

Prosjektrapport



Automato

Teknologi for 3D-avbildning av tomatplanter – en nøkkel til fremtidige systemer for tilnærmet helautomatisk robotisert stell og høsting av tomatplanter.

PROSJEKTNR	VERSJON	DATO
302005016-2	1.0	2021-12-15

FORFATTER(E)
Elling Ruud Øye, Morten Bondø, John Reidar Mathiassen

OPPDRAGSGIVER(E)	OPPDRAGSGIVERS REF.
Landbruksdirektoratet	2019/40581, Agros 120377

GRADERING	ANTALL SIDER OG VEDLEGG:
Åpen	17

SAMMENDRAG

Prosjektets formål har vært å definere en teknologi for 3D-avbildning av tomatplanter, som har potensiale til å identifisere og lokalisere tomater, blader, greiner, knopper, skudd og blomsterklaser i 3D – i tilstrekkelig nøyaktighet til at en robot kan stelle og høste en tomatplante. Prosjektets finansieringskilde er Forskningstilne midlene for jordbruk og matindustri.

Avbildning i drivhus har vist seg å være utfordrende. Belysning, tett og uoversiktlig plantemasse som gir vanskelig tilgjengelighet og måten drivhusene er tradisjonelt sett er organisert, gjør avbildning utfordrende. I prosjektet har flere forskjellige 3D-kameraer blitt testet for å avbilde forskjellige deler av tomatplanten. Det har vist seg at få maskinsynsystemer alene oppfyller alle behov for å kunne gi en robot nok informasjon til å kunne stelle og høste en tomatplante, men ved å kombinere flere systemer har man likevel kommet frem til løsninger som sammen gir gode nok data til at oppgaver som å identifisere og lokalisere tomater, blader, greiner, skudd og blomster er gjennomførbare.

Kombinasjonen som ble testet er en Azure Kinect satt sammen med en stereokamera-rigg bestående av to fargekamera med høy oppløsning. Kombinasjonsriggen ble testet hos Mære Landbruksskole.



HOVEDFORFATTER
Elling Ruud Øye

GODKJENT AV
Kirsti Greiff

Historikk

VERSJON	DATO	VERSJONSBEKRIVELSE
1.0	2021-12-15	Avsluttende rapport

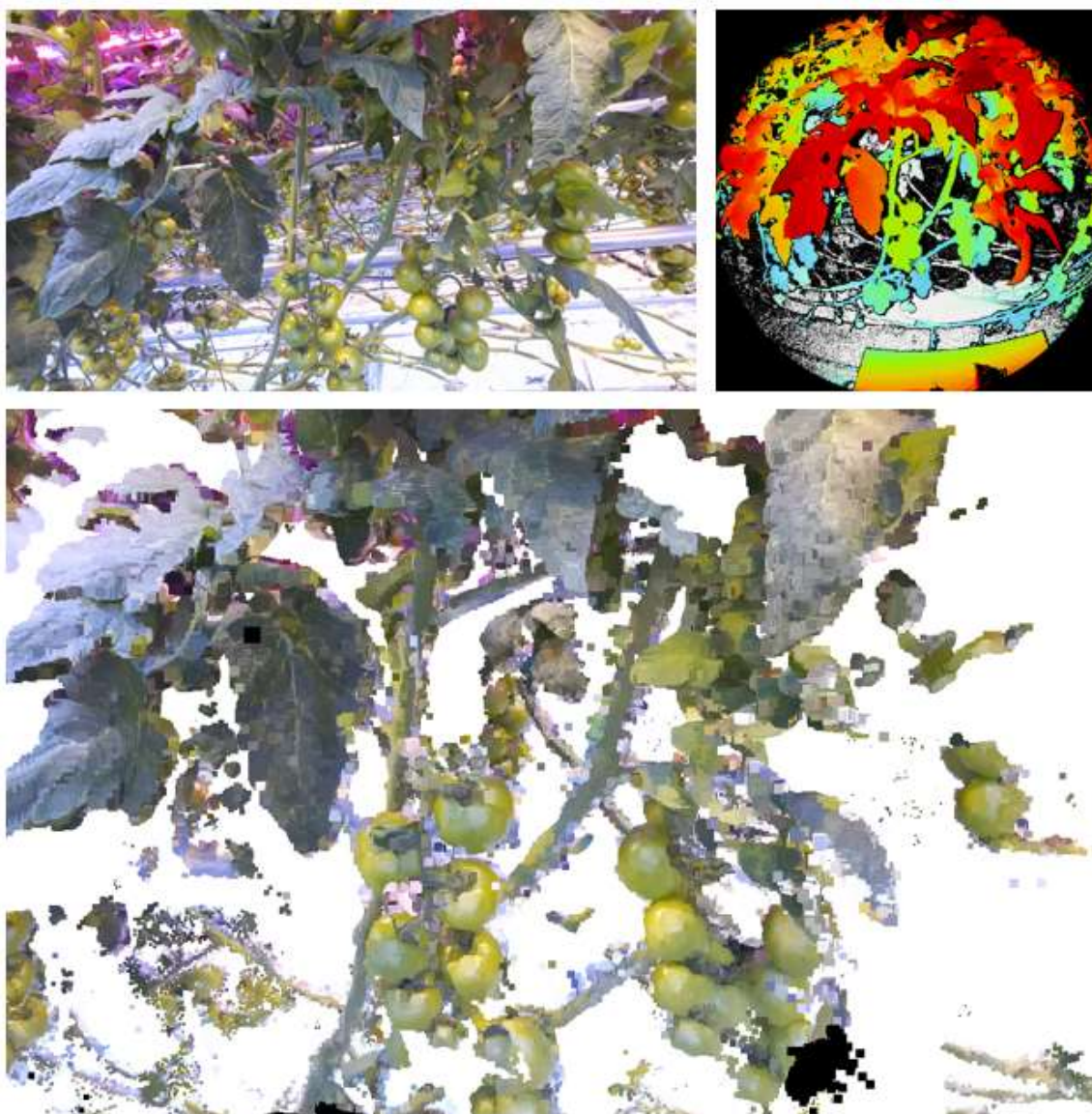
Innholdsfortegnelse

1	Introduksjon	4
2	Test av forskjellige sensorer	5
2.1	Generelt om avbildning i drivhus	5
2.2	Test med Zivid One+ Medium i drivhus.	5
2.3	Test med RealSense L515.....	7
2.4	Test med Azure Kinect	7
2.5	Test av stereoskopirigg med speilreflekskamera.....	9
2.6	Test av kombinasjonsoppsett med Kinect og stereokamera.....	11
3	Diskusjon	14
3.1	Generelt om avbildning i drivhus	14
3.2	Vurdering av kamera- og maskinsynsystem	15
4	Konklusjon	17
5	Veien videre	17

1 Introduksjon

Det har i prosjektet blitt testet mange forskjellige avbildningsmetoder med fokus på å avbilde de forskjellige bestanddelene på tomatplanten. Bestanddelene som har blitt fotografert er modne og umodne tomater, stengel uten bladvekst, bladvekst, tyver (uønskede skudd) og blomster.

På de forskjellige kameraene har det så langt det har latt seg gjøre blitt tatt bilde av de forskjellige bestanddelene for å vurdere hvor godt utstyret er egnet for å gi en avbildning som kan brukes til automatisk analysering, og for å vurdere om målingene kan brukes for å styre en robot som skal gjøre arbeidsoperasjoner på det som observeres. Å identifisere bestanddelene av interesse krever gjerne gode fargebilder, mens for å kunne styre en robot til å utføre arbeidsoperasjonene gjerne krever 3D-avbildning. Slik vet robotsystemet hvor de observerte bestanddelene befinner seg. Dette oppnås med sensorer som gir såkalte RGBD-bilder. Dette vil si at man henter inn avstanden pikslene i bildet har til kameraet i tillegg til farge. Figur 1 viser et eksempel på slike bilder fra en Azure Kinect.



Figur 1: Øverst til venstre - fargebilde. Øverst til høyre – dybdebilde hvor fargen forteller hvor nærme, eller fjern, hvert punkt er fra kameraet. Nederst – punktsky med punkter som inneholder fargeinformasjon og posisjon i et x, y, z-koordinatsystem

Siden målet med dette prosjektet er å *definere en teknologi for 3D-avbildning av tomatplanter, som har potensiale til å identifisere og lokalisere tomater, blader, greiner, knopper, skudd og blomsterklaser i 3D – i tilstrekkelig nøyaktighet til at en robot kan stelle og høste en tomatplante*, har vi valgt å teste et sett med forskjellige sensorer. Ved å visuelt inspisere resultatene skal vi avgjøre om sensorene er i stand til å både gi gode nok fargebilder til at man kan gjenkjenne viktige deler av tomatplanten, samtidig som 3D-informasjonen er tilstrekkelig til at man kan styre en robot eller andre aktuatorer til å utføre en nøyaktig operasjon på disse. Det er naturlig å tenke seg at ingen sensor er i stand til å alene kunne fylle samtlige krav for alle operasjoner, så vurderingen vil i tillegg se på hvilke sensorer som kan utfylle hverandre.

2 Test av forskjellige sensorer

Sensorene / teknologiene som ble testet under dette forsøket er følgende:

- Zivid One+ Medium
- Azure Kinect
- RealSense L515
- Stereo-rigg med 2 stk. Nikon D800e speilreflekskameraer
- Kombinasjonsoppsett med Azure Kinect og stereokamera

All avbildning har skjedd hos Mære Landbruksskole.

2.1 Generelt om avbildning i drivhus

Mære Landbruksskole har et ganske typisk layout i drivhuset, der tomatplantene vokser i parvise rader, med en senter-avstand på 1,2 meter og en plantehøyde på omtrent 4 meter. Dette gjør at åpen avstand mellom radene er ca. 60 cm, og man kan observere hver enkelt plante kun fra en side (se Figur 2). Dette setter begrensninger på hvordan man kan avbilde plantene og man er avhengig av et nokså bredt synsfelt for å se. Noen planter er likevel plassert slik at man kan avbilde på lenger avstand, så dette ble det også gjort forsøk på.

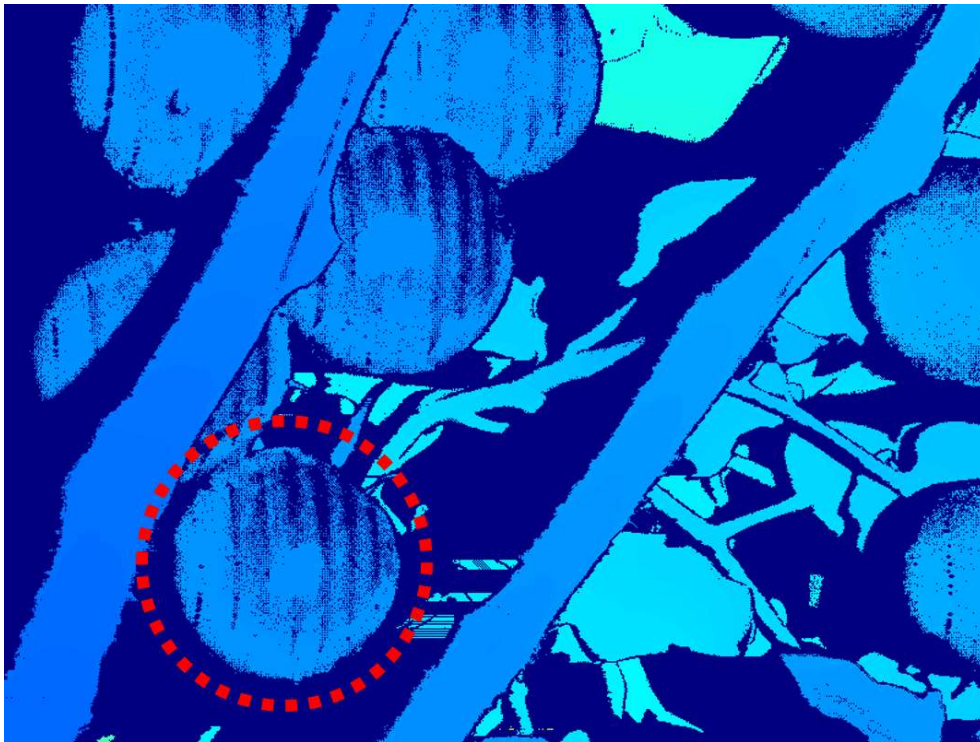
2.2 Test med Zivid One+ Medium i drivhus.

Zivid leverer industrielle RGBD-kameraer for forskjellige applikasjoner. Kameraene gir god oppløsning og nøyaktighet, men krever ofte at objektene står stille over en kort periode. I områder med krevende belysning (som et drivhus) kan denne perioden med fordel justeres opp til 1,5 – 2 sekund. Under testene ble kameraet plassert på et stativ og manuelt flyttet til forskjellige posisjoner foran en tomatplante. Det ble under avbildning registrert at lysforholdene i drivhuset sammen med tomatenes blanke overflate gjorde det problematisk å få gode avlesninger av dybde på selve tomatene.



Figur 2: Oppsett av maskinsynsensor mellom to rader med tomatplanter

Resten av planten, som blader og stilker, synes godt og får gode avlesninger for dybde (se Figur 3). For å motvirke den dårlige avlesningen av deler av selve tomatene, ble flere bilder (og punktskyer) fra kameraet sydd sammen til en større 3D-modell av planten. Slik kompenseres man for enkelte dårlige målepunkt og ender like vel opp med en svært nøyaktig modell av hele planten (se Figur 4).



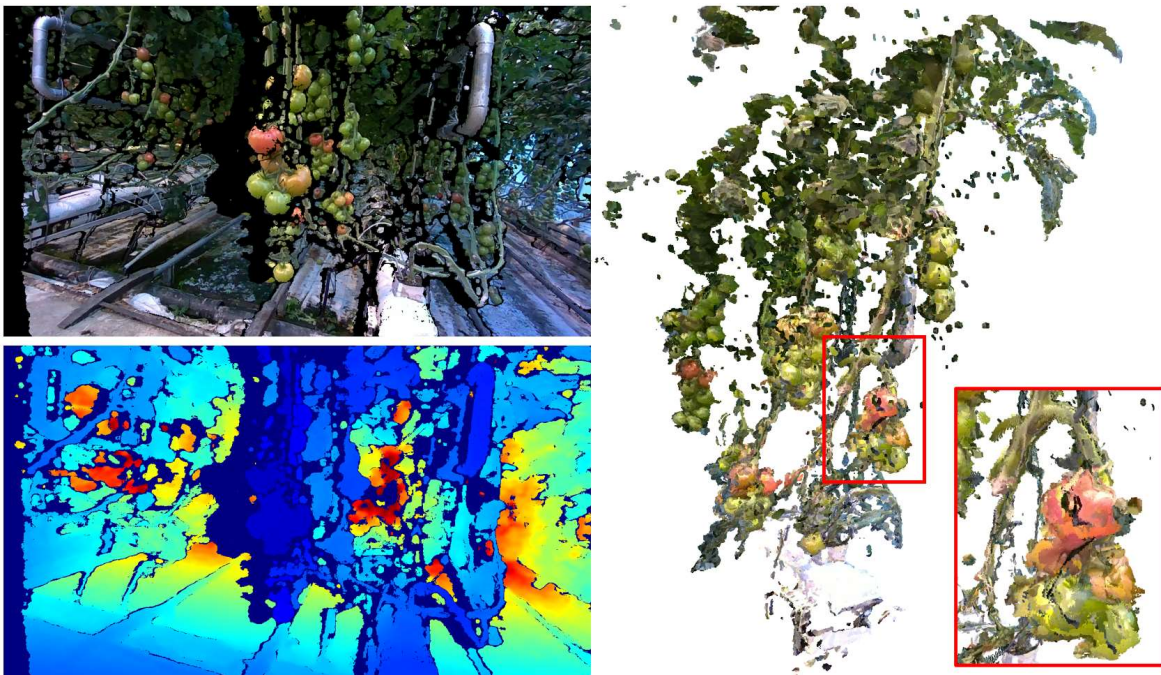
Figur 3: Dybdebilde fra Zivid-kamera. Rød ring rundt eksempel av tomat med dårlig dybdeavlesning.



Figur 4: Punktsky av tomatplante bestående av flere bilder tatt av Zivid-kamera. Rødt vindu viser utsnitt hvor detaljer på både stilk og tomater er godt synlig.

2.3 Test med RealSense L515

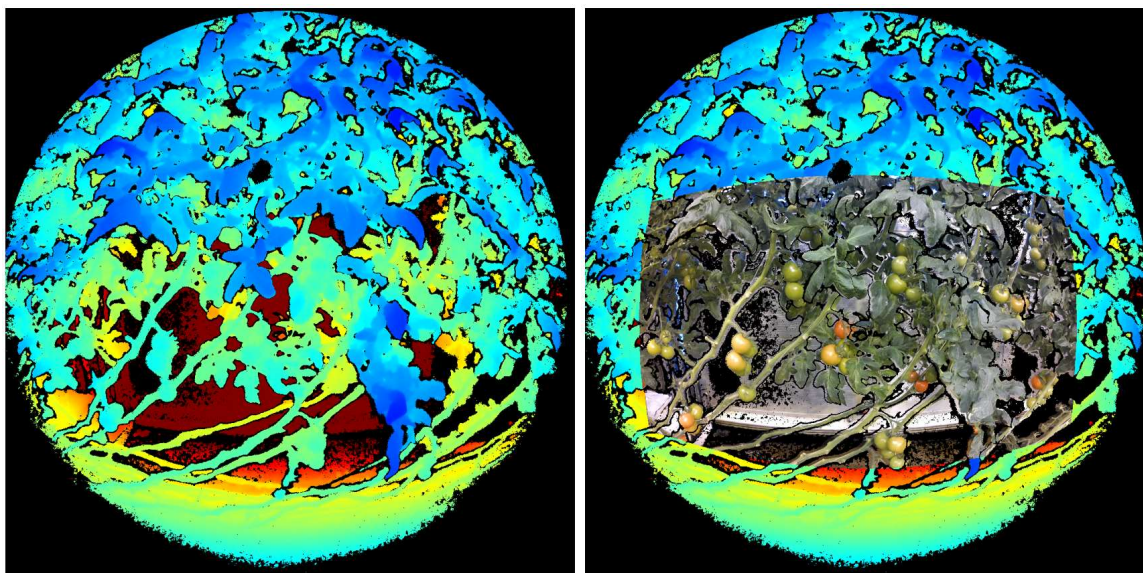
RealSense L515 er et 3D-kamera som kombinerer en LiDAR-skanner og et fargekamera. LiDAR-en henter ut 3D-målinger for scenen ved å sende ut mange punktmålinger med laser på svært kort tid, og fargekameraet brukes til å hente ut fargeinformasjon fra de samme områdene. Dette er et RGBD-kamera i en prisklasse mer tilpasset en lavkostløsning eller løsning med mange kamera. Nøyaktigheten og oppløsningen kan ikke måle seg med mer profesjonelle kameraer som Zivid, men L515 er tilgjengelig kjappere, og tåler derfor mer bevegelse i bildet. Ulempen med LiDAR-kameraet er at målingene er preget av en god del støy. Dersom man ser på dybde-bildet nederst til venstre i Figur 5, ser det ut til at man får gode avlesninger for hele scenen. Det dybdebildet ikke viser er at selv om man får avlesning i nesten hele bildet, så er disse avlesningene preget av mye målestøy. Dersom man ser på området uthevet i rødt fra punktskyen til høyre i Figur 5, ser man at formen på tomatklasen er svært forvrengt. Bildene viser at man med dette kameraet kan detektere store objekt som bladverk og stilk, men at deteksjoner som krever mer nøyaktighet (som deteksjon av enkelt-tomater i en klase eller tyv/skudd) blir vanskelig fordi detaljene drukner i målestøy.



Figur 5: Eksempelbilder fra RealSense L515: Øverst til venstre – dybdebilde med fargeinformasjon. Nederst til venstre dybdebilde med avstandsmåling. Til høyre – sammensydd punktsky.

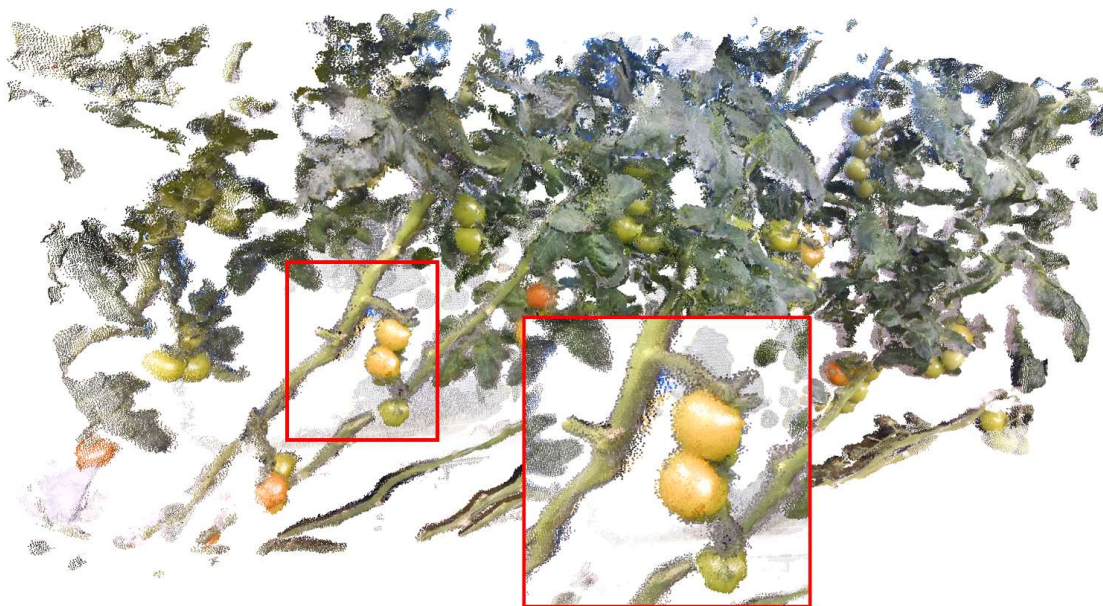
2.4 Test med Azure Kinect

Azure Kinect benytter et 3D-kamera med vidvinkel-linse sammen med et fargekamera med et noe mindre synsfelt for å til sammen gi både farge og 3D-data fra scenen den observerer. Det forholdsvis store synsfeltet gir muligheten for å få data på en stor del av planten med få bilder. Samtidig fører fargekameraets mindre synsfelt til at ikke all 3D-data får fargeinformasjon. Dette er illustrert i Figur 6.



Figur 6: Til venstre - dybdebilde av tomatplante med Azure Kinect.
Til høyre – område av dybdebildet som også er dekket av fargekameraet.

Det er likevel slik at selv det mindre området som både har farge- og 3D-informasjon, fremdeles er betydelig større enn det de andre kameraene i testen. Derfor er fremdeles Kinect et godt alternativ for rask oversikt over planten. Riktignok med en del lavere nøyaktighet og oppløsning enn Zivid-kameraet.



Figur 7: Punktsky av tomatplante bestående av flere bilder tatt med Azure Kinect.
Rødt vindu viser utsnitt hvor detaljer på både stilk og tomater er godt synlig.

2.5 Test av stereoskopirigg med speilreflekskamera

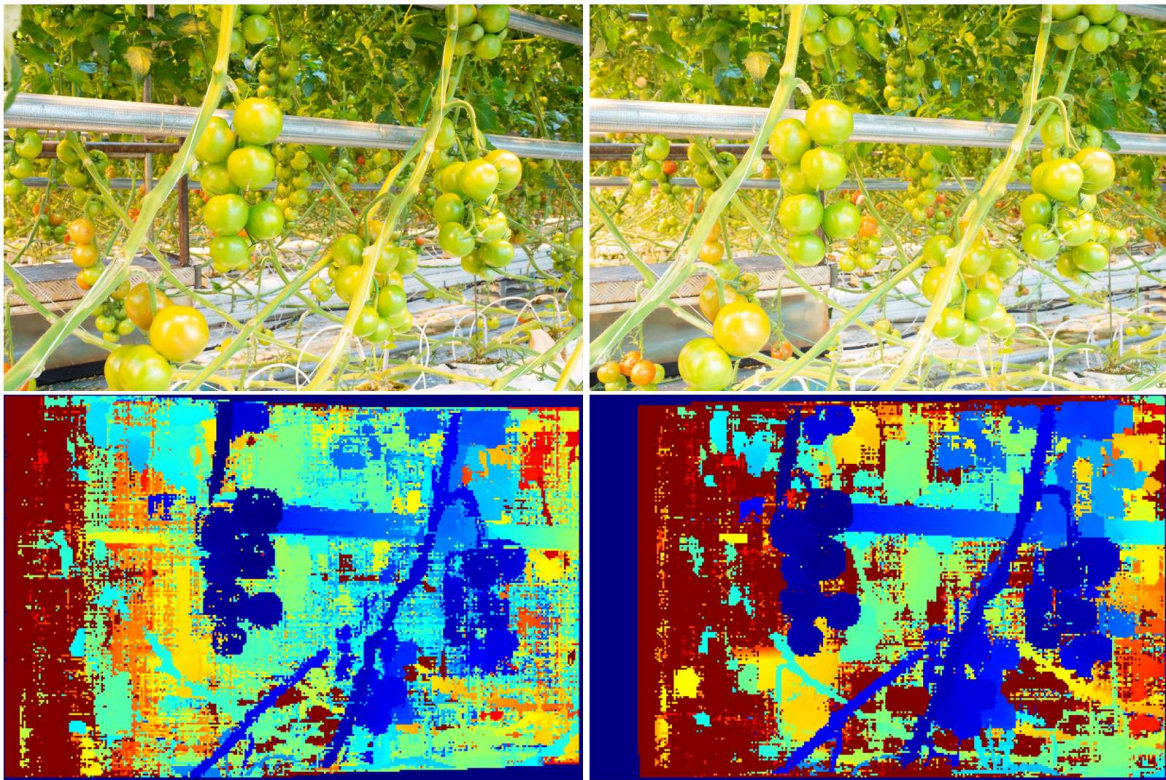
Det ble lagd en rigg med to Nikon D800e speilreflekskameraer. Disse har 36.3MP oppløsning og tar 14-bits råformatbilder. De ble brukt med to identiske 35mm F1.8 fastobjektiver. Disse ble montert så nærme som mulig, kameraene med kollimert synsfelt på én meters avstand. 2D-bildekvaliteten fra disse viste seg å være svært god, som forventet. 3D-måling fra dette oppsettet viste seg å være problematisk. På grunn av den fysiske størrelsen på kameraene blir de plassert med så stor avstand at perspektivforandringer mellom de to bildene som tas blir så stor at stereomatching er vanskelig. De store bildesensorene gjør også at dybdeskarpheten fra disse bildene blir begrenset, slik at det som observeres bør ha mest mulig lik avstand til kameraene.



Figur 8: Bilde av avbildning med stereo-rigg bestående av to Nikon D800e speilreflekskameraer

Eksempel på resultat fra stereo-matching kan sees i Figur 9. De nederste bildene viser et såkalt "disparity-map" for hvert bilde. Dette viser avstanden algoritmen tror det er mellom samme punkt i hvert i bilde. Denne avstanden brukes til å beregne 3D ved hjelp av triangulering. Som man kan se er denne matchingen ikke perfekt.

Det var derfor hensiktsmessig å heller se på et sett av bilder (som vist i Figur 10) der en fotogrammetrialgoritme bruker flere bilder tatt fra forskjellige vinkler. Dette gjør at en del av støyen fra punkt-matchingen forsvinner siden man kan filtrere ut punkt som bare matcher på et par av bildene. Bruken av stereokamera er likevel hensiktsmessig fordi man kan detektere enkelt-punkt (som f.eks. en blomst) i begge bildene med høy nøyaktighet. Dersom man først er sikker på at et punkt i venstre og høyre peker på samme objekt, er trianguleringen fra stereo-synet svært nøyaktig og kan gi posisjoner med nøyaktighet helt ned til under en millimeter.



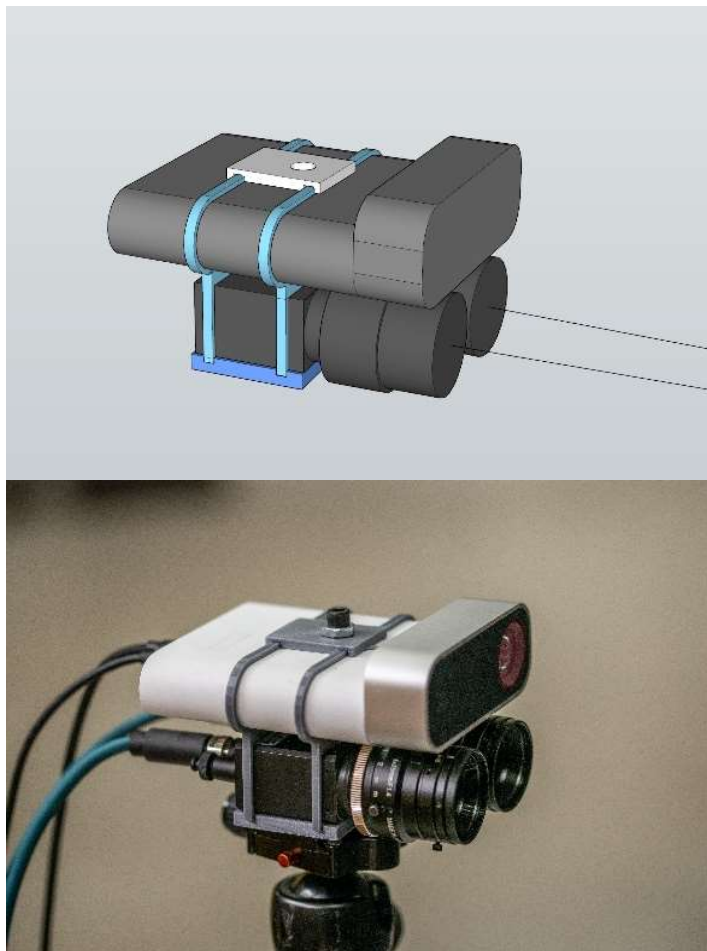
Figur 9: Stereo-matching.
 Øverst – fargebilder fra venstre og høyre kamera.
 Nederst tilhørende disparity-maps som viser stereo-matching mellom hver piksel i bildene over.



Figur 10: Fotogrammetri.
 Til venstre – bildene som brukes av fotogrammetrialgoritmen.
 Til høyre – den resulterende punktskyen

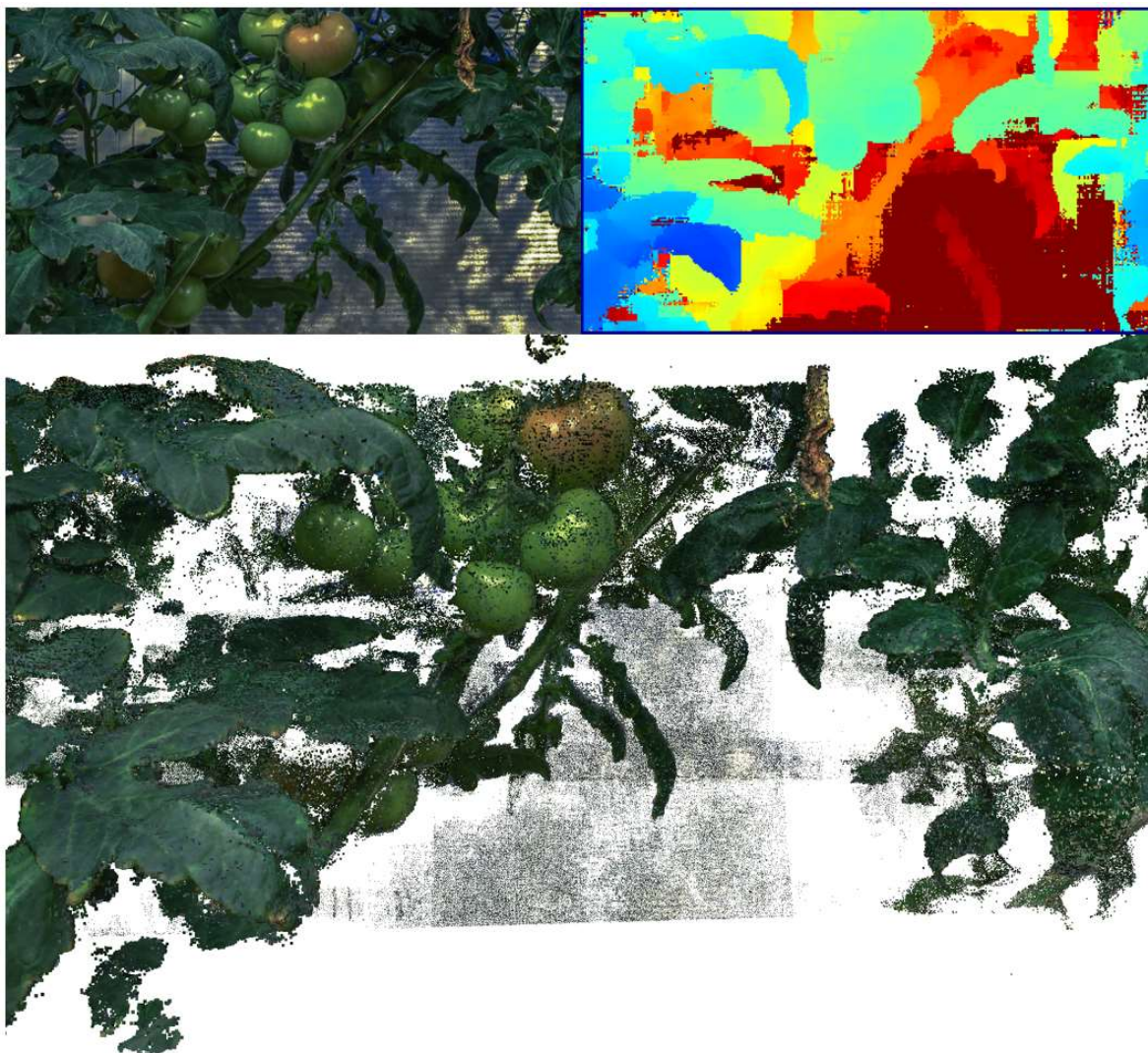
2.6 Test av kombinasjonsoppsett med Kinect og stereokamera

Ut fra de foregående forsøkene og resultatene ble det vurdert hva som var mest lovende og hvordan oppsettene kunne bli utbedret for å få best mulig bilder. Blant annet ble det vurdert forskjellige måter for å oppnå mindre perspektivforandringer på en stereorigg. Det ble vurdert løsninger der kamera flyttes en kort avstand mellom hvert bilde. Endelig løsning ble likevel å sette opp to kamera med mindre avstand mellom seg som tar bilder synkront. Det ble kjøpt inn to Lucid Triton TRI162S-CC kameraer som er bare 29x29mm store, noe som gjør at de kan plasseres svært nærme hverandre. Disse har høy oppløsning på 16.2MP og tar fargebilder. Disse to kameraene ble satt sammen i et oppsett som var kollimert noe nærmere enn speilrefleksriggen og med 25mm linser som gir et lite synsfelt, men høy presisjon på 3D-målinger og ser mange detaljer på plantene. Se Figur 11 for 3D-modell og bilde av det ferdige oppsettet. Disse kameraene gir en bedre dybdeskarpheit enn oppsettet med speilrefleks og vil derfor være egnet til å detektere ting som befinner seg både litt lengre unna og nærmere kameraet enn fokuspunktet. Siden stereokameraene gir et så lite synsfelt ble de satt sammen med en Azure Kinect, som i innledende tester viste seg å være god for å få oversiktsbilder. Formålet med denne kombinasjonen er at Kinect-kamera skal kunne benyttes i et robotsystem for å navigere og finne interessepunkter og stereokameraene kan avbilde interessepunktene med høy presisjon for å kunne gjøre gode beslutninger og styring av roboten for å utføre arbeidsoppgavene.



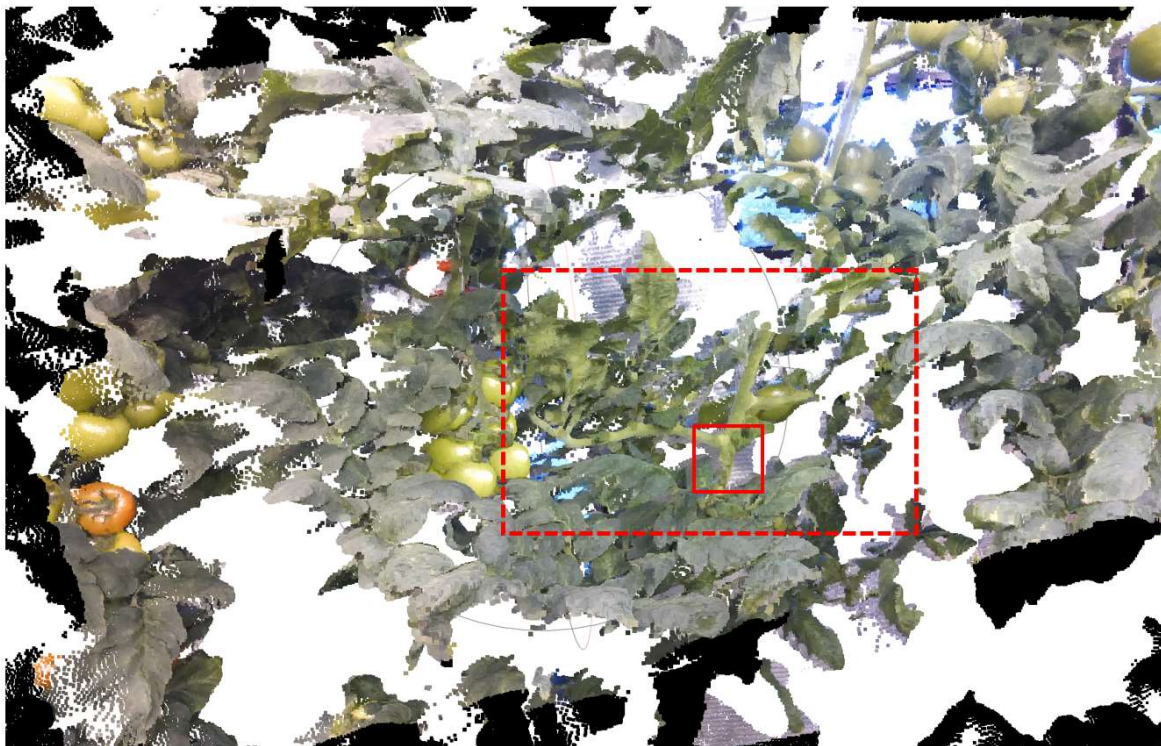
Figur 11: Kombinasjonsoppsett av Azure Kinect og stereokamera

Den noe mindre avstanden mellom kameraene i stereo-delen viste seg å gi noe bedre matching av punkter fra de forskjellige vinklene. Sammenligner man "disparity-map" fra Figur 9 og Figur 12, ser man at bildet i Figur 12 er noe mindre støyete og inneholder en del mer detaljer. Det er likevel slik at de beste 3D-modellene først oppnås ved å sy sammen flere bilder enn ett stereopar på samme måte som det ble gjort i den første stereo-riggen. Resultatet, som kan sees nederst i Figur 12 gir svært god nøyaktighet på store deler av planten som er i synsfeltet.

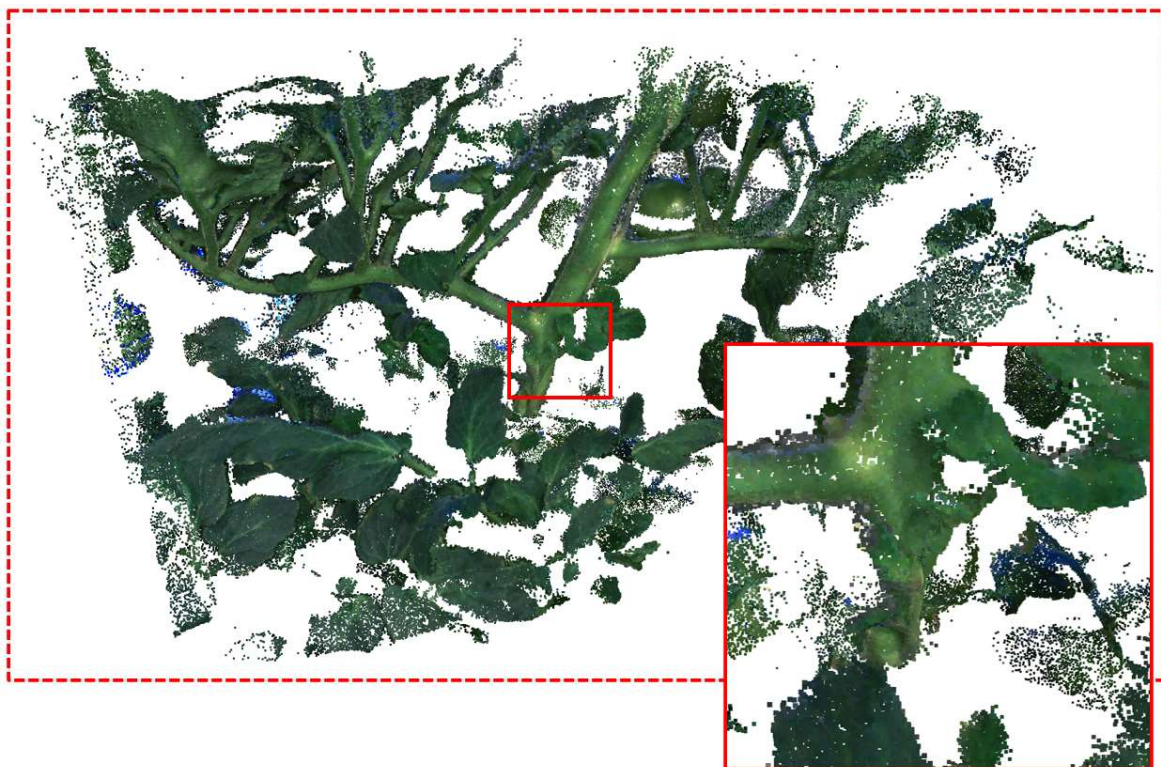


*Figur 12: Stereobilde og punktsky.
 Øverst til venstre – enkeltbilde fra stereorigg
 Øverst til høyre – disparity map fra samme kamera
 Nederst – resulterende punktsky*

Som en siste case ble en tyv (uønsket skudd i en forgrening i planten) avbildet med kamerariggen fra forskjellige vinkler. Målet var å teste nøyaktighet både fra Azure Kinect og stereo-riggen og hvordan disse bildene kan brukes sammen. I Figur 13 ser man først oversiktsbildet fra Azure Kinect (i hovedsak bare den delen av punktskyen som kameraet i tillegg dekker med fargekameraet. I Figur 13 er så synsfeltet til stereoriggen lagt inn som et rødt rektangel med stiplede linjer. Videre er området hvor tyven kan observeres markert med et rødt rektangel med hele linjer. I Figur 14 vises den detaljerte punktskyen fra stereokameraene (rektangelet med stipla linje i Figur 13). I tillegg er et forstørret vindu uthevet rundt tyven (rødt rektangel i både Figur 13 og 14). Som man kan se fra figurene får man både et godt oversiktsbilde i tillegg til en detaljert 3D-modell av det man prøver å ta bilde av.



Figur 13: Punktsky fra Azure Kinect. Rødt rektangel med stipla linje viser området som er synlig for stereorigg (hele bildet i Figur 14). Rødt rektangel med hele linjer viser området med tyv som også er fremhevet i Figur 14.



Figur 14: Punktsky fra stereorigg. Tilsvaret området markert med rødt rektangel med stipla linjer i Figur 13. Fremhevet bilde av tyv i rødt rektangel med hele linjer.

3 Diskusjon

3.1 Generelt om avbildning i drivhus

Lysforholdene for avbildning i et drivhus er ofte nokså utfordrende. På dagtid kan det være veldig mye lys fra solen, noe som kan gi store kontraster mellom over og undersiden av bladene, og veldig stor forskjell i lysintensitet avhengig av høyden man avbilder. Store deler av året benyttes det også forskjellig kunstig belysning som skaper ekstra utfordringer for kamerasystemer. Høytrykksnatriumlamper (HPS) er fortsatt vanlig å bruke. Moderne versjoner av disse har drivere med elektronisk forkobling, ofte med frekvenser rundt 20kHz, som gir relativt stabil belysning. Ordinære ballaster som gjør at lysene pulserer på 50Hz er likevel ofte brukt, slik som hos Mære Landbruksskole. Lysintensiteten på disse varierer betydelig, noe som krever at kamera enten er synkronisert med lysene, eller bruker eksponeringstid som unngår problemer med variasjonen i intensitet. I stereooppsett bør kameraene være synkronisert, og helst også synkronisert med lyset.

Det blir også mer og mer vanlig å benytte LED-belysning. Noen av disse kan ha spesielle fargesammensetning som er optimalisert for det plantene trenger og kan ha ufullstendig fargespektrum for avbildning. Blandet sammen med dagslys kan de gjøre fargedeteksjon utfordrende.

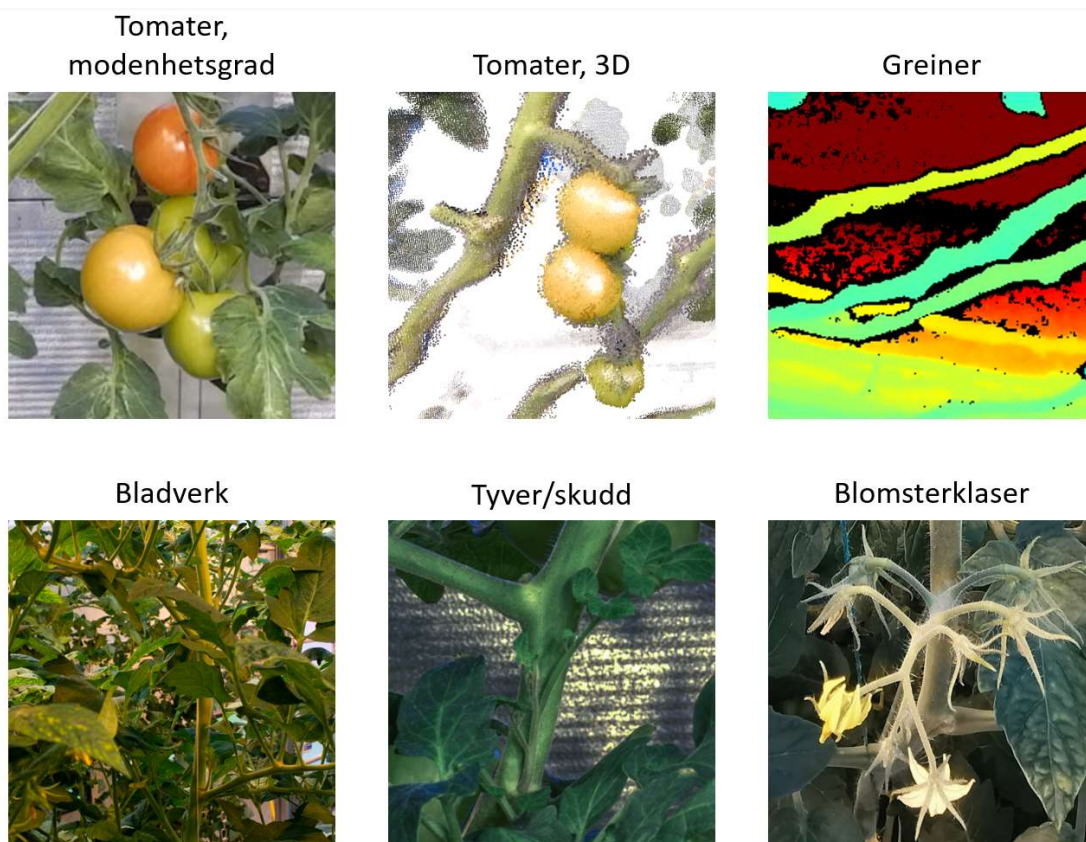
Det ble ikke lagt vekt på å finne oppsett eller utstyr som tar spesielt hensyn til ekstrabelysning i drivhusene. Dette er ting som kan løses i et endelig oppsett. Det kan f.eks. synkroniseres ved at kamerasystem og belysningen i drivhuset er koblet sammen slik at man kan tilpasse avbildningstidspunktet eller kanskje til og med skru av f.eks. LED-belysning i 1/100-sekund mens avbildningen utføres.

Miljøet i drivhuset for øvrig kan skape utfordringer for et kamerasystem. I tomatdrivhus er det ofte relativt tørt og spraying og direkte vannsprut er ikke vanlig. Det kan likevel forekomme kondens og drypping når det er kaldt ute eller i forbindelse med kjøleutstyr som er montert. Hvis det er drivhus der det benyttes spraying av plantene bør dette kunne styres slik at det ikke er aktivt der kamerasystemet opererer. Kamerasystem bør designes slik at det ikke tar skade av noen få dråper med vann, og at linsene er beskyttet litt slik at det ikke kan legge seg dråper på glassene.

En annen generell utfordring er avbildning av deler av planten som er skjult av seg selv. F.eks. dersom mye bladverk dekker deler av tomatplanten. For avbildningsforsøket i dette prosjektet har dette blitt løst ved å ta bilder fra mange forskjellige vinkler, men dette er ikke alltid nok for å få oversikt av hele planten. Man kan tenke seg at en robotløsning i tillegg må være i stand til å fysisk flytte unna deler av planten (blad og greiner) slik at kameraet kommer til og får tatt de nødvendige bildene. Dette perspektivet bør være med når man skal finne fram til en passende robotplattform. For at en robot skal være i stand til å dette må den f.eks. ha tilgang på to armer som kan opereres samtidig.

3.2 Vurdering av kamera- og maskinsynsystem

For å vurdere de forskjellige maskinsynløsningene som har blitt testet tar vi utgangspunkt i seks forskjellige objekter som vist i Figur 15 og hvor godt maskinsynsystemet passer til å detektere, måle eller vurdere disse.



Figur 15: Relevante objekter for vurdering av maskinsyn

I tillegg til disse spesifikke objektene har vi også vurdert to mer generelle kategorier: Robotnavigasjon og 3D-målinger. For robotnavigasjon er det viktig at kameraet kan reagere forholdsvis raskt samtidig som det kan få overblikk over hele den relevante scenen det skal observere. Kameraer som bare fokuserer på en liten del av scenen samtidig som det trenger lang tid for å prosessere bildene scorer f.eks. dårligere her. For 3D-målinger ser vi på hvor god 3D-målingene generelt er fra kameraet og hvorvidt 3D-informasjonen er preget av målestøy eller manglende avlesninger. Vurderingene av disse kriteriene er basert på alle bildene som er samlet inn fra flere omganger med avbildning hos Mære Landbruksskole og vurderingen er utført av forskere som har mange års erfaring med å utvikle og integrere maskinsynsystem både i industri- og forskningsprosjekt. Resultatene av denne vurderingen oppsummeres i Figur 16. Hver rute er farget etter hvor godt kameraet fungerer til den spesifikke oppgaven. Rødt betyr at systemet fungerer dårlig eller ikke i det hele tatt. Gult betyr at det fungerer delvis, men ikke optimalt. Grønt betyr at systemet fungerer godt til denne oppgaven.

	Generelt		Spesifikke objekter					
	Robot-navigasjon	3D-målinger	Tomater, modenhets-grad	Tomater, 3D	Greiner	Bladverk	Tyver/skudd	Blomster-klaser
Zivid One+ Medium	Grønt	Grønt	Grønt	Gult	Grønt	Grønt	Grønt	Grønt
Realsense L515	Grønt	Gult	Gult	Gult	Gult	Gult	Rødt	Rødt
Azure Kinect	Grønt	Grønt	Grønt	Gult	Grønt	Grønt	Gult	Gult
Speilrefleks stereorigg	Rødt	Gult	Grønt	Grønt	Gult	Gult	Grønt	Grønt
Kombinasjonsrigg med Kinect og stereo	Grønt	Grønt	Grønt	Grønt	Grønt	Grønt	Grønt	Grønt

Figur 16: Oppsummering av hvor godt hvert maskinsynsystem kan fungere til de relevante oppgavene.

Rødt: systemet fungerer dårlig eller ikke i det hele tatt

Gult: systemet kan fungere til dels, men ikke optimalt.

Grønt: systemet kan fungere godt.

Zivid One+ Medium fungerte svært godt til nesten alle oppgaver. Den eneste utfordringen var å få gode avlesninger av hele tomaten fra bare en vinkel. Men dette er en utfordring som de aller fleste kamerasystemene hadde.

RealSense L515 fungerte bare delvis til de aller fleste oppgavene. Hovedårsaken til dette er mye støy i 3D-målingene. For oppgaver som trenger høy nøyaktighet, som å finne tyver og blomster, blir målestøyen for stor.

Azure Kinect fungerer noe bedre enn *Realsense L515* på de fleste områder. Men også denne sliter med å få gode målinger av tomat uten mye målestøy. Mindre objekter som tyver og blomster er noe vanskeligere pga. lavere nøyaktighet enn f.eks. *Zivid*, men likevel bedre enn hos *Realsense L515*.

Speilrefleks stereorigg gjør det bedre på å detektere mindre, spesifikke objekter som tyver og skudd. Samtidig har man svært god fargeoppløsning som er gunstig for deteksjon av modenhet. Videre får man, noe overraskende, ganske gode 3D-målinger av tomat gitt at man er nærme nok til å kunne hente ut detaljert tekstur. Til robotnavigasjon og generelle 3D-målinger egner systemet seg derimot ikke like godt til. Dette kommer av lang prosesseringstid og en del målestøy når man ikke bare leter etter konkrete enkeltobjekter, men skal matche hele scenen.

Kombinasjonsrigg med Kinect og stereo dekker i all hovedsak alle de forskjellige kategoriene godt. Det er selvsagt også her utfordringer med enkeltbilder og enkeltobjekter innenfor en kategori, men alt i alt gir kombinasjonen av disse systemene svært gode resultat på alle åtte vurderingskriterier.

Det kan til slutt nevnes at flere mulige kombinasjoner av de fire forskjellige maskinsynsystemene også kan gi de samme resultatene. For eksempel vil en kombinasjon av *Azure Kinect* og *Zivid*-kameraet også gi gode resultater på nesten alle kategorier. Fordelen med en slik rigg er at den bruker mindre tid på prosessering sammenliknet med stereo-riggen. Samtidig vil man ikke nødvendigvis få like gode fargebilder fordi justering av fargekameraet på *Zivid*-systemet også må ta hensyn til å få best mulig 3D-målinger. Lang prosesseringstid er dessuten ikke nødvendigvis et like stort problem for disse oppgavene som i mange andre industrier, siden roboten som skal stelle plantene kan bevege seg mellom plantene gjennom hele dagen. Man kan f.eks. se for seg at roboten veksler mellom å ta nøyaktige bilder av store deler av planten for å så håndtere andre oppgaver mens en ekstern datamaskin analyserer de tidligere bildene.

Et annet alternativ er å kombinere *Zivid*-kameraet med stereoriggen. Ulempen med dette systemet er at man ikke får det samme oversiktsbildet som *Azure Kinect* gir, samtidig får man to svært nøyaktige kamerasystem som kan utfylle hverandre godt på nærbilder av de forskjellige objektene.

4 Konklusjon

Avbildning i drivhus har vist seg å være utfordrende. Belysning, tett og uoversiktlig plantemasse som gir vanskelig tilgjengelighet og måten drivhusene er tradisjonelt sett er organisert gjør avbildning utfordrende. I prosjektet har flere forskjellige 3D-kameraer blitt testet for å avbilde forskjellige deler av tomatplanten. Det har vist seg at få maskinsynsystemer alene oppfyller alle behov for å kunne gi en robot nok informasjon til å kunne stelle og høste en tomatplante, men ved å kombinere flere systemer har man likevel kommet frem til løsninger som sammen gir gode nok data til at oppgaver som identifisere og lokalisere tomater, blader, greiner, skudd og blomster er gjennomførbare.

Kombinasjonen som ble testet er en Azure Kinect satt sammen med en stereokamera-rigg bestående av to fargekamera med høy oppløsning. I tillegg er to alternative løsninger beskrevet basert på resultat fra hvert individuelle system: Azure Kinect kombinert med Zivid One+ Medium, samt Zivid One+ Medium kombinert med en stereokamera-rigg bestående av to fargekamera med høy oppløsning. Kombinasjonsriggen ble testet hos Mære Landbruksskole og resultatene fra prosjektet ble presentert for de på Mære Landbruksskole som har vært involvert i prosjektet i et avslutningsmøte i desember 2021.

Prosjektet har nådd de målene som er beskrevet i søknaden: Definere en teknologi for 3D-avbildning av tomatplanter, som har potensiale til å identifisere og lokalisere tomater, blader, greiner, knopper, skudd og blomsterklaser i 3D – i tilstrekkelig nøyaktighet til at en robot kan stelle og høste en tomatplante. Prosjektet har gitt en tydelig forståelse for utfordringene med å avbilde tomatplanter i 3D og resultatene fra dette prosjektet gjør det mulig å gå videre med å utforme et realistisk konsept for et robotisert system for stell og høsting av tomatplanter.

5 Veien videre

I prosjektsøknaden er forventet resultater formulert slik:

En tydelig forståelse for utfordringene med å avbilde tomatplanter i 3D. Konsekvensen av den forståelsen er at det blir mulig å utforme et realistisk konsept for et robotisert system for stell og høsting av tomatplanter. Dette konseptet skal realiseres i et fremtidig forskerprosjekt som det vil søkes om innen utgangen av 2020¹.

Denne rapporten har vist at det er mulig å komme frem til maskinsynløsninger som har potensialet til å kunne gjøre en robot i stand til å utføre flere oppgaver innenfor stell og høsting av en tomatplante. Basert på dette er det derfor naturlig å se på hvordan disse resultatene best kan tas videre. Under avslutningsmøtet, hvor resultatene ble presentert for Mære Landbruksskole, ble det gitt gode tilbakemeldinger fra deltagerne hvor man så for seg at et slikt system kunne benyttes i den daglige driften.

Kompetansen hos de ansatte på Mære Landbruksskole har vært viktig for spissingen av dette prosjektet, og kommer til å bli enda viktigere for en eventuell videreføring der man skal begynne å teste robotsystemer som skal interagere med tomatplanten. Flexibiliteten Mære Landbruksskole har vist under en ganske krevende tid med tanke på smittevern og nedstenginger har også vært avgjørende for at prosjektet var mulig å gjennomføre.

¹ Søknaden ble sendt inn før COVID-19, og prosjektet har fått utsettelse til slutten av 2021. Videreføring blir derfor naturlig først i 2022.