

Niko Hyytiä

# Markkeriton liikkeenkaappaus

Tietotekniikan kandidaatintutkielma

21. toukokuuta 2018

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

**Tekijä:** Niko Hyytiä

**Yhteystiedot:** niko\_hyytia@hotmail.com

**Ohjaaja:** Antti-Jussi Lakanen

**Työn nimi:** Markkeriton liikkeenkaappaus

**Title in English:** Markerless motion capture

**Työ:** Kandidaatintutkielma

**Sivumäärä:** 21+0

**Tiivistelmä:** Tämä kirjallisuuskatsaus käsittelee markkeritonta liikkeenkaappaus-tekniikkaa ja vertaa sitä muihin perinteisiin liikkeenkaappaustekniikoihin. Markkeritonta liikkeenkaappausta on tutkittu jo paljon, mutta sitä ei ole pystytty täysin hyödyntämään osana arkipäivän teknologiaa. Siksi tässä katsauksessa yritetään selvittää tekijöitä, jotka mahdollistavat markkerittoman liikkeenkaappauksen. Katsauksesta käy ilmi, kuinka RGB-kuvaa ja infrapunasäteillä saatavaa syvyyskuvaa hyödynnetään, jotta saadaan kuvasta kolmiulotteista tietoa. Tähän on käytetty suurimmaksi osaksi Microsoftin Kinect-laitetta. Tämä tekniikka on vasta kehitysvaiheessa, mutta sillä on potentiaalia kehittyä tehokkaaksi osaksi jokapäiväistä teknologiaa mm. lääketieteessä ja koulutuksessa.

**Avainsanat:** markkeriton liikkeenkaappaus, rgb-d kamera, kinect, infrapuna

**Abstract:** This review explains markerless motion capture and compares it to other more traditional motion capture techniques. Lots of studies have been conducted in markerless motion capture but so far there haven't been found efficient way to use it in daily technology. That's why this review aims to explain how it is possible to capture data without any markers. The review explains how RGB-picture and depth image, captured by infrared lasers, can be used to extract three-dimensional data of the picture or video. Most of the studies have used Microsoft Kinect for their research. Markerless motion capture is still in development but it has great potential to become an efficient factor in our daily technology, in example on medical field and

training in various fields.

**Keywords:** markerless motion capture, rgb-d camera, kinect, infrared

## **Kuviot**

Kuvio 1. Infrapunakameran toiminta ensimmäisessä Microsoft Kinect -versiossa.

Lähde: Khoshelham & Elberink, (2012) ..... 6

## Sisältö

1	JOHDANTO .....	1
2	MARKKERITTOMAN LIIKKEENKAAPPAUKSEN TEORIA .....	3
	2.1 Liikkeenkaappausdatan tulkitseminen .....	3
	2.2 Infrapuna liikkeenkaappauksessa .....	5
	2.3 Virtuaalinen luuranko .....	6
3	MARKKERITTOMAN LIIKKEENKAAPPAUKSEN SOVELLUKSET.....	8
	3.1 Liikunta .....	8
	3.2 Lääketiede .....	9
	3.3 Koulutus ja opetus .....	11
4	POHDINTAA.....	13
	KIRJALLISUUTTA .....	15

# 1 Johdanto

Liikkeenkaappausta (engl. motion capture) on viime aikoina tutkittu enenevässä määrin, mutta vielä ei ole löytynyt parhaaksi ja tehokkaaksi koettua tapaa kaapata ja tulkita liikkeestä saatavaa dataa (Zhang, 2012). Liikettä voidaan kaapata kohteeseen asennetuilla sensoreilla, erilaisilla kiintopisteillä kuten markkereilla tai liikkeenkaappaukseen soveltuvilla kameroilla jolloin ei tarvitse liittää kuvattavaan kohteeseen erityisiä oheislaitteita. Kullakin näistä on hyvät ja huonot puolensa ja niiden soveltuvuus riippuu hyvin paljon siitä, mihin käyttötarkoitukseen liikedataa halutaan hyödyntää. Käyttötarkoituksia liikedatalle on paljon ja sitä voidaan soveltaa lähes kaikkeen, mihin liittyy kappaleen liikettä.

Ennen liikkeenkaappausteknologian yleistymistä, piirrosanimaatioiden tekeminen oli hyvin samankaltaista kuin nykyään liikkeenkaappausdatan luominen. Piirrosanimaatioissa piirrettiin käsin jokainen kuva, jotka kuvattiin ihmisen silmälle sopivalla kuvataajuudella ja saatiin aikaan tuhansista kuvista koostuva elokuva. Suurin osa kuvista oli muuten samanlaisia lukuunottamatta hahmojen liikkeitä tarkoittaen, että elokuvaa varten jouduttiin piirtämään todella suuria määriä kuvia. Tänä päivänä nämä tuhannet kuvat voidaan korvata kopioimalla aktorien liikkeitä liikkeenkaappauksen avulla ja viemällä ne tietokoneelle, jossa niiden päälle lisätään tekstuurit. Tämä on muuttanut olennaisesti nykypäivän animaatioteknologian prosessia ja luontitapaa. Samalla liikkeenkaappaus on löytänyt jalansijaa myös muilta aloilta, kuten kuntoutuksessa (Metcalf et al., 2013) ja opetuksessa (Alexiadis et al., 2011).

Tässä kirjallisuuskatsauksessa tarkastellaan markkerittoman liikkeenkaappauksen pääpiirteitä ja sitä, miksi se on hyvä tapa kaapata liikettä. Tutkielman tavoitteena on luoda selkeä kuva menetelmän teoriasta ja eri osasista, jotka mahdollistavat liikkeen reaaliaikaisen digitalisoinnin esimerkiksi tietokoneelle. Kyseessä on vasta nuori tekniikka, joka on alati tutkimuksen alaisena. Siksi käsitellään myös niitä tekijöitä, joiden vuoksi markkeriton liikkeenkaappaus ei vielä ole arkipäivää. Luvussa 2 käydään läpi menetelmän teoriaa ja esitellään teknisiä vaatimuksia joita markkeriton

liikkeenkaappaus vaatii sekä muita tärkeitä tekijöitä liikkeenkaappauksen onnistumiseksi. Aliluvussa 2.1 puhutaan yleisellä tasolla markkerittoman liikkeenkaappauksen mahdollistavasta teknologiasta, aliluvussa 2.2 katsotaan kuinka kuvasta saadaan syvyystietoa infrapunalla ja aliluvussa 2.3 muodostettavasta virtuaalisesta luurankomallista. Luvussa 3 käydään läpi eri tieteenaloja ja esimerkkejä, missä markkeritonta liikkeenkaappausta on pystytty hyödyntämään ja mihin nykypäivän aloille sillä olisi hyviä edellytyksiä edetä. Luku 4 sisältää yleistä pohdintaa aiheen pohjalta. Laitteiston osalta keskitytään pääasiassa käsittelemään Microsoft Kinectillä toteutettua markkeritonta liikkeenkaappausta. Se on ollut tutkimuksissa selkeästi käytetyin laite sillä sen ominaisuudet sekä hinta ovat olleet kaikkein sopivimmat tutkimustarkoitukseen. Vaikka tämän katsauksen tavoite ei olekaan esitellä Kinectiä, on silti tärkeää ymmärtää miten Kinect mahdollistaa markkerittoman liikkeenkaappauksen, sillä pääpiirteittäin markkeriton liikkeenkaappaus perustuu näihin asioihin. Tätä käsitellään enemmän luvussa 2.2.

## 2 Markkerittoman liikkeenkaappauksen teoria

Markkeriton liikkeenkaappaus perustuu ideaan, jossa halutaan kaapata liikedataa henkilöstä luonnollisessa tilassa. Erityisiä vaatteita tai muita oheislaitteita ei prosessiin tarvita. Markkeritonta liikkeenkaappausta suoritetaan yleensä RGB-D-kameroilla (Red-Green-Blue-Depth). RGB-D-kamera on kamerakokonaisuus, johon kuuluu kaksi kameraa: RGB-kamera sekä syvyyskamera. RGB-kamera kuvaa perinteistä värikuvaa ja syvyyskamera toimii sensorina, joka laskee kuvan pikselien etäisyyttä kameraan muodostaen pikselien etäisyyttä kuvaavan harmaasävytteisen kuvan. Kohteen syvyyskuvaamiseen on nykypäivänä olemassa useita erilaisia laitteita. Esimerkiksi vesistön syvyyttä mitataan kaikuluotaimella ja mm. nykypäivän henkilöautot on varustettu peruutustutkilla. Markkerittomassa liikkeenkaappauksessa käytetään hyväksi infrapunasäteitä. Syvyyskamerassa oleva projektori lähettää infrapunasäteitä kuvattavalle alueelle, jotka heijastuvat infrapunakameralle. Näiden tietojen avulla voidaan erotella aktori, eli kuvattava kohde, itsenäiseksi kuvaksi ja käsitellä sitä samalla kun tausta jätetään pois prosessoinnista. Saadusta datasta prosessoidaan aktorin ruumiinosien orientaatio ja luodaan virtuaalinen luurankomalli, joka voidaan loppuvaiheessa yhdistää RGB-kuvan aktoriin (da Motta et al., 2017). Täten pystytään erottamaan haluttu kohde ympäristöstä ja yksityiskohtaisesti seuraamaan kohdetta. Yleisimmin tutkimuksissa on käytetty Microsoftin Kinect-kameraa (Zhang, 2012), jossa on molemmat RGB- ja syvyyskamera. Vaihtoehtoisesti tutkimusta voidaan suorittaa myös laitteilla, joissa ei ole sisäänrakennettuna infrapunakameraa. Tällöin tulee olla ulkoinen infrapunakamera ja kamerat on asetettava tutkimuksen kannalta oikein, jotta saadaan oikeanlaista dataa.

### 2.1 Liikkeenkaappausdatan tulkitseminen

Microsoft Kinectin RGB-D-kameralla saadaan kolmenlaista hyödynnettävää dataa: RGB-kuva, syvyyskuva sekä algoritmin avulla laskettavissa oleva virtuaalinen luuranko aktorista (Shotton et al., 2011). RGB- ja syvyyskuva ovat molemmat matriisimuodossa. Syvyyskuvamatriisissa jokaiselle pikselille on annettu arvo 0 tai 1, joten



matriisi toimii eräänlaisena maskina RGB-kuvamatriisille. Tässä tapauksessa 0 tarkoittaa ettei kyseinen pikseli ole osa siluettia ja arvo 1 tarkoittaa pikselin kuuluvan siluettiin. Vastaavasti RGB-kuvamatriisissa on jokaiselle pikselille oma väriarvonsa. Laitteen kalibrointi on tärkeää, sillä RGB-kuvan ja syvyyskuvan pikselien tulee täsmätä. Edellä mainituille kahdelle matriisille voidaan suorittaa laskutoimitus jolla saadaan uusi matriisi, jossa on mukana vain ne pikselit joiden syvyysmatriisin arvo oli 1. (Shotton et al., 2011). Alla on esitetty kaava, jolla matriisit kerrotaan keskenään. Matriisi A on RGB-matriisi, matriisi B on syvyyskuvan matriisi ja matriisi Z on näiden matriisien tulo.

$$z_{i,j} = a_{i,j} * b_{i,j} | i, j \in Z | 1 \leq i \leq m | 1 \leq j \leq n \quad (2.1)$$

$$A_{m,n} = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} B_{m,n} = \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ b_{2,1} & \dots & b_{2,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} Z_{m,n} = \begin{pmatrix} z_{1,1} & \dots & z_{1,n} \\ z_{2,1} & \dots & z_{2,n} \\ \vdots & \ddots & \vdots \\ z_{m,1} & \dots & z_{m,n} \end{pmatrix}$$

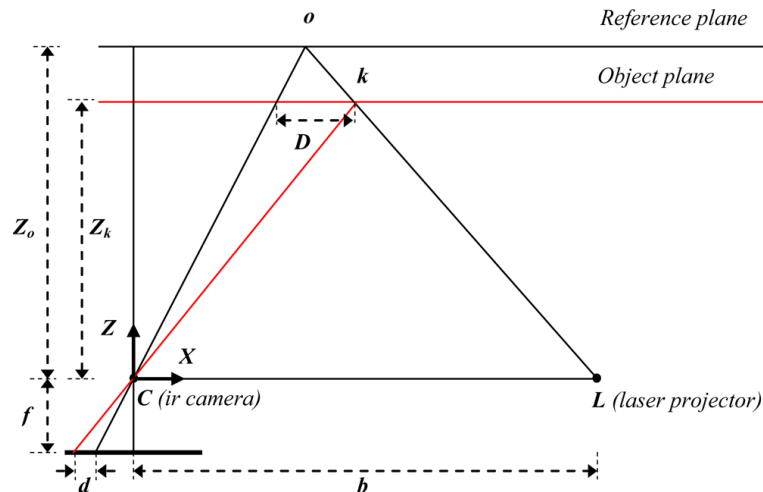
Esimerkiksi käytettäessä Microsoftin Kinectiä, videon jokainen kuva sisältää edellä mainitut matriisit. Nämä siirretään kameran omaan puskeriin yhtenäisenä pakettina ja siirrytään käsittelemään seuraavaa kuvaa. Kun videon kuvat on käsitelty ja kuva kerrallaan lisätty puskeriin, siirretään puskeriin tallennetut tiedot kovalevyille. Tämän jälkeen Kinectiä ei enää tarvita vaan halutulla tietokoneohjelmalla voidaan suorittaa haluttuja toimenpiteitä. Esimerkiksi Microsoft Kinectin mukana on mahdollista saada erillinen sovelluksien kehityspaketti, jolla voidaan luoda omia ohjelmia ja käydä dataa läpi vaatimusten mukaisesti. (Zhang, 2012). Kun liikkeenkaappauksen tiedot ovat kovalevyllä, voidaan aloittaa sen prosessointi. Puhuttaessa markkerittomasta liikkeenkaappauksesta, on tärkeää erottaa haluttu kohde taustaympäristöstään. Syvyyskuva luo siluetin aktorista, RGB-kuva liittää värit siluetin päälle ja tekstuurit aktorin taustalta voidaan poistaa matemaattisesti matriisiyhdistäällä. Virheiltä ei myöskään markkeriton liikkeenkaappaus välty. Joissakin tutki-

muksissa esiintyi virheitä kuvattavasta datasta esimerkiksi jos henkilö oli tekemisissä toisen objektin kanssa ja ruumiinosat peittyivät objektien tai muiden ruumiinosien taakse (da Motta et al., 2017).

## 2.2 Infrapuna liikkeenkaappauksessa

Luvussa 2.1 esitetyn matemaattisen kaavan matriisissa B on harmaasävytteinen kuva, joka on saatu syvyyskameralla. Esimerkiksi Kinectissä on syvyyskuvan luontia varten infrapunakamera ja infrapunaprojektori. Projektori lähettää infrapunasäteitä, jotka kulkevat projektorissa olevan hilan läpi ja jakautuvat kohdealueelle. Microsoftin vuonna 2010 julkaisemassa ensimmäisessä Kinect v1-mallissa säteet jakautuvat alueelle tietyllä kaavalla muodostaen nk. pistepilven, joka muodostuu infrapunapisteistä kuvattavan alueen pinnalla. Tätä kutsutaan ”structured light”-nimiseksi tavaksi, sillä se lähettää tietyn mallin mukaan säteitä kuvattavalle alueelle (Lau, 2013). Säteet muodostavat pisteitä osumiinsa pintoihin ja heijastuvat siitä infrapunakameraan. Infrapunakameralla tutkitaan pisteitä ja niiden naapureiden välisiä etäisyyseroja ja vertaillaan niiden suhteellisia eroja siihen malliin, jolla projektori on infrapunasäteitä lähettänyt. Tämän avulla voidaan laskea pisteiden välisiä syvyyseroja ja esittää harmaansävyisistä pikseleistä muodostuva syvyyskuva. Mitä tummempi sävy pikselillä on, sitä lähempänä kameraa kohde on. Joissakin tilanteissa on mahdollista ettei voida määrittellä syvyysväriä, esimerkiksi jos osa kuvasta on liian kaukana kamerasta tai jos pinta heijastaa huonosti infrapunasäteitä. Tällöin pikselit näkyvät täysin valkoisina tai mustina. Kuviossa 1 on esitetty kuinka infrapunaprojektori käytännössä toimii, mutta siihen ei tämän enempää tässä katsauksessa perehdytä.

Vuonna 2013 julkaistun Xbox One-pelikonsolin yhteydessä myytiin myös uutta versiota Kinectistä, jonka eräs muutoksista on sen infrapunakameran ja infrapunaprojektorin toiminta. Sen sijaan, että syvyysmittaus toimisi ”structured light”-periaatteella, käytetään uudessa Kinect-versiossa vapaasti suomennettuna lentoaikamenetelmää (engl. time of flight). Tällä menetelmällä lähetetään projektorista infrapunavälähdyksiä kohdealueelle tietyllä kellotaajuudella. Infrapunakamera jakaa jokai-



Kuvio 1. Infrapunakameran toiminta ensimmäisessä Microsoft Kinect -versiossa. Lähde: Khoshelham & Elberink, (2012)

sen pikselin kahteen osaan. Tietyin kellotaajuksin, se sammuttaa aina toisen puolen pikselistä, jolloin kyseinen pikselin puolikas ei voi vastaanottaa projektorista lähtevän valon fotoneja. Pikselien puolikkaat toimivat päinvastoin eli vain toinen on kerrallaan päällä ja toinen pikselin puolikas on poissa päältä. (Lau, 2013). Infrapunakamera voi tällä tavalla seurata kuinka nopeasti eri pikselien säteet liikkuvat verraten sitä projektorin kellotaajuuteen. Markkerittoman liikkeenkaappauksen kannalta on huomioitava muista valonlähteistä aiheutuvat infrapunäsäteet, jotka saattavat vaikuttaa kuvauksen lopputulokseen. Esimerkiksi auringon säteet voivat vääristää ulkona otettuja syvyysmittauksia.

### 2.3 Virtuaalinen luuranko

Kuten luvussa 2.1 mainittiin, osa liikkeenkaappauksista on henkilön tunnistaminen kuvasta ja hänen asentonsa. Vaikka tietäisimmekin vain RGB-datan ja syvyysdatan samasta kuvasta, voimme jo tällä tiedolla eristää kuvattava henkilö tausta. Tämä yksinään ei vielä riitä, jos halutaan esimerkiksi luoda virtuaalinen hahmo kuvan perusteella. Tarvitaan tietoa kuvassa olevan henkilön raajojen ja ruumiinosien asennoista ja orientaatioista. Tätä varten on tarpeellista luoda virtuaalinen luurankomalli. Ensimmäiseksi on tarpeellista selvittää kuvassa olevan henkilön ruumiin-

osat. Apuna tässä käytetään jo valmiita datajoukkoja, joihin saatua kuvaa voidaan verrata. Esimerkiksi Kinectiin kuuluu ohjelma, joka osaa luoda luurangon henkilöstä halutulla määrällä niveliä (engl. joint). Kinect v2 voi löytää 25 nivelkohtaa ihmisen ruumiista, mutta ohjelman käyttäjä voi valita kuinka monta tällaista nivelkohtaa on tarpeellista käyttää. (da Motta et al., 2017)

Saadulla ruumiinosien tiedolla voidaan luoda luurankomalli ja nähdä kuvattavan henkilön orientaatio digitaalisesti. Kun tiedetään ruumiinosien sijainnit, voidaan laskea kunkin nivelen keskipiste. Keskipisteisiin voidaan sijoittaa pallot kuvaamaan luurankomallin niveliä. Tällä tavalla saadaan vain kaksiulotteinen koordinaatti jokaiselle nivelelle. Tarvitaan vielä kolmas koordinaatti  $z$ , jotta sijainti voidaan antaa tietokoneelle kolmiulotteisena ja aktorista saadaan jotain mielenkiintoista dataa käytettäväksi. Syvyyskoordinaatti saadaan kaavalla, jossa verrataan haluttua nivelen sijaintia johonkin toiseen niveleen, joka jo tunnetaan. Usein tällaisena nivelenä käytetään lantiota, joka määritellään ns. perusnivelenä. Aiemmin kuvasta etsittyjen ruumiinosien ja nivelten perusteella, voidaan laskea perusnivelen  $x$ - ja  $y$ -koordinaatit kameran koordinaatistossa. Kinect laskee kuvasta selkärangan alaosan avulla  $z$ -koordinaatin. Täten saadaan kaksi koordinaattia  $z$ :lle, joista valitaan toinen riippuen henkilön orientaatiosta. Saatua niveltä käytetään seuraavan nivelen haussa rekursiivisesti kunnes kaikki halutut nivelet on käyty läpi. Nivelet käydään läpi niin, että saadun nivelen ja laskettavan nivelen välillä on luu, jolloin edetään luurangon hierarkiassa järjestyksessä eteenpäin. (da Motta et al., 2017)

### 3 Markkerittoman liikkeenkaappauksen sovellukset

Markkerittomalle liikkeenkaappaukselle löytyy paljon hyötykäyttöä vaikka sitä ei ole vielä täysin pystytty hyödyntämään sen nuoren iän takia. On tärkeää saavuttaa riippumattomuus erilaisista markkereista ja sensoreista, jotta liikkeenkaappausta voitaisiin yleisesti hyödyntää monenlaisissa arkisissa käyttökohteissa. Tämän päivän markkerittomassa liikkeenkaappauksessa käytetään vielä hyvin paljon visuaalisesti esimerkiksi tietokoneen näytöltä saatavaa palautetta, jota yritetään yhdistellä ja soveltaa markkerittomasta liikkeenkaappauksesta saatavan datan kanssa. Siksi ehkä hyödyllisin käyttökohde markkerittomalle liikkeenkaappaukselle on lajitekniikan opetus erilaisissa liikuntalajeissa tai muissa tekniikkaa vaativissa töissä (Xing et al., 2012). Sillä on myös edellytyksiä kehittyä esimerkiksi robotiikassa osaksi robotin liikkeentunnistusta hyvin samalla tavalla kuin ihmisen silmä toimii. Markkerittoman liikkeenkaappauksen onkin todettu olevan seuraava askel liikkeenkaappausteknologiassa (Field et. al., 2009).

#### 3.1 Liikunta

Alexiadis et al., (2011) tutkivat markkerittoman liikkeenkaappauksen käyttämistä tanssin etäopetuksessa. Kuvauslaitteena käytettiin Kinect-kameraa. Heidän kuvittelemassa tilanteessa jonkin tanssin opettaja luo videon, jossa opettaja esittelee tanssin tekniikkaa internetissä. Tanssiopetusta katsova henkilö voi valita mitä tanssin askeleita hän opettelee. Opettajan tanssia on kuvattu Kinect-kameralla ja myös oppilaan tanssia tulisi kuvata jollain RGB-D kameralla. Oppilas yrittää tehdä omat tanssiliikkeensä ja -askeleensa kuten opettaja. Suorituksen jälkeen oppilaan liike siirtyy digitaalisesti oppimisjärjestelmään, jossa tämän virtuaalista minä-hahmoa verrataan samanaikaisesti opettajan hahmoon. Järjestelmä laskee pisteitä sen perusteella, miten oppilas on onnistunut jäljittelemään opettajan liikkeitä. Tutkimuksessa käytettiin OpenNI-ajureita, joilla saatiin muodostettua kohdetta seuraava luurankomalli 17:llä virtuaalinivelellä. Pisteitä laskettiin seuraamalla nivelten liikenopeutta ja sijaintia. Liikkeenkaappauksella saadut tulokset korreloivat hyvin vertailupisteisiin

nähdän eikä suuria virheitä tapahtunut ellei laitteen kalibroinnissa ollut ongelmia tai käyttäjä katosi ruudulta. Tutkimus osoitti, että markkeritomalla liikkeenkaappauksella saatava data voi antaa riittävän hyvää dataa, jotta vastaavanlaisia toimintoja voidaan internetin kautta tarjota tehokkaasti varsinkin opetettaessa lajitekniikkaa.

## 3.2 Lääketiede

Liikunnan ja urheilun lisäksi markkeritonta liikkeenkaappausta pystytään hyödyntämään lääketieteessä ja potilaan kuntouttamisessa. Metcalf et al., (2013) ottivat asian esille tutkimuksessaan, jossa he tutkivat markkeritonta liikkeenkaappausta osana potilaan käden kuntouttamisprosessia. Tutkimuksessaan he asettivat yhden Kinect-kameran kuvaamaan pöytää ylhäältä kohti tutkittavan henkilön kättä, joka on asetettu pöydän päälle. Tutkimuksessa käsi pidettiin joko sormet levitettynä pöydän päällä tai muodostaen pihtiote peukalolla ja sormenpäillä. Ensimmäiseksi kameran tulee tunnistaa käden tärkeimmät kohdat, jotta voidaan ymmärtää käden asento. Kun käsi ja sormet on tunnistettu, voidaan algoritmilla seurata sormien asentoa ja kulmia.

Johtopäätöksenä tutkimuksessa todettiin, että markkerittomalla liikkeenkaappauksella lääkärit pystyisivät antamaan etäyhteydellä kuntoutusohjeita ja -palautetta sen sijaan, että potilaan tulisi käydä kontrollikäynneillä lääkärin vastaanotolla. Tämänkaltaisella järjestelmällä kuntoutuksesta vastaava lääkäri saisi saman tiedon potilaan kuntoutuksen edistymisestä kuin kasvokkain. Samalla se toimisi potilaalle reaaliaikaisena palautteena liikkeestä ja potilas osaa mahdollisesti itse korjata liikkeitään hoidonmukaisiksi. Markkerittomaan liikkeenkaappaukseen perustuva potilaan seuranta olisi myös käytännöllisistä syistä paljon tehokkaampi, sillä se ei vaatisi potilaalta muita kuin itse kamerasta aiheutuvia kustannuksia ja kuvaustilaksi kelpaisi mikä tahansa luonnollinen ympäristö. (Metcalf et al., 2013)

Tutkimuksessaan Li et al., (2018) tutkivat markkerittoman liikkeenkaappauksen ja katseenseurantalaitteen (engl. eye tracker) samanaikaista käyttöä dyspraksian eli

motoristen toimintojen koordinaatiohäiriön toteamisessa. Koordinaatiohäiriö vaikeuttaa henkilön motoristen kykyjen oppimista ja niiden hallitsemista. Lääkäri tai alan ammattilainen voi todeta häiriön, mutta se vaatii tarkkaa tutkittavan henkilön seuraamista, sillä tutkimusprosessi on melko lyhytkestoinen. Tutkijalta voi helposti jäädä huomaamatta esimerkiksi silmän liike ratkaisevalla hetkellä, mikäli hän keskittyy liikaa esimerkiksi käden liikkeisiin. Tätä varten tutkimuksen suorittaneet aloittivat aiheen tutkimisen, jotta saataisiin seurattua jokaista koordinaatiohäiriön kannalta tärkeää ruumiinosaa samanaikaisesti ja täten luotua mahdollinen yhteys, mikäli tällainen häiriö henkilöllä olisi olemassa. Tutkimus jaettiin kahteen osaan, joista ensimmäisessä seurattiin käsien ja silmien liikettä ja vertailtiin näitä. Toisessa osassa seurattiin ja vertailtiin koko vartalon ja silmien liikettä. Tutkimuksessa käytettiin kahta Microsoftin Kinect-kameraa, jotka kuvasivat kuvattavaa kohdetta eri suunnista sekä katseenseurantalaitetta. Näin pystyttiin saamaan dataa kuvattavan henkilön molemmilta puolilta ja välttämällä tärkeiden ruumiinosien peittyminen muiden ruumiinosien taakse.

Tutkimuksessa luotiin ohjelma, joka pystyi integroimaan henkilön virtuaalisen luurankon Kinectillä saatavaan RGB-kuvaan sekä näyttämään vihreällä merkillä henkilön katseen kohde. Tätä varten valittiin toinen kuvaavista Kinect-kameroista ”pääkameraksi”, jonka koordinaatistoon vihreä merkki vietiin. Tutkimusta varten suoritettiin kaksi testiä, joista ensimmäisessä asetettiin nappuloita oikeisiin reikiin. Kyseisessä testissä keskityttiin vain silmien ja käsien koordinaation seurantaan. Toisessa testissä seurattiin koko ruumiin liikettä, kun kuvattava henkilö yritti heittää hernepussia tietylle alustalle. Tutkimuksessa pystyttiin erottamaan jokaisesta videon kuvasta henkilön luurankomalli ja tarkastelemaan suoritusta hyvin yksityiskohtaisella tasolla. Luodun ohjelman ja datan analysointimenetelmän koettiin olevan hyvä aloitus markkerittoman liikkeenkaappauksen hyödyntämiseen käsi-ruumis-koordinaation ja siihen liittyvien sairauksien tai häiriöiden tutkimisessa. (Li et al., 2018)

### 3.3 Koulutus ja opetus

Tämän hetkiselällä tekniikalla ja tiedolla, markkeritonta liikkeenkaappausta hyödynnetään parhaiten koulutustarkoituksessa. Edellä mainituissa tutkimustapauksissa eräs ominaisuus on ollut samanaikaisesti saatava palaute kuvattavan henkilön toiminnasta. Vaikka koulutustarkoitukseen voidaankin käyttää esimerkiksi kameralista älypuhelinta, ei sillä saatavista kuvista ja videoista voida erotella muuta tietoa kuin mitä ihmissilmä havaitsee. Sen sijaan markkerittomalla liikkeenkaappauksella on pyritty esittämään tarkkaa tietoa koulutukseen liittyvän harjoituksen tuloksesta. Markkerittoman liikkeenkaappauksen hyödyntämistä onkin tutkittu mm. ensiapuun liittyvän koulutuksen apuna (Loconsole et al., 2016) sekä pianonsoitossa (Payeur et al., 2014).

Tutkimuksessaan Loconsole et al., (2016) loivat ohjelman, joka kertoo ensiapua suorittavalle tämän paineluelvytyksen tahdin ja laadun. Vertailtavina laitteina käytettiin Microsoftin Kinect v1:tä ja Creative Sens3D:tä, joista molemmat ovat RGB-D-kameroita. Tässä käytettiin hyödyksi markkeritonta liikkeenkaappausta, joka keskittyi seuraamaan kahta aluetta potilaasta: potilaan keskivartaloa ja liikkeenkaappauslaitteella luotua ns. kiinnostusalueetta, jolla ensiapua tekevän kädet oletetaan olevan suorituksen aikana. Potilaan keskivartaloa seuraava alue seuraa painallusten aiheuttamaa alueen muutosta. Toisin sanoen seurataan kuinka potilaan rinta laskee ja nousee painalluselvytyksen aikana. Tämä mahdollistaa suorituksen seuraamisen ja jokainen painallus voidaan luokitella sen syvyyden ja tahdin mukaan. Toinen alue, nk. kiinnostusalue, sisältää potilaan kädet. Painalluselvytyksessä kädet ovat yhdessä ja tästä syystä se saattoi häiritä kameroiden tuloksia, joten tutkimuksessa päädyttiin seuraamaan elvytystä suorittavan rystysiä. Tutkimuksessa käytettiin apuna myös markkereita mutta vain tuomaan tarkkaa vertailutietoa, johon molemmat RGB-D-kameroiden saamat tiedot verrattiin, kun selvitettiin näiden tarkkuutta rintakehän ja käden sijainneissa. Tästä selvisi, että Kinect v1 on luotettavampi kuin Sens3D, sillä sen keskimääräinen sijaintivirhe käden seurannassa oli 4mm kun Sens3D:llä se oli 7mm. Etenkin harjoituskäyttöön 4mm:n eron on täysin hyväksyttävä. Tutkimuksen myötä varmistuttiin siitä, että RGB-D-kameralla voidaan



tässä käyttötapauksessa erottaa käsi ja sen asento sekä suorittaa haluttuja toimenpiteitä.

Markkeritonta liikkeenkaappausta on käytetty myös tutkimaan pianon soitossa tapahtuvaa liikettä. Tähän on syntynyt tarvetta, kun pianonsoiton opettajat ja alan asiantuntijat ovat todenneet ryhdin olevan tärkeä elementti soitossa. Heidän mielestään tätä olisi tärkeä tutkia, sillä huonolla ryhdillä voi olla vaikutus tahtiin ja puhtaaseen soittamiseen. Tutkimuksessa tutkimustilaan asennettiin Kinect-sensori, joka keskittyi soittajan ruumiin ryhtiin ja käsien liikkeeseen. Tutkimuksessa tutkittava koehenkilö soitti pianoa kolmella, liioitellusti, väärällä tavalla. Hän soitti vielä neljännen kerran normaalilla soittotavalla, jota käytettiin vertailupohjana (Payeur et al., 2014). Tutkimuksessa kävi ilmi, että Kinect pystyy seuraamaan liikettä ja luokittelemaan siitä saatavaa informaatiota hyvin vaikkakin toisen käden liikkeestä ei saatu paljoa tietoa sen jäädessä muun ruumiin taakse piiloon. Vaikka tutkimuksessa käy ilmi Kinectin hyödyllisyys liikkeen seuraamisessa ja luokittelussa, todettiin samalla että sillä pitäisi suorittaa vielä enemmän erilaisia testejä, joissa toistettaisiin liikkeitä erilaisin ruumiillisin variaatioin.

## 4 Pohdintaa

Markkeriton liikkeenkaappaus on herättänyt tällä vuosituhanalla erittäin paljon mielenkiintoa. Sen nuoresta iästä johtuen markkeritonta liikkeenkaappausta on päästy kunnolla tutkimaan vasta 2010-luvulla, kun kamera- ja sensoriteknologia ovat kehittyneet riittävälle tasolle. Myös markkerittoman liikkeenkaappauksen edullisuus ja helppokäyttöisyys tekee sen käytöstä vaivattomampaa varsinkin kuluttajien näkökulmasta. Koska liikkeenkaappausta voi suorittaa vain yhdellä RGB-D-kameralla, eikä se vaadi muita teknisiä edellytyksiä, voidaan markkeritonta liikkeenkaappausta suorittaa melkein missä vain. Tietyissä tilanteissa on kuitenkin syytä miettiä tarkasti haluttavan liiketiedon laatua. On esimerkiksi kannattavampaa käyttää liikettä seuraavia sensoreita tai markkereita ruumiinosien liikkeiden seuraamiseen sillä se takaa paljon yksityiskohtaisemmat ja tarkemmat tulokset. Joissakin tutkimuksissa onkin käytetty markkereihin pohjautuvaa liikkeenkaappausta vertailudatan luontiin, johon markkerittomalla liikkeenkaappausdatalla voidaan peilata tuloksien tarkkuutta.

Microsoftin Kinect on ollut osana monessa liikkeenkaappaustutkimuksessa. Sen valmistus on lakkautettu, mutta teknologian kehittyessä vastaavien kameroiden hinnat ovat laskeneet eikä liikkeenkaappauslaitteistoon investoiminen ole enää suuri este, johon vain hyvin varakkailta yrityksillä ja toimijoilla on resursseja päästä käsiksi. Markkereihin ja sensoreihin perustuvat liikkeenkaappaustekniikat vaativat kuitenkin oikeanlaisen valmistautumisen, tavarat sekä tilat, mikä tekee siitä paljon hankalampaa ja kalliimpaa kuin markkeriton liikkeenkaappaus. Siksi markkerittomalla liikkeenkaappauksella on suuri potentiaali kehittyä tehokkaammaksi menetelmäksi seurata ja kaapata liikettä kuin muut liikkeenkaappausmenetelmät. Varsinkin koulutustarkoitukseen markkeriton liikkeenkaappaus olisi loistava lisälaite, jolla voi paremmin analysoida omaa suoritustaan. Myös yksilöurheilussa ammattilaisten suorituksia välitettäisiin internettiin, josta lajiharrastajat pystyvät vertailemaan omia suorituksiaan ammattilaisten vastaaviin suorituksiin.

Aihe on kiinnostanut paljon tutkijoita, mutta teknologia ei ole vielä edennyt riittä-

vän pitkälle, että markkeritonta liikkeenkaappausta voitaisiin hyödyntää yleisesti paremmin kuin muita liikkeenkaappausmenetelmiä. Microsoft Kinectin myötä markkerittomalle liikkeenkaappaukselle on syntynyt vahva pohja, johon tulevat tutkimukset, laitteistot ja ohjelmistot voivat perustua ja edistää liikkeenkaappauksen tehokkuutta ja tarkkuutta. On totta, ettei markkeriton liikkeenkaappaus sovellu tarkkuutta vaativiin tehtäviin, mutta kaikki sovelluskohteet eivät tarkkaa tarkkuutta vaadi. Tämä kirjallisuuskatsaus on herättänyt paljon uusia kysymyksiä ja ideoita markkerittoman liikkeenkaappauksen mahdollisesta jatkotutkimuksesta. Olisi mielenkiintoista selvittää, kuinka markkeritonta liikkeenkaappausta saadaan hyödynnettyä esimerkiksi robotiikassa ja videopeleissä.

## Kirjallisuutta

- Alexiadis, D. S. & Kelly, P. & Daras, P. & O'Connor, N. E. & Boubekeur, T. & Mousa, M. B. 2011. *Evaluating a dancer's performance using kinect-based skeleton tracking*. Teoksessa Proceedings of the 19th ACM international conference on Multimedia (MM '11), ACM, 659-662, New York, NY. DOI: 10.1145/2072298.2072412
- Butkiewicz, T. 2014. *Low-cost coastal mapping using Kinect v2 time-of-flight cameras*. Teoksessa 2014 Oceans - St. John's, 1-9, St. John's, NL. DOI: 10.1109/OCEANS.2014.7003084
- Ceseracciu, E. & Sawacha, Z. & Cobelli, C. 2014. *Comparison of Markerless and Marker-Based Motion Capture Technologies through Simultaneous Data Collection during Gait: Proof of Concept*. PLoS ONE, 9(3). DOI: 10.1371/journal.pone.0087640
- da Motta, E. S. & Sementille A. C. & Aguilar, I. A. 2017. *Development of a Method for Capturing Human Motion Using a RGB-D Camera*. Teoksessa 2017 19th Symposium on Virtual and Augmented Reality (SVR), 97-106, Curitiba. DOI: 10.1109/SVR.2017.21
- Field, M. & Stirling, D. & Naghdy, F. & Pan, Z. 2009. *Motion capture in robotics review*. Teoksessa 2009 IEEE International Conference on Control and Automation, 1697-1702, Christchurch. DOI: 10.1109/ICCA.2009.5410185
- Khoshelham, K. & Elberink, S. O. 2012. *Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications*. Teoksessa Sensors, 12(2), 1437-1454. DOI: 10.3390/s120201437
- Lau, D. 2013. *The Science Behind Kinects or Kinect 1.0 versus 2.0*. URL: <http://www.gamasutra.com/blogs/DanielLau/20131127/205820/>
- Li, S. & Li, B. & Zhang, S. & Fu, H. & Lo, W. & Yu, J. & Sit, C. H. P. & Li, R. 2018. *A markerless visual-motor tracking system for behavior monitoring in DCD assessment*. Teoksessa 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 774-777, Kuala Lumpur. DOI: 10.1109/APSIPA.2017.8282139
- Loconsole, C. & Frisoli, A. & Semeraro, F. & Stroppa, F. & Mastronicola, N. & Filipeschi, A. & Marchetti, L. 2016. *RELIVE: A Markerless Assistant for CPR Training*.

- Teoksessa IEEE Transactions on Human-Machine Systems, vol. 46, no. 5, 755-760. DOI: 10.1109/THMS.2016.2586756
- Metcalf, C. D. & Robinson, R. & Malpass, A. J. & Bogle, T. P. & Dell, T. A. & Harris, C. & Demain, S. H. 2013. *Markerless Motion Capture and Measurement of Hand Kinematics: Validation and Application to Home-Based Upper Limb Rehabilitation*. Teoksessa IEEE Transactions on Biomedical Engineering, vol. 60, no. 8, 2184-2192. DOI: 10.1109/TBME.2013.2250286
- Payeur, P. & Naschimento, G. M. G. & Beacon, J. & Comeau, G. & Cretu, A. M. & D'Aoust, V. & Charpentier, M. A. 2014. *Human gesture quantification: An evaluation tool for somatic training and piano performance*. Teoksessa 2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) Proceedings, 100-105, Richardson, TX. DOI: 10.1109/HAVE.2014.6954339
- Shotton, J. & Fitzgibbon, A. & Cook, M. & Sharp, T. & Finocchio, M. & Moore, R. & Kipman, A. & Blake, A. 2011. *Real-time human pose recognition in parts from single depth images*. Teoksessa CVPR 2011, 1297-1304, Providence, RI. DOI: 10.1109/CVPR.2011.5995316
- Xing, T. & Yu, Y. & Zhou, Y. & Du, S. 2012. *Markerless Motion Capture of Human Body Using PSO with Single Depth Camera*. Teoksessa 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, 192-197, Zurich. DOI: 10.1109/3DIMPVT.2012.21
- Ye, M. & Yang R. 2014. *Real-Time Simultaneous Pose and Shape Estimation for Articulated Objects Using a Single Depth Camera*. Teoksessa 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2353-2360, Columbus, OH. DOI: 10.1109/CVPR.2014.301
- Zhang, Z. 2012. *Microsoft Kinect Sensor and Its Effect*. Teoksessa IEEE MultiMedia, vol. 19, no. 2, 4-10. DOI: 10.1109/MMUL.2012.24