

Alexi Järvelä

3D-KONENÄKÖ TEOLLISUUSROBOTIL- LA TEHTÄVÄSSÄ KAPPALEENKÄSIT- TELYSSÄ

Tekniikan ja luonnontieteiden tiedekunta
Kandidaatintyö
Tarkastaja: Niko Siltala
Toukokuu 2024

TIIVISTELMÄ

Alexi Järvelä: 3D-konenäkö teollisuusrobotilla tehtävässä kappaleenkäsittelyssä
Kandidaatintyö
Tampereen yliopisto
Tekniikan kandidaatin tutkinto-ohjelma, Konetekniikka
Toukokuu 2024

Tässä opinnäytetyössä tutkitaan erilaisia kolmiulotteisia kuvantamismenetelmiä, menetelmillä saadun informaation käsittelyä ja menetelmien käyttämistä kappaleenkäsittelysovelluksissa. Kolmiulotteisen tiedon saamiseen tarvittavat työkalut ovat yleistyneet teollisuuden erilaisissa sovelluksissa menetelmien kehittyessä ja tarvittavien komponenttien hintojen laskiessa. Monissa kappaleenkäsittelysovelluksissa kaksiulotteinen tieto on riittävä, mutta joissakin sovelluksissa saatetaan tarvita myös syvyystietoa.

Aluksi työssä käydään läpi erilaisia kolmiulotteisia kuvantamismenetelmiä, kuten stereonäkö, rakenteellinen valaistus, lentoaikamenetelmä ja lyhyesti myös muita harvinaisempia menetelmiä, sekä sitä kuinka erilaisilla menetelmillä saadaan laskettua syvyysinformaatio. Menetelmät ovat jaoteltu passiivisiin ja aktiivisiin menetelmiin, joista jälkimmäisissä tarvitaan ulkopuolista valaistusta. Eri menetelmillä on omat vahvuutensa ja heikkoutensa teollisessa ympäristössä ja käytetty menetelmä tuleekin valita sovelluskohtaisesti.

Tavallisesti kuvantamismenetelmillä voidaan tuottaa esimerkiksi pistepilvi tai etäisyyskuva, ja sen käsittely sisältää haastavia laskennallisia ongelmia. Pistepilvet saattavat sisältää esimerkiksi hajapisteitä, jotka voivat vaatia suodatusta. Lisäksi pistepilvestä täytyy tunnistaa kappaleet, niiden asento sekä paikka käyttäen hyödyksi erilaisia rekisteröintialgoritmeja ja tunnettuja Computer-aided design (CAD) -malleja. Tässä työssä rekisteröintialgoritmeista käsitellään vain iteratiivinen lähin piste (engl. Iterative Closest Point, ICP) -algoritmia sen yleisyyden takia, mutta myös muita menetelmiä on olemassa suuri määrä.

3D-konenäköä käytetään erityisen paljon teollisuudessa esimerkiksi kasasta poiminta -sovelluksissa, joissa poimittavat kappaleet ovat laatikossa satunnaisessa järjestyksessä toistensa päällä. Tällaisissa sovelluksissa on tärkeää tietää mikä kappale on päällimmäisenä ja miten osat ovat asettuneet toistensa päälle, jolloin syvyystieto on välttämätön. Olennaista on myös kappaleen poiminnan ja liikeradan suunnittelu, sekä se että missä kamera sijaitsee robottiin nähden.

Avainsanat: 3D-konenäkö, kappaleenkäsittely, robotiikka, bin picking, pistepilvi

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

SISÄLLYSLUETTELO

1. JOHDANTO	1
2. 3D-KONENÄKÖ	3
2.1 Kameratyypit ja kameran malli	3
2.2 3D-kuvantamisen menetelmät	5
2.2.1 Stereonäkö	5
2.2.2 Rakenteellinen valaistus	8
2.2.3 Lentoaikamenetelmä	9
2.2.4 Muita menetelmiä	10
3. MALLIN MUODOSTAMINEN JA KAPPALEEN TUNNISTAMINEN	12
3.1 3D-mallin esittäminen	12
3.2 Kappaleen tunnistaminen	13
4. KONENÄKÖ ROBOTIN OHJAUKSESSA	15
5. ANALYYSI JA POHDINNAT	18
6. YHTEENVETO	19
LÄHTEET	20

LYHENTEET JA MERKINNÄT

2D	Kaksiulotteinen
3D	Kolmiulotteinen
LCD	engl. Liquid Crystal Display
CAD	engl. Computer-aided design
CCD	engl. Charge-Coupled Device
CMOS	engl. Complementary Metal Oxide Semiconductor
ICP	engl. Iterative Closest Point
ToF	engl. Time-of-Flight
PCA	engl. Principal Component Analysis
PVFH	Piirrehistogrammi
RANSAC	engl. Random Sample Consensus

1. JOHDANTO

Perinteisesti teollisuusrobotit ohjelmoidaan toistamaan tietynlaisia liikeratoja. Perinteinen menetelmä on kuitenkin herkkä muutoksille, eikä se sovellu ympäristöön, jossa kohdekappale sijaitsee satunnaisessa paikassa tai kohdekappaletta ei tiedetä ennalta. Konenäön avulla robotista saadaan älykkäämpi ja kohdekappale pystytään tunnistamaan sekä paikantamaan. (Gan et al. 2011)

Konenäköä on hyödynnetty erilaisissa teollisuuden sovelluksissa ja robottien ohjauksessa jo vuosikymmenten ajan. Iso osa sovelluksista käyttää kuitenkin kaksiulotteista konenäköä ja kolmiulotteinen konenäkö on alkanut saada enemmän huomiota vasta viime vuosina. Kolmiulotteinen konenäkö eroaa kaksiulotteisesta konenäöstä siten, että kuvattavasta kohteesta saadaan kaksiulotteisen kuvan sijaan kolmiulotteinen malli. Kolmiulotteisen mallin avulla kohteesta saadaan mitat, asento ja etäisyys kamerasta.

Kaksiulotteinen konenäkö on riittävä tapauksissa, joissa kohdekappale sijaitsee esimerkiksi kuljettimella tietyllä korkeudella tai laatikossa sille tarkoitettussa lokerossa. Mikäli kohdekappale on esimerkiksi laatikossa päällekkäin muiden kappaleiden kanssa, satunnaisessa järjestyksessä, tarvitaan kappaleen paikantamiseen ja asennon määrittämiseen kolmiulotteista konenäköä. Kyseisestä tapauksesta käytetään englanniksi nimitystä bin picking. (Herakovic 2010)

Kolmiulotteinen konenäkö soveltuu hyvin muuttuvaan ja ennalta määrittelemättömään ympäristöön. Käsiteltäviä kappaleita ei tarvitse asettaa tarkasti tiettyyn asentoon tai tietylle tasolle, mikä helpottaa esimerkiksi kokoonpanossa erilaisten konfiguraatioiden valmistamista.

Tässä työssä tutkitkaamme erityyppisiä 3D-kuvantamismenetelmiä, miten saatu data käsitellään, sekä kuinka eri kuvantamis- ja datankäsittelymenetelmiä pystytään hyödyntämään kappaleen tunnistuksessa. Työn tavoitteena on saada lukijalle teoreettinen ymmärrys eri menetelmistä helpottamaan mahdollisten toteutusvaihtoehtojen suunnittelua. Työ vastaa seuraaviin tutkimuskysymyksiin:

1. Mitä eri 3D-kuvantamismenetelmiä robotisoiduissa poimintasovelluksissa hyödynnetään?
2. Mitä robotisoitujen poimintasovellusten toteutuksessa ja kameratyyppejä valitessa tulee ottaa huomioon?

Tutkimusmenetelmänä toimii kirjallisuuskatsaus, jossa lähteinä on käytetty tutkimusartikkeleita, konferenssijulkaisuja sekä kirjoja. Kirjallisuuskatsauksen aineisto haettiin hyödyntäen Tampereen yliopiston hakupalvelu Andoria sekä Google Scholaria käyttäen keskeisimpinä hakusanoina "3D computer vision" ja "bin picking". Palveluista löytyneiden kirjojen lähdeluetteloiden ja tekstin perusteella haettiin lisälähteitä. Lähteissä ei rajattu vanhoja materiaaleja pois sillä monet teoreettisen osuuden lähteet ovat vielä ajankohtaisia. Käytännön sovelluksista haluttiin löytää myös uudempia lähteitä.

Luvussa 2 käydään läpi yleisimpiä menetelmiä, joita käytetään kolmiulotteisen informaation saamiseksi kohteesta, sekä teoriaa niiden toiminnasta. Luvussa 3 käsitellään kappaleen esittämiseen ja tunnistamiseen liittyviä toimenpiteitä. Luvussa 4 on esitetty robotin konenäköpohjaista ohjausta ja sovelluksia, joissa kolmiulotteista konenäköä voidaan hyödyntää. Luvussa 5 on esitettyjen menetelmien analyysia ja pohdintaa. Luvussa 6 on yhteenveto.

2. 3D-KONENÄKÖ

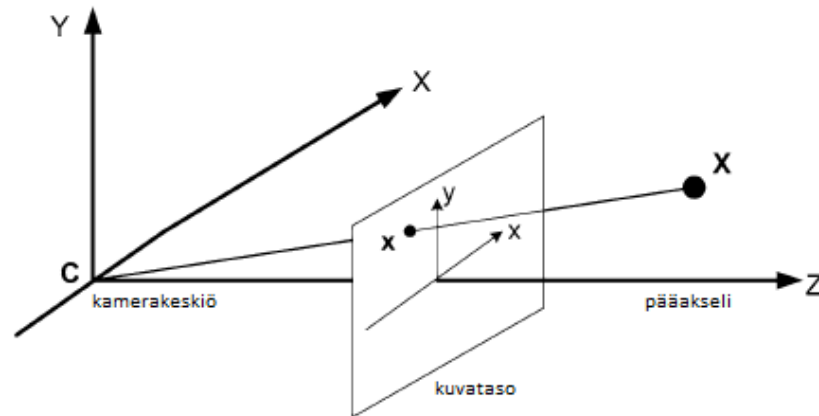
Tässä luvussa käsitellään kuvantamiseen käytettäviä kameratyyppejä, kameramallia sekä sen matemaattista taustaa. Myöhemmin käydään läpi erilaisia menetelmiä, joilla syvyysinformaatio saadaan laskettua otetuista kuvista.

2.1 Kameratyypit ja kameran malli

Suurin osa kameroista käyttää joko charge-coupled device (CCD) tai complementary metal-oxide semiconductor (CMOS) -kennoa, joiden avulla valo muutetaan sähköiseksi signaaliksi (Pears et al. 2012). CCD- ja CMOS-kennot ovat jaettu pieniin suorakulmion muotoisiin sensoreihin, jotka vastaavat kuvan pikseleitä ja fotonin osuessa niihin ne muodostavat sähkövarauksen. CCD-kennossa syntyvä sähkövaraus on pieni ja se tallennetaan myöhempää käsittelyä varten, kun taas CMOS-kennossa varaus vahvistetaan.

Gan et al. (2011) mukaan CMOS-kennot ovat herkempiä valaistuksen aiheuttamille häiriöille, niillä otetut kuvat saattavat olla rakeisempia CCD-kennoihin verrattuna, CMOS-kennot ovat halvempia ja käyttävät vähemmän virtaa. CMOS-kennojen kuvanlaatu on kuitenkin kehittyneet vuosien saatossa huomattavasti ja usein uudet kamerrat ovatkin CMOS-kennollisia.

Kameramallilla kuvataan kolmiulotteisen maailman koordinaattien projisoitumista kaksiulotteiselle kuvatasolle. Kuvauksessa syvyysinformaatio menetetään ja sen selvittämiseksi voidaan käyttää myöhemmin käsiteltyjä menetelmiä. Yksinkertaisin kameramalli on niin sanottu neulansilmäkameramalli (Hartley et al. 2014). Se sopii kuvaamaan hyvin esimerkiksi CCD-kameran toimintaa, mutta soveltuu myös muille kameratyypeille.



Kuva 1. Havainnollistava kuva neulansilmäkameramallista, jossa kuvataso on tuotu kamerakeskiön eteen. Mukailten Pears et al 2011, s.42, kuva 2.4.

Kuvassa 1 on havainnollistettu neulansilmäkameramallia, joka kuvaa kolmiulotteisen pisteiden muodostumista kameras kuvatasolle. Kameramalli on olennainen työkalu eri kuvantamismenetelmissä laskettaessa syvyysinformaatiota saadusta kuvasta. Seuraavaksi esiteltyjä kaavoja tarvitaan muun muassa kameras kalibrointiin, jotta etäisyys saadaan laskettua kuvasta esimerkiksi stereonäköön sekä strukturoituun valoon perustuvilla menetelmillä ja erilaiset linssien vääristymät voidaan ottaa huomioon.

Avaruuden piste $X = (X Y Z)^T$ piirtyy kuvatason pisteeseen $x = (x y)^T$, jossa kamerakeskiön ja pisteen X välinen jana lävistää kuvatason. Kuvaus $X \mapsto x$ voidaan laskea kaavalla

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} fX/Z \\ fY/Z \end{pmatrix}, \quad (1)$$

jossa f on kamerakeskiön ja kuvatason välinen matka eli polttoväli (Gan et al. 2011). Todellisuudessa kuvataso sijaitsee kamerakeskiön takana, mutta kuvissa se on piirretty joskus kamerakeskiön etupuolelle niin sanotuksi virtuaaliseksi kuvatasonksi.

Kuvaus voidaan esittää myös homogeenisessa matriisimuodossa yhtälöllä

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = PX = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (2)$$

missä P on projektiomatriisi ja λ on skaalauskerroin (Gan et al. 2011). Kamerakoordinaattien muuntamiseksi pikselikoordinaatteihin voidaan käyttää kalibrointimatriisia

$$K = \begin{bmatrix} fm_x & s & x_0 \\ 0 & fm_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

jossa m_x ja m_y ovat pikseleitä pituusyksikköä kohti ja s on vinouma, joka on normaalissa tapauksessa nolla (Hartley et al. 2004). Pisteen muuttaminen maailmakoordinaatista kamerakoordinaatistoon voidaan tehdä ulkoiset kameraparametrit eli rotaation ja translaation sisältävällä matriisilla $[R|t]$ (Hartley et al. 2004). Projektiomatriisi saa tällöin muodon

$$P = K[R|t]. \quad (4)$$

2.2 3D-kuvantamisen menetelmät

Kolmiulotteisessa konenäössä käytettävät menetelmät voidaan jakaa aktiivisiin ja passiivisiin menetelmiin (Taylor et al. 2006). Aktiivisissa menetelmissä käytetään omaa valonlähdettä, jolla heijastetaan kuvattavalle pinnalle kuvio. Passiivisissa menetelmissä omaa säteilynlähdettä ei tarvita, vaan ympäristön valaistus ja kohteessa olevat piirteet riittävät. Yleisimmät tavat saada kolmiulotteinen informaatio passiivisella menetelmällä on ottaa kuvia kohteesta hieman eri kulmista käyttäen yhtä liikkuvaa tai useampaa paikallaan olevaa kameraa. Myös yhden paikallaan olevan kameran käyttö on mahdollista, mutta siinä käytettävät menetelmät eivät ole kovin nopeita tai robusteja. (Pears et al. 2012)

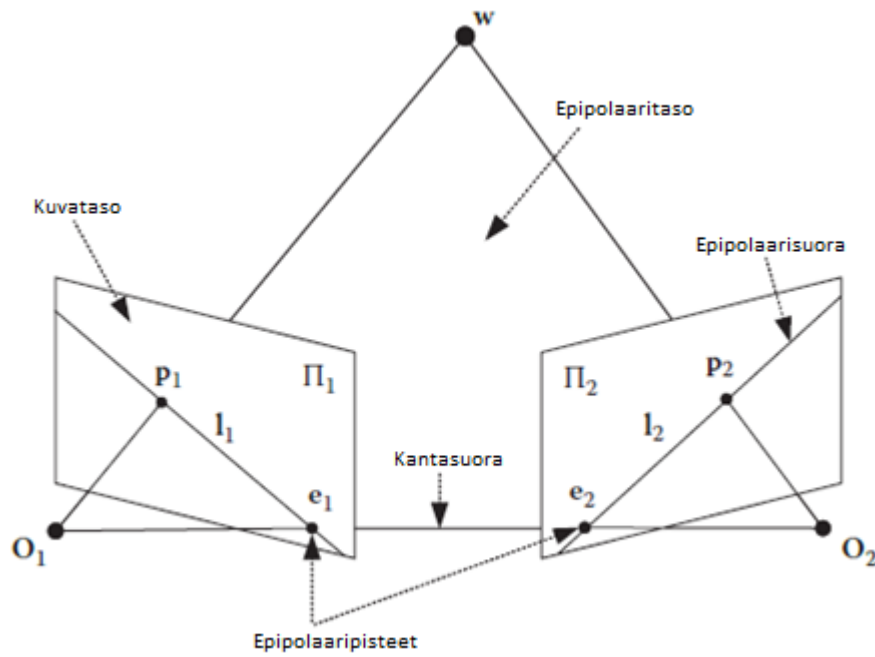
2.2.1 Stereonäkö

Stereonäössä kahdesta eri suunnasta otetusta kaksiulotteisesta kuvasta lasketaan kohteesta kolmiulotteinen informaatio. Syvyysinformaation saamiseksi kuvasta tulee ratkaista niin sanottu vastaavuusongelma, jossa molemmista kuvista tulisi löytää vastaavat kuvapisteen kohteesta. Vastaavien kuvapisteiden eli vastinpisteiden löytämisen jälkeen syvyysinformaatio saadaan vastinpisteiden koordinaattien perusteella. (Song 2013) Vastinpisteiden ero on kääntäen verrannollinen kohteen etäisyyteen (Woodfill et al. 2004). Mitä kauempana vastinpisteet siis ovat toisistaan kuvakoordinaateissa, sitä lähempänä kohde on.

Vastinpisteiden löytäminen on laskennallisesti hyvin raskasta, ja niiden etsimistä rajoitetaan tavallisesti tietylle alueelle kuvissa. Kahden vierekkäisen kameran järjestelmässä voidaan käyttää epipolaarista rajoitusta, jossa vastaavien kuvapisteiden haku rajoittuu yksiulotteiselle suoralle tehden vastaavuusongelman ratkaisemisesta huomattavasti tehokkaamman ja helpomman (Song 2013).

Kuvassa 2 on havainnollistettu epipolaarisen geometrian käyttöä stereonäköjärjestelmässä. Epipolaaripisteet e_1 ja e_2 muodostuvat kuvatasolle niihin kohtiin, missä kame-

roiden keskipisteiden välinen suora leikkaa kuvatasot. Epipolaaritaso on kameroiden keskipisteiden O_1 ja O_2 , sekä pisteen w välillä oleva taso, ja kuvapisteen p_1 ja p_2 muodostuvat O_1 , O_2 sekä w välisten suorien leikkaamiselle kuvatasoille. Epipolaarisuorat l_1 ja l_2 sijaitsevat kuvatasoilla epipolaari- ja kuvapisteen välillä. Vastinpisteitä voidaan etsiä epipolaarisuorilta sen sijaan, että tarvitsisi käydä läpi koko kuvatasoa.



Kuva 2. Epipolaarigeometria stereonäköjärjestelmässä. Mukailten lähteestä Song 2013, sivu 5, kuva 1.2.

Stereonäköjärjestelmässä kamerat ovat tyypillisesti käännetty hiukan toisiaan kohti, mikä aiheuttaa sen, että epipolaarisuorat sijaitsevat erilaisissa kulmissa kuvatasojen välillä (Pears et al. 2012). Epipolaarisuorat saadaan samoille vaakariveille kuvissa rektifioimalla, jolloin vastinpisteiden etsiminen ja kolmiomittaus helpottuvat. Rektifioinnin jälkeen kolmiomittaus tehdään kuvassa 3 näkyvien suureiden perusteella. Koordinaatiston origon ollessa oikean kameran kamerakeskiössä, saadaan kuvatasolla olevien vastinpisteiden x_c ja x'_c paikat laskettua yhtälöstä

$$\begin{pmatrix} x_c \\ x'_c \end{pmatrix} = \frac{f}{Z} \begin{pmatrix} X \\ X+B \end{pmatrix}, \quad (5)$$

missä f on polttoväli ja B on kamerakeskiöiden välinen etäisyys.

Pisteiden x_c ja x'_c erotuksesta saadaan poikkeama:

$$\mathbf{d} = x_c - x'_c = \frac{fB}{Z}. \quad (6)$$

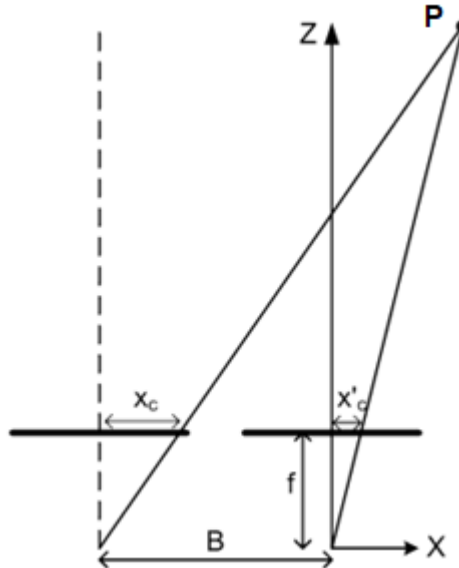
Kun poikkeama tiedetään, voidaan pisteen P koordinaatit laskea kaavoilla

$$Z = \frac{fB}{d}, \quad (7)$$

$$X = \frac{Zx'_c}{f}, \quad (8)$$

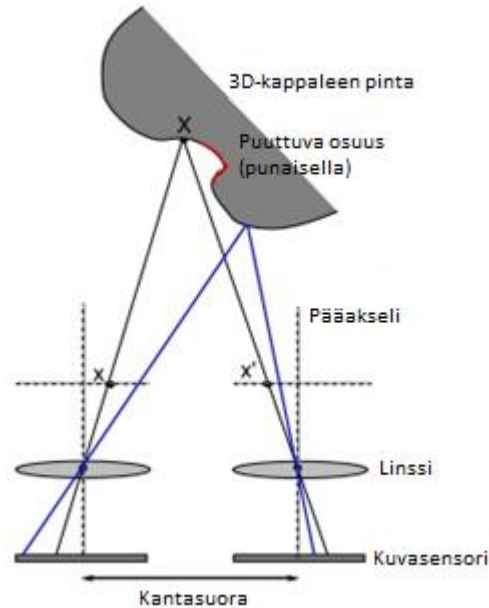
$$Y = \frac{Zy'_c}{f}, \quad (9)$$

Missä y'_c on oikeanpuoleisen kuvan vastinpisteen pystykoordinaatti.



Kuva 3. Kolmiomittaus stereonäköjärjestelmässä. Mukailten lähteestä Pears et al. 2012, s. 76, kuva 16.

Passiivista stereonäköä käytettäessä kuvattavassa kohteessa täytyy olla riittävästi yksityiskohtia, jotta vastinpisteitä pystyttäisiin löytämään tarpeeksi. Menetelmä ei siis sovi kaiken tyyppisille kappaleille tai ainakin tarkkuus saattaa kärsiä, jos kappaleesta ei löydy tarpeeksi vastinpisteitä. Myöskään kuvan 4 mukaisissa kappaleissa, joissa molemmat kamerat eivät näe kohteista samoja osia, täydellistä mallia ei saada rakennettua.



Kuva 4. Stereonäössä esiintyvä ongelma, jossa molemmat kamerat eivät näe kohdekappaleesta samoja osia. Mukailten lähteestä Pears et al 2012, s. 12, kuva 3.

2.2.2 Rakenteellinen valaistus

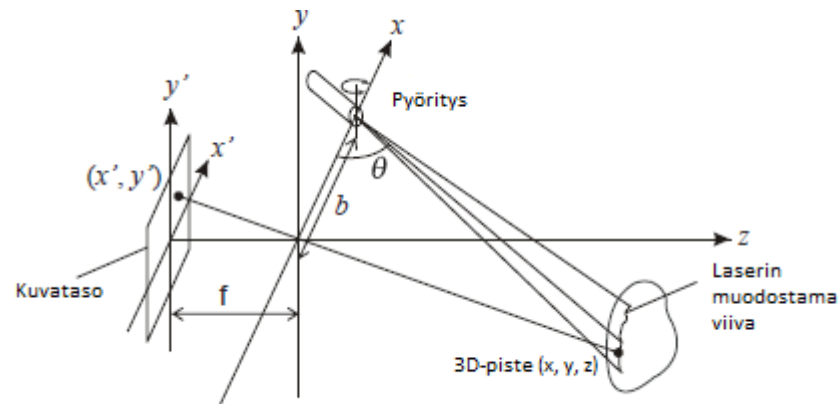
Rakenteellisessa valaistuksessa yhden tai useamman kameran lisäksi on projektori, jolla heijastetaan useita valokuvioita kuvattavan kohteen pinnalle. Kyseessä on siis aktiivinen kuvausmenetelmä. Kuvion avulla vältetään passiivisessä stereonäössä esiintyvältä ongelmalta, jossa vastinpisteitä ei löydy riittävästi pinnan yksityiskohtien puutteen takia (Gan et al. 2011). Heijastettavaa ja kuvattua kuviota vertaamalla pystytään kuvattavasta kohteesta muodostamaan kolmiulotteinen malli (Salvi et al. 2010).

Heijastettava kuvio voi olla esimerkiksi yksittäinen piste, viiva, viivajoukko, ristikko tai pistejoukko. Yksittäistä pistettä ja viivaa käytettäessä joudutaan kuviota liikuttamaan pinnalla mekaanisesti sekä ottamaan useampia kuvia, mikä hidastaa mallin muodostamista huomattavasti. Menetelmästä käytetäänkin usein nimitystä laserskannaus tai laserkolmiomittaus ja se luokitellaan usein eri tekniikaksi kuin rakenteellinen valaistus.

Laserskannauksen toimintaa on esitetty kuvassa 5, jossa viivalaser on projisoitu kohteeseen ja kolmiulotteisen maailman pisteet kuvautuvat kameramallin mukaisesti kuvatasolle. Kolmiulotteisen pisteen $(X Y Z)^T$ koordinaatit saadaan laskettua kaavalla

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{b}{f \cot \theta - x'} \begin{pmatrix} x \\ y \\ f \end{pmatrix}, \quad (10)$$

missä f on polttoväli, b on kameran ja valaisimen välinen etäisyys, θ on projektorin ja kuvatason välinen kulma, sekä x ja y ovat kuvatasolle projisoituneen pisteen koordinaatit (Franca et al. 2005).



Kuva 5. Laserskannaukseen perustuva rakenteellisen valaistuksen menetelmä. Mukailen lähteestä Franca et al. 2005, s. 2, kuva 2.

Mekaaninen skannauksen välttämiseksi voidaan käyttää kaksiulotteista kuviota (Fofi et al. 2004). Kaksiulotteisen kuvion muodostamiseen on kehitetty useita eri kodifikaatiomenetelmiä, jotka perustuvat muun muassa aikamultipleksaukseen, spatiaaliseen naapurustoon tai suorakoodaukseen helpottamaan vastinpisteiden löytämistä (Salvi et al. 2003).

Valokuvion muodostamiseen käytetään yleensä joko laseria tai liquid crystal display (LCD) -projektoria. LCD-projektorin käytettäessä etuna on käytettävän kodifikaatiomenetelmän helppo vaihtaminen. Menetelmää voidaan vaihtaa esimerkiksi väreihin perustuvan koodauksen ja Gray-koodauksen välillä. Laseria käytettäessä voidaan käyttää näkyvän valon lisäksi myös ihmissilmälle näkymätöntä infrapunavaloa, jonka kamera pystyy kuitenkin havaitsemaan.

Rakenteellista valaistusta käytettäessä ongelmia aiheuttavat erityisesti peilimäiset pinnat ja varjot (Nayar et al. 2012). Peilimäisistä pinnoista projisoitu valo saattaa heijastua muihin suuntiin kameran sijasta, mikä heikentää kuvauksen laatua (Song et al. 2013). Varjot aiheutuvat samalla tavoin kuin stereonäössä kuvan 4 mukaisesti oleva puuttuva kohta. Ne ovat siis seurausta kameran ja valonlähteen välisestä etäisyydestä.

2.2.3 Lentoaikamenetelmä

Lentoaikamenetelmässä (engl. Time of Flight, ToF) kohteen etäisyys mitataan valon kulkemasta matkasta joko suoralla tai epäsuoralla menetelmällä heijastamalla valonsäde kuvattavalle pinnalle (Jang et al. 2013). Suorassa menetelmässä kohteen etäi-

syys lasketaan valonsäteen lähtö- ja havaitsemisajan erotuksen perusteella. Koska valonnopeus tiedetään, saadaan valon kulkema matka laskettua kaavalla

$$d = \frac{cT}{2}, \quad (11)$$

missä c on valonnopeus ja T on aika. Epäsuorassa menetelmässä matka voidaan laskea lähtevän ja tulevan välisen sini- tai pulssimoduloidun valon vaihe-eron avulla (Jang et al. 2013). Vaihe-eroon perustuvaa menetelmää käytettäessä etäisyys saadaan laskettua kaavalla

$$d = \frac{c}{2f} \frac{\varphi}{2\pi}, \quad (12)$$

missä c on valonnopeus, φ on vaihe-ero ja f on modulointitaajuus. Mitä suurempaa modulointitaajuutta käytetään, sitä tarkemmin etäisyys saadaan mitattua (Yoon et al. 2011).

Lentoaikamenetelmää käyttävät kamerat ovat usein edullisia ja vähän virtaa käyttäviä sekä niiden suorituskyky on parantunut huomattavasti vuosien saatossa. Stereonäön kärsiessä vastaavuusongelmasta ja rakenteellisen valaistuksen ollessa herkkä ympäröivälle valaistukselle lentoaikamenetelmää käyttävä kamera voi olla hyvä vaihtoehtona monissa sovelluksissa. (Paradez et al. 2023).

2.2.4 Muita menetelmiä

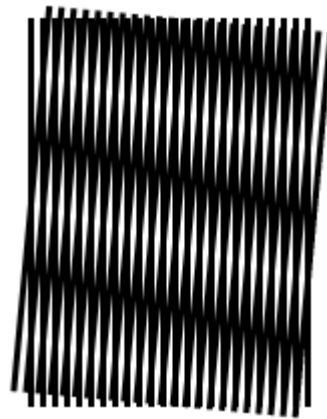
Muita kolmiulotteisen informaation saamiseksi käytettäviä konenäkömenetelmiä on muun muassa fotometrinen stereo, Moiré-ilmiöön perustuvat menetelmät, muoto tekstuurista (engl. shape from texture) ja muoto tarkennuksesta (engl. shape from focus). Kaikki näistä menetelmistä eivät kuitenkaan sovellu kappaleenkäsittelysovelluksiin erityisen hyvin sillä esimerkiksi etäisyysinformaatio jää joillain menetelmillä saamatta.

Fotometrisessä stereossa paikallaan olevalla kameralla kahdesta tai useammasta eri valaistuksella otetusta kuvasta voidaan arvioida pinnan suunta kussakin pisteessä (Tagare et al. 1991). Paikallaan olevan kameran ansiosta vastinpisteet on erittäin helppo löytää, sillä ne sijaitsevat kaikissa kuvissa samoissa kohdissa. Menetelmässä on kuitenkin paljon heikkouksia kappaleenkäsittelyyn sopimisen kannalta ja ongelmia tuottaa muun muassa tarve valaistuksen vaihtelulle jokaisen mittauksen aikana. Menetelmällä ei myöskään saada mitattua etäisyyttä kohteeseen.

Kuvan 6 tyyppinen Moiré-kuvio syntyy, kun kaksi samanlaista viivastoa asetetaan päällekkäin eri kulmaan toisiinsa nähden. Moiré-ilmiötä voidaan hyödyntää kolmiulotteises-

sa konenäössä ja siihen perustuvat menetelmät voidaan jakaa varjo- ja projektitekniikoihin (Chen et al. 2000).

Projektitekniikkaan perustuvassa Moiré-menetelmässä kameran ja valonlähteen edessä on omat hilansa. Kohteen pinnalle heijastuu siis vain viivakuviot, mutta kameran edessä oleva hila avulla saadaan kameran kuvatasolle Moiré-kuvio. Varjotekniikkaan perustuvassa Moiré-menetelmässä kohteen eteen asetetaan yksi hila, jolloin sama hila aiheuttaa sekä varjostuksen kohteeseen että kuvion kameralle. Kohteen pinnanmuodot vääristävät varjokuvioita ja tätä voidaan hyödyntää 3D-mittauksessa.



Kuva 6. Moiré-kuvion muodostamiseen käytettävät hilat.

Muoto tekstuurista menetelmässä kuvattavassa kohteessa täytyy olla tekstuuri tai merkintöjä, joiden vääristymiä voidaan käyttää muodon laskemiseen (Loh et al. 2005). Tämä rajoittaa olennaisesti menetelmän käyttöä sovelluksissa, joissa kuvattavissa kappaleissa ei tällaisia ominaisuuksia ole.

Muoto tarkennuksesta menetelmässä syvyystieto lasketaan ottamalla kohteesta useita kuvia kameran eri tarkennusasetuksilla (Pertuz et al. 2012). Menetelmällä kuvattavasta kohteesta muodostetaan usein etäisyyskuva, jota käsitellään lisää seuraavassa luvussa.

3. MALLIN MUODOSTAMINEN JA KAPPALEEN TUNNISTAMINEN

Kuvatusta kohteesta 3D-informaatio voi olla erityyppisissä muodoissa, kuten pistepilvenä, etäisyyskuvana tai polygoneina (Pears et al. 2012). Malli voi vaatia esikäsittelyä ennen kuin sitä voidaan käyttää kappaleen tunnistamiseen. Kappaleen sekä sen asennon tunnistaminen ovat olennaisessa osassa myöhemmin käsiteltävissä kappaleenkäsittelysovelluksissa.

3.1 3D-mallin esittäminen

Yksi yleisimmistä esitystavoista on pistepilvi. Pistepilvi koostuu joukosta pisteitä kolmiulotteisessa avaruudessa. Pistepilven pisteet ovat kuvatun kohteen pinnan koordinaatteja ja piste voidaan esittää x , y ja z koordinaattien avulla muodossa $p_i = \{x_i, y_i, z_i\}$. Jolloin n määrästä pisteitä koostuva pistepilvi voidaan esittää muodossa $P = \{p_1, p_2, \dots, p_n\}$.

Toinen yleinen tapa 3D-informaation esittämiseen on etäisyyskuva. Etuna pistepilveen on muun muassa vähäisempi tilantarve, sillä syvyyskuvassa jokaiselle kuvapikselille on vain etäisyystieto (Bletterer et al., 2016). Etäisyyskuva on tyypillinen esitystapa muun muassa stereonäkö- ja muoto tarkennuksesta -menetelmissä, mutta esitystapaa voidaan hyödyntää hyvin myös aktiivisissa menetelmissä.

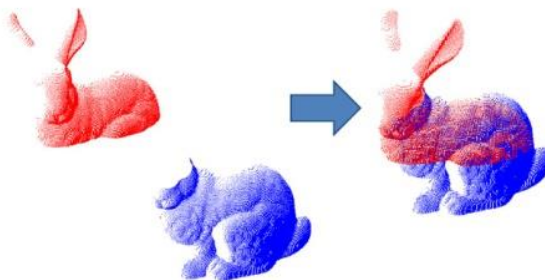
Erilaiset häiriötekijät, kuten kamerassa esiintyvä kohina, linssin vääristymät, valaistus ja kuvien epätäydellisyys aiheuttavat malliin usein virheellisiä mittauspisteitä eli hajapisteitä, jotka saattavat haitata mallin käsittelyä erityisesti kuvapohjaisissa menetelmissä (Wolff et al. 2016). Hajapisteet ovat yleisempiä passiivisissa, kuin aktiivisissa menetelmissä. Hajapisteitä voidaan suodattaa esimerkiksi pääkomponenttianalyysillä (PCA), neuroverkoilla tai Random Sample Consensus (RANSAC) -menetelmällä (Sanchez et al., 2020; Fischler et al., 1981). Toiset rekonstruointi- ja rekisteröintimenetelmät ovat herkempiä hajapisteille kuin toiset, eikä suodatusta tarvitse välttämättä tehdä, vaan hajapisteet käsitellään muilla tavoin. Rekonstruoinnilla tarkoitetaan pintamallin muodostamista pistepilvestä. Kappaleenkäsittelysovelluksissa pintamallin muodostaminen ei kuitenkaan ole välttämätöntä, vaan pistepilveä voidaan verrata esimerkiksi suoraan kuvattavasta kappaleesta olevaan Computer-aided design (CAD) -malliin.

3.2 Kappaleen tunnistaminen

Teollisuuden poimintasovelluksissa kappaleesta on tavallisesti olemassa CAD-malli, johon mallia voidaan verrata sekä tunnistaa kappaleen asento tarttumapisteiden määrittämistä varten. Erilaisia menetelmiä kappaleen tunnistamiseen on kehitetty huomattavia määriä, ja niitä ovat muun muassa opetuspohjaiset menetelmät ja sovituspohjaiset menetelmät (Hinterstoisser et al. 2011).

Opetuspohjaiset menetelmät tarvitsevat tyypillisesti paljon opetusdataa sekä esiopeutuksen. Esimerkiksi Deping et al. (2019) esittämässä osionäkökulman piirrehistogrammi (engl. Partition Viewpoint Feature Histogram, PVFH) on opetuspohjainen menetelmä, jossa CAD-malleista on muodostettu osittaisia pistepilviä simuloimaan 3D-kameran ottamaa kuvaa, joita voidaan käyttää opetusdatana.

Sovitusmenetelmiin kuuluvassa rekisteröinnissä tarkoituksena on sovittaa kaksi eri pistepilveä toisiinsa löytämällä niistä vastinpisteet, jonka avulla voidaan muodostaa kappaleesta yksi pistepilvi eri suunnista otetuista kuvista. Rekisteröintimenetelmiä voidaan hyödyntää kappaleen tai sen asennon tunnistamisessa. Kuvassa 7 on havainnollistettu rekisteröintiä.



Kuva 7. Pistepilvi ennen ja jälkeen rekisteröinnin (Pears et al. 2012, s. 243, kuva 6.14).

Yksi yleisimmistä menetelmistä rekisteröinnin toteuttamiseen on iteratiivinen lähin piste (engl. Iterative Closest Point, ICP) -algoritmi (Mitra et al. 2004). ICP-algoritmin ensimmäisen version kehittivät Chen ja Medioni (Chen et al 1991) sekä hieman myöhemmin myös Besl ja McKay (Besl et al. 1992), mutta sen pohjalta on tehty monia muunnoksia, kuten iteratiivinen lähin yhteensopiva piste (engl. Iterative Closest Compatible Point) ja ICP käyttäen muuttumattomia piirteitä (engl. ICP using invariant features) (Sharp et al. 2002).

ICP-algoritmissa minimoidaan kahden pistepilven vastinpisteiden etäisyyttä. Jokaisen lähdepistepilven pisteen etäisyyttä kohdepistepilvessä olevaan vastinpisteeseen minimoidaan, kunnes ne kohtaavat ja pistepilvet ovat kohdakkain. ICP-algoritmissa on joidakin tunnettuja ongelmia, jotka pitää ottaa huomioon rekisteröitymisen onnistumisen

takaamiseksi. Pistepilvien täytyy olla tarpeeksi lähellä toisiaan, jotta algoritmi ei jäisi jumiin paikalliseen minimiin. Tämä ongelman voidaan ratkaista asettamalla pistepilvet lähelle toisiaan ennen ICP-algoritmin suoritusta. Perinteinen ICP-algoritmi on myös herkkä hajapisteille, mutta ne voidaan poistaa jo ennen sovitus esimerkiksi aiemmin mainituilla suodatusmenetelmillä.

Barrow et al. (1997) esittämä viisteensovitusmenetelmä (engl. Chamfer matching) on rekisteröintimenetelmä, jota suositetaan yleisesti sovelluksissa, joissa kappaleen 3D-malli on tiedossa (Choi et al. 2012). Menetelmässä malleissa olevia reunoja yritetään sovittaa toisiinsa. Viisteensovitusta voidaan pitää ICP-menetelmän joustavampana ja tehokkaampana versiona, ja se sietää hyvin puuttuvaa dataa, mutta hajapisteitä viisteensovituksen kanssa tulee välttää (Van Herk, 2009).

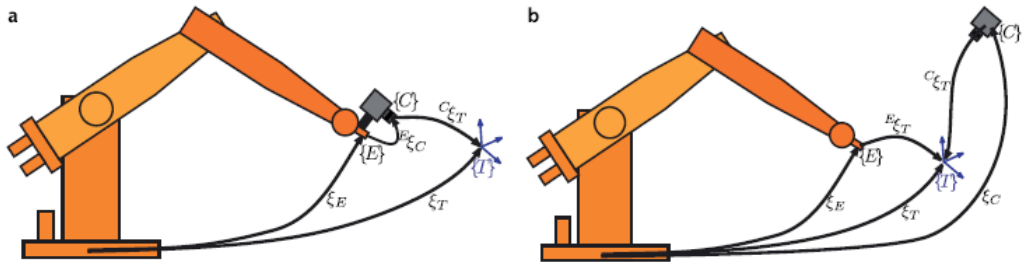
4. KONENÄKÖ ROBOTIN OHJAUKSESSA

Konenäköpohjaisessa robotin ohjauksessa robotin liikkeet suoritetaan kuvasta saatujen tietojen perusteella sen sijaan, että käytettäisiin ennalta määritettyjä liikeratoja. Konenäköä käytettäessä robottia voidaan ohjata joko avoimella tai suljetulla ohjauksella (Gan et al. 2011).

Avoimeen silmukkaan perustuvassa ohjauksessa kohteesta otetaan kuva, analysoidaan se ja liikutetaan robotti haluttuun pisteeseen tehtyjen mittausten perusteella. Kuvaus ja liikeradan laskeminen sekä robotin toimiminen ovat siis erillisiä tapahtumia, eikä reaaliaikaista havainnointia ympäristöstä tapahdu. Avoimen silmukan käyttö onnistuu vain kontrolloidussa ympäristössä, jossa yllättäviä muutoksia ei tapahdu. Halutun toiminnon aikaansaamiseksi, täytyy ympäristön pysyä muuttumattomana kuvauksen jälkeen.

Suljettuun silmukkaan perustuvassa ohjauksessa kohteen kuvaus ja analysointi tehdään uudelleen jokaisessa robotin ohjaussilmukassa. Koska havainnointi tapahtuu reaaliaikaisesti, pystytään ympäristössä tapahtuviin muutoksiin reagoimaan. Suljetun silmukan menetelmää kutsutaan englanniksi nimellä visual servoing.

Visual servoing-menetelmissä kamera on tyypillisesti sijoitettu joko robotin käsivarteen tai johonkin kiinteään paikkaan robotin työtilassa (Hutchinson et al. 1996). Kuvassa 8 on esitetty eri konfiguraatioiden käyttöä. Järjestelmässä, jossa kamera on sijoitettu robotin käsivarteen, eli niin sanotussa eye-in-hand-konfiguraatiossa kameran ja robotin työkalun välinen etäisyys ja asento ovat tiedossa ja kamera kuvaa vain kohdetta. Työtilassa sijaitessa kamera kuvaa sekä kohdetta, että robotin työkalua. Kameran sijaitessa työtilassa myös kohde pysyy muuttumattomana toisin kuin eye-in-hand-konfiguraatiossa, jossa kameran ottama kuva muuttuu robotin liikkuesssa. Käytettävä konfiguraatio saattaa vaikuttaa olennaisesti valittavaan konenäkömenetelmään.



Kuva 8. Eye-in-hand-konfiguraatio ja työtilakonfiguraatio (Corke 2017, s. 455, kuva 15.1.).

Teollisessa tuotannossa osat saattavat tulla laatikoissa, jotka sisältävät monenlaisia osia ja ne saattavat olla satunnaisessa järjestyksessä toistensa päällä. Satunnaisessa järjestyksessä olevien osien poimimiseksi robotilla täytyy selvittää niiden paikka sekä asento, tunnistaa osat ja määrittää tarttumispisteet. Satunnaisesta kasasta poiminnasta käytetään englanniksi nimitystä random bin picking. Satunnaisen kasasta poiminnan ansiosta valmistusjärjestelmistä saadaan joustavampia, sillä erilaisia osia voi olla laatikoissa samaan aikaan, sekä uudenlaisia osia voidaan tuoda järjestelmään nopeasti ilman aikaa vievää robotin opettamista. 3D-konenäkökameroiden hintojen lasku on myös auttanut osaltaan satunnaisen kasasta poiminta -sovellusten yleistymistä.

Ensimmäisiä kasasta poiminta -sovelluksia esitteli Horn et al. (1983), jossa hyödynnettiin fotometristä stereota. Stereonäköä ja eye-in-hand konfiguraatiota ovat hyödyntäneet esimerkiksi Hema et al. (2007). Martinez et al. (2015) mukaan monet aiemmat kasasta poiminta -sovellukset eivät ole olleet kuitenkaan tarpeeksi robusteja teolliseen tuotantoon, ja he esittelevät oman stereonäköä hyödyntävän järjestelmänsä. Chen et al. (2018) esittävät useampaan työtilassa olevaan rakenteelliseen valaistukseen perustuvan kasasta poimimisjärjestelmänsä, jossa hyödynnetään muiden esikäsittelymenetelmien lisäksi muun muassa RANSAC-menetelmää hajapisteiden suodatukseen, sekä ICP-algoritmia asennon tunnistamisessa.

Kleeberger et al. (2019) käyttävät tutkimuksessaan kolmea erilaista 3D-konenäkökameraa sekä ICP-algoritmia rekisteröintiin tutkiakseen minkälainen vaikutus kameran herkkyydellä on kasasta poimimistehtävästä suoriutumiseen. Tutkimuksessa käytettyjä kameroita ovat 12450 € maksanut strukturoituun valoon perustuva Phoxi 3D scanner L, 208 € maksanut stereonäköön perustuva Intel RealSense Depth Camera D415 sekä 200 € maksanut lentoaikamenetelmään perustuva Microsoft Kinect v2 sensor. Kameroilla saatujen pistepilvien laadussa oli huomattavia eroja ja Microsoft Kinect osoittautuikin käyttökelvottomaksi kyseisessä testitilanteessa. Kaksi jäljelle jäänyttä kameraa sen sijaan suoriutuivat testistä lähestulkoon yhtä hyvin tulosten ollessa hiekan kalliimman Phoxi 3D scannerin eduksi. On kuitenkin huomioitava tutkimuksessa

käytettyjen 3D-konenäkökameroiden laatu ja tarkkuus, joten tästä ei voida tehdä suoria päätelmiä esimerkiksi eri menetelmien paremmuudesta toisiinsa nähden.

Kasasta poiminta -sovelluksissa kuvaamisen ja kappaleen tunnistamisen jälkeen täytyy valita mikä kappale nostetaan sekä määrittää kappaleen tarttumispisteet robotin tarttujalle. Tyypillinen tapa valita tartuttava kappale on ottaa päällimmäinen ensin, mutta valinta ei ole aina selkeä, jos monta kappaletta on päällimmäisenä, päällimmäinen kappale on osittain jonkun toisen kappaleen alla, tai robotin tarttuja osuisi laatikon reunoihin tai viereisiin kappaleisiin poimittaessa (Dupuis et al. 2008).

Kappaleen tunnistamisen ja valinnan lisäksi robotin liikkeen suunnittelu tulisi tehdä siten, että törmäyksiä ympäristön kanssa ei tapahtuisi, muille osille ei tulisi vahinkoa kappaletta poimittaessa, eivätkä kappaleet jumittuisi toisiinsa kiinni. Liikerata vaikuttaa myös poiminnan nopeuteen sekä sykliin, mitkä ovat kriittisiä saada mahdollisimman tehokkaiksi teollisissa sovelluksissa. Liikeradan määrittely voidaan tehdä joko käsin valittuja ja kovakoodattuja ratoja käyttäen tai käyttäen automaattista liikeradan suunnittelua (Ellekilde et al. 2013). Kameran sijoittelulla onkin olennainen osa siinä, että edellä mainittuja tavoitteita voidaan saavuttaa. Monissa sovelluksissa ympäristöön sijoitettu kamera voi olla riittävä tarpeellisen informaation saamiseksi, mutta joissain sovelluksissa kuvan ottaminen useammasta kulmasta voi olla tarpeen.

5. ANALYYSI JA POHDINNAT

Ensimmäiseen tutkimuskysymykseen vastauksena on, että nykyaikaiset kaupalliset teolliseen tarkoitukseen käytettävät 3D-konenäkökamerat vaikuttaisivat käyttävän useimmiten joko lentoaikamenetelmää, strukturoitua valoa tai stereonäköä. Esimerkiksi fotometristä stereota on käytetty ensimmäisiä kasasta poiminta -sovelluksia kehitettäessä 1980-luvulla, mutta uudemmissa se on harvinaisuus. Muiden menetelmien käyttö saattaa olla myös todella vaikeaa käytännön kappaleenkäsittelysovelluksissa.

Toiseen tutkimuskysymykseen työssä löydettiin useita huomioonotettavia asioita. Kameratyyppiä valitessa on tärkeää ottaa huomioon, että millainen kuvattava kappale on, millaiset pinnanmuodot siinä on ja onko siinä heijastavia pintoja. Stereonäköä käytettäessä voi olla esimerkiksi vaikeaa löytää vastinpisteitä kaikenlaisista kappaleista ja ympäristön valaistukseen pitää olla sopiva, kun taas lentoaikamenetelmää käyttävässä kamerassa ympäristön valaistuksella saattaa olla pienempi merkitys. Kameran sijoitteluun ja valaistuksen toteutukseen on myös syytä kiinnittää huomiota sovellusta toteutettaessa. Pelkkä käytetty kuvantamismenetelmä ei takaa onnistunutta lopputulosta, vaan myös itse kameran laadulla on suuri vaikutus sen muodostaman mallin hyvyyteen. Muodostetun mallin pitää olla riittävän hyvälaatuinen kappaleen tunnistamista varten.

Erilaiset poimintasovellukset kuten satunnainen kasasta poimiminen voivat säästää aikaa ja rahaa teollisissa sovelluksissa sekä tuovat joustavuutta tuotantoon. Sovellusten toteuttaminen vaatii kuitenkin tarkkaa suunnittelua ja erilaisten ongelmakohtia aiheuttavien seikkojen huomioonottamista, joita työssä on käsitelty.

Työssä käsiteltiin melko vähäisesti itse poimintasovelluksen toteutukseen liittyviä asioita kuten kappaleen valintaa, liikeradan suunnittelua, erilaisten kuvantamismenetelmien vertailua erilaisilla kappaleilla sekä valaistuksilla, tai uusinta alan tekniikka. Esitellyt menetelmät ovat pintaraapaisu tunnetuimmista 3D-kuvantamismenetelmillä saadun informaation käsittelyyn käytetyistä menetelmistä, mutta työ antaa kuitenkin hyvät valmiudet syventyä aiheisiin lisää.

6. YHTEENVETO

Tässä työssä käsiteltiin erilaisia menetelmiä syvyysinformaation saamiseksi 3D-kameralla, kappaleen tunnistamista kuvasta 3D-mallista vertaamalla sitä tunnettuun CAD-malliin sekä mihin näitä menetelmiä voidaan käyttää teollisissa sovelluksissa. Nykyaikaiset kaupalliset 3D-konenäkökamerat perustuvat usein joko strukturoituun valaistukseen, lentoaikamenetelmään tai stereonäköön. Eri menetelmien etuja toisiinsa nähden erilaisissa sovelluksissa täytyy arvioida kuitenkin tapauskohtaisesti. Suuri vaikutus on esimerkiksi kuvattavan kappaleen pinnan muodoilla, heijastavuudella sekä ympäristön valaistuksella.

Kappaleen tunnistamiseen ja pistepilven suodattamiseen on olemassa valtava määrä erilaisia menetelmiä ja uusia kehitellään jatkuvasti. Tässä työssä käsitellyistä rekisteröintiin käytettävästä ICP-algoritmista on esimerkiksi jatkokehitetty monia eri variaatioita. Kuvapohjaisissa menetelmissä kuten stereonäössä hajapisteiden suodatus voi olla tärkeämpää kuin esimerkiksi lentoaikamenetelmään perustuvassa 3D-kuvantamismenetelmässä. Hajapisteiden suodatukseen työssä on mainittu eri menetelmiä kuten pääkomponenttianalyysi, neuroverkot ja RANSAC. Myös käytetty tunnistamistapa vaikuttaa siihen onko tunnistaminen herkkä hajapisteille.

Kasasta poiminta -sovellusten ansiosta valmistusjärjestelmistä voidaan saada joustavampia sekä voidaan säästää aikaa ja rahaa. 3D-kuvantamismenetelmät ovat olennaisessa osassa tällaisissa sovelluksissa, ja työssä käytiin läpi erilaisia huomioonotettavia asioita kasasta poiminta -sovelluksen toteutuksessa, kuten kameran sijoittelu ja poimitavan kappaleen valinta, sekä miten eri kuvantamis- ja tunnistusmenetelmiä on hyödynnetty eri käytännön sovelluksissa.

LÄHTEET

Barrow, H.G., Tenenbaum, J.M., Bolles, R.C., Wolf, H.C. (1977). Parametric correspondence and chamfer matching: two new techniques for image matching. IJCAI'77: Proceedings of the 5th international joint conference on Artificial intelligence, Vol. 2, pp. 659-663.

Besl, P.J., McKay, N. D. (1992). A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, pp. 239-256.

Bletterer, A., Payan, F., Antonini, M., Meftah, A. (2016). Point Cloud Compression using Depth Maps. Electronic Imaging. 1-6. 10.2352/ISSN.2470-1173.2016.21.3DIPM-397.

Chen, F., Brown, G.M., Song, M. (2000). Overview of 3-D shape measurement using optical methods. Optical Engineering, Vol. 39(1), pp. 10-22.

Chen, Y. -K., Sun, G. -J., Lin, H. -Y., Chen, S. -L. (2018). Random Bin Picking with Multi-view Image Acquisition and CAD-Based Pose Estimation, 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2218-2223, doi: 10.1109/SMC.2018.00381.

Chen, Y., Medioni, G. (1991). Object modeling by registration of multiple range images. Proceedings. IEEE International Conference on Robotics and Automation, Vol. 3, pp. 2724-2729.

Choi, C., Christensen, H. (2012). 3D Textureless Object Detection and Tracking: An Edge-based approach. IEEE International Conference on Intelligent Robots and Systems. 10.1109/IROS.2012.6386065.

Corke, P. (2017). Robotics, Vision and Control. Fundamental Algorithms in MATLAB®. Springer-Verlag Berlin Heidelberg.

Deping L., Ning L., Yulan G., Xiaoming W., Jin X. (2019). 3D object recognition and pose estimation for random bin-picking using Partition Viewpoint Feature Histograms, Pattern Recognition Letters, Vol. 128, pp. 148-154, ISSN 0167-8655.

Dupuis, D., Léonard, S., Baumann, M., Croft, E, Little, J.J. (2008). Two-Fingered Grasp Planning for Randomized Bin-Picking.

Ellekilde, L-P., Petersen, H. G. (2013). Motion planning efficient trajectories for industrial bin-picking. The International Journal of Robotics Research.

Fofi, D., Sliwa, T. Voisin, Y. (2004). A comparative survey on invisible structured light. Machine Vision Applications in Industrial Inspection XII, 90.

Franca, J. G. D. M., Gazziro, M. A., Ide, A. N., Saito, J. H. (2005). A 3D scanning system based on laser triangulation and variable field of view. IEEE International Conference on Image Processing 2005, pp. 1-425.

- Gan, Z., Tang, Q. (2011). *Visual Sensing and its Applications*. Springer Berlin Heidelberg.
- Gold, S., Rangarajan, A., Lu, C-P., Suguna, P., Mjolsness, E. (1998). New algorithms for 2d and 3d point matching. pose estimation and correspondence. *Pattern Recognition*, Vol. 31, Is. 8, pp. 1019-1031.
- Gordon, G. Buck, R. (2004). Tyxz DeepSea High Speed Stereo Vision System. *Computer Vision and Pattern Recognition Workshop. CVPRW '04. Conference on*, Washington, DC, USA, 2004, pp. 41.
- Hartley, R., Zisserman, A. (2004). *Multiple View Geometry in computer vision*. Cambridge University Press.
- Hema, C.R., Radhakrishnamurthy, P., Murugesapandian, N. & Ramachandran, S., Yaacob, S. (2007). *Stereo Vision System for A Bin Picking Adept Robot*.
- Herakovic, N. (2010). *Robot Vision in Industrial Assembly and Quality Control Processes*. Robot Vision, Ude, A. (Ed.), InTech.
- Hinterstoisser, S., Holzer, S., Cagniart, C., Ilic, S., Konolige, K., Navab, N., Lepetit, V. (2011). Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes. *International Conference on Computer Vision*.
- Hutchinson, S., Hager, G. D., Corke, P. I. (1996). A Tutorial on Visual Servo Control. *IEEE Transactions on Robotics and Automation*, Vol. 12, no. 5, pp. 651-670.
- Jang, J., Hwnag, S., Park, K. (2013). Design of Indirect Time-of-Flight Based Lidar for Precise Three-Dimensional Measurement Under Various Reflection Conditions. *SENSORDEVICES 2013, The Fourth International Conference on Sensor Device Technologies and Applications*.
- Kleeberger, K., Landgraf, C., Huber, M. F. (2019). Large-scale 6D Object Pose Estimation Dataset for Industrial Bin-Picking, 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2573-2578, doi: 10.1109/IROS40897.2019.8967594.
- Loh, A. M. E., Hartley, R. (2005). Shape from non-homogeneous, non-stationary, anisotropic, perspective texture. *British Machine Vision Conference 2005, BMVA*, Vol. 1., pp. 69-78.
- Lopez Paredes, A. Song, Q., Conde, M. H. (2023), Performance Evaluation of State-of-the-Art High-Resolution Time-of-Flight Cameras, in *IEEE Sensors Journal*, Vol. 23, no. 12, pp. 13711-13727, 15 June15, doi: 10.1109/JSEN.2023.3273165.
- Martinez, C., Boca, R., Zhang, B., Chen, H., Nidamarthi, S. (2015). Automated bin picking system for randomly located industrial parts, *IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, pp. 1-6, doi: 10.1109/TePRA.2015.7219656.
- Mitra, N., Gelfand, N., Pottmann, H., Guibas, L. (2004). Registration of Point Cloud Data from a Geometric Optimization Perspective. *ACM International Conference Proceeding Series*. 71, pp. 23-32.
- Nayar, S. K., Gupta, M. (2012). Diffuse Structured Light. *IEEE International Conference on Computational Photography*, pp. 1-8.

- Pears, N., Liu Y., Bunting P. (2012). 3D Imaging, Analysis and Applications, Springer London.
- Pertuz, S., Puig, D., García, M. (2012). Analysis of focus measure operators in shape-from-focus. *Pattern Recognition*. 46. 10.1016/j.patcog.2012.11.011.
- Salvi, J., Fernandez, S., Pribanic, T., Llado, X. (2010). A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, Vol. 43, Iss. 8, pp. 2666-2680.
- Salvi, J., Pagès J., Batlle, J. (2003). Pattern codification strategies in structured light systems. *Pattern Recognition*, Vol. 37, Iss. 4, pp. 827-849.
- Sanchez, J., Florence D., David C., Florent D., Laurent T., Paul C. (2020). Robust normal vector estimation in 3D point clouds through iterative principal component analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 163, pp. 18-35, ISSN 0924-2716
- Sharp, G.C., Lee, S.W., Wehe, D. K. (2002). ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 90-102.
- Song, Z. (2013). *Handbook Of 3D Machine Vision: Optical Metrology And Imaging*, CRC Press, Boca Raton.
- Tagare, H.D., deFigueiredo, R.J.P. (1991). A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, Iss. 2, pp. 133-152.
- Taylor, G., Kleeman, L. (2006). *Visual Perception and Robotic Manipulation*. Springer Berlin Heidelberg.
- Van Herk, M. (2009). Chapter 35 - Image Registration Using Chamfer Matching, Editor(s): ISAAC N. BANKMAN, *Handbook of Medical Image Processing and Analysis (Second Edition)*, Academic Press, pp. 591-603, ISBN 9780123739049, doi: 10.1016/B978-012373904-9.50045-3.
- Wang, J., Wilson W. J. (1992). 3D relative position and orientation estimation using Kalman filter for robot control. *Robotics and Automation. Proceedings.*, 1992 IEEE International Conference on, Vol. 3, pp. 2638-2645.
- Wilson, W. J., Williams Hulls, C. C., Bell, G. S. (1996). Relative end-effector control using Cartesian position based visual servoing. *IEEE Transactions on Robotics and Automation*, Vol. 12, no. 5, pp. 684-696.
- Wolff K., Kim, C., Zimmer, H., Schroers, C., Botsch, M., Sorkine-Hornung, O., Sorkine-Hornung, A. (2016). Point Cloud Noise and Outlier Removal for Image-Based 3D Reconstruction, 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, pp. 118-127, doi: 10.1109/3DV.2016.20.
- Yoon, H., Song, H., Park, K. (2011). A phase-shift laser scanner based on a time-counting method for high linearity performance. *Review of Scientific Instruments*, 82.