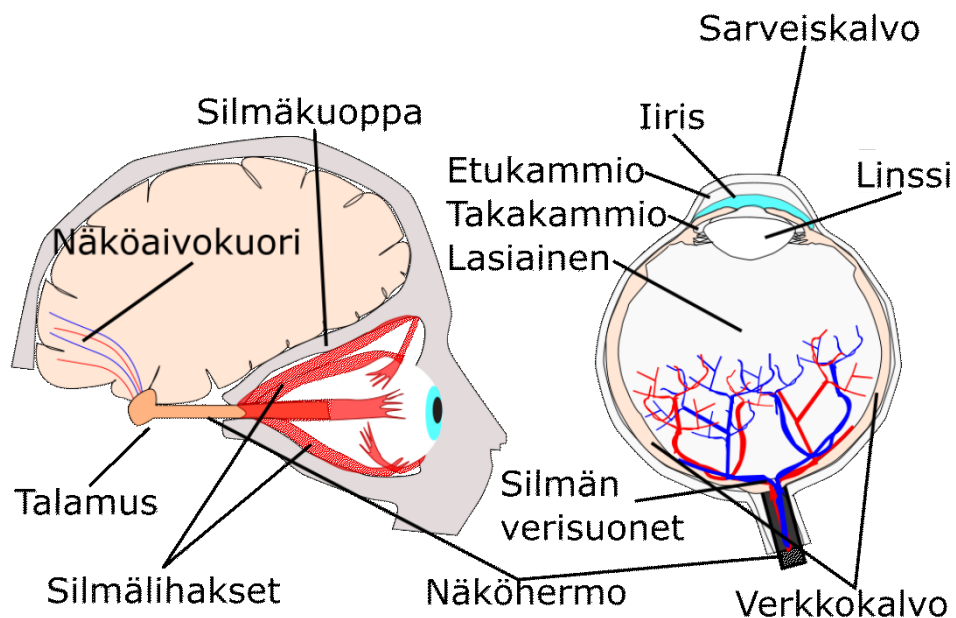
2. TAUSTA

Yhä useammat robotiikan sovellukset on tarkoitettu toimimaan jatkuvassa yhteistyössä ihmisen kanssa. Tästä johtuen ihmisen ja robotin vuorovaikutusta tutkitaan yhä enemmän. Tutkimusalana se on hyvin monipuolinen: Aiheeseen liittyy monialaisesti muun muassa tekoäly, ihmisen käyttäytyminen ja psykologia. Sosiaalisen robotin kehitystä lähestytään yleensä ihmisen käyttäytymistä tutkimalla. Sen kehittäminen vaatii hyvää ymmärrystä ihmisen biologiasta ja ihmisten välisestä vuorovaikutuksesta.

2.1. Ihmisen näköjärjestelmä

Ihmisen näköaisti perustuu ympäristöstä heijastuvan näkyvän valon havainnointiin ja se on ihmisen pääasiallinen informaatiolähde ympäröivästä maailmasta. Näköjärjestelmän rakenteita ovat luisiin silmäkuoppiin sijoittuneet silmämunat, silmälihaksat, näköhermosto ja aivojen näköaivokuori, jotka nähdään kuvassa 1 yhdessä silmän rakenteen kanssa.

Silmämunien näköaistimuksen mahdollistavat osat ovat silmän etuosassa oleva läpinäkyvä sarveiskalvo, valon määrän mukaan supistuva iiris ja iiriksen keskellä sijaitseva musta pupilli, jonka takana on valoa taittava linssi. Silmämunan keskiosa koostuu silmälle muodon antavasta nesteellä täyttyneestä lasiainesta, joka on yksi kolmesta nestekammioista silmässä. Toinen kammioista on sarveiskalvon ja iiriksen välissä sijaitseva etukammio ja kolmas iiriksen ja linssin väliin sijoittunut takakammio. Itse näköaistimus syntyy silmän takaosassa sijaitsevalla verkkokalvolla, jolta näköinformaatio kulkee näköhermoa pitkin näköaivokuorelle [7 s.473-475].



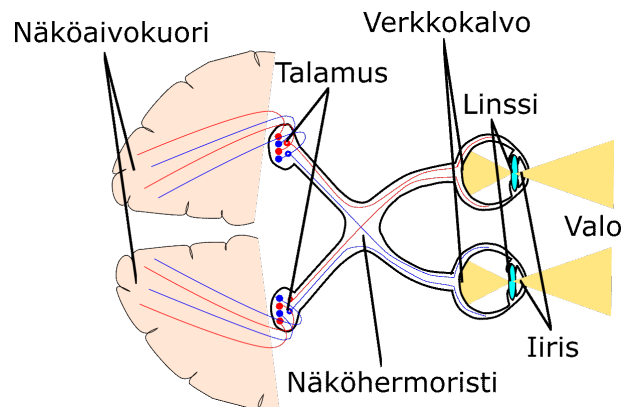
Kuva 1. Näköjärjestelmän rakenteet

2.1.1. Näkö

Näkeminen on kyky havainnoida ympäristöä havaitsemalla sen kappaleista heijastuvaa valoa 380 - 740 nanometrin aallonpituuksilla eli näkyvän valon alueella. Eliöillä tämän kyvyn toteuttaa näköaisti, joka koostuu ihmisillä ja useilla muilla nisäkkäillä silmästä, näköradasta, sekä aivojen näkökeskuksesta [8, 9].

Silmässä näkeminen tapahtuu kuvan 2 osoittamalla tavalla. Silmän etuosassa iiris säätelee linssille päätyvää valomäärää ympäristön valoisuuden mukaan. Kirkkaassa ympäristössä iiris supistuu ja linssille päätyvän valon määrä vähenee. Hämärissä ympäristössä iiris rentoutuu ja linssille päästetään enemmän valoa. Linssin tehtävä on taittaa sille päätyvä valo silmän takaosan verkkokalvolle, niin ettei sille heijastuva näkymä ole epätarkka. Tässä silmä onnistuu venyttämällä tai supistamalla linssiä niin että katseen kohteen valonsäteiden leikkauspiste on verkkokalvolla.

Verkkokalvolla olevat reseptorit yhdessä verkkokalvon hermosolujen kanssa muuntavat niihin osuvien fotonien energian hermoimpulsseiksi. Nämä hermoimpulsit kulkevat näkörataa pitkin näköhermoristin kautta näköjärjestelmän takaossa sijaitsevalle talamukselle ja edelleen näköaivokuorelle, jossa näköaistimus tapahtuu [7 s.483-484].



Kuva 2. Ihmisen näköaistimuksen toiminta

2.1.2. Silmän liikkeet

Ihmisen molempiin silmiin on kiinnittynyt kuusi lihasta, jotka ohjaavat niiden liikkeitä silmän keskipisteen ympäri [10, 11]. Silmät kykenevät nopeisiin katsekohtan muutoksiin, jotka kestävät 20-200 millisekuntia riippuen kulmasta, jolla katsekohta muuttuu [12]. Nopeiden liikkeiden lisäksi silmät pystyvät seuraamaan sulavasti näkökentässä liikkuvaa objektiä [13]. Ihmiset myös kykenevät säilyttämään katseensa yhdessä pisteessä automaattisesti kun he kääntävät päätänsä mihin tahansa suuntaan. Tätä kutsutaan vestibulo-okulaariseksi refleksiksi ja se on seurausta sisäkorvan tuottamasta liikkeen informaatiosta, joka aiheuttaa silmissä pään liikkeelle vastakkaisuuntaisen liikkeen [14]. Silmän keskittyessä yhteen paikkaan tapahtuu silmässä koko ajan pientä liikettä, joka estää silmän reseptorien tottumisen samaan ärsykkeeseen [15, 16]. Tämä on erityisesti huomattavissa katsekontaktissa.

2.2. Tekoäly ja konenäkö

Tekoäly on suuressa roolissa monissa nykyajan teknologioissa. Sitä käytetään muun muassa robottiajoneuvoissa, puhetunnistuksessa, logistisessa suunnittelussa, sekä roskapostin suodattamisessa [17]. Tekoälyn etuna tavanomaisiin menetelmiin verrattuna on kyky oppia ja mukautua eri tilanteisiin.

Kirjallisuudessa tekoäly on ollut olemassa käsitteenä jo antiikin ajoista lähtien [18], mutta vasta 1950-luvulta lähtien tekoälyä on voitu soveltaa tarpeeksi tehokkaiden laskennallisten laitteiden myötä. Alan Turingin julkaisu [19] filosofialehdessä *Mind* oli erityisen merkittävä tekoälyn historian kannalta. Julkaisu esitti muun muassa tunnetun Turingin testin, jolla mitataan koneen kykyä osoittaa ihmismäistä käyttäytymistä. Tekoäly on edistynyt huimaa vauhtia viime vuosikymmeninä, mutta aiheesta on vielä paljon opittavaa.

Russellin ja Norvigin [17] mukaan tekoälyn käsittelyyn on neljä lähestymistapaa, riippuen käyttötarkoituksesta. Näitä lähestymistapoja ovat tekoälyt, jotka

- ajattelevat ihmismäisesti,
- ajattelevat rationaalisesti,
- toimivat ihmismäisesti tai
- toimivat rationaalisesti.

Näistä rationaalisesti toimiva lähestymistapa on konenäkö toteuttamiseen sopivin. Rationaalisesti toimivan tekoälyn lähestymistavan mukaan tekoäly on tutkimusta agenteista jotka vastaanottavat havaintoja ympäristöstä sensoreilla ja toimivat niiden mukaan aktuaattoreilla [20, 17]. Rationaalisen agentin tavoite on päätyä parhaaseen lopputulokseen, tai vähintäänkin parhaaseen odotettuun tulokseen. Agentin suorituskykyä kuvataan jollakin suorituskykymitalla, joka kertoo kuinka hyvin agentti saavutti halutun tuloksen. Esimerkiksi konenäkössä mittana käytetään tunnistamisen tarkkuutta: hyvä agentti tunnistaa kaikki tai vähintäänkin useimmat halutuista esineistä kuvassa.

Toiminnan rationaalisuus kullakin hetkellä riippuu neljästä asiasta:

1. Suorituskykyä kuvaava mitta,
2. toiminnot mitä agentti voi suorittaa,
3. agentin aikaisempi tieto ympäristöstä ja
4. agentin havaintosekvenssi ajan hetkeen asti.

Näistä saadaan rationaalisen agentin määritelmä [17]:

Rationaalinen agentti valitsee jokaiselle mahdolliselle havaintosekvenssille toiminnon, jonka oletetaan maksimoivan sen suorituskykymitan havaintosekvenssin ja muun sisäisen tiedon mukaan, mitä agentilla on.

Kasvojentunnistuksen tapauksessa agentin aikaisempi tieto ja havainnot ovat tarpeellisia vain, jos halutaan, agentin muistavan kasvoja ja kykenemään tunnistamaan eri henkilöitä.

2.3. Katsekontakti

Sanaton kommunikaation käsittää 60-70 prosenttia kaikesta ihmisten välisestä kommunikaatiosta [21, 22, 23]. Uudemmissa tutkimuksissa on kuitenkin todettu, ettei sanattomasti välitetyn tiedon määrää pystytä tarkasti määrittämään, koska nyky-yhteiskunnassa ihmiset kommunikoivat kirjoittaen huomattavasti enemmän kuin ennen ja todellinen osuus kaikesta kommunikaatiosta on pienempi kuin 50 prosenttia [24, 25]. Kuitenkin kirjoitettu kommunikaatio on rajoittuneempi viestintäkeino, mikä johtaa helpommin väärinkäsityksiin. Esimerkiksi perinteisesti äänenpainotusta käyttävät ilmaisut, kuten sarkasmi, ovat haasteellisia täysin kirjoitetussa kommunikaatiossa [26].

Katsekontakti on osa tätä ihmisten välistä sanatonta kommunikointia [27], ja sitä pidetään tärkeimpänä kanssakäynnin keinona osoittaa toisille kiinnostusta, huomiota ja osallistumista [28, 29]. Katsekontaktin puuttumista taas pidetään merkinä välinpitämättömyydestä [28]. Eckman Hammerin mukaan tärkein tekijä katsekontaktilla kommunikoidessa on sen kesto, sillä sen on todettu olevan suoraan verrannollinen kommunikoivien henkilöiden läheisyyteen [30, 23]. Katsekontaktin keston ja yleisyyden lisäksi tärkeinä pidetään katseen liikkumista kasvoilla, pupillien laajentumista ja silmän räpäytyksien tiheyttä [27, 29, 31].

Erityisesti kädellisillä eläimillä, kuten simpansseilla katsekontakti käsitetään yleensä uhaksi [32]. Tämä pätee myös ihmisillä, jotka tulkitsevat pitkittyneen intensiivisen katsekontaktin epämuikavaksi tai uhkaavaksi riippuen katsojan etäisyydestä [33, 34]. Estääkseen uhkaavan vaikutelman syntymisen ihmiset yleensä siirtelevät katsettaan vuorovaikutuksen kohteen ja muiden kohteiden välillä [34].

2.4. Ihmisen ja robotin vuorovaikutus

Ihmisen ja robotin välisen yhteistyön yleistymisen esimerkiksi kotihoidossa, lääketieteessä ja ajoneuvoissa on johtanut siihen, että robottien on kyettävä saumattomaan vuorovaikutukseen ihmisen kanssa [35]. Saumaton vuorovaikutus vaatii robotilta kykyä lukea ihmisen viestintää sekä kykyä tuottaa ihmiselle ymmärrettävää viestintää. Edellä mainittujen ominaisuuksien onnistunut toteuttaminen on tärkeää sillä robotin viestinnän epätäydellisyys saattaa aiheuttaa pelkoa niiden potentiaalisissa käyttäjissä [36].

Robotin ulkonäkö vaikuttaa ihmisen ja robotin väliseen vuorovaikutukseen. Ihmisen kaltainen robotti yhdistetään useammin olioon, jolla on inhimillisiä ajatuksia tai tunteita, kuin robotti, jolla on pelkistetty mekaaninen ulkonäkö [37]. Ihmismäinen, mutta kuitenkin elämättömäksi tunnistettava robotti saattaa kuitenkin aiheuttaa ihmisessä ns. uncanny valley -ilmiön. Uncanny valley -ilmiössä ihmisen myönteinen ja empaattinen suhtautuminen robottia kohtaan kasvaa, kun robotin inhimillisiä piirteitä ja käyttäytymistä lisätään, mutta jossain vaiheessa saavutetaan raja jonka jälkeen ihminen saattaa kokea robotin todella inhottavaksi. Robotin inhimillisiä piirteitä voidaan kuitenkin vielä kehittää, jolloin ihminen saattaa käyttäytyä robottia kohtaan samalla lailla kuin käyttäytyisi ihmistä kohtaan [3, 38].

Myös robotin käytöksellä on vaikutusta ihmisen ja robotin välisessä kanssakäynnissä. Ihmismäisillä eleillä kuten nyökkäilyllä tai käsillä tehtävillä liikkeillä on myönteinen vaikutus ihmiseen vuorovaikutustilanteessa [39]. Wiesen ja

Abubshaitin tutkimuksessa todetaan, että robotin käytöksellä on suurempi vaikutus siihen, että miten ihminen seuraa robotin katsetta kuin ulkonäöllä. Kun robotti käyttäytyy ennakoitavasti ja luotettavasti niin ihminen seuraa robotin katsetta tarkemmin, jolloin robotti voi kiinnittää ihmisen huomion asiaan, jota se katsoo [37].

Katsekontaktilla on myös merkittävä rooli luodessa ihmisen ja robotin välistä vuorovaikutusta. Robotin katseen on tutkittu muuttavan ihmisten käyttäytymistä ja suhtautumista robottiin vuorovaikutustilanteessa [40, 37]. Katsekontaktin avulla ihminen kokee olevansa osallinen keskustelua, mikä johtaa vahvempaan vuorovaikutukseen. Ihmiselle on siis hyvä luoda tunne, että robotti katsoo häntä.

Sosiaalisen robotin pitää osata tunnistaa ihmisen tunteet ja persoona, jotta se voi käyttäytyä oikein erilaisissa tilanteissa. Rossin ja kumppaneiden tutkimuksessa todetaan, että ihmisen persoona ja ihmisen tekemä aktiviteetti vaikuttavat siihen että millä etäisyydellä ihminen haluaa robotin pysyvän [41]. Tunteiden käsittely on tärkeä osa sosiaalisen robotin toimintaa. Robotin toimiva tunteiden käsittely mahdollistaa luonnollisemman ja ymmärrettävämmän kommunikaation [42]. Vastapuolen tunteiden perusteella voidaan esimerkiksi määrittää, mitä robotin kannattaa tehdä seuraavaksi.

Vaikka sosiaalisille roboteille on yhä kasvava tarve esimerkiksi terveydenhuollossa ja vanhustenhoidossa, ei niiden käyttö ole kuitenkaan yleistynyt [43, 44]. Euroopan komission rahoittamassa mielipidetutkimuksessa jopa 61 prosenttia vastaajista ei usko siihen että robotti pystyisi tekemään heidän töitään tulevaisuudessa [45]. Nämä asiat johtuvat todennäköisesti siitä, että robottien konenäössä, päätöksenteossa ja ihmisten kanssa kommunikoinnissa on vielä paljon kehitettävää [46].

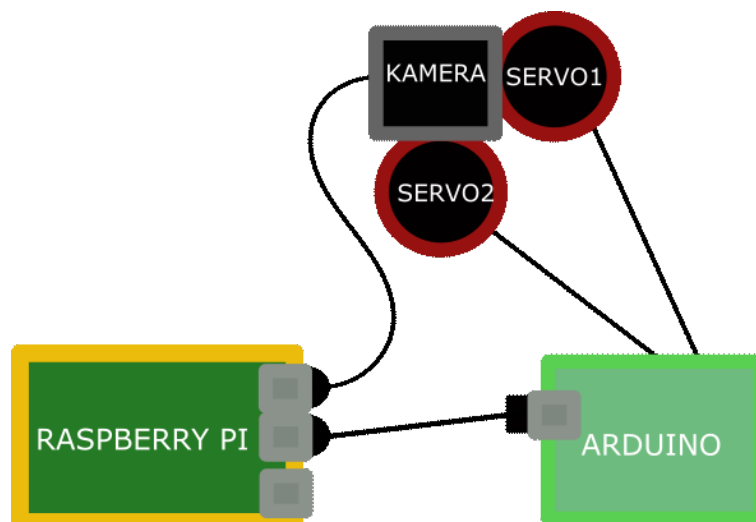
3. TOTEUTUS

InMoov-Robotin tulee kyetä katsomaan sen näkökentässä olevia henkilöitä silmiin. Tärkeää on myös välissä katsoa sivuun vuorovaikutuksen kohteista, sillä useimmat ihmiset voivat pitää pitkittynyttä katsekontaktia uhkaavana. Jotta katseesta tulisi mahdollisimman ihmismäinen, on sen noudatettava kappaleessa 2.1.2 käsiteltyjä silmän liikkeen määritelmiä. Robotin silmän tulisi kyetä keskittämään katse nopeasti katsottavaan henkilöön ja katseen ollessa fiksoitunut, tulisi silmässä esiintyä pientä liikettä. Lisäksi robotin pään asennon muutokset eivät saa muuttaa katseen kohdistamista.

Tässä työssä saavutettiin näistä vaatimuksista osa-alue, jossa robotti keskittää katseensa näkökentässään olevan ihmisen kasvoihin ja pitää sen siinä pään liikkeistä huolimatta. Toteutus tehtiin luomalla alusta, joka tunnistaa kameralla kasvot ja ohjaa robotin katseen niitä kohti reaaliajassa. Robotin silmässä sijaitsevan USB-kameran videostriimi avataan Raspberry Pi:llä ja OpenCV-kirjaston mahdollistamalla kasvontunnistuksella saadaan lista kuvassa olevien kasvojen keskikohtien koordinaateista. Koordinaattilista lähetetään sarjayhteyden kautta Arduinolle. Arduinolle kirjoitettu koodi parsii koordinaattitiedot lähetetystä viestistä ja ohjaa robotin silmiin kytkettyjä servoja niiden perusteella.

3.1. Laitteisto

Työn laitteisto koostuu Raspberry Pi 3 Model B:stä, siihen USB:lla kytketyistä Microsoft LifeCam HD-3000 web-kamerasta ja Arduino UNO:sta. Arduinon ulostuloon on lisäksi kiinnitetty kaksi servoa, jotka kykenevät kääntämään kameraa x- ja y-tasoissa. Näistä Raspberry Pi 3 Model B olisi helposti korvattavissa millä tahansa Debian- pohjaisella Linux PC:llä, jos esiintyy tarvetta suuremmalle laskentateholle.



Kuva 3. Kokonaiskuva laitteistosta

3.1.1. Kamera

Kamerana toimii Microsoft LifeCam HD-3000 web-kamera, koska se on yhteensopiva käytetyn InMoov-robotin silmäkenteen kanssa. Taulukko 1 osoittaa kameran tuetut resoluutiot, kuvasuhteet ja suurimmat kuvataajuudet.

Taulukko 1. Kameran tukemat resoluutiot ja kuvataajuudet.

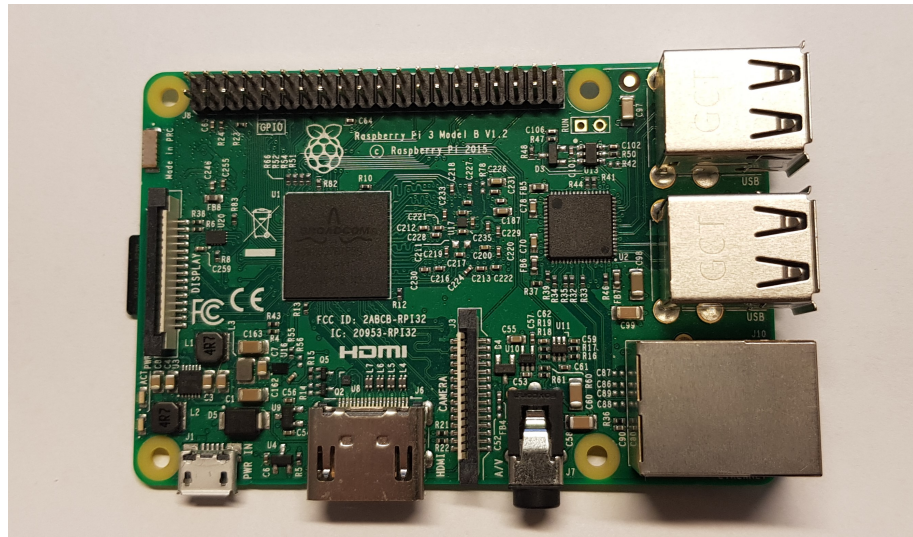
Resoluutio	Kuvasuhde	Kuvataajuus
1280x800	16:10	10fps
1280x720	16:9	30fps
960x544	16:9	30fps
800x600	4:3	30fps
800x448	16:9	30fps
640x480	4:3	30fps
640x360	16:9	30fps
424x240	-	30fps
352x288		30fps
320x240	4:3	30fps
176x144	-	30fps
160x120	4:3	30fps



Kuva 4. Microsoft LifeCam HD-3000 web-kamera alkuperäisessä kuoressaan, purettuna ja asennettuna InMoov-robotin silmään.

3.1.2. Raspberry Pi 3 Model B

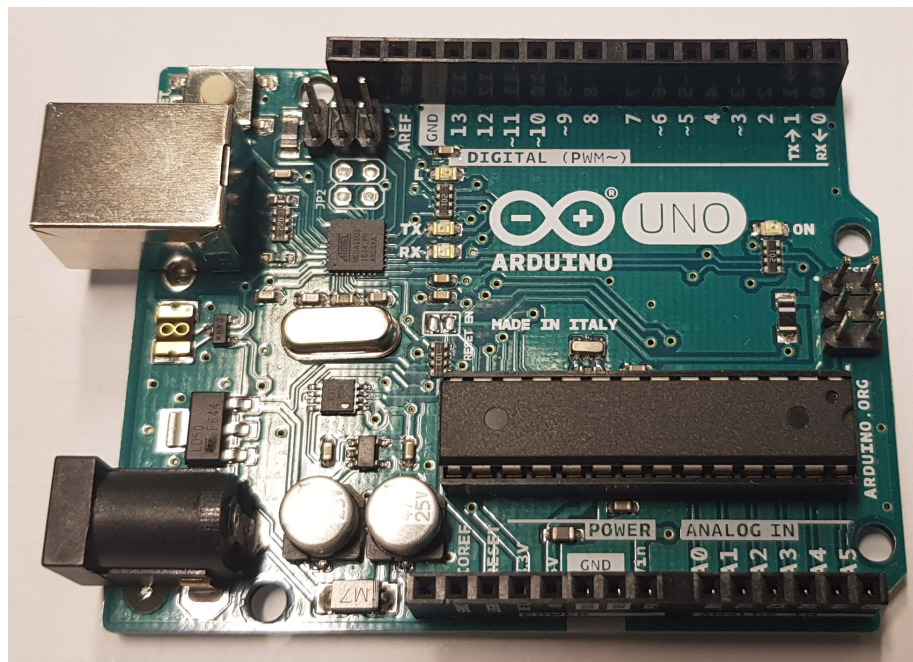
Raspberry Pi 3 Model B on yhden piirilevyn tietokone, joka on Raspberry Pi sarjan tietokoneiden neljäs iteraatio. Tietokoneessa on 64-bittinen Broadcom BCM2837 1.2 GHz (gigahertsiä) neliydin prosessori ja sen voi kytkeä 2.4 GHz ja 5 GHz WiFi-verkkoihin. Laite on mahdollista bootata SD-kortilta, USB-muistitikulta tai verkosta ethernet kaapelin kautta. Neljä USB-porttia mahdollistavat USB-laitteiden liittämisen ja muita laitteita voidaan liittää GPIO, HDMI, DSI, CSI ja Ethernet-liitinten avulla.



Kuva 5. Raspberry Pi 3 Model B V1.2

3.1.3. Arduino UNO

Arduino Uno rev3 on mikrokontrolleripiirilevy, joka pohjautuu ATmega328P 8-bittiseen AVR mikrokontrolleriin. Mikrokontrollerissa on 32 kB (kilotavua) ohjelmitavaa flash-muistia ja se toimii 16 Mhz (megahertsiä) kelloaajuudella. Arduino Unossa on 14 digitaalista I/O (Input/Output) pinniä, joista kuusi voi tuottaa pulssinleveysmoduloitua (PWM) signaalia, sekä kuusi analogista I/O pinniä. Laitetta voi ohjelmoida Arduino IDE:llä (Integrated Development Environment) USB-kaapelin kautta.



Kuva 6. Arduino UNO

3.1.4. Servot

Silmien liikuttamiseen käytettiin Dynamixelin XL320 -servomootoreita. XL320-servoja ohjataan TTL (Time To Live)-kommunikaatioprotokollaa käyttäen, mikä mahdollistaa servojen ketjutuskytken ja erillisen ohjauksen servoille asetetun tunnistearvon mukaan. Servojen suositeltu käyttöjännite on 7.4V (Voltia) ja niiden maksimivääntömomentti on 0.39 Nm (newtonmetriä).

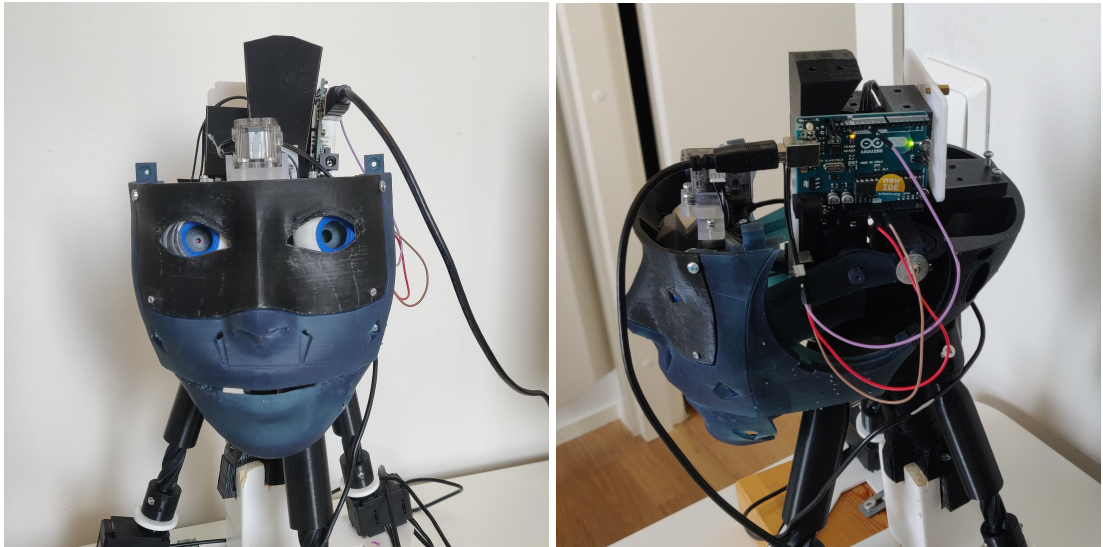


Kuva 7. Dynamixel XL320-servomootorit

3.1.5. InMoov-robotti

InMoov [47] on 3D-tulostettava luonnollisen kokoinen humanoidirobotti, jonka on kehittänyt ranskalainen kuvanveistäjä ja suunnittelija Gael Langevin. InMoov on täysin avoimen lähdekoodin projekti, ja osien 3D-tulostettavat mallit on vapaasti saatavilla Creative Commons lisenssin (CC-BY-NC) alla. Robotti on suunniteltu kontrolloitavaksi Arduino-mikrokontrollereilla.

InMoov-robotti kattaa koko kehon päästä jalkoihin, mutta projektissa käytössä oli pelkkä pää ja kaula. Päähen kuuluu kuusi servo-moottoria: kaksi silmien liikuttukseen, yksi pään kääntämiseen ja kolme niskan liikuttukseen. Näistä vain silmien liikuttukseen käytettäviä servoja käytettiin. Päässä on lisäksi Arduino Uno, sekä valmis paikka Raspberry Pi:lle.



Kuva 8. Työssä käytetty InMoov-robotin pää.

3.2. Kasvontunnistus

Kasvontunnistukseen käytettiin avoimen lähdekoodin OpenCV [48]- kirjaston AdaBoost [49]-kasvontunnistusluokittelijaa. Alustana toimi Raspberry Pi ja ohjelmointikielenä Python3. Kasvontunnistusta havainnollistetaan kuvassa 9. Tunnistus tapahtuu ohjelman silmukassa seuraavalla tavalla:

1. Ohjelmalle annetaan sen käynnistyessä käytettävä resoluutio ja luokittelija komentoriviargumentteina.
2. Silmukkaan siirryttäessä otetaan kuva videostriimistä, joka on aloitettu ohjelman käynnistyessä.
3. Muunnetaan otettu kuva mustavalkoiseksi, jotta se soveltuu luokittelijan käsiteltäväksi.
4. Luokittelija tunnistaa mustavalkokuvasta kasvat ja palauttaa kasvojen keskikohtien koordinaatit listana. Mikäli kasvoja ei löydy, palautetaan tyhjä lista.
5. Koordinaattien perusteella voidaan piirtää otetun kuvan päälle neliö, jotta nähdään onko tunnistus onnistunut. Tämä toiminnallisuus ei kuitenkaan ole ohjelman normaalissa suorituksessa käytössä sen aiheuttaman viiveen vuoksi.
6. Koordinaattilistasta muodostetaan viesti, joka lähetetään sarjayhteyden välityksellä arduinolle. Arduinon vastatessa viestiin siirrytään takaisin silmukan alkuun. Jos arduino ei vastaa viestiin, siirrytään silmukan alkuun 1 sekunnin jälkeen.



Kuva 9. Esimerkki kasvontunnistuksen toteutuksesta, kun kasvot ovat 1 metrin päässä kamerasta. Ylimmässä alkuperäinen kuva. Keskimmäisenä kuva mustavalkomuunnoksen jälkeen ja alimpana kuva, josta kasvontunnistajan antamiin koordinaatteihin on piirretty sininen neliö.

3.2.1. Kameran asetusten valinta

Kamerana käytettiin robotin silmään sijoitettua Microsoft LifeCam HD-3000 web-kameraa, jonka testatut kuvaformaattit olivat 1280x800, 1280x720, 640x480, 320x240 ja 160x120. Nämä kaikki resoluutiot nopeudella 10 fps. Kameran kuvakulma on sivusuunnassa 60 astetta.

Alustavassa testauksessa tutkittiin, millä resoluutioilla ja asetuksilla kasvontunnistus on mahdollista 1-2 metrin päässä kamerasta optimaalisissa olosuhteissa. Optimaalisilla olosuhteilla tarkoitetaan tässä tapauksessa tilannetta, jossa kamera näkee yhden kasvot, jotka ovat keskellä kuvaa, selkeää taustaa vasten ja runsaassa valossa. Alustavassa testauksessa kokeiltiin kasvontunnistusta eri kameran asetuksilla ja havaittiin automaattisen valotusajan vaikuttavan negatiivisesti kuvan laatuun, joten valotusaika asetettiin manuaalisesti kymmeneen millisekuntiin.

3.2.2. Luokittelijan testaus

Kasvojentunnistusta varten harkittiin kolmea luokittelijavaihtoehtoa: Haar cascade, LPB (Local Binary Patterns) ja LPBi (LPB improved). Luokittelijan valinta tehtiin testauksen perusteella. Taulukon 2 kasvot sisältävissä testeissä yksi henkilö istui testausetäisyydellä tyhjää valkoista seinää vasten, niin että kasvot olivat suoraan kameraa päin. Muissa testeissä ei kameran edessä ollut kasvoja. Testauksessa ei huomioitu, olivatko löydetty kasvot oikealla kohtaa. Huomioon otettiin vain löytyikö kuvatilanteesta kasvot vai ei. Kasvontunnistuksen todettiin kuitenkin myöhemmässä testauksessa löytävän kasvot oikealta kohtaa kuvan 9 osoittamalla tavalla.

Testattaviksi kameran resoluutioksi valittiin 1280x800 ja 640x480, koska tätä pienemmillä resoluutioilla tehdyllä kasvontunnistuksella ei alustavassa testauksessa saavutettu merkittäviä nopeusparannuksia tarkkuuden kärsiessä huomattavasti. Taulukossa 2 esitetyt mittaustulokset saatiin suorittamalla jokaiselle resoluutio, luokittelija ja kuvatilanne -yhdistelmällä 100 koordinaattien lähetystä ennen keskimääräisten arvojen tallennusta. Lähetysten välinen aika kertoo ajan kasvokoordinaattien lähetysten välillä.

Kuvan resoluution ollessa 1280x800 on kaikkien luokittelijoiden lähetysten välinen aika välillä 1,3 - 2,1 sekuntia. Koska robotin halutaan katsovan sen näkökenttään ilmestyvää ihmistä silmiin mahdollisimman nopealla reaktioajalla, voitiin resoluutiolla 1280x800 toimiva tunnistus sulkea pois liian hitaana. Resoluutiolla 640x480 lähetysten välinen aika tippuu välille 0,44 - 1,23 sekuntia.

Haar cascade löysi kasvot sadan prosentin tarkkuudella molemmilla testatuilla resoluutioilla neljään metriin saakka. Sen voitiinkin todeta olevan luokittelijoista tarkin kasvontunnistuksessa, mutta myöskin noin 0.5 sekuntia hitaampi kuin LBP tai LPBi. Sekä LPB:n, että LPBi:n tunnistuskyky heikkeni neljän metrin etäisyydellä huomattavasti. Kun kasvoja tunnistettiin sotkuista kasvotonta taustaa vasten pienemmällä resoluutiolla, esiintyi testatuista kuvista vääriä positiivisia kasvontunnistuksia Haar cascade -luokittelijalla 5% ja LPB-luokittelijalla 14%. LPBi-luokittelijalla vääriä positiivisiä ei ilmentynyt.

Testauksen perusteella valittiin luokittelijaksi LPBi ja resoluutioksi 640x480. Perusteina olivat toteutuksen vaatima nopeus ja LPBi-luokittelijan vähäinen väärien positiivisten kasvontunnistusten määrä.

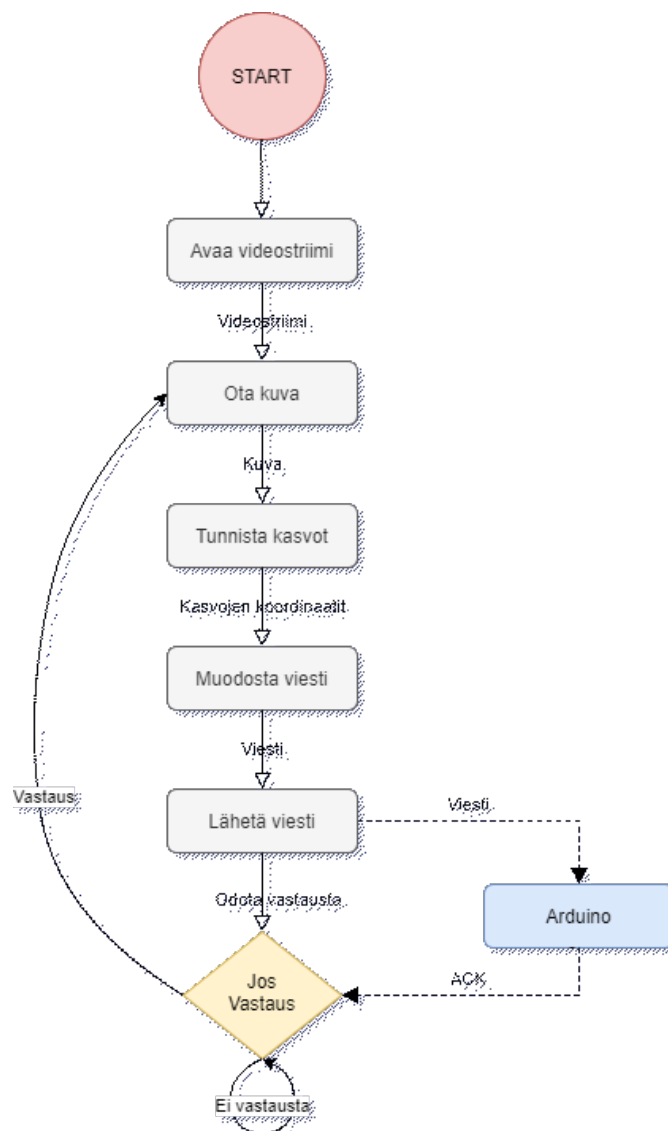
Taulukko 2. Kasvontunnistuksen suorituskyky eri luokittelijoilla ja resoluutioilla.

Ei kasvoja kameran edessä: Tyhjä seinä			
Luokittelija	Resoluutio	Aika lähetysten välillä	Väärät positiiviset kasvontunnistukset %
Haar	1280x800	1.81 s	0%
	640x480	0.79 s	0%
LBP	1280x800	1.57 s	0%
	640x480	0.51 s	0%
LBP improved	1280x800	1.31 s	0%
	640x480	0.44 s	0%
Ei kasvoja kameran edessä: Sotkuinen huone			
Luokittelija	Resoluutio	Aika lähetysten välillä	Väärät positiiviset kasvontunnistukset %
Haar	1280x800	2.88 s	0%
	640x480	1.23 s	5%
LBP	1280x800	1.59 s	40%
	640x480	0.54 s	14%
LBP improved	1280x800	1.71 s	0%
	640x480	0.57 s	0%
Etäisyys kasvoista: 1 metri			
Luokittelija	Resoluutio	Aika lähetysten välillä	Positiiviset kasvontunnistukset %
Haar	1280x800	1.96 s	100%
	640x480	0.94 s	100%
LBP	1280x800	1.62 s	100%
	640x480	0.57 s	100%
LBP improved	1280x800	1.52 s	100%
	640x480	0.52 s	100%
Etäisyys kasvoista: 2 metriä			
Luokittelija	Resoluutio	Aika lähetysten välillä	Positiiviset kasvontunnistukset %
Haar	1280x800	2.07 s	100%
	640x480	0.98 s	100%
LBP	1280x800	1.62 s	100%
	640x480	0.56 s	100%
LBP improved	1280x800	1.51 s	100%
	640x480	0.54 s	100%
Etäisyys kasvoista: 4 metriä			
Luokittelija	Resoluutio	Aika lähetysten välillä	Positiiviset kasvontunnistukset %
Haar	1280x800	2.07 s	100%
	640x480	0.96 s	100%
LBP	1280x800	1.57 s	100%
	640x480	0.55 s	41%
LBP improved	1280x800	1.52 s	100%
	640x480	0.52 s	26%

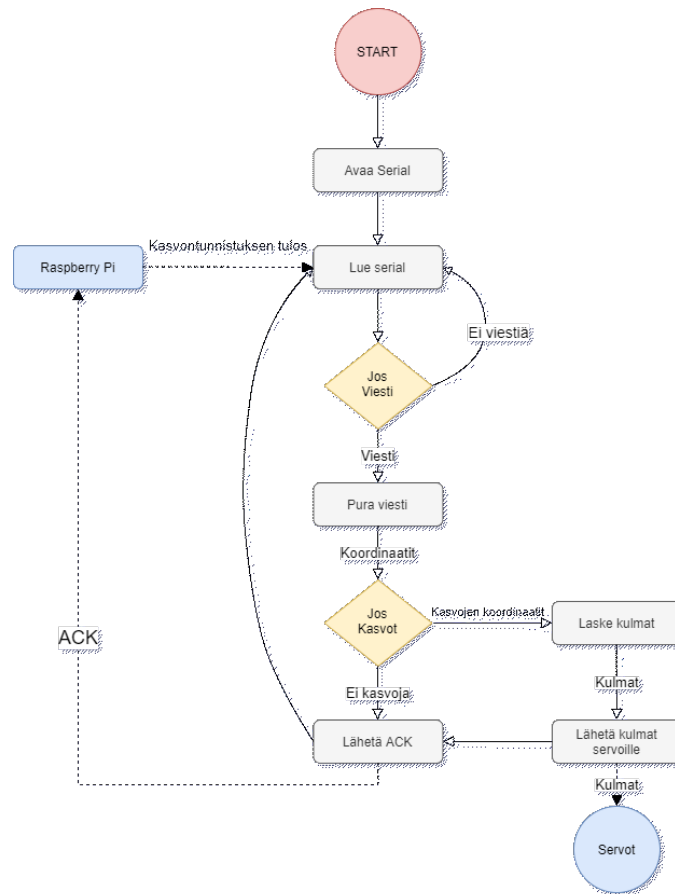
3.3. Sarjakommunikaatio ja ohjelmasilmukka

Kasvojen koordinaattitiedot siirtyvät USB-kaapelilla Raspberry Pi:ltä Arduinolle. Lähetystä varten luodaan listasta koordinaatteja merkkijono, joka sisältää käsitellyn kuvan resoluution ja tunnistettujen kasvojen koordinaatit. Merkkijono päätetään merkkeihin \n, joiden avulla vastaanottaja osaa lopettaa sarjayhteyden lukemisen oikeaan aikaan.

Arduino odottaa Raspberry Pi:ltä sarjayhteydellä lähetettyä viestiä. Viestin vastaanottamisen jälkeen resoluutio ja koordinaatit parsitaan merkkijonosta ja käytetään kulmien laskemiseen. Raspberry Pi odottaa, että Arduino lähettää vastauksen ennen kuin ottaa uuden kuvan. Mikäli vastausta ei saada sekunnin kuluessa, jatkaa ohjelma suoritusta normaalisti. Ohjelman suoritusta esitellään Raspberry Pi:llä kuvassa 10 ja Arduinolla kuvassa 11.



Kuva 10. Kaavio ohjelman suorituksesta Raspberry Pi:llä.



Kuva 11. Ohjelman suoritus arduinolla.

3.4. Servojen ohjaus

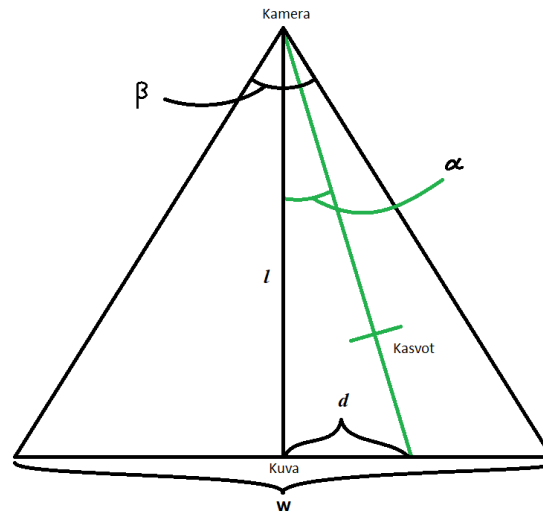
Arduinon vastaanottettua ja purettua Raspberry Pi:ltä saadut kasvojen koordinaattitiedot lasketaan ensimmäisten kasvojen tiedoista tarvittava silmien kulmien muutos, jotta silmät saadaan käännettyä kohti kasvoja. Kulmien laskuperiaate kuvaillaan kuvassa 12. Kameran kulmien ja kasvojen kulmien erot lasketaan yhtälön 1 mukaan:

$$\alpha = \arctan\left(\frac{d}{l}\right) \quad (1)$$

Yhtälössä d on kasvojen etäisyys akselilla kuvan keskikohdasta ja l on saadaan yhtälöstä 2:

$$l = \frac{w/2}{\tan(\beta/2)} \quad (2)$$

Yhtälössä w on kuvan resoluutio leveyssuunnassa ja β on kameran kuvakulma leveyssuunnassa (60°).



Kuva 12. Kulmien laskuperiaate.

Kameran ja kasvojen kulmien erot α lisätään kameran kulmiin, minkä jälkeen tuloksena saadut kulmat rajoitetaan etukäteen määritettyihin minimi- ja maksimiarvoihin. Kulmat muunnetaan servoille annettaviin arvoihin ja lähetetään GPIO-pinnistä serial-yhteydellä servoille. Servojen ohjaamiseen käytettiin Arduinolla hackerspace-adelaide XL320 [50] kirjastoa. Linjanopeutena käytettiin 115200 baudia.

Silmät kykenevät liikkumaan sivusuunnassa noin 30 astetta ja pystysuunnassa noin 20 astetta. Silmien erillistä liikuttelua harkittiin, mutta toteutuksen yksinkertaisttamiseksi ja suoritustehon puutteen vuoksi ajatus hylättiin. Erikseen liikkuvilla silmillä katseen kohdistus erityisesti lähietäisyydellä olisi luonnollisemman näköistä, sillä kumpikin silmä kohdistaisi kohti kasvoja.

4. POHDINTA

Työssä kehitettiin onnistuneesti toteutus, joka mahdollistaa robottipään katsekontaktin ihmisen kanssa. Rajoitetusta laskentatehosta johtuen kasvojen tunnistamisetäisyys jäi lyhyehköksi ja silmien siirtojen välinen aika on liian pitkä luonnollisten silmän liikkeiden toteuttamiseksi. Laskentatehon lisääminen, ohjelmiston optimointi ja erinäisten ominaisuuksien lisääminen toisivat helposti lisää luonnollisuutta robotille. Jatkokehityksessä toteutuksen parantaminen tähän mennessä toteutetun toiminnallisuuden pohjalle pitäisi olla helppoa.

4.1. Tulokset

Raspberry Pi 3 Model B V1.2:lla saatiin robotin kasvojentunnistuksen suorituskyvyksi noin 3 kertaa sekunnissa, sekä kasvojen löydyttyä servojen liikutukseen kului noin 0,3 sekuntia, laskien suorituskyvyn noin 1,7 kertaan sekunnissa. Katse seuraa kasvoja siis hieman tökkivästi, mikä on katsottavasta usein hieman häiritsevää. Tunnistaminen onnistuu hyvin noin 2-3 metrin päähän hyvässä valaistuksessa. Huonossa valaistuksessa ja silmän reunojen peittäessä ei robotti usein tunnista kasvoja. Erityisesti silmien ollessa kääntyneenä kohti ulkoreunaa voi silmän reuna peittää osaa kuvasta ja vaikeuttaa tunnistusta.

Luotu toteutus onnistuu luomaan tunteen katsekontaktista. Kuitenkin robotin tuijottava katse ja silmien pätkivä liike laskevat realistisen katsekontaktin tunnetta. Lisäksi toinen silmä ei aina osoita täysin samaan suuntaan, mikä vei myös uskottavuutta robotin katsekontaktilta. Kuitenkin robottia testatessa useimmat ihmiset kokivat robotin häiritseväksi tai pelottavaksi, mikä johtuu todennäköisesti uncanny valley -ilmiöstä.

4.2. Toiminnalliset parannukset

Lisäämällä toiseenkin silmään kamera sekä silmille erillinen ohjaus saataisiin parannettua katsekontaktin realistisuutta, sillä kumpikin silmä saataisiin itsenäisesti osoittamaan kohti kasvoja. Toteutuksessa toinen, kameraton silmä osoittaa usein hieman kieron, eikä siten luo tunnetta ihmismäisestä katsekontaktista. Kaksi kameraa vaatisi kuitenkin myös kaksinkertaisen kasvojentunnistuksen, jos haluttaisiin saada selville kasvojen etäisyys. Tämän johdosta myös tarvittava laskentateho tuplaantuisi. Yksi Raspberry Pi ei enää olisi riittävä. Erikseen liikkuvat silmät taas puolestaan vaatisivat silmien liikutusmekanismin uusimista sekä toista servomootoria toiselle silmälle. Pystysuunnan liikkeelle riittää yksi servomoottori, joka ohjaa kumpaakin silmää, sillä pystysuunnassa silmät eivät merkittävästi osoita eri suuntiin. Vaihtoehtoisesti olemassaolevaa mekanismia voitaisiin yrittää säätää siten, että toinen silmä ei osoittaisi kieron.

Kasvojentunnistuksen suorituskykyä voisi nostaa käyttämällä Raspberry Pi:n sijaan tehokkaampaa laitteistoa, sekä optimoimalla ohjelmistoa käyttämällä pythonin sijaan ohjelmointikielenä C:tä. Suorituskykyä nostamalla saataisiin parannettua katsekontaktin luonnollisuutta ja sulavuutta, sekä mahdollistettua toisen kameran

käyttö hyväksyttävillä nopeuksilla. Servojen liikutuksessa käytettyjä viiveitä voidaan myös optimoida. Raspberry Pi:n ja Arduinon välistä sarja-yhteyttä on mahdollista optimoida poistamalla vastauksen odottaminen ja sen sijaan lähettämällä jatkuvasti kasvodataa ja varmistamalla että Arduino käyttää vain viimeisintä, ajankohtaista dataa.

Useiden kasvojen käsittelyä varten robotille täytyisi luoda keino, jolla se yhdistää samat kasvot kuvien välillä. Kasvojen yhdistämistä varten voi esimerkiksi luoda koordinaattien perusteella toimivan logiikan. Toinen, raskaampi ja laskentatehoa vaativampi vaihtoehto olisi käyttää kasvot tunnistavaa luokittelijaa, joka kykenisi erottamaan kasvot toisistaan ja yhdistämään ne eri kuvien välillä.

Nykyisessä toteutuksessa robotin silmät eivät reagoi pään kääntämiseen tai niskan liikkeisiin, toisin kuin ihmisen silmät, jotka kääntyvät automaattisesti pään liikettä vastaan. Tätä käyttäytymistä voisi myös toteuttaa robotille olettaen, että niskan ja pään liikkeet saadaan Arduinolle. Päättä voisi myös kääntää, jos katseen kohde poistuu näköpiiristä.

Robotin tuijottavaa katsetta saisi luonnollisemmaksi suunnittelemalla katseenohjausalgoritmin, joka ohjaa robotin katsetta. Nykyisessä toteutuksessa robotti tuijottaa vain eteenpäin mikäli se ei havaitse kasvoja. Kehitetyssä versiossa robotti voisi tarkkailla ympäristöään eli katsella ympärilleen löytääkseen uudet kasvot mikäli se ei ole hetkeen havainnut kasvoja. Myös katsekontaktin aikana tapahtuvaan käyttäytymiseen voisi tehdä muutoksia. Ihmiset kokevat pitkäkestoisen tuijotuksen häiritseväksi, joten säätämällä robotti katsomaan välillä muualle saisi sen miellyttävyyttä parannettua. Useiden kasvojen käsittelyn voi myös ottaa algoritmista huomioon niin, että robotti päättää mitä kasvoja se haluaa katsoa.

Uncanny valley -ilmiön vuoksi useimmat ihmiset kokivat robotin pelottavaksi. Ilmiötä voisi lieventää muokkaamalla robotin ulkonäköä. Helpompi ja toimivampi lähestymistapa olisi muokata robotti vähemmän ihmismäiseksi, jolloin ilmiö ei olisi niin vahva. Robotin ulkonäköä voisi muuttaa esimerkiksi "lelumaistemaksi", jolloin se olisi ulkonäöltään miellyttävämpi. Toinen lähestymistapa olisi tehdä robotti niin ihmismäisen näköiseksi ja ihmismäisesti käyttäytyväksi, että uncanny valley -ilmiötä ei enää ilmenisi niin merkittävästi. Tämä olisi kuitenkin huomattavasti vaikeampi, sillä ihmismäisyyden riittävän onnistunut luominen robotille on hyvin vaikeaa toteuttaa.

4.3. Jatkokehitys

Projektissa luotu kasvojentunnistus ja katsekontaktisovellus luo pohjaa monenlaisille mahdollisuuksille. Vaikka uncanny valley -ilmiö ja tekniset haasteellisuudet vaikeuttavat humanoidirobottejen soveltamista arkielämässä, on kuitenkin ihmismäisille koneille paljon mahdollisia käyttötarkoituksia.

Ihmisten välinen kommunikointi on useimmiten molemminpuolista. Tästä johtuen jotta robotin käyttäytymisestä saisi ihmismäistä ja uncanny valley -ilmiön vaikutus hävytettyä, täytyisi sen pystyä reagoimaan ihmisten käyttäytymiseen. Tätä varten täytyy robotin pystyä lukemaan ihmisten ilmeitä, mielentiloja ja muita visuaalisia viestejä. Katseen ohjaaminen ja kasvojen tunnistaminen antaa hyvän pohjan kehitykselle erityisesti ilmeiden tulkintaa kohti.

5. AJANKÄYTTÖ

Työskentely tapahtui pääosin itsenäisesti erityisesti COVID-19 pandemian aiheuttamien rajoitusten vuoksi. Lisäksi projektia käytiin läpi usein etätapaamisissa. Taulukossa 3 on esitelty ryhmän työhön käyttämä aika. Ryhmän jäsenet käyttivät projektiin suurin piirtein yhtä paljon aikaa.

Taustatutkimukseen osallistuivat kaikki ryhmän jäsenet itsenäisesti. Tutkimuksen työmäärä jakautui tasaisesti kaikille jäsenille. Työn teknisessä toteutuksessa kasvontunnistuksesta vastasi Lahti ja servo-ohjauksesta Luukkonen. Saari avusti molemmissa tapauksissa.

Kaikki osallistuivat työn kirjoitukseen saman verran. Johdanto kirjoitettiin yhteistyönä ennen koronapandemian alkua. Tausta-osioon ihmisen ja robotin vuorovaikutuksesta kirjoitti Saari. Luukkonen kirjoitti tekoälystä ja konenäöstä. Lahti kirjoitti ihmisen näköjärjestelmästä, näöstä ja silmän liikkeistä. Lahti ja Saari kirjoittivat katsekontakti-osion yhdessä.

Toteutuksessa laitteisto-osion kirjoittivat kaikki yhdessä. Kasvontunnistuksesta, sekä luokittelijan testauksesta kirjoitti Lahti. Servojen ohjauksesta kirjoitti Luukkonen, sarjakommunikaatiosta ja ohjelmasilmuksista kirjoittivat Lahti ja Luukkonen yhdessä. Pohdinnasta kirjoittivat Saari ja Luukkonen yhteistyönä.

Taulukko 3

Ajankäyttö		
Nimi	Tunnit	Prosentit
Aleksi Lahti	194h	34%
Antti Luukkonen	188h	33%
Jouni Saari	185h	33%

Linkki GitHub-repositorioon, joka sisältää lähdekoodin: <https://github.com/AnatinttiL/Kandiproju>

6. YHTEENVETO

Tässä työssä esiteltiin ihmisten ja robotin vuorovaikutukseen liittyviä asioita yleisellä tasolla sekä ihmisten kanssa toimivien robottejen vaatimuksia ja haasteita. Haasteisiin kuuluu muun muassa niin sanottu uncanny valley -ilmiö, jossa osittain ihmismäinen robotti vaikuttaa ihmisestä pelottavalta ja häiritsevältä, jossa ihminen kokee osittain ihmismäisen robotin pelottavaksi ja häiritseväksi.

Työssä toteutettiin kasvojentunnistuksen avulla toimiva robotti, jonka tarkoituksena on luoda katsekontakti näkemäänsä ihmiseen. Toteutus koostui InMoov-robottipäädystä, sen silmiä ohjaavista servomooottoreista, Arduino Uno-mikrokontrollerista sekä Raspberry Pi-tietokoneesta. Kasvojentunnistus toteutettiin Raspberry Pi:llä ja Arduinoa käytettiin servojen kontrollointiin.

Testeissä robotin katsekontakti osoittautui toimivaksi rajoitetusta laskentatehosta huolimatta. Tehokkaampi laitteisto mahdollistaisi katsekontaktin kehittämisen luonnollisemmaksi ja sulavammaksi. Toteutusta voitaisiin kehittää myös optimoimalla ohjelmistoa sarjakommunikaation ja katsekontaktin osalta.

7. VIITTEET

- [1] Meeussen W., Wise M., Glaser S., Chitta S., McGann C., Mihelich P., Marder-Eppstein E., Muja M., Eruhimov V., Foote T., Hsu J., Rusu R.B., Marthi B., Bradski G., Konolige K., Gerkey B. & Berger E. (2010) Autonomous door opening and plugging in with a personal robot. Teoksessa: 2010 IEEE International Conference on Robotics and Automation, ss. 729–736.
- [2] He B., Xu S., Zhou Y. & Wang Z. (2018) Mobility properties analyses of a wall climbing hexapod robot. *Journal of Mechanical Science and Technology* 32, ss. 1333–1344. URL: <https://doi.org/10.1007/s12206-018-0237-2>.
- [3] Mathur M.B. & Reichling D.B. (2016) Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition* 146, ss. 22 – 32. URL: <http://www.sciencedirect.com/science/article/pii/S0010027715300640>.
- [4] Emery N. (2000) The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience I & Biobehavioral Reviews* 24, ss. 581 – 604. URL: <http://www.sciencedirect.com/science/article/pii/S0149763400000257>.
- [5] Admoni H. & Scassellati B. (2017) Social eye gaze in human-robot interaction: A review. *J. Hum.-Robot Interact.* 6, s. 25–63. URL: <https://doi.org/10.5898/JHRI.6.1.Admoni>.
- [6] Prashanth N, Karthik Ravi K M, Reddy P.K., Sunil Kumar H S & Majumdar J. (2017) Minimalist mechanical design and characterization of an android for human robot interaction. Teoksessa: 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), ss. 77–84.
- [7] Ihmisen fysiologia ja anatomia, WSOY, Helsinki. 18. uud. p. p.
- [8] Land M.F. & Fernald R.D. (1992) The evolution of eyes. *Annual Review of Neuroscience* 15, ss. 1–29. URL: <https://doi.org/10.1146/annurev.ne.15.030192.000245>, PMID: 1575438.
- [9] Zimmer C. (2012) Our strange, important, subconscious light detectors. *Discover Magazine*. Retrieved , ss. 02–18.
- [10] DEMER J.L. (2002) The orbital pulley system: A revolution in concepts of orbital anatomy. *Annals of the New York Academy of Sciences* 956, ss. 17–32. URL: <https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-6632.2002.tb02805.x>.
- [11] Richmond F., Johnston W., Baker R. & Steinbach M. (1984) Palisade endings in human extraocular muscles. *Investigative ophthalmology & visual science* 25, ss. 471–476.

- [12] Fischer B. & Ramsperger E. (1984) Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research* 57, ss. 191–195.
- [13] Krauzlis R.J. (2005) The control of voluntary eye movements: New perspectives. *The Neuroscientist* 11, ss. 124–137. URL: <https://doi.org/10.1177/1073858404271196>, PMID: 15746381.
- [14] Barr C.C., Schultheis L.W. & Robinson D.A. (1976) Voluntary, non-visual control of the human vestibulo-ocular reflex. *Acta Oto-Laryngologica* 81, ss. 365–375. URL: <https://doi.org/10.3109/00016487609107490>.
- [15] Rolfs M. (2009) Microsaccades: small steps on a long way. *Vision research* 49, ss. 2415–2441.
- [16] Casile A., Victor J.D. & Rucci M. (2019) Contrast sensitivity reveals an oculomotor strategy for temporally encoding space. *eLife* 8, s. e40924.
- [17] Russell S.J.k. (2016) *Artificial intelligence : a modern approach*. Prentice Hall Series in Artificial Intelligence, Pearson Education, Harlow, third edition. global edition p.
- [18] Buchanan B.G. (2005) A (very) brief history of artificial intelligence. *Ai Magazine* 26, ss. 53–53.
- [19] Turing A.M. (1950) Computing machinery and intelligence. *Mind* 59, ss. 433–460. URL: <http://www.jstor.org/stable/2251299>.
- [20] Russell S.J. & Norvig P. (2016) *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.
- [21] Mehrabian A. & Ferris S.R. (1967) Inference of attitudes from nonverbal communication in two channels. *Journal of consulting psychology* 31, s. 248.
- [22] Fontenot K.A. (2014), *Nonverbal communication and social cognition*.
- [23] Hogan K. (2003) *Can't get through: eight barriers to communication*. Pelican Publishing.
- [24] WAHLER H. (2012), *93% of all communication is nonverbal? correcting a common misconception of the mehrabian studies*, e-book.
- [25] Yaffe P. (2011) The 7% rule: fact, fiction, or misunderstanding. *Ubiquity* 2011, ss. 1–5.
- [26] Thompson D. & Filik R. (2016) Sarcasm in Written Communication: Emoticons are Efficient Markers of Intention. *Journal of Computer-Mediated Communication* 21, ss. 105–120. URL: <https://doi.org/10.1111/jcc4.12156>.
- [27] MCCORNACK S. (2019) *CHOICES CONNECTIONS: an Introduction to Communication*. BEDFORD BKS ST MARTIN'S, 2nd p., 138 s.

- [28] Davidhizar R. (1992) Interpersonal communication: a review of eye contact. *Infection Control & Hospital Epidemiology* 13, ss. 222–225.
- [29] Argyle M. & Dean J. (1965) Eye-contact, distance and affiliation. *Sociometry* , ss. 289–304.
- [30] Weiten W., Dunn D.S. & Hammer E.Y. (2014) *Psychology applied to modern life: Adjustment in the 21st century*. Cengage Learning.
- [31] Chambers S. (2003) Use of non-verbal communication skills to improve nursing care. *British journal of nursing* 12, ss. 874–878.
- [32] Odent M. (1986) Primal health. *Medicine* 329, ss. 1246–1253.
- [33] Albas D.C. & Albas C.A. (1989) Meaning in context: The impact of eye contact and perception of threat on proximity. *The Journal of social psychology* 129, ss. 525–531.
- [34] Senju A. & Johnson M.H. (2009) The eye contact effect: mechanisms and development. *Trends in cognitive sciences* 13, ss. 127–134.
- [35] Vignolo A., Powell H., McEllin L., Rea F., Sciutti A. & Michael J. (2019) An adaptive robot teacher boosts a human partner’s learning performance in joint action. *Teoksessa: 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, ss. 1–7.
- [36] Sciutti A., Mara M., Tagliasco V. & Sandini G. (2018) Humanizing human-robot interaction: On the importance of mutual understanding. *IEEE Technology and Society Magazine* 37, ss. 22–29.
- [37] Abubshait A. & Wiese E. (2017) You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human–robot interaction. *Frontiers in Psychology* 8, s. 1393. URL: <https://www.frontiersin.org/article/10.3389/fpsyg.2017.01393>.
- [38] Mori M. (2017) The uncanny valley: The original essay by masahiro mori. *IEEE Robots & .*
- [39] Carter E.J., Mistry M.N., Carr G.P.K., Kelly B.A. & Hodgins J.K. (2014) Playing catch with robots: Incorporating social gestures into physical interactions. *Teoksessa: The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, ss. 231–236.
- [40] Zhang Y., Beskow J. & Kjellström H. (2017) Look but don’t stare: Mutual gaze interaction in social robots. *Teoksessa: A. Kheddar, E. Yoshida, S.S. Ge, K. Suzuki, J.J. Cabibihan, F. Eyssel & H. He (toim.) Social Robotics, Springer International Publishing, Cham*, ss. 556–566.
- [41] Rossi S., Staffa M., Bove L., Capasso R. & Ercolano G. (2017) User’s personality and activity influence on hri comfortable distances. *Teoksessa: A. Kheddar, E. Yoshida, S.S. Ge, K. Suzuki, J.J. Cabibihan, F. Eyssel & H. He (toim.) Social Robotics, Springer International Publishing, Cham*, ss. 167–177.

- [42] Perez-Gaspar L.A., Caballero-Morales S.O. & Trujillo-Romero F. (2016) Multimodal emotion recognition with evolutionary computation for human-robot interaction. *Expert Systems with Applications* 66, ss. 42 – 61. URL: <http://www.sciencedirect.com/science/article/pii/S0957417416304468>.
- [43] Vänni K.J. & Salin S.E. (2017) A need for service robots among health care professionals in hospitals and housing services. Teoksessa: A. Kheddar, E. Yoshida, S.S. Ge, K. Suzuki, J.J. Cabibihan, F. Eyssel & H. He (toim.) *Social Robotics*, Springer International Publishing, Cham, ss. 178–187.
- [44] Kondo I. (2019) Frailty in an aging society and the applications of robots. *Japanese Journal of Comprehensive Rehabilitation Science* 10, ss. 47–49.
- [45] Special eurobarometer 427: Autonomous systems. http://data.europa.eu/euodp/en/data/dataset/S2018_82_4_427_ENG.
- [46] He W., Li Z. & Chen C.L.P. (2017) A survey of human-centered intelligent robots: issues and challenges. *IEEE/CAA Journal of Automatica Sinica* 4, ss. 602–609.
- [47] Inmoov. <http://inmoov.fr/>.
- [48] Opencv. <https://opencv.org/>.
- [49] Freund Y. & Schapire R.E. (1995) A decision-theoretic generalization of on-line learning and an application to boosting. Teoksessa: P. Vitányi (toim.) *Computational Learning Theory*, Springer Berlin Heidelberg, Berlin, Heidelberg, ss. 23–37.
- [50] Hackerspace adelaide xl320 servo library. <https://github.com/hackerspace-adelaide/XL320>.