

# Bruk av autostereoskopisk 3D i videosamtaler

**Sindre Ruud Grønningen**  
**Håkon Smeplass**

Master i kommunikasjonsteknologi  
Oppgaven levert: Juni 2010  
Hovedveileder: Leif Arne Rønningen, ITEM  
Biveileder(e): Tor Andre Kvanli, Tandberg  
Stein Ove Eriksen, Tandberg



# Oppgavetekst

Sammenlign brukeropplevelsen av en autostereoskopisk multiview-skjerm med brukeropplevelsen av en vanlig skjerm. Skjermene er tenkt å brukes til en videosamtale mellom to eller flere personer, avhengig av størrelsen på skjermen. Brukeropplevelsen skal kartlegges ved hjelp av kvalitative metoder.

Videoinnholdet til multiview-skjermen kan filmes av et array av 5 kameraer. Finn fram til best mulig oppsett av kameraene ut fra parametre som plassering av kameraene, avstand mellom kameraene, vinkler mellom kameraene og avstand fra kameraene til deltaker. Finn en generell sammenheng slik at det kan skrives spesifikasjoner for kameraoppsett til ulike scenarioer. Lag en slik spesifikasjon for et kontormiljø.

Vurder et oppsett som tar i bruk flere skjermer for å dekke en større del av synsfeltet. Finn hvordan multiplexingen av views må gjøres for at en kan oppleve 3D over flere skjermer plassert som et hjørne.

Videoinnholdet som testpersonene skal sammenligne kan være et opptak, ettersom sanntids overføring av multiview-video ligger utenfor omfanget av denne oppgaven.

Oppgaven gitt: 15. januar 2010  
Hovedveileder: Leif Arne Rønningen, ITEM



# Sammendrag

I oppgaven vurderes det om autostereoskopisk 3D er egnet å bruke i videosamtaler.

Bruk av 3D-teknikker for å forbedre den opplevde kvaliteten av video er i dag mer aktuelt enn noen gang før, og 3D får stadig nye bruksområder. I telepresence-systemer er målet å gjøre illusjonen av at menneskene du snakker med sitter i samme rom så realistisk som mulig. Viktigheten av øyekontakt gjør samtidig bruk av 3D-briller ganske uaktuelt. Det har derfor vært naturlig å se på om bruk av autostereoskopisk 3D er egnet til å forbedre den opplevde kvaliteten og realismen i en videosamtale. For å vurdere dette har vi gjennom å etablere matematiske sammenhenger og gjøre praktiske forsøk kommet fram til hvilke faktorer som er viktige for å lage autostereoskopisk 3D av god kvalitet. Gjennom erfaringene med å lage 3D har vi kommet fram til hvilke muligheter og begrensninger denne 3D-teknikken har, og vi har vurdert i hvilken grad autostereoskopisk 3D er egnet for ulike scenarier. Vi har også gjort en direkte sammenligning av opplevelsen av en 2D-video og 3D-video ved hjelp av kvalitative metoder.

Gjennom erfaringene vi har fått og resultatene av forsøk vi har gjort, har vi kommet fram til at autostereoskopisk 3D kan øke den opplevde kvaliteten av videosamtaler. Dybdeeffekten gjør at ansiktsuttrykk og kroppsspråk blir tydeligere, samtidig som brukeren opplever større innlevelse. Illusjonen av at personen på skjermen faktisk sitter der i virkeligheten forsterkes. Samtidig har autostereoskopisk 3D noen klare begrensninger, og er preget av å være en umoden teknologi. Begrenset innsynsvinkel, dårlige overganger mellom visningsvindu og lang optimal avstand begrenser friheten til seeren mye. På kamasiden er det utfordrende å finne kameraoppsett som er egnet på både kort og lang avstand og det er vanskelig å synkronisere kameraene.

Vi har likevel tro på at ettersom teknologien utvikler seg, utfordringer blir løst og begrensningene i autostereoskopisk 3D blir mindre, vil dette bli en veldig aktuell teknologi å bruke i telepresence-systemer.



# Forord

Denne oppgaven er det avsluttende prosjektet i sivilingeniørutdannelsen kommunikasjonsteknologi, og er utført i samarbeid av masterstudentene Sindre Grønningen og Håkon Smeplass. Oppgaven er gitt av Tandberg og institutt for telematikk, NTNU, og har vært gjennomført på Caruso 3D-laboratorium.

Vi vil gjerne takke våre veiledere, professor Leif Arne Rønningen ved ITEM og Stein Ove Eriksen, Tor Andre Kvanli og Espen Berger i Tandberg for kyndig veiledning og motivasjon. Vi vil også takke Solveig Bergstrøm og Sverre Smeplass for gjennomlesing og tilbakemeldinger, samt Pål Sæther og Asbjørn Karstensen i IT-drift ved ITEM for hjelp med utstyr på laboratoriet.

Sindre Grønningen vil takke venner fra datasalene “Tulipan” og “Tulipan 2.0” for mange gode stunder og utallige “headshots”. En stor takk rettes til foreldre, søsken og slektninger for motivasjon og støtte gjennom hele studietiden.

Håkon Smeplass vil gjerne takke foreldre, venner fra Gløshaugen og Studentersamfundet og ikke minst Solveig Bergstrøm for støtte og motivasjon gjennom studietiden.

*Trondheim, 17.06.2010*

---

Sindre Grønningen

---

Håkon Smeplass





# Innhold

<b>Forord</b>	<b>vii</b>
<b>1 Introduksjon</b>	<b>1</b>
<b>2 Forskningsdesign</b>	<b>3</b>
2.1 Forskningsmetoder . . . . .	3
2.1.1 Beskrivende metoder . . . . .	3
2.1.2 Eksperimentelle metoder . . . . .	4
2.2 Valg av metode . . . . .	4
2.2.1 Kvalitativt intervju . . . . .	4
2.2.2 Eksperimentell forskning . . . . .	6
<b>3 Videosamtaler og telepresence</b>	<b>9</b>
3.1 Begrepene videosamtale og telepresence . . . . .	9
3.2 Teknologi . . . . .	10
3.3 Bruksområder . . . . .	12
3.4 Utfordringer . . . . .	12
<b>4 Kvalitet i 3D</b>	<b>15</b>
4.1 Persepsjon . . . . .	15
4.2 Kontrollerbare parametre . . . . .	16
4.3 Andre faktorer . . . . .	18
4.4 3D-teknikker . . . . .	18
4.4.1 Polarisasjon . . . . .	18
4.4.2 Lukkerbriller . . . . .	19
4.4.3 Autostereoskopi . . . . .	19
<b>5 3D-situasjonen i dag</b>	<b>21</b>
5.1 Stereoskopi . . . . .	21
5.2 Autostereoskopi . . . . .	22

<b>6</b>	<b>Autostereoskopi</b>	<b>23</b>
6.1	Historie . . . . .	23
6.2	Visningsvindu . . . . .	25
6.3	Head tracking . . . . .	26
6.4	Multiview . . . . .	27
6.5	Virkemåte . . . . .	28
6.5.1	Parallaksebarriere . . . . .	29
6.5.2	Lentikulær separasjon . . . . .	32
6.6	Multiview autostereoskopi sammenlignet med stereoskopi . . .	33
6.7	Ubehag . . . . .	34
<b>7</b>	<b>Utstyr og programvare</b>	<b>37</b>
7.1	Bakgrunn for teknologivalg . . . . .	37
7.2	Skjermer . . . . .	39
7.2.1	Separasjon i bildet . . . . .	40
7.2.2	Krav til beskjæring . . . . .	40
7.2.3	Autostereoskopi . . . . .	40
7.2.4	Tekniske detaljer . . . . .	41
7.3	Kamera . . . . .	41
7.4	Kamerafeste . . . . .	42
7.5	Programvare . . . . .	43
7.5.1	Opptak . . . . .	43
7.5.2	Redigering . . . . .	43
7.5.3	Avspilling . . . . .	44
<b>8</b>	<b>Kameraoppsett</b>	<b>45</b>
8.1	Konvergerende kameraer . . . . .	45
8.2	Parallele kameraer . . . . .	47
8.3	Ulike metoder til ulike scenarier . . . . .	49
<b>9</b>	<b>Formler og utregninger</b>	<b>51</b>
9.1	Beskjæring . . . . .	51
9.1.1	Nødvendig beskjæring for parallelle kamera . . . . .	52
9.1.2	Nødvendig beskjæring for konvergerende kamera . . . . .	52
9.2	Separasjon . . . . .	55
9.2.1	Parallele kamera . . . . .	55
9.2.2	Konvergerende kamera . . . . .	56
9.2.3	Grader . . . . .	56
9.3	Avstander . . . . .	57
<b>10</b>	<b>Forsøk</b>	<b>61</b>

10.1	Kalibrering av kamera . . . . .	61
10.1.1	Parallele kamera . . . . .	61
10.1.2	Konvergerende kamera . . . . .	62
10.2	Måling av optimal avstand og vinkel til skjerm . . . . .	62
10.3	Måling av separasjon i bildet . . . . .	64
10.4	Tidlige bildeforsøk . . . . .	65
10.5	Sammenligning av 2D og 3D . . . . .	65
10.5.1	Bakgrunn . . . . .	65
10.5.2	Gjennomføring . . . . .	65
10.5.3	Resultater . . . . .	67
10.5.4	Feilkilder . . . . .	68
10.5.5	Konklusjoner . . . . .	68
10.6	3D-i-3D . . . . .	69
<b>11</b>	<b>Diskusjon</b>	<b>71</b>
11.1	Generelle betraktninger . . . . .	71
11.2	Utfordringer . . . . .	72
11.2.1	Programvare og maskinvare . . . . .	72
11.2.2	Oppløsning og beskjæring . . . . .	76
11.2.3	Visninger . . . . .	78
11.3	Relevans for videosamtaler . . . . .	79
11.3.1	3D som neste steg i telepresence . . . . .	79
11.3.2	Autostereoskopi i telepresence . . . . .	80
<b>12</b>	<b>Scenarier</b>	<b>83</b>
12.1	Videotelefon . . . . .	83
12.2	Møterom . . . . .	84
12.3	Videosamtalerom . . . . .	85
12.4	Optimalt oppsett . . . . .	86
<b>13</b>	<b>Konklusjon</b>	<b>89</b>
	<b>Bibliografi</b>	<b>99</b>
	<b>Tillegg</b>	<b>99</b>
<b>A</b>	<b>Utregninger</b>	<b>101</b>
A.1	Konstanter og variabler . . . . .	101
A.2	Absolutt minimumsavstand . . . . .	102
A.3	Nødvendig beskjæring parallelle kamera . . . . .	102
A.4	Nødvendig beskjæring for konvergerende kamera . . . . .	103
A.5	Separasjon i bildet for parallelle kamera . . . . .	105

A.6	Separasjon i bildet for konvergerende kamera . . . . .	105
-----	--	-----

# Tabeller

7.1	Spesifikasjoner for skjermen på 24 tommer fra Tridality. . . . .	41
7.2	Spesifikasjoner for skjermen på 42 tommer fra Tridality. . . . .	41



# Figurer

3.1	Eksempel på forskjellige endepunkter som kan delta i en videosamtale gjennom et felles nettverk . . . . .	10
3.2	Eksempel på en telepresence-løsning fra Tandberg . . . . .	11
4.1	Eksempel på perspektivendring ved bevegelsesparallakse . . . .	16
4.2	Figuren viser fokuserings-/konvergerings-problemet. Her konverterer øynene mot objektet på avstand A, mens de fokuserer på skjermen som har en avstand B. . . . .	17
6.1	Patentet til Frederick E. Ives: “The Parallax Stereogram” . . . .	24
6.2	Visningsvinduene som skapes av visningene . . . . .	25
6.3	Visningsvinduene er formet som diamanter. Man vil oppleve 3D så lenge øynene befinner seg i hvert sitt visningsvindu. . . .	26
6.4	Figuren viser hva man kaller skjermene etter hvor mange visninger de har. . . . .	28
6.5	Figuren viser vinduer som repeteres ut til sidene. . . . .	30
6.6	Figuren viser prinsippene fra filming til visning på skjerm. . . .	31
6.7	Figuren viser prinsippet bak parallaksebarriere. . . . .	31
6.8	Figuren viser prinsippet bak lentikulær separasjon. . . . .	32
6.9	En lentikulær linse er satt sammen av mange små sylindriske linser . . . . .	33
6.10	Figuren viser den geometriske fordreiningen som oppstår i stereoskopiske systemer. . . . .	34
7.1	24-tommer skjerm fra Tridelity . . . . .	39
7.2	42-tommer skjerm fra Tridelity . . . . .	39
7.3	Et Panasonic HDC-HS300 ble brukt til å filme i 2D . . . . .	42
7.4	Spesiellaget kamerafeste for nøyaktig kalibrering av kameraene	42
7.5	Kamerafestet plassert på et vanlig kamerastativ . . . . .	43
8.1	Fem konvergerende kameraer, 50 grader bildevinkel . . . . .	46
8.2	Keystoning . . . . .	47

8.3	Fem parallelle kameraer, 50 grader bildevinkel . . . . .	48
9.1	Nødvendig beskjæring i prosent for konvergerende kameraer som funksjon av avstanden til objektet. Avstanden til brennpunktet er 130 centimeter. . . . .	54
9.2	Nødvendig beskjæring i prosent for konvergerende kameraer som funksjon av avstanden til objektet. Avstanden til brennpunktet er 260 centimeter. . . . .	54
9.3	Nødvendig beskjæring i prosent for parallelle kameraer som funksjon av avstanden til objektet. . . . .	55
9.4	Separasjonen i bildet som funksjon av avstanden, korrigert for nødvendig beskjæring. Kameraene er konvergerende og avstanden til brennpunktet er fastsatt til 130 centimeter. . . . .	57
9.5	Separasjonen i bildet som funksjon av avstanden, korrigert for nødvendig beskjæring. Kameraene er konvergerende og avstanden til brennpunktet er fastsatt til 260 centimeter. . . . .	57
9.6	Separasjonen i bildet som funksjon av avstanden, korrigert for nødvendig beskjæring. Parallelle kameraer. . . . .	58
9.7	Separasjonen i bildet som funksjon av avstanden mellom kameralinsene. Kameraene er konvergerende, avstanden til brennpunktet er 130 centimeter og avstanden til objektet er 100 cm. . . . .	59
9.8	Separasjonen i bildet som funksjon av avstanden mellom kameralinsene for parallelle kameraer. Avstanden til objektet er 335 cm. . . . .	60
10.1	Skjermen viser én farge, altså én visning, for hvert øye ved optimal avstand . . . . .	63
10.2	Skjermen viser flere farger (visninger) sett litt fra siden . . . . .	64
11.1	Skjermskudd som viser synkroniseringsfeil, samt høyde- og vinkelforskjeller . . . . .	73
11.2	Skjermskudd av programmet CrystalDiskMark . . . . .	74
11.3	Skjermskudd av programmet CrystalDiskMark . . . . .	74
11.4	Skjermskudd av programmet Resource Monitor, prosessorbelastning . . . . .	75
11.5	Skjermskudd av programmet Resource Monitor, diskbruk . . . . .	75
11.6	Skjermskudd av programmet Resource Monitor, minnebruk . . . . .	75
11.7	Skjermskudd som viser en horisontal strek tvers over bildet. Skjermskuddet er tatt ved at TriView er brukt på en vanlig skjerm, og viser dermed også hvordan TriView fordeler visningene på den autostereoskopiske skjermen. . . . .	77



# Definisjoner

<b>Autostereoskopi</b>	Teknikk for å vise 3D som ved hjelp av ulike metoder sender bilder i forskjellige retninger slik at øynene mottar forskjellige bilder uten behov for spesielle briller.
<b>Avstand mellom kamera</b>	Avstanden mellom sentrum av kameralin-sene.
<b>Bevegelsesparallakse</b>	Når man beveger seg sidelengs, vil objek-ter i dybden ha ulik relativ bevegelse og man ser dem fra et nytt perspektiv.
<b>Bilde</b>	Ordet bilde blir i denne oppgaven brukt om det som vises på skjermen, og om-handler både levende og stille bilder.
<b>Binokulære hint</b>	Når hjernen bruker det at øynene ser gjenstander fra litt ulikt perspektiv til å oppfatte dybde.
<b>Brennpunkt</b>	Punktet hvor tenkte linjer rett ut fra ka-meraene krysser.
<b>Cross talk</b>	Informasjon fra en visning havner i en an-nen visning.
<b>Deltaker</b>	En person som er med i videosamtalen. Ser på en skjerm og blir filmet av kamera.
<b>Dybdevirkning</b>	Opplevelsen av at bildet er i tre dimen-sjoner.
<b>Ghosting</b>	Forstyrrelse i et 3D-bilde som følge av cross talk. Ser ut som objekter i bildet dupliseres og overlapper hverandre eller får uklare kanter.

<b>Head tracking</b>	Teknikk som finner ut hvor hodet til seeren befinner seg i forhold til skjermen. Justerer bildet slik at seeren alltid får se riktig visning.
<b>Keystone-feil</b>	Kameraene har ulikt perspektiv, dette skaper vertikal parallakse og feil langs sidekantene av bildet.
<b>Konvergerende kamera Kvalitet</b>	Kameraene peker inn mot et brennpunkt. I denne oppgaven brukes ordet kvalitet om den opplevde kvaliteten.
<b>Monokulære hint</b>	Hjernen tolker informasjon fra ett øye for å oppfatte dybde.
<b>Multiview</b>	At det blir brukt flere enn to visninger. Autostereoskopi i denne oppgaven er alltid multiview.
<b>Parallele kamera</b>	Siktlinjer fra kameraene er parallelle, avstanden til brennpunktet blir uendelig.
<b>Positiv og negativ parallakse</b>	Områdene henholdsvis bak og foran brennpunktet, sett fra kameraet.
<b>Seer</b>	Person som ser på skjermen.
<b>Separasjon</b>	Avstanden mellom siktlinjene fra to kameraer ved siden av hverandre i en gitt avstand fra kameraene.
<b>Separasjon i bildet</b>	Avstanden mellom to visninger av samme objekt på skjermen.
<b>Stereoskopisk 3D</b>	3D-teknikker der spesielle briller brukes for å fordele visninger på øynene.
<b>Sweetspot</b>	En sweetspot er en posisjon hvor man opplever korrekt 3D.
<b>Telepresence</b>	Videosamtaler av veldig høy kvalitet hvor tilstedeværelse er nøkkelord.
<b>Videosamtale</b>	En videosamtale er en sanntids video- og lydoverføring mellom to eller flere geografisk spredte parter med det formål å kommunisere. Når vi snakker om videosamtaler i denne oppgaven tenker vi på telepresence.

**Visning**

Fra engelsk view. Bildet fra en autostereoskopisk skjerm består av flere visninger hvor en visning inneholder bildeinformasjon fra ett kamera, ment til ett øye. Denne blir sendt i en spesifikk retning.

**Visningsvindu**

Når et øye mottar en distinkte visning, befinner øyet seg i et visningsvindu



# Kapittel 1

## Introduksjon

Vår forståelse av oppgaven er at vi skal gjøre en vurdering av om 3D kan brukes for å øke kvaliteten og realismen av en videosamtale. For å gjøre dette må vi først finne ut hvordan man lager så bra autostereoskopisk 3D som mulig ut fra tilgjengelig utstyr. Deretter må vi lage 3D-videoer som lar seg spille av på en autostereoskopisk multiview-skjerm og sammeligne opplevelsen av disse videoene med tilsvarende videoer i 2D ved hjelp av kvalitative metoder.

Oppgavetekstens del om å ta i bruk flere skjermer ble det dessverre ikke anledning til, ettersom de to siste skjermene vi trengte for å gjennomføre forsøk med flere skjermer kom svært sent.

I de første kapitlene går vi gjennom det teoretiske grunnlaget for resten av oppgaven: Hva som kjennetegner kvalitative metoder i forskningssammenheng, en gjennomgang av de tekniske aspektene ved autostereoskopi, hva som kjennetegner god 3D og hvilke egenskaper som kjennetegner en videosamtale. Det neste hovedtemaet er grunnlaget for å lage 3D: Hvordan kameraene kan settes opp og hva som kjennetegner de ulike måtene, en oversikt over utstyret vi har brukt og hvordan egenskapene til utstyret har påvirket valgene vi har tatt og til slutt en oversikt over alle utregningene og formlene vi har kommet fram til som beskriver matematiske sammenhenger for kameraoppsett.

Den neste hovedbolken omhandler det sammenlignende forsøket som ble gjort, og et diskusjonskapittel hvor erfaringene fra arbeidet med å lage 3D og forsøket blir drøftet. I det samme kapitlet blir det også drøftet i hvilken grad autostereoskopisk 3D er egnet for videokonferanser, sett i lys av mulighetene og begrensningene denne typen 3D har. Etter dette presenterer vi noen aktuelle scenarier hvor det kan være aktuelt å bruke autostereoskopisk 3D og drøfter om og hvordan det kan brukes for å gi en best mulig opplevelse.

Helt til slutt presenteres hovedkonklusjonene i et eget kapittel.

# Kapittel 2

## Forskningsdesign

### 2.1 Forskningsmetoder

I denne seksjonen vil vi beskrive metoder som kan være aktuelle for dette prosjektet. I vitenskapelige studier har vi to hovedtyper metoder: Beskrivende metoder og eksperimentelle metoder [1].

#### 2.1.1 Beskrivende metoder

Beskrivende metoder kan ha en kvalitativ eller en kvantitativ tilnærming. Kvantitative tilnærminger kan likevel inneholde kvalitative elementer og motsatt [2]. Kvalitet er egenskapene eller karaktertrekkene ved noe, mens kvantitet betegner mengden av egenskapene eller karaktertrekkene. Nøkkelord for kvalitative tilnærminger er dybde- og detaljforståelse og innhold [3]. Kvalitative tilnærminger er innholdssøkende og forsøker å forklare et fenomens karakter. Observasjon og intervju er vanlige teknikker som benyttes. Man går her gjerne i dybden på en liten undersøkelsesgruppe, i motsetning til de kvantitative tilnærmingene som fokuserer mer på antall og mengde. Kvantitative tilnærminger er innholdsstyrte og har ofte som mål å fastslå mengden av et fenomens karakter. Teknikker som benyttes er gjerne spørreskjema og strukturerte intervju. I strukturerte intervju er både spørsmål og rekkefølge planlagt på forhånd. Dataene som hentes inn i slike studier kan analyseres ved hjelp av statistisk analyse og representeres deretter gjerne med grafer og tabeller. Disse resultatene svarer på hypoteser og kan generaliseres når utvalget er representativt og stort nok [4]. Tallene og mengdene som samles inn i kvantitative metoder byttes i kvalitative studier ut med notater, tekster

og bilder. Disse analyseres og resulterer i informasjon som kun er gyldig i den konteksten studiet ble gjennomført i. Videre, generelle konklusjoner vil kun være hypoteser.

Ved valg av tilnærming er det nødvendig å vurdere problemstilling og studiets formål [5]. Kvantitative studier brukes når man har god oversikt over hva man ser etter og svarer på en eller flere hypoteser. Alle aspekter ved studiet planlegges grundig før gjennomførelsen [6]. Kvalitative studier er mer utforskende og kjennetegnes av mindre planlegging. I intervjuer gir dette for eksempel rom for oppfølgingsspørsmål fra både respondent og informant<sup>1</sup>, og intervjuet kan dermed fort ende opp med å bevege seg i nye retninger. Kvalitative tilnærminger er anbefalt brukt i tidlige faser av forskningsprosjekter for å danne hypoteser som eventuelt kan testes med kvantitative tilnærminger i senere faser.

### 2.1.2 Eksperimentelle metoder

Eksperimentelle metoder skiller seg fra beskrivende metoder ved at man i eksperimentelle metoder kan kontrollere forsøkssituasjonene ved å manipulere visse variabler [1]. Man har derfor muligheten til å undersøke årsaks- og virkningsforhold, eller kausale relasjoner, mellom variabler i eksperimentelle metoder. Dette kan gjøres ved å variere variabler av interesse og la andre variabler være konstante.

## 2.2 Valg av metode

I denne seksjonen vil vi se på hvilke forskningsmetoder vi har valgt å bruke for å svare på oppgavens problemstilling.

### 2.2.1 Kvalitativt intervju

I oppgaveteksten heter det at vi skal sammenligne brukeropplevelsen av en autostereoskopisk multiview-skjerm med brukeropplevelsen av en vanlig skjerm. Dette kan gjøres gjennom beskrivende metoder med en kvalitativ eller kvantitativ tilnærming. Oppgaven går ut på å vurdere spesifikt for en

---

<sup>1</sup>Personen som utfører intervjuet kalles respondent og personen som intervjues kalles informant



videokonferansesammenheng om autostereoskopisk 3D gir en bedre brukeropplevelse enn 2D med tilgjengelige skjermer. Vi ser altså på brukeropplevelsen i en gitt kontekst, noe som tilsier at metodene vi skal bruke bør være kvalitative. Vi ønsker å komme fram til konklusjoner rundt hvilke egenskaper ved 3D som fungerer og ikke. Vi ønsker oss med andre ord ikke nødvendigvis svar på spørsmål som “er 2D eller 3D best?”. Dette tilsier at forsøkspersonene bør ha god kjennskap til 3D og dermed ha evnen til å observere detaljer og egenskaper ved bildene som kan være vanskelig å oppdage for andre. Relevansen av autostereoskopisk 3D i videokonferanser er i tillegg lite studert fra før, og en kvalitativ tilnærming er derfor et naturlig valg. Vi er i stor grad kvalifiserte til å vurdere 2D mot 3D selv og behøver derfor ikke et stort antall personer for å kunne komme med gode konklusjoner. Delvis strukturerte intervju av et lite antall fagpersoner vil gi oss mange fine innspill som vil berike vår diskusjon om relevansen av autostereoskopisk 3D i videokonferanser. Et delvis strukturert intervju benytter seg av en intervjuguide bestående av en liste med temaer og tilhørende spørsmål man ønsker å komme innom. Intervjuet vil foregå ansikt-til-ansikt og det vil åpne med generelle spørsmål før det blir stilt mer inngående spørsmål. Vi vil unngå ledende spørsmål og passe på at alle spørsmålene er relevante i forhold til problemstillingen.

### **Etiske vurderinger**

I forsøk som involverer personer er det viktig å følge god forskningsetikk [7]. Man skal respektere forskningspersonenes krav på privatliv og integritet. I Norge stilles det tre grunnleggende krav i forskningsetikken: informert samtykke, krav på privatliv og krav på å bli gjengitt korrekt [8]. De involverte personene i forsøkene forbundet med denne oppgaven vil informeres om hensikten med oppgaven og hva som skal gjøres med resultatene. Deltakelse er deretter frivillig. Det vil heller ikke bli stilt personlige spørsmål, noe som heller ikke har noen hensikt i denne oppgaven. Forsøkspersonene har rett på anonymisering, men dette ble ikke krevd av forsøkspersonen i dette tilfellet. Sitater vil gjengis korrekt og i korrekt sammenheng.

### **Reliabilitet og validitet**

Reliabilitet og validitet forbindes gjerne med kvantitative metoder, men bør også vurderes ved bruk av kvalitative metoder. Reliabilitet sier noen om påliteligheten og repeterbarheten til forsøkene [1]. I vårt intervju vil vi unngå

ledende spørsmål og informanten skal ha gode fagkunnskaper. Dette vil styrke påliteligheten til forsøket.

Validitet er overensstemmelsen mellom det man ønsker å måle og det man faktisk måler. Validitet sier altså noe om hvor gyldige resultatene og slutningene fra forsøkene er. Resultatene fra det kvalitative intervjuet vårt vil ikke være allment gjeldende. De vil kun gjelde for den konteksten forsøket foregår i. Det vil derimot være mulig å trekke generelle slutninger basert på resultatene, men disse vil kun være antagelser.

### 2.2.2 Eksperimentell forskning

Det står i oppgaveteksten at vi skal finne fram til best mulig oppsett av kameraene og finne generelle sammenhenger slik at det kan skrives spesifikasjoner for kameraoppsett til ulike scenarier. Dette skal gjøres ut fra variabler som plassering av kameraene, avstand mellom kameraene, vinkler mellom kameraene og avstand fra kameraene til deltaker. En naturlig framgangsmåte for å løse oppgaven er prøving og feiling gjennom praktiske forsøk, altså eksperimentell forskning. Etter hvert som vår forståelse av hvordan variablene påvirker hverandre og opplevelsen av 3D øker, kan vi finne nye matematiske sammenhenger og formler som vi så kan teste i praksis. Testene kan gjennomføres ved å holde noen av parametrene konstante og variere andre. Forsøkene vil foregå i 3D-laboratoriet Caruso på NTNU.

#### Etiske vurderinger

Det er ingen spesielle etiske hensyn å ta ved testing av kameraoppsett i et laboratoriemiljø.

#### Reliabilitet og validitet

Påliteligheten til resultatene våre avhenger blant annet av at måleinstrumentene og målemetodene våre er reliable. Vi benytter oss av kameraer og målebånd. Avstandene vi måler i forsøkene kan ikke sies å være helt nøyaktige. Avviket mellom målt avstand og virkelig avstand fra kameraer til objekt eller objekt til vegg kan være fra null til et par centimeter. Målt avstand mellom kameraene kan sies å være millimeterpresist med ganske stor nøyaktighet, grunnet

## 2.2. VALG AV METODE

---

den korte avstanden man måler. Kameraene vil stilles parallelt eller konvergerende mot et brennpunkt ved hjelp av manuelle metoder. Dette ville vært mer nøyaktig med automatisk kalibrering.

En stor del av resultatene vil være matematiske formler. Forutsetninger for at disse er gyldige er at vi har forstått hvordan opptak av autostereoskopisk 3D fungerer og at vi har regnet riktig. Videre konklusjoner vil i stor grad basere seg på disse resultatene, og avhenger derfor av at disse er riktige.



# Kapittel 3

## Videosamtaler og telepresence

Dette kapitlet gir en innføring i begrepene videosamtale og telepresence. Deretter presenteres bruksområder med spesiell vekt på jobbsammenheng før det tar for seg teknologiene og utstyret som gjør det hele mulig. Vi skal ikke gå inn på alle detaljene som det må tas hensyn til i utviklingen av et telepresence-rom. Det vil kun gis en oversikt som gir leseren et godt grunnlag for å forstå hvorfor og hvordan 3D kan bidra til å øke kvaliteten i videosamtaler. Til slutt vil vi se på noen utfordringer knyttet til telepresence.

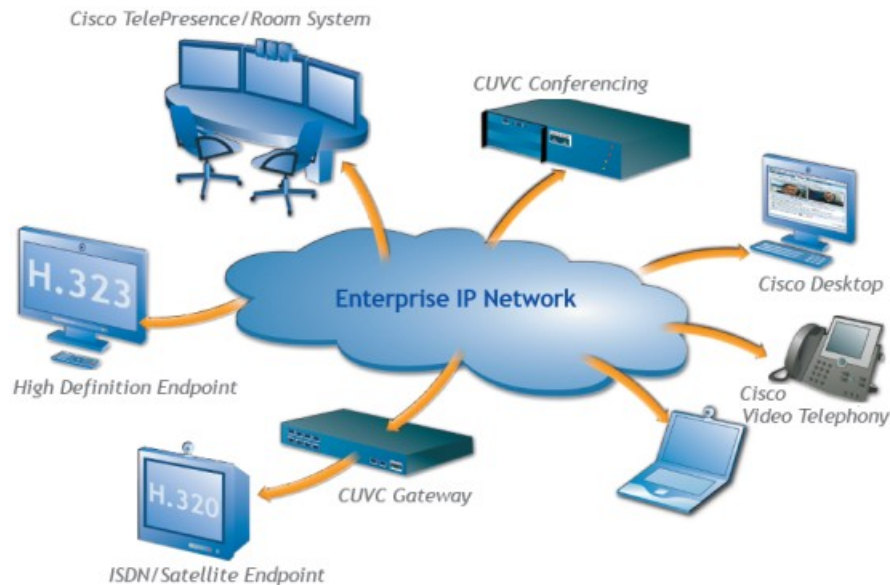
### 3.1 Begrepene videosamtale og telepresence

En videosamtale er en sanntids<sup>1</sup> video- og lydoverføring mellom to eller flere geografisk spredte parter med det formål å kommunisere. En videosamtale med kun to deltakere kalles gjerne punkt-til-punkt, mens en samtale med flere deltakere kalles multi-punkt [9]. Partene i en videosamtale kan delta via mobiltelefoner, bærbare PC-er eller dedikerte systemer som vist i figur 3.1 [10]. I tillegg til å vise bilde og lyd kan man dele dokumenter, presentasjoner og lignende.

Telepresence er et mye brukt ord i forbindelse med videosamtaler. Telepresence er satt sammen av ordene tele (fjern) og presence (tilstedeværelse) og brukes i denne oppgaven om videosamtaler av veldig høy kvalitet, både teknisk og designmessig. Det ultimate målet med telepresence-systemer er at man skal få følelsen av at personene man snakker med er tilstedeværende i rommet selv om de befinner seg et sted som er forskjellig fra rommets fysiske

---

<sup>1</sup>Noe som foregår med én gang, uten merkbar forsinkelse



**Figur 3.1:** Eksempel på forskjellige endepunkter som kan delta i en video-samtale gjennom et felles nettverk

posisjon. Med andre ord etterstreber man i slike systemer den naturlige og realistiske samtalen man opplever ansikt-til-ansikt, slik at denne kan erstattes. Telepresence karakteriseres av sylskarpe bilder, førsteklasses lyd, naturlig lyssetting og strenge krav til tidsforsinkelser. Sylskarpe bilder i dag vil si bilder i full HD<sup>2</sup> og bedre. Når det gjelder lyd, må denne komme fra posisjonen til den som snakker og med et volum som ville vært naturlig i en samtale ansikt til ansikt. I denne sammenheng er plassering av høyttalere og mikrofoner, samt akustikken i rommet kritisk. Riktig lyssetting gjør følelsen av at personene man snakker med befinner seg i samme rom større ved blant annet å gi en følelse av dybde i bildet [12]. For å oppnå disse kravene plasseres ofte telepresence-systemer i dedikerte rom, som Tandberg sine telepresencerom er eksempler på [13]. Figur 3.2 viser hvordan et slikt rom kan se ut [14].

## 3.2 Teknologi

Et videosamtalsesystem må ha en del komponenter på plass for å kunne brukes. Disse inkluderer blant annet utstyr for å spille inn og spille av lyd og video. For å spille inn video brukes kameraer, webkameraer, mobilkameraer

<sup>2</sup>Full HD har en oppløsning på 1920×1080 og er vanlig i dag [11].



**Figur 3.2:** Eksempel på en telepresence-løsning fra Tandberg

og lignende. For å vise video brukes TV-skjermer, PC-skjermer, mobilskjermer og lignende. Lyd tas opp med en eller flere mikrofoner og spilles av med en eller flere høyttalere. Et digitalt nettverk trengs for å overføre lyd og video til partene man snakker med. Lyd er ikke noe problem å overføre på grunn av den relativt lave båndbredden som trengs. De store mengdene videodata som skal sendes krever derimot stor båndbredde og ville brukt veldig lang tid på å overføres uten komprimering [15]. For å komprimere dataene benyttes en kodek. En kodek består av en enkoder som konverterer analoge bildesignaler til en digital datafil og en dekoder som tolker datafilen slik at den kan spilles av [16]. Siden videosamtaler foregår i sanntid er kravene til tidsforsinkelser i lyd og bilde strenge. Samtidig ønsker man så høy kvalitet som mulig. Kompresjonsforholdet er derfor en viktig faktor. Desto mer dataene komprimeres, desto raskere kan de sendes over nettverket, men dessverre vil komprimering i noen tilfeller også redusere kvaliteten på bildet. Noen komprimeringsteknikker krever også mye prosesseringskraft når videoen skal dekodes. H.264<sup>3</sup> er en standard som brukes i flere telepresence-systemer i dag [18]. Denne kodeken er nesten tapsfri<sup>4</sup> og leverer bilder av høy kvalitet med rask overføring [20].

En punkt-til-punkt videosamtale er ikke vanskeligere å sette opp enn en vanlig telefonsamtale. Multi-punkt gjør derimot det hele mye mer komplisert.

---

<sup>3</sup>H.264 er en industristandard for videokompresjon [17].

<sup>4</sup>Tapsfri komprimering innebærer at de originale dataene kan bli rekonstruert eksakt fra de komprimerte [19].

Alle parter skal ha muligheten til å se og høre alle de andre eller noen av de andre partene. Alle deltakerne kan ikke ha direkte kontakt med hver og en av de andre samtidig. Et felles kontaktpunkt benyttes derfor for å koordinere samtalen. Dette kontaktpunktet kalles Video Conferencing Multipoint Control Unit (MCU). Alle datastrømmer passerer gjennom denne enheten og sendes ut til hver deltaker. MCU-en kan ligge på en PC eller, for større systemer, på en ekstern server [9].

### 3.3 Bruksområder

Gjennom 90-tallet gikk videosamtalsystemer fra å være svært dyre til å bli tilgjengelige for allmennheten til en overkommelig pris. De blir mer og mer brukt i dag og antall brukere fortsetter å øke ettersom teknologien utvikler seg og gjør systemene enda bedre og billigere. Bruksområdene er mange og inkluderer alt fra møter i jobbsammenheng til kommunikasjon ved hjelp av tegnspråk for døve og stumme [21, 22]. Videosamtaler benyttes spesielt når visuell informasjon er en viktig del av samtalen, når de som skal snakke sammen ikke fysisk kan komme til samme sted og når reisekostnader og tid er av betydning.

Det er vanlig å si at cirka 90% av vår kommunikasjon er ikke-verbal, altså kroppsspråk, holdning og tonefall [23]. Dette tallet stammer fra et studie rundt en veldig begrenset situasjon og kan ikke sies å gjelde generelt [24]. Uansett er ikke-verbal kommunikasjon en sentral del av en samtale og videosamtaler er av denne grunn som regel ofte å foretrekke framfor for eksempel telefonsamtaler, e-post og lignende. Spesielt i jobbsammenheng er det viktig at man får fram budskapet sitt og unngår misforståelser som følge av dårlig kommunikasjon. Det å ha muligheten til å sette opp jobbmøter med kolleger og kunder som er spredt over store avstander på kort varsel, kan spare bedrifter for store reisekostnader og tidsbruk. I tillegg vil energibruk og forurensing bli lavere gjennom redusert reisevirksomhet [25].

### 3.4 utfordringer

Et problem nesten alle videosamtalsystemer støter på, er at man ikke klarer å simulere naturlig øyekontakt med den man prater med. Øyekontakt er en sentral del av den ikke-verbale kommunikasjonen i en samtale. I jobbsammenheng er det spesielt viktig med øyekontakt når man prater med kunder



### 3.4. UTFORDRINGER

---

[26]. God øyekontakt skaper tillit, mens dårlig øyekontakt ofte assosieres med fiendtlighet og andre negative holdninger.

I en studie testet man behovet for øyekontakt i et telepresencesystem [27]. To videostrømmer ble vist fram for deltakerne: Én hvor man opplevde øyekontakt og én hvor man ikke opplevde øyekontakt. Deltakerne ble så bedt om å identifisere forskjellen i bildene, samt peke ut hvilken video de foretrakk. 93% av de spurte oppdaget forskjellen i bildene og 93% uttrykte at de ville foretrukket å kommunisere med videoen med øyekontakt. Det er ingen tvil om at telepresence-systemer vil dra fordel av naturlig øyekontakt og dermed nærme seg å kunne tilby følelsen av en naturlig samtale.

I de fleste telepresencesystemer i dag oppnår man derimot ikke øyekontakt [27]. utfordringen er å plassere kameraene slik at man ser rett inn i dem samtidig som man ser direkte på skjermen. Det er utarbeidet flere forslag til løsninger, men ingen som tilbyr like god øyekontakt som man opplever ansikt-til-ansikt med alle deltakerne. Digital Video Enterprise<sup>5</sup> sin løsning er å plassere kameraer bak en gjennomsiktig skjerm [29]. Bildet som vises på den gjennomsiktige skjermen er reflektert fra en skjerm som befinner seg under denne. Telcordia<sup>6</sup> eier en patent på øye-til-øye videosamtaler ved å bruke bakprojeksjon på et lerret og plassere kameraene bak lerretet [31]. Sony benytter seg av stråledelere<sup>7</sup> for å oppnå øye-kontakt [33]. Tandberg bruker flere skjermer med inverterte kameraer [34]. Med dette menes at man speilvender rekkefølgen av bildene under avspillingen slik at man vil få øyekontakt med de som er på skjermen rett ovenfor seg. Man vil ikke oppleve naturlig øyekontakt med personer på de andre skjermene.

En annen utfordring innen telepresence er å gjenskape dybdefølelsen vi opplever i virkeligheten. I den virkelige verden ser vi hverandre i tre dimensjoner, mens alle telepresencesystemer i dag kun viser personene i 2D. Noen systemer, som Sony 3D Telepresence og løsningen til Digital Video Enterprise, skaper en unaturlig 3D-effekt ved å speile 2D-bilder på en gjennomsiktig skjerm og blende bakgrunnen i bildet med bakgrunnen i rommet. Bedre 3D, og dermed en sterkere følelse av tilstedeværelse, kan oppnås ved bruk av andre 3D-teknikker.

Bevissthet rundt egen framtoning kan være et problem hos noen under videosamtaler. Ikke alle klarer å oppføre seg naturlig når de vet at de blir filmet og at dette kanskje også tas opp [35]. Dette er noe man kan venne seg til

---

<sup>5</sup>Digital Video Enterprise er et selskap som lager telepresence-systemer [28]

<sup>6</sup>Telcordia driver med forskning og utvikling innen telekommunikasjon [30]

<sup>7</sup>Stråledelere er speil eller prizmer som deler lys i to ved å reflektere deler av strålen [32].

etter hvert. I tillegg kan det være fort gjort å glemme at man er på kamera i gode telepresencesystemer.

# Kapittel 4

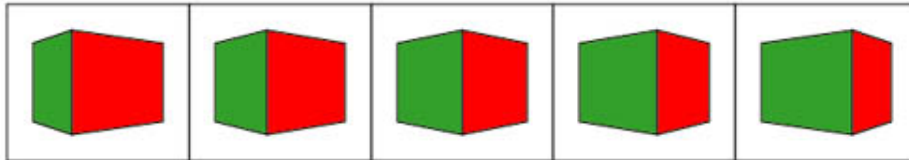
## Kvalitet i 3D

Høsten 2009 skrev vi en prosjektoppgave som omhandler bruk av 3D for å lage virtuelle læringsmiljø for musikere og scenekunstnere [36]. I denne oppgaven konkluderte vi med at bruk av 3D er en faktor som kan være med å øke subjektiv oppfatning av bildekvalitet og følelsen av realisme. Det er dermed nærliggende å tro at 3D kan øke oppfatningen av kvalitet og realisme også i andre bruksområder, i dette tilfellet videosamtaler. For at dette skal kunne være tilfelle er vi avhengige av at 3D-opplevelsen er av en tilfredsstillende kvalitet. Vi skal gå gjennom hvilke parametre som er avgjørende for at et 3D-bilde kan oppfattes som bra, samt de vanligste teknikkene for å vise 3D.

### 4.1 Persepsjon

Det menneskelige synet bruker en rekke hint for å oppfatte dybdedimensjonen i omgivelsene. Disse kan deles i kategoriene monokulære og binokulære hint [37]. De monokulære hintene tar bare i bruk informasjon fra ett øye og tolker denne for å avgjøre hvor i dybdedimensjonen et objekt befinner seg. Én metode er å sammenligne størrelsen på kjente objekter. Dersom det ene av to like store objekter ser ut til å være større, blir det tolket til å være nærmere. Dersom objekter vokser eller krymper, blir de tolket til å bevege seg henholdsvis mot deg og fra deg. Dersom en gjenstand overlapper en annen, blir dette tolket som at den overlappende gjenstanden er den nærmeste. Et annet viktig hint er bevegelsesparallaksen. Når mennesket beveger seg sidelengs, blir objektenes relative bevegelser tolket. Objekter som beveger seg lite i forhold til horisonten blir tolket til å være langt borte, mens objekter som beveger seg mye i forhold til horisonten blir tolket til å være nært. Perspektiv-

endring er også en del av denne tolkningen, jamfør figur 4.1 [38]. Gjenstander som har liten relativ rotasjon og perspektivendring når mennesket beveger seg sidelengs i forhold til objektet blir tolket til å være langt borte.



**Figur 4.1:** Eksempel på perspektivendring ved bevegelsesparallakse

De binokulære hintene utnytter at mennesket har to øyne. Når mennesker ser på et objekt konvergerer øynene mot dette punktet. Ettersom hjernen vet avstanden mellom øynene og vinkelen øynene står i, kan den triangulere og regne ut avstanden til objektet.

## 4.2 Kontrollerbare parametre

Kunstige 3D-bilder blir laget ved å sende ulike bilder til menneskets høyre og venstre øye. Disse bildene etterligner det høyre og venstre øye ville sett dersom de befant seg i scenariet 3D-bildet etterligner. For å kunne lage et kunstig 3D-bilde som er så troverdig at seeren oppfatter skjermbildet til å faktisk være i tre dimensjoner, er det viktig å bruke de monokulære og binokulære hintene riktig. For at den som ser på skal kunne bruke de binokulære hintene, er det viktig at bildene øynene får se er så like virkeligheten som mulig. Et problem som er viktig å løse når man skal lage 3D-bilder, er å sørge for at bildene til høyre og venstre øye ikke flyter over i hverandre. I virkeligheten ser høyre og venstre øye forskjellige bilder, men i de fleste 3D-teknikker vil noe informasjon fra én visning treffe begge øynene. Det er derfor svært viktig at det ikke er andre forskjeller mellom bildene enn de som skal være der på grunn av kameraenes ulike perspektiv.

Den viktigste kontrollerbare parameteren er separasjonen i bildet. Denne kan forklares slik: Når bilder fra to kameraer som står ved siden av hverandre legges oppå hverandre, vil et objekt i bildet vises to ganger. Avstanden mellom disse objektene kalles separasjonen i bildet og er et binokulært hint. Hvor stor separasjonen i bildet er avgjøres av kameraoppsettet: Avstanden mellom kameraene, avstanden til objektet separasjonen måles for og hvor mye kameraene eventuelt konvergerer. Dybdefølelsen man opplever av et 3D-bilde

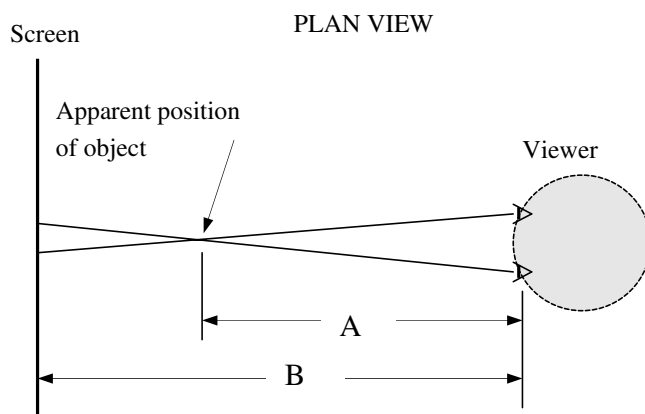
## 4.2. KONTROLLERBARE PARAMETRE

---

avgjøres av variasjonen i separasjonen i bildet for objekter med ulik avstand til kameraene. Det er viktig å passe på når man lager 3D at den maksimale separasjonen i bildet ikke er for stor. Dette kan oppleves som ubehagelig for seeren og kan med enkelte 3D-teknikker gi store feil i bildet. Dersom den maksimale separasjonen i bildet er for lav, eller forskjellene mellom maksimal og minimal separasjon er små, vil bildet oppleves som flatt og med lite dybde. Hvor stor separasjonen i bildet optimalt bør være varierer mellom de ulike 3D-teknikkene, og vil også variere med bildets størrelse.

For at 3D-bildet ikke skal oppleves som ubehagelig, er det viktig at de binokulære og monokulære hintene ikke virker motstridende for hjernen. Hvis et objekt ser ut som det befinner seg foran skjermen, og samtidig er overlappet av kanten av skjermen, mottar hjernen motstridende hint og vil ha trøbbel med å plassere objektet i dybden. Dette gir ubehag over tid og bør unngås.

Et typisk problem for kunstig 3D er det som kalles fokusering-/konvergeringsproblemet. Til vanlig vil øynene våre fokusere og konvergere på samme punkt. Når derimot et objekt vises i 3D på en skjerm, vil øynene konvergere mot punktet hvor hjernen oppfatter at objektet befinner seg, samtidig som øynene fokuserer på skjermen som vist i figur 4.2 [39]. Hvis et objekt befinner seg foran eller bak skjermen vil altså fokuspunktet og konvergeringspunktet være ulike, og dette fører til ubehag og er et av de største problemene med 3D-skjermer i dag. Det er derfor viktig å unngå overdreven dybde når man lager 3D.



**Figur 4.2:** Figuren viser fokuserings-/konvergeringsproblemet. Her konverterer øynene mot objektet på avstand A, mens de fokuserer på skjermen som har en avstand B.

I 4.1 nevnes bevegelsesparallaksen som et viktig element i vårt dybdesyn.

Dette er et fenomen det ikke er mulig å gjenskape med alle 3D-teknikker, først og fremst fordi det kreves at et objekt er filmet med ulike perspektiv for at man skal kunne klare å se det fra ulike perspektiv på skjermen. De faktorene som avgjør hvorvidt bevegelsesparallaksen er realistisk gjenskapt er overgangene mellom bildene fra kameraene som har filmet, antallet kameraer og om det nye perspektivet seeren får når han eller hun beveger seg faktisk stemmer med avstanden han eller hun har beveget seg.

### 4.3 Andre faktorer

Det finnes en rekke andre faktorer som bidrar til at opplevelsen av 3D blir god, som ikke nødvendigvis har så mye med dybdefølelsen i seg selv å gjøre. En viktig faktor er hvor stor del av synsfeltet skjermen dekker. Desto større andel av synsfeltet som blir dekket, desto større innlevelse og følelse av tilstedeværelse [40]. En annen viktig faktor er den objektive bildekvaliteten: Oppløsning, fargedybde, kontrast og antall bilder per sekund.

### 4.4 3D-teknikker

Stereoskopi er en samlebetegnelse på teknikker som krever briller, mens auto-stereoskopi lar deg se 3D uten bruk av briller. For å gi brukeren en illusjon av at bildet de ser har en dybde dimensjon, har alle 3D-teknikker én ting til felles. De må sørge for at høyre og venstre øye mottar ulike bilder. De tre vanligste teknikkene gjør dette litt ulikt.

#### 4.4.1 Polarisasjon

Denne teknikken bruker filter som bare slipper gjennom lys med riktig polarisering. Filtrene er laget parvis slik at lyset som slipper gjennom det ene filteret ikke slipper gjennom det andre. Seeren bruker briller som har ulike filtre for høyre og venstre øye. Ved å bruke filtre med motsatt polarisering ved skjermen, kommer lyset fra høyre bildekanal bare til høyre øye og tilsvarende for venstre bildekanal. Filtringen ved skjermen kan skje enten ved at bildene blir sendt med dobbel bildefrekvens og med ulik polarisering for anethvert bilde, eller man kan projisere høyre og venstre bildekanal på samme skjerm med ulik polarisering [41].

### 4.4.2 Lukkerbriller

Lukkerbriller er spesielle briller som stenger for lyset som kommer inn til høyre og venstre øye for annethvert bilde. Disse bildene er synkronisert med skjermen som viser venstre og høyre bildekanal i annethvert bilde. På den måten ser øynene ulike bilder [42].

### 4.4.3 Autostereoskopi

I autostereoskopi skjer fordelingen av bildet til høyre og venstre øye i selve skjermen ved at bilder sendes i ulike retninger. Denne teknikken blir forklart ytterligere i kapittel 6.





# Kapittel 5

## 3D-situasjonen i dag

Dette kapittelet gir en oversikt over populariteten til 3D-teknikker og hvilke bruksområder de nylig har fått og snarlig vil få.

### 5.1 Stereoskopi

3D har opplevd økt popularitet de siste årene [36]. Spesielt er det stereoskopisk 3D som har dominert markedet. Bruksområder spenner fra spill og annen underholdning til arbeid for eksempel innen medisin og oljebransjen [43]. Bruk av 3D i kinosaler er godt etablert, og flere og flere filmer lages i både 2D- og 3D-versjoner. De fleste kinoer som tilbyr 3D benytter seg av polariserte briller, jmfør 4.4.1. 3D-TV-er er også på full fart ut på markedet, og i Norge kan du nå kjøpe slike TV-er fra blant annet Samsung, Panasonic og Philips i vanlige nettbutikker [44]. 3D-TV-ene på markedet er tilgjengelige i mange størrelser og prisklasser [45]. For å bruke disse TV-ene trenger man enten polariserte briller eller lukkerbriller, avhengig av teknologien produsentene bruker [46]. 3D var den mest omtalte teknologien under CES 2010<sup>1</sup> hvor det foruten 3D-skjermer ble vist fram 3D-kameraer med full HD. Det er utviklet en 3D-standard for BluRay, og Playstation 3 vil støtte 3D-spill i løpet av sommeren 2010 [48, 49, 50]. At YouTube<sup>2</sup> og Google Earth<sup>3</sup> nå støtter 3D er andre gode indikasjoner på at 3D er i vinden om dagen og at det er kommet for å bli [51, 52].

---

<sup>1</sup>Consumer Electronics Association (CES) er en årlig utstilling hvor det vises fram forbrukerteknologi [47]

<sup>2</sup>YouTube er verdens største videonettsted [51].

<sup>3</sup>Google Earth er en svært populær digital globus [52].

## 5.2 Autostereoskopi

I kinosammenheng, hvor man sitter stille og som regel har nesten full fokus på filmen, er briller akseptert og mangelen på bevegelsesparallakse, jamfør 4.1, er ikke et problem [53]. Når 3D-opplevelsen flyttes inn i hjemmet derimot, vil brukerne ha helt andre krav enn i en kinosal. I en slik situasjon er man vant til å kunne bevege seg fritt og full oppmerksomhet er ikke alltid viet TV-skjermen. Briller blir sett på som en stor hindring og det er usikkert hvor godt 3D-TV-er som krever briller vil slå igjennom. Det er derfor ikke usannsynlig at autostereoskopiske skjermer vil bli en større suksess enn skjermer som krever briller. Autostereoskopi kan sees på som neste steg etter stereoskopi, og et søk på “autostereoscopic” i Google Trends<sup>4</sup> viser en økende interesse for autostereoskopi i 2009 og 2010 [55]. Et søk på “autostereoscopic” og “stereoscopic” i Google Insights<sup>5</sup> viser likevel at det fortsatt er mye større interesse for stereoskopi [57]. Autostereoskopiske 3D-TV-er er allerede vist fram av produsenter som Mitsubishi, Haihatsu og Philips, og forventes å komme på markedet om ikke lenge [45] [58]. Under CES 2010 kunne 3D Eye Solutions<sup>6</sup> vise fram autostereoskopiske skjermer som viste filmer konvertert fra 2D til 3D [47]. Etter hvert vil de kunne konvertere direktesendte TV-sendinger i 2D til 3D på direkten.

Autostereoskopi har flere bruksområder hvor 3D-briller er ugunstig å bruke. For eksempel gjelder dette reklameplakater som er et marked hvor løsninger basert på bruk av briller ikke kan brukes [58] [60]. Mobiltelefon skjermer er et annet område hvor autostereoskopisk teknologi trekker det lengste strået. En mobiltelefon med 3D-skjerm er allerede lansert i Japan [61]. Nintendo sjokkerte mange ved å kunngjøre at deres neste håndholdte spillkonsoll, med foreløpig navn Nintendo 3DS, vil støtte spill i 3D uten behov for briller [62]. Denne skal etter planen lanseres allerede i løpet av første kvartal i 2011 [63].

Selv om diverse produkter basert på autostereoskopi kommer på markedet, er det viktig å huske på at teknologien har sine utfordringer og at det fortsatt er uenighet rundt hvordan man lager de beste skjermene. Det finnes også 3D-teknologier med mye større potensiale enn autostereoskopi, men disse er fortstatt på prototypestadiet. Et eksempel på hva vi kan forvente å oppleve i framtiden er følbare hologrammer [64].

---

<sup>4</sup>Google Trends viser søketrender over tid [54]

<sup>5</sup>Google Insights lar deg blant annet sammenligne søketrender over tid [56]

<sup>6</sup>3D Eye Solutions er et selskap som utvikler 3D-medier [59]

# Kapittel 6

## Autostereoskopi

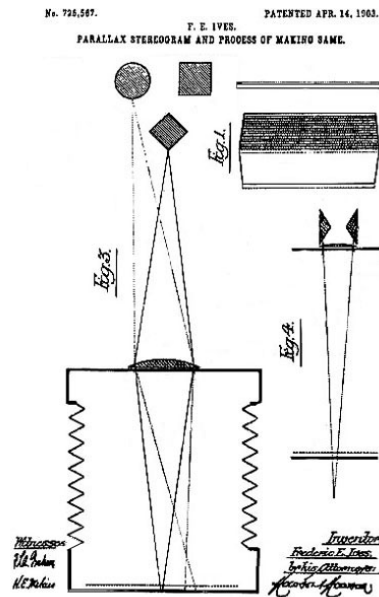
I dette kapitlet vil vi gå gjennom autostereoskopiens historie og hvordan det fungerer, samt se på fordeler og ulemper med autostereoskopi sammenlignet med stereoskopi.

### 6.1 Historie

En autostereoskopisk metode ble demonstrert allerede i 1692 [65]. Det var den franske maleren G. A. Bois-Clair som prøvde seg på en enkel utgave av det som senere er blitt kjent som barriere-teknikken, som vi vil gå nærmere inn på i 6.5.1. I 1903 ble fotografisk film tatt i bruk og patentert av Frederick E. Ives under navnet “Parallax Stereogram”, se figur 6.1.

Sønnen til Frederick E. Ives, Herbert E. Ives, forsket videre på emnet fram til tidlig på 1940-tallet. Han eksperimenterte blant annet med oppsett med flere kameraer og autostereoskopisk film. Et av hans viktigste bidrag var hans kameraer som kunne filme flere visninger samtidig. Disse la grunnlaget for “skannende” kamerasystemer som fortsatt er i bruk i dag. De viktigste bidragsyterne for slike systemer inkluderer Douglas Winnek fra USA og Marrice Bonnet fra Frankrike. Kamerasystemene kalles gjerne Bonnet-style-kameraer.

Tidlig på 1900-tallet ble det også foreslått andre autostereoskopiske metoder. Professor Gabriel M. Lippmann byttet ut barrieren fra barriere-teknikken med en serie av linser i 1908. Teknikken betegnes i dag som integral-teknikken etter Lippmanns originale navn “La Photographie Integral”. Teknikken har deretter blitt kraftig forbedret, spesielt de siste 30 årene. I 1978 benyttet



Figur 6.1: Patentet til Frederick E. Ives: “The Parallax Stereogram”

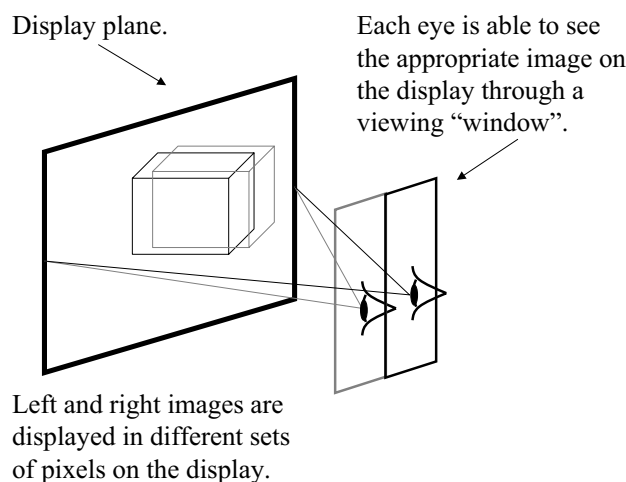
man for første gang PC-er for å digitalt flette sammen flere PC-genererte 2D-bilder for å skape 3D-bilder. Selv om integral-teknikken har potensiale, har den enda ikke opplevd den kommersielle suksessen som for eksempel lentikulær-teknikken opplever akkurat nå.

Lentikulær-teknikken forenkler integral-teknikken ved å ta i bruk lentikulære linsematriser<sup>1</sup>. Disse er optisk like skjermen i barriere-teknikken, og man kunne derfor ta utgangspunkt i forskning som var gjort tidligere. Mange forskere jobbet videre med lentikulær-teknikken i 1930-årene, men det var ikke før på 60-tallet at man opplevde virkelig rask fremgang. Man oppdaget det store potensialet autostereoskopi hadde i reklamebransjen, og i 1964 ble det for første gang tatt i bruk i en utgave av magasinet Look Magazine. De neste tjue årene ble en rekke produkter produsert før interessen falt. På slutten av 80-tallet da PC-er begynte å bli vanlig tok interessen seg opp igjen og i dag, med alt fra ny linseteknologi til høyere bildeoppløsning, opplever autostereoskopi en ny vår.

<sup>1</sup>En lentikulær linse består av en matrise med mange små linser som forstørrer forskjellige bilder avhengig av posisjonen man ser fra [66]

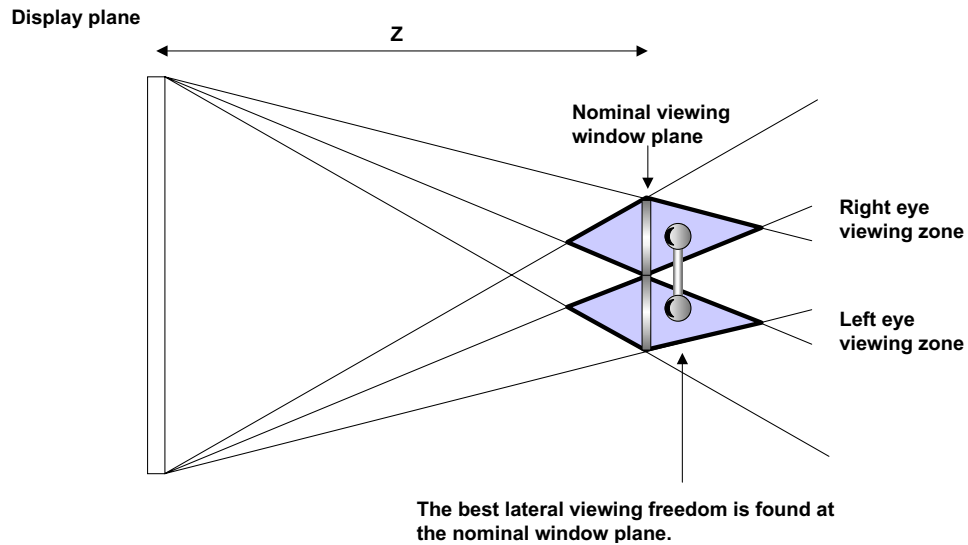
## 6.2 Visningsvindu

En visning (engelsk: view) er et sentralt begrep innen autostereoskopiteorien. Det er definert som det todimensjonale bildet skjermen sender i en bestemt retning [67]. For å kunne oppleve 3D trengs det minimum to slike visninger som rettes i forskjellige retninger slik at øynene mottar forskjellige bilder. Visninger danner det som kalles visningsvinduer (engelsk: viewing window) som er et annet grunnleggende prinsipp i virkemåten til autostereoskopiske skjermer. Når vi ser på vanlige 2D-skjermer er vi vant til å kunne befinne oss i mange ulike posisjoner og fortsatt se et bilde av høy kvalitet. Ved bruk av autostereoskopiske skjermer begrenses denne friheten betraktelig. Her må de forskjellige bildene på skjermen sendes separat til øynene ved bruk av for eksempel barriere-teknikken. Vi vil dermed få visningsvinduer hvor ett øye optimalt bare ser ett bilde. For å se et korrekt 3D-bilde er det nødvendig at pupillene befinner seg innenfor hvert sitt vindu. Gjennom disse vinduene vil øynene se litt forskjellige bilder og man vil dermed oppleve 3D, jamfør figur 6.2 [68]. Sett ovenifra er visningsvinduene formet som diamanter og øynene kan bevege seg fritt innenfor disse både i høyde-, bredde- og lengderetningen som vist i figur 6.3.



**Figur 6.2:** Visningsvinduene som skapes av visningene

Posisjoner som gir korrekte 3D-bilder kalles gjerne sweetspots [69]. Øynene i figur 6.3 befinner seg med andre ord i en sweetspot [68]. Sweetspots oppgis gjerne i centimeter fra skjermen og avstanden du kan bevege øynene sidelengs. Beveger man seg utenfor disse områdene vil bildekvaliteten avta helt til man eventuelt nærmer seg en ny sweetspot. Desto lengre man kan bevege seg



**Figur 6.3:** Visningsvinduene er formet som diamanter. Man vil oppleve 3D så lenge øynene befinner seg i hvert sitt visningsvindu.

sidelengs uten at bildekvaliteten avtar, desto mer komfortabelt blir det å se på skjermen. Denne bredden begrenses av avstanden mellom øynene våre som har et gjennomsnitt på cirka 63 mm [70].

I noen tilfeller vil man kunne akseptere den begrensede friheten autostereoskopiske skjermer gir til å bevege seg. Et eksempel på dette er mobiltelefonbruk. Det er sjeldent at man ser på en mobiltelefonskjerm fra siden. I andre situasjoner vil den begrensede friheten være forstyrrende og også direkte uakseptabel. For at autostereoskopiske skjermer skal bli akseptert i hjemmet, er det naturlig å tro at de må oppfylle samme egenskaper som 2D-skjermene man er vant til. Mulige løsninger for å gi brukerne mer bevegelsesfrihet foran skjermen blir presentert i de to neste delkapitlene: Bruke head tracking eller øke antallet visninger.

### 6.3 Head tracking

For å forbedre bevegelsesfriheten til seeren, kan man bruke en teknikk som kalles head tracking. Ved å spore brukernes posisjon til enhver tid, kan riktige visninger, og kun disse, hele tiden rettes mot hvor øynene til seeren befinner seg. På denne måten utvides området man kan befinne seg i og optimale visningsforhold oppleves i hele rekkevidden [71, 72]. I tillegg kan man med

denne metoden vise høyoppløselig 3D som nå er blitt vanlig i stereoskopiske systemer. Det er også på sin plass å nevne at head tracking ikke nødvendigvis er noen mirakelkur for autostereoskopiske systemer. Som det meste av teknologi som brukes i forbindelse med autostereoskopi, finnes det utfordringer og problemer også ved head tracking [73].

## 6.4 Multiview

Problemet med at man må befinne seg innenfor et smalt visningsvindu for å se bildet korrekt, kan delvis løses ved å bruke flere visninger og på den måten få flere visningsvinduer. En ekstra effekt av dette er at det blir mulig å gjenskape bevegelsesparallaksen. Det vil si at man får muligheten til å se et objekt fra et nytt perspektiv når man beveger seg sidelengs som vist i figur 4.1. Dette vil ikke gjelde når man beveger seg opp og ned. Autostereoskopiske skjermer med mellom tre og flere hundre visninger betegnes som multiview.

Selv om det for mange anvendelsesområder vil lønne seg å ha så mange visninger som mulig, har mange visninger en stor ulempe. Ettersom en skjerm har et fast antall bildepunkter, blir antallet bildepunkter per visning lik antallet bildepunkter på skjermen delt på antallet visninger. En skjerm med full HD, altså en oppløsning på  $1920 \times 1080$ , har cirka to millioner piksler. Hvis denne har fem visninger vil hver av disse kun ha rundt 400 000 piksler. Det tilsvarer en oppløsning på for eksempel  $854 \times 480$ , også kjent som WVGA<sup>2</sup>. Dette er samme oppløsning som brukes på mange mobiltelefoner i dag [75]. Dette er et viktig ankepunkt for autostereoskopiske skjermer med mange visninger, men det foreligger ingen klare svar på hva som er det beste kompromisset mellom mange visninger og høyest mulig oppløsning per visning.

En annen ulempe med å ha mange visninger er at antallet kamera som må brukes må tilsvare antallet visninger. Å bruke mange kamera er upraktisk av flere årsaker, først og fremst er det dyrt og det kompliserer også kalibrering og kameraoppsett. Det er mulig å redusere antallet kamera ved å bruke dybdekart. Dybdekartene muliggjør generering av virtuelle visninger ut fra så få som én eller to videostrømmer [76]. Selv om man kan redusere antallet kamera med bruk av dybdekart, vil bruk av virtuelle visninger aldri gi like gode resultater som å bruke et kamera.

Multiview-skjermer er som regel LCD-skjermer kombinert med optiske ele-

---

<sup>2</sup>WVGA er navnet på skjermoppløsninger som inkluderer blant annet  $800 \times 480$ ,  $848 \times 480$  og  $854 \times 840$  [74]

menter. En vellykket måte å lage høykvalitets autostereoskopiske skjermer på, som blant annet Sharp har sett nærmere på, er å bruke to LCD-TV-er og rette lyset fra den ene til det venstre øyet og fra den andre til det høyre øyet [77]. På den måten vil man få full oppløsning på hvert øye. Ulempene er at man får ekstra kostnader i form av en ekstra skjerm i tillegg til at løsningen er lite brukervennlig. Dette gjør den uegnet til masseproduksjon og kommersiell bruk og vi skal derfor ikke gå nærmere inn på den her. Det er utviklet mange flere teknikker, men de to vi skal se nærmere på er parallaksebarriere og lentikulær separasjon. Disse ansees som hovedteknikkene og brukes i de fleste autostereoskopiske skjermer i dag [78].

## 6.5 Virkemåte

Grunnprinsippet i autostereoskopi er det samme som i vanlig stereoskopi. Øynene presenteres for to forskjellige bilder og hjernen oppfatter bildet som at det er i tre dimensjoner, jamfør 4.1. Bildene viser de samme objektene, men fra ulike vinkler, og på den måten etterligner man synet slik det fungerer i virkeligheten. Stereoskopiske teknikker benytter hjelpemidler ved øynene i form av spesielle briller for å levere ulike bilder til venstre og høyre øye, mens autostereoskopiske teknikker benytter optiske elementer på skjermen for å få til det samme. Man deler gjerne inn skjermtyper som ikke krever briller slik: binokulære, multiview, holoforme og holografiske. Det som skiller disse fra hverandre er antall visninger de viser, som vist i figur 6.4 [79].

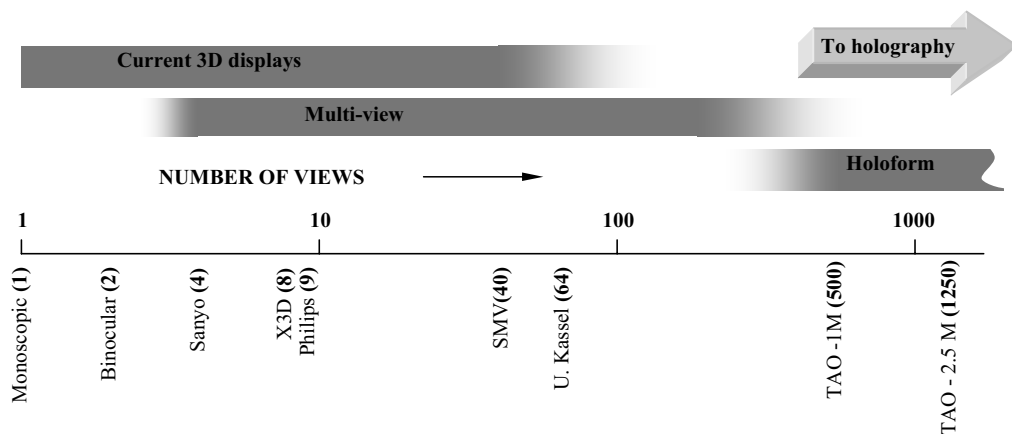


Figure 1. Multiple View 3D Display Spectrum

**Figur 6.4:** Figuren viser hva man kaller skjermene etter hvor mange visninger de har.



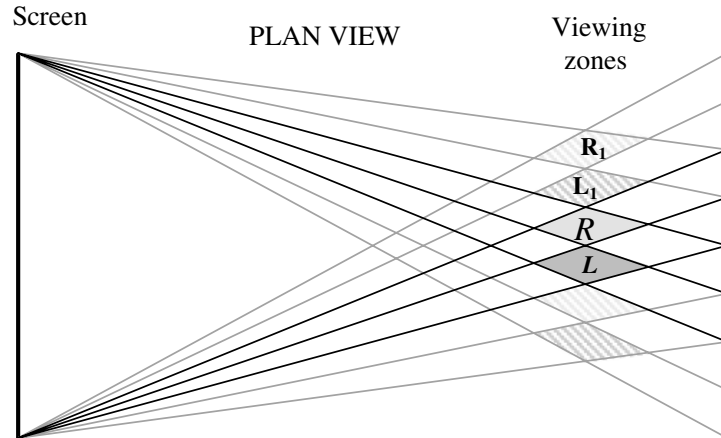
Multiview-skjermer med så mange visninger at man opplever nær kontinuerlig bevegelsesparallakse, jamfør 4.1, og hvor det ikke er forskjell på punktet øynene konvergerer mot og hvor de fokuserer, jamfør 4.2, kalles holoforme. The Telecommunications Advancement Organisation (TAO) i Japan har kommet fram til at et stort antall visninger vil la øyenene fokusere på samme distanse som de konvergerer. Lignende resultater er lagt fram av Heinrich Hertz Institut som har uttalt at det trengs typisk 20 visninger per øyeavstand for å oppnå jevn bevegelsesparallakse [80]. TAO setter kravet til ca 32 visninger per øyeavstand. Totalt antall visninger vil variere med bredden man skal kunne bevege seg i, og vil fort bli ganske stort. Neste steg etter holoforme skjermer er å gjenskape objekter nøyaktig slik som de framkommer i virkeligheten. Dette er kun mulig med holografiske teknikker.

### 6.5.1 Parallaxebarriere

Autostereoskopiske skjermer som tar i bruk parallaxebarriere har en enkel struktur bestående av en flatskjerm og en parallaxebarriere. Flatskjermene er som regel av typen LCD. Skjermer med parallaxebarriere kan ha to eller flere visninger. I en skjerm med to visninger, har du to bilder som er filmet fra litt forskjellige vinkler. Disse konverteres til flere elementer og plasseres annenhver gang i kolonner på piksler i skjermen. For å rette lyset fra elementene av det høyre bildet mot det høyre øyet og lyset fra elementene av det venstre bildet mot det venstre øyet, plasseres en barriere foran skjermen. Små åpninger i en ellers ugjennomsiktig barriere gjør at lyset fra pikslene rettes i riktig retning. Resultatet er visningsvinduer, jamfør 6.2, for venstre og høyre bilde som vist i figur 6.5 [39].

Du vil få den optimale bildekvaliteten ved å ha høyre øye i  $R$  og venstre øye i  $L$ . Visningsvinduer av dårligere kvalitet ( $R1$  og  $L1$ ) vil repeteres ut til begge sidene fra midtposisjonen på grunn av geometrien i skjermoptikken. Fordelen med dette er at det kan være flere enn én seer hvis den optiske kvaliteten er god nok. Ulempen er at hvis venstre og høyre øye befinner seg i henholdsvis  $R$  og  $L1$ , vil den opplevde dybden være motsatt av den virkelige dybden; man får det som kalles et pseudoskopisk bilde. Dette gjør det vanskelig, spesielt for personer med lite erfaring med å se 3D, å avgjøre om de ser korrekt 3D eller ikke. Det er foreslått flere måter å løse dette på, for eksempel head tracking.

Prinsippene fra filming av objekt til visning på skjerm er vist i figur 6.6 [81]. Se foreløpig bort i fra at bildet viser et multiview-oppsett. Prinsippet til parallaxebarriere er vist mer detaljert i figur 6.7 [68]. I figur 6.7 ser man venstre ( $L$ ) og høyre ( $H$ ) bilde liggende annenhver gang bortover skjermen

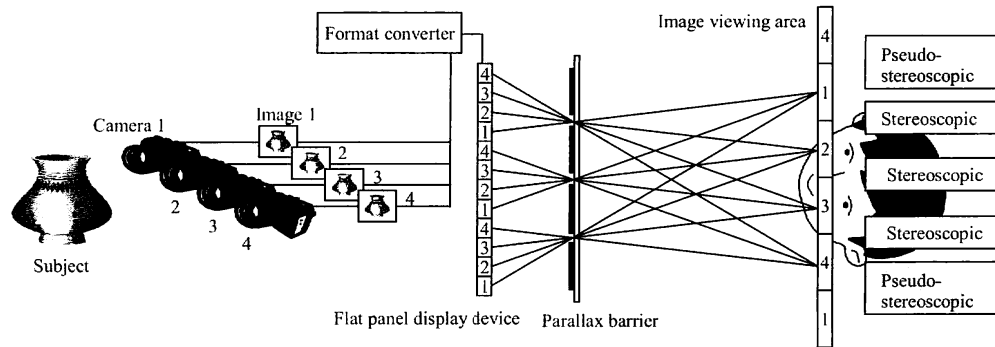


**Figur 6.5:** Figuren viser vinduer som repeteres ut til sidene.

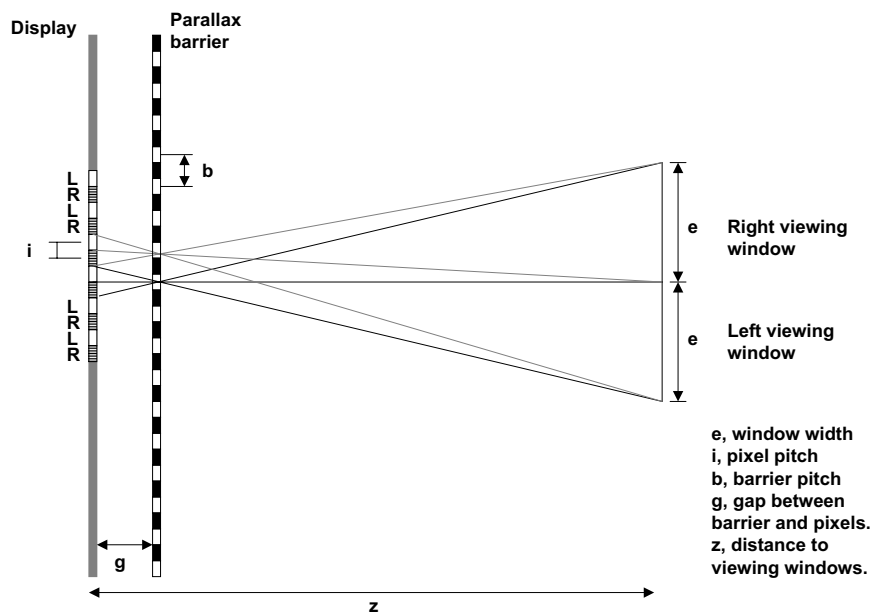
med pikselavstand  $i$ . Avstanden fra pikslene til barrieren,  $g$ , bestemmes av tykkelsen på LCD-skjermen. Optimal avstand fra piksler til seeren er  $z$ . Bredden på visningsvinduene varierer med avstanden fra skjermen og betegnes  $e$ . Bredden av barriereåpningen er betegnet  $b$ . Ved hjelp av geometri finner man sammenhenger mellom disse variablene som kan brukes i design av skjermer for blant annet å bestemme optimal seeravstand. Optimal seeravstand er bestemt av  $e$  som typisk er satt til gjennomsnittlig øyeavstand på cirka 63 mm, og  $i$  og  $g$  som bestemmes av LCD-skjermen [68]. Optimal seeravstand begrenses med andre ord av teknologien for LCD-skjermer. Typisk seeravstand på en 24-tommer og en 42-tommer i dag er i dag henholdsvis 70 cm og 335 cm, jamfør 7.2. I noen tilfeller kan det være ønskelig med enda kortere avstand, for eksempel for å oppnå økt følelse av tilstedeværelse. Det forskes på å lage tynnere LCD-skjermer som vil gi mindre  $g$ , som igjen vil kunne minske optimal seeravstand [82].

Multiview-versjoner av skjermer med parallaksebarriere fungerer på samme måte som skjermer med to visninger, bare med flere forskjellige bilder. En skjerm med  $n$  visninger har  $n$  antall bilder lagt ved siden av hverandre gjentatte ganger i samme rekkefølge som kameraposisjonene, se figur 6.6. Slike skjermer vil vise pseudoskopiske bilder når øynene er plassert i visning 1 og  $n$ . Desto større antall visninger man har, desto større område kan man bevege seg i før man ser pseudostereoskopiske bilder. Ulempen med flere visninger er at bildekvaliteten reduseres. De tilgjengelige pikslene på skjermen må fordeles på  $n$  bilder. Det vil si at hvert bilde, og dermed hver visning, vil ha en oppløsning på en  $n$ -te-del av den totale skjermopløsningen. Ønsker man  $n$  visninger med oppløsning på  $r \times s$ , må skjermen ha en oppløsning på  $n \times r \times s$

## 6.5. VIRKEMÅTE



Figur 6.6: Figuren viser prinsippene fra filming til visning på skjerm.

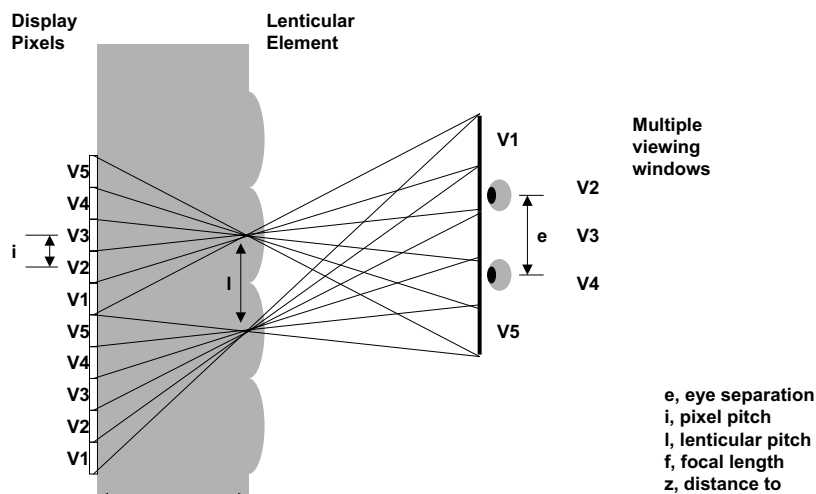


Figur 6.7: Figuren viser prinsippet bak parallaksebarriere.

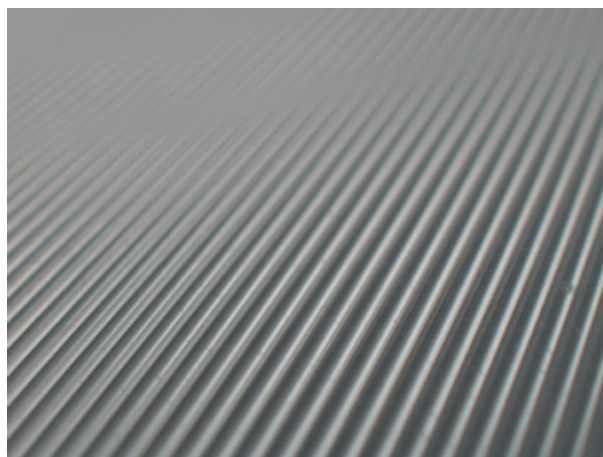
[83]. Oppløsningen reduseres kun i horisontal retning hvis barrieren står parallelt med pikslene. Ved å plassere barrieren litt på skrå balanserer man tapet av oppløsning i begge retninger [84]. Når barrieren er rotert i forhold til pikslene minimerer man også Moiré-effekten, som gir et interferensmønster mellom barrieren og pikslene [71].

### 6.5.2 Lentikulær separasjon

Konseptet i skjermer som benytter lentikulær separasjon er akkurat det samme som i skjermer som benytter seg av parallaksebarriere. Et optisk element sender lyset fra pikslene i forskjellige retninger slik at øynene, når de er plassert riktig, kun mottar ett bilde hver. Ved parallaksebarriere gjøres dette ved at uønskede deler av lyset blokkeres av en barriere. Ved lentikulær separasjon reflekteres lyset i riktig retning ved hjelp av en lentikulær linse plassert foran pikslene som vist i figur 6.8 [68]. Den lentikulære linsen er satt sammen av mange små sylindriske linser, ordnet vertikalt i forhold til pikslene og vil da se ut som i figur 6.9 [66]. Linsene fungerer etter prinsippet om lysbrytning for å vise kun deler av det underliggende bildet til øynene basert på øynenes posisjon i forhold til linsene.



Figur 6.8: Figuren viser prinsippet bak lentikulær separasjon.

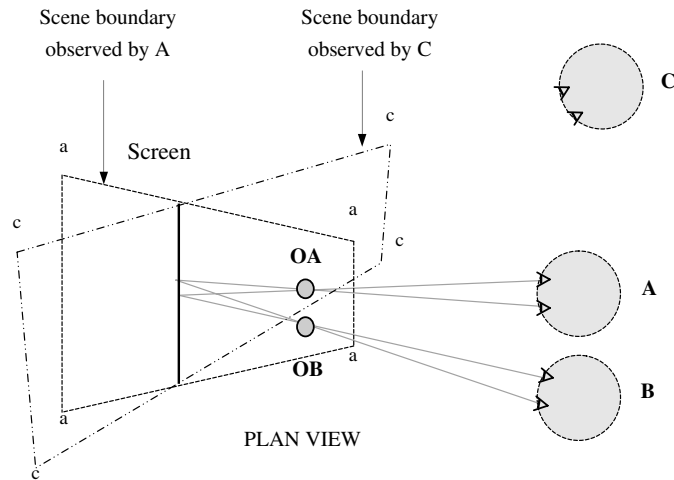


**Figur 6.9:** En lentikulær linse er satt sammen av mange små sylindriske linser

### 6.6 Multiview autostereoskopi sammenlignet med stereoskopi

Den mest åpenbare forskjellen mellom autostereoskopi og stereoskopi er at med autostereoskopi behøver man ikke bruke briller. I tillegg vil man kunne oppleve bevegelsesparallaksen, jamfør 4.1, ved bruk av multiview autostereoskopiske skjermer. Dette er to gode argumenter for å velge autostereoskopi. Et tredje argument er at noe som kalles geometrisk fordreining, vanlig i stereoskopiske systemer, kan elimineres. Geometrisk fordreining vises som fordreining i dybden av bildet og feilaktig rotasjon av bildet. Dette er vist i figur 6.10 [39]. Dybdefordreining gjør at dybden øker med avstanden til skjermen. En person i posisjon  $A$  kan for eksempel se objekter innenfor rektangelet “a,a,a,a”. Fra posisjon  $C$  vil man se de samme objektene, men innenfor linjene “c,c,c,c”. Feilaktig rotasjon er det som skjer når man ikke har bevegelsesparallakse. Et objekt i bildet vil ikke stå i ro eller kunne observeres fra flere vinkler mens man beveger seg rundt det. I stedet vil objektet flytte seg slik at det alltid har samme posisjon i forhold til seeren og skjermen. Dette er vist med objektene  $OA$  og  $OB$ , samt posisjonene  $A$  og  $B$ . Dette kan løses med head tracking i stereoskopiske systemer og uten head tracking i autostereoskopiske systemer.

En av ulempene med autostereoskopiske skjermer er at det å lage visningsmateriale er mye mer komplisert enn ved stereoskopiske systemer. I stereoskopiske systemer bruker man to kameraer. For å lage materiale som kan vises



**Figur 6.10:** Figuren viser den geometriske fordreiningen som oppstår i stereoskopiske systemer.

på en multiview autostereoskopisk skjerm, trengs det like mange kameraer som antallet visninger. Eventuelt kan man bruke færre kameraer og bruke et dybdekart til å generere virtuelle mellomliggende visninger, jamfør 6.4. Desto flere kamera som må kalibreres og synkroniseres, desto mer avansert og vanskelig blir det å gjøre opptak til autostereoskopiske multiview-skjermer. Opptakene må også settes sammen til et format som skjermen forstår og kan vise. Etersom bildene må etterbehandles, stiller det høyere krav til maskinvaren når levende bilder skal vises i sanntid.

## 6.7 Ubehag

Fokusering-konvergering-problemet, nevnt i 4.2, er et problem for autostereoskopiske skjermer, som for de fleste 3D-teknikker. Som nevnt i 6.5, kan dette løses i autostereoskopi ved å vise veldig mange visninger. Det er også forsket på løsninger på dette problemet til stereoskopiske systemer [85]. Ingen av løsningene til verken autostereoskopi eller stereoskopi i dag lar seg gjøre på en billig og enkel måte.

Kvaliteten på visningsvinduene som produseres har stor betydning for den opplevde kvaliteten av autostereoskopiske skjermer. Spesielt er det viktig å prøve å unngå et fenomen som kalles cross talk [68]. Ved cross talk ser man doble bilder og dette refereres ofte til som ghosting. Ved lave verdier av cross

## 6.7. UBEHAG

---

talk oppleves bildet som uskarpt, mens ved høye verdier kan 3D-effekten forsvinne helt. Cross talk oppstår når et øye mottar deler av en eller flere visninger som det egentlig ikke skal se. Dette kan skyldes alt fra oppsett av kameraene som brukes ved opptak, til dårlig behandling av bildekanalene i skjermelektronikken. Uten perfekt avgrensede bildekanaler, vil øyet motta deler av andre visninger når det befinner seg i et visningsvindu. Dette er et kjent problem i dagens skjermer og også noe vi erfarte i arbeidet med våre autostereoskopiske skjermer. Cross talk oppstår også hvis man finner seg i soner mellom sweetspots.





# Kapittel 7

## Utstyr og programvare

I dette kapittelet går vi gjennom bakgrunnen for valgene vi tok for bruk av teknologi og utstyr for å lage 3D-video. Videre blir utstyret presentert, og alle egenskapene disse har som påvirker hvordan 3D-film bør lages for å få til best mulig kvalitet blir vurdert.

### 7.1 Bakgrunn for teknologivalg

Årsaken til at vi har valgt å fokusere på autostereoskopiske skjermer når vi skal se på 3D i videosamtaler, ligger i videosamtalens natur. Autostereoskopiske skjermer har fordelen av at man ikke trenger spesielle briller for å se på dem, i motsetning til mange andre 3D-teknologier, jamfør 6.6. Når vi ser på bruk av 3D i videosamtaler er dette for å kunne tilby en økt følelse av realisme i videosamtalen [36]. Økt realisme i videosamtalen vil forsterke deltakernes følelse av at dette er en samtale med mennesker som befinner seg i samme rom. Videosamtaler blir ofte brukt som erstatning for møter og samtaler der mennesker møtes ansikt til ansikt, jamfør 3.1. I samtaler mellom mennesker er øyekontakt viktig, og å måtte bruke spesielle briller vil ødelegge mye av denne kontakten, jamfør 3.4. I tillegg burde ikke deltakerne i en videosamtale være i stand til å se mellomliggende teknologi og hjelpemidler som gjør kommunikasjonen mulig [76]. Dette svekker naturligheten av samtalen betraktelig, og bruk av spesielle briller er dermed uaktuelt for videosamtaler. Det er derfor naturlig å se på autostereoskopiske skjermer når en skal vurdere bruk av 3D i videosamtaler.

For best mulig å etterligne verden slik vi opplever den, burde man i video-

samtaler forsøke å gjenskape bevegelsesparallaksen, slik at deltakerne opplever riktig perspektiv av personene de snakker med. En person som kommer gående inn i konferanserommet vil dermed også oppleve ulike perspektiver i overensstemmelse med hans posisjon i forhold til skjermen. Bevegelsesparallakse vil også gjøre det enklere å oppfatte retningen av håndbevegelsene til personene på skjermen. Dette er viktig for å ivareta den naturlige følelsen av samtalen når for eksempel en deltaker retter seg mot eller peker mot en annen deltaker. For å oppnå bevegelsesparallakse med en autostereoskopisk skjerm, må man benytte en skjerm av typen multiview, jamfør 6.4. Denne typen skjermer støtter også flere brukere samtidig, noe som er et krav i videosamtaler med flere enn to personer. Vi vil derfor se på skjermer med multiview i vår vurdering av 3D i videosamtaler.

Selv om autostereoskopisk 3D er mer aktuelt nå enn noen gang før, jamfør 5, er det ikke enkelt å anskaffe en autostereoskopisk skjerm. Vårt valg, for det meste basert på leveringstid, falt på en 42-tommer skjerm fra det tyske selskapet Tridality. Fra før hadde vi tilgjengelig en 24-tommer skjerm fra samme selskap. Disse to skjermstørrelsene gir gode muligheter for å se på ulike typer videosamtaler med to eller flere deltakere. Begge disse har fem visninger, jamfør 6.2.

Kameraene vi har valgt å bruke er valgt ut fra at de gir utmerket bildekvalitet størrelsen tatt i betraktning. Etttersom avstanden mellom kameraene er viktig, har smale kamera gitt oss muligheten til å gjøre forsøk med både liten og stor kameraavstand. Valg av antall kameraer er tatt med hensyn at på skjermene vi har hatt tilgjengelig støtter fem visninger. Det har derfor vært naturlig å benytte seg av fem kameraer, ett for hver visning. Å bruke ett kamera for hver visning gir en del utfordringer. Minimumsavstanden mellom kameraene bestemmes av størrelsen de har, og i noen tilfeller er det ønskelig med mindre avstand enn dette tillater. I tillegg kan kameraantallet bli veldig stort etter hvert som man får skjermer med flere visninger. Disse problemene kan løses ved hjelp av et dybdekart, jamfør 6.4. Dette kan være en bedre løsning for videosamtaler, men dybdekart ligger utenfor omfanget av denne oppgaven.

I forbindelse med autostereoskopiske skjermer kan det også være naturlig å se på head-tracking, jamfør 6.3. Grunnet mangel på utstyr og begrenset tid har vi valgt å ikke gå nærmere inn på head tracking i denne oppgaven.

I valg av programvarene har vi brukt programmer som er anerkjente og markedsledende innen sine felt [86, 87].

## 7.2 Skjermer



**Figur 7.1:** 24-tommer skjerm fra Tridality



**Figur 7.2:** 42-tommer skjerm fra Tridality

Vi har hatt to autostereoskopiske multiview-skjermer på 24 og 42 tommer tilgjengelige. Begge har fem visninger, er produsert av den tyske produsenten Tridality og bruker en teknikk som kalles skrå parallaksebarriere, jamfør 6.5.1 [88]. Utover størrelsen og noen tekniske detaljer har de mange egenskaper felles. Vi har også hatt en ordinær 46-tommer Samsung-skjerm med full HD-oppløsning og god bildekvalitet som har blitt brukt til å vise 2D-video.

### 7.2.1 Separasjon i bildet

I følge en epost<sup>1</sup> fra Tridelity, må separasjonen i bildet ikke overstige 1,5% av den totale bredden av bildet. Dersom den gjør det, vil det oppstå ghosting, jamfør 6.7. I følge dokumentasjonen fra Tridelity som følger med skjermene, oppgis det også at dersom man lager et triangel mellom to visninger av samme objektet på skjermen og seeren, kan ikke vinkelen ved seeren være mer enn 1,5 grader før det oppstår ghosting.

### 7.2.2 Krav til beskjæring

En viktig egenskap ved skjermene er at en visning ikke kan inneholde bildeinformasjon som ikke finnes i en annen visning uten at det oppstår ghosting. En autostereoskopisk skjerm kan altså ikke vise noe fra et kamera som ingen av de andre kameraene ser uten at det oppstår feil i bildet. Bildet den autostereoskopiske skjermen viser er altså i praksis snittet av det kameraene fanger. Bildene fra kameraene må dermed beskjæres før de kan brukes på den autostereoskopiske skjermen. Hvor mye bildet fra et kamera må beskjæres for å ta bort det de andre kameraene ikke ser, er avhengig av kameraoppsett.

### 7.2.3 Autostereoskopi

For å spille av video på skjermene, anbefaler Tridelity deres egen programvare TriView. Denne krever at videoene er satt sammen i et spesielt matrisformat. Programmet fordeler deretter bildepunktene fra matrisen på skjermen slik at videoene havner i hver sin visning. I utgangspunktet fungerer skjermene med alle videoavspillere, men for å få videoene i riktig visning, må man bruke TriView. Tridelity ønsket ikke å oppgi hvilke bildepunkter som tilhører hvilke visninger, slik at en kan bruke annen programvare for avspilling.

---

<sup>1</sup>Epost mottatt 16.03.2010 fra Marco Lopes, Software Development Manager, Tridelity Display Solutions GmbH

### 7.2.4 Tekniske detaljer

<b>Tridality 24-tommer</b>	
Oppløsning, panel	1920×1200
Oppløsning, matrise	1920×1500
Oppløsning, visning	960×600
Aspektrate	16:10
Optimal avstand, oppgitt	70 cm
Optimal avstand, målt	Ikke målt
Målt innsynsvinkel	Ikke målt

**Tabell 7.1:** Spesifikasjoner for skjermen på 24 tommer fra Tridality.

<b>Tridality 42-tommer</b>	
Oppløsning, panel	1920×1080
Oppløsning, matrise	1936×1360
Oppløsning, visning	968×544
Aspektrate	16:9
Optimal avstand, oppgitt	335 cm
Optimal avstand, målt	360 cm
Målt innsynsvinkel	ca 45 grader

**Tabell 7.2:** Spesifikasjoner for skjermen på 42 tommer fra Tridality.

## 7.3 Kamera

De fem kameraene vi har brukt til å filme 3D er av typen Tandberg Precision-HD [89]. Disse er oppgitt til å ha en bildevinkel på 50 grader, og kameraene har en fysisk bredde på 3,8 cm. Den maksimale oppløsningen er 1280×720 bildepunkter, og kameraene tar opp 30 bilder i sekundet. Kameraene kobles til PC ved hjelp av USB, og alle opptak må gjøres ved hjelp av programvare på PC. Styring av fokus, farger og lysstyrke gjøres automatisk i kameraene. Kameraet som ble brukt til å lage 2D-video er et Panasonic HDC-HS300, forøvrig samme type kamera som ble brukt i prosjektoppgaven vår [36]. Dette gjør opptak i 1920×1080 med meget god bildekvalitet [90].



**Figur 7.3:** Et Panasonic HDC-HS300 ble brukt til å filme i 2D

## 7.4 Kamerafeste

Kamerafestet er spesiallaget for kameraene av verkstedet hos institutt for elektronikk og telekommunikasjon, NTNU. Se figur 7.4. Kamerafestet ble konstruert slik at det kan festes på et vanlig kamerastativ som vist i figur 7.5.



**Figur 7.4:** Spesiallaget kamerafeste for nøyaktig kalibrering av kameraene



**Figur 7.5:** Kamerafestet plassert på et vanlig kamerastativ

## 7.5 Programvare

### 7.5.1 Opptak

For å gjøre opptakene har vi brukt programmet AMCap, som er et opptaksprogram for USB-kamera [91]. Vi har brukt en prøveversjon, men denne har ingen begrensninger utover at vi ikke kan gjøre lengre opptak enn 20 minutter. Til enkoding har vi brukt kodeken HuffYUV, en svært rask, tapsfri kodek med åpen kildekode. HuffYUV kan lastes ned for seg, men vi brukte den som en del av kodek-pakken FFDSHow [92].

### 7.5.2 Redigering

Til å redigere videoene har vi brukt programmet Adobe Premiere Pro, utgave CS4 og mot slutten CS5 [93]. Dette programmet lar oss synkronisere og klippe videoene, samt plassere dem i formatet avspillingsprogramvaren krever. Til enkoding av den ferdigredigerte videoen har vi brukt Adobe Media Encoder, som følger med Premiere Pro. Som kodek her har vi valgt Windows Media Video 9. Denne er én av få kodeker med støtte for den spesielle oppløsningen på videomatriksen. Den er dessverre ikke tapsfri, men vi har satt

alle innstillinger til maks kvalitet. Det lot seg ikke gjøre å bruke FFDSHOW til å gjøre enkoding av video fra Premiere Pro.

### **7.5.3 Avspilling**

Med skjermene fulgte det et proprietært avspillingsprogram kalt TriView som vi har brukt til å spille av videoer med. TriView har støtte for mange ulike antall visninger, men en skjerm med fem visninger vil uansett aldri kunne vise flere enn fem.



# Kapittel 8

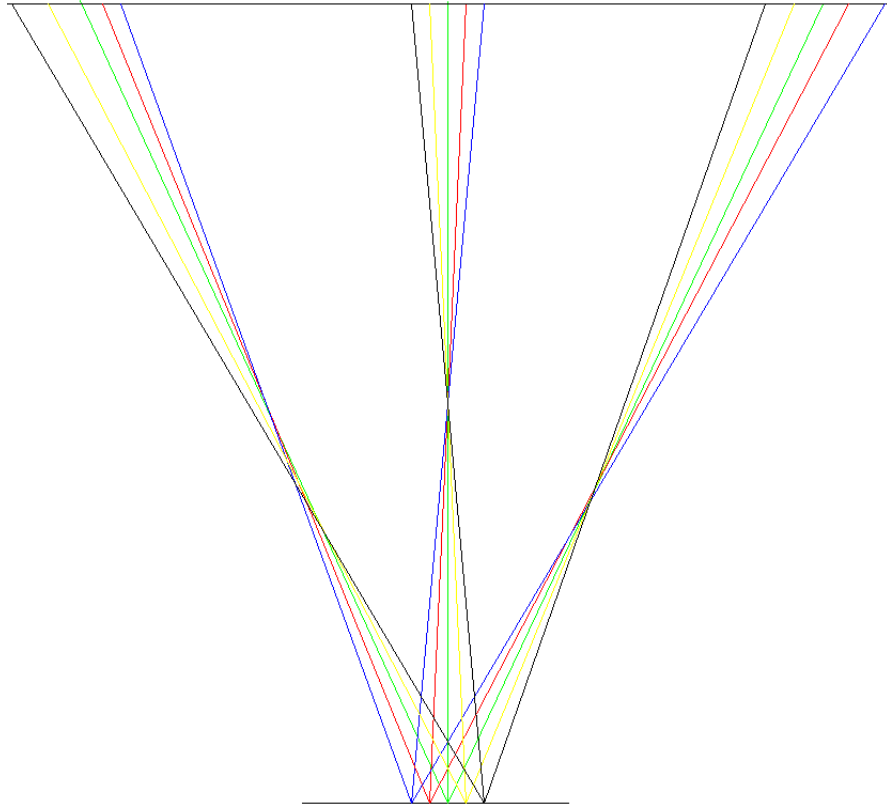
## Kameraoppsett

Kameraoppsettene som brukes for å lage 3D-video kan deles inn i to hovedkategorier: Parallelle og konvergerende kameraer. Begge disse metodene har sine fordeler og ulemper, og disse må vurderes grundig før en kan avgjøre hva som er mest egnet for å filme video som skal spilles av på en autostereoskopisk skjerm.

### 8.1 Konvergerende kameraer

Med konvergerende kameraer menes at kameraene settes ved siden av hverandre langs en linje der kameraene peker inn mot et felles punkt, brennpunktet. Hvor stor separasjonen i bildet kan bli med konvergerende kameraer, avgjøres av fire faktorer. Disse er kameraenes bildevinkel, avstanden mellom sentrum av kameranlinsene, avstanden fra det midterste kameraet til brennpunktet og avstanden fra kameraene til nærmeste og bakerste objekt, for eksempel en vegg. Se figur 8.1. Når kameraene konvergerer inn mot et punkt, reduseres separasjonen inn mot dette punktet.

Det er viktig at bildet som skal vises på skjermen ikke inneholder noe fra ett kamera som ingen av de andre kameraene ser, jamfør 7.2. Hvor mye bildet må beskjæres, er avhengig av avstanden mellom kameraene, avstanden til brennpunktet, avstanden fra kameraene til nærmeste objekt og avstanden til det fjerneste synlige objektet. Dersom brennpunktet befinner seg rett foran en vegg, trenger ikke bildet beskjæres like mye som når det er objekter i bildet langt bak brennpunktet. Desto mer kameraene konvergerer, desto mer må bildene beskjæres. Hvor mye hvert bilde må beskjæres på hver side vil dermed



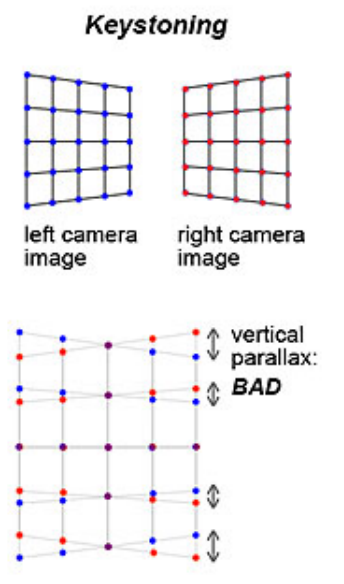
**Figur 8.1:** Fem konvergerende kameraer, 50 grader bildevinkel

være en funksjon av de samme faktorene som avgjør hvor stor separasjonen er. Ettersom aspektraten i bildet, altså forholdet mellom bildets høyde og bredde, må være 16:9, må bildet beskjæres like mye i høyden som i bredden. Å beskjære et bilde mye vil i praksis være det samme som å zoome inn mye, ettersom bildet har samme størrelse på skjermen uansett. Så dersom kameraene konvergerer mye, vil zoomingen bli betydelig. Ettersom bildet har samme oppløsning i visningen uavhengig av hvor mye det er beskåret på forhånd, vil beskjæring i praksis øke separasjonen i bildet. Dersom objektet som skal filmes er nært kameraene, og kameraene må konvergeres for å minke separasjonen i bildet, vil beskjæringen som må gjøres øke separasjonen i bildet igjen. Hvor nære kameraene brennpunktet kan være, er altså begrenset. En bør også ta i betraktning at dersom beskjæringen gjør at bildet får lavere oppløsning enn det den autostereoskopiske skjermen krever, må bildet blåses opp, og dette vil gå ut over bildekvaliteten.

Med konvergerende kameraer vil det også være slik at separasjonen i bildet varierer med avstanden fra brennpunktet. Desto større avstand, desto større

separasjon. Dette gjør at konvergerende kameraer er uegnet for scenarioer med mye dybde dersom konvergeringen er stor.

Et annet aspekt ved konvergerende kameraer, er at bildene er utsatt for bildeforstyrrelser som skyldes fenomenet som kalles “keystoning” [94]. Kort oppsummert er keystoning en bildeforvrenging som oppstår som en følge av at kameraene har litt ulikt perspektiv, dette er vist i figur 8.2 [95]. Denne forvrengingen kalles vertikal parallakse og gjør at objekter i kantene av bildet får en vertikal forskyvning mellom visningene. Denne forstyrrelsen kan reduseres ved å beskjære bildet, men vil alltid være der med konvergerende kameraer.

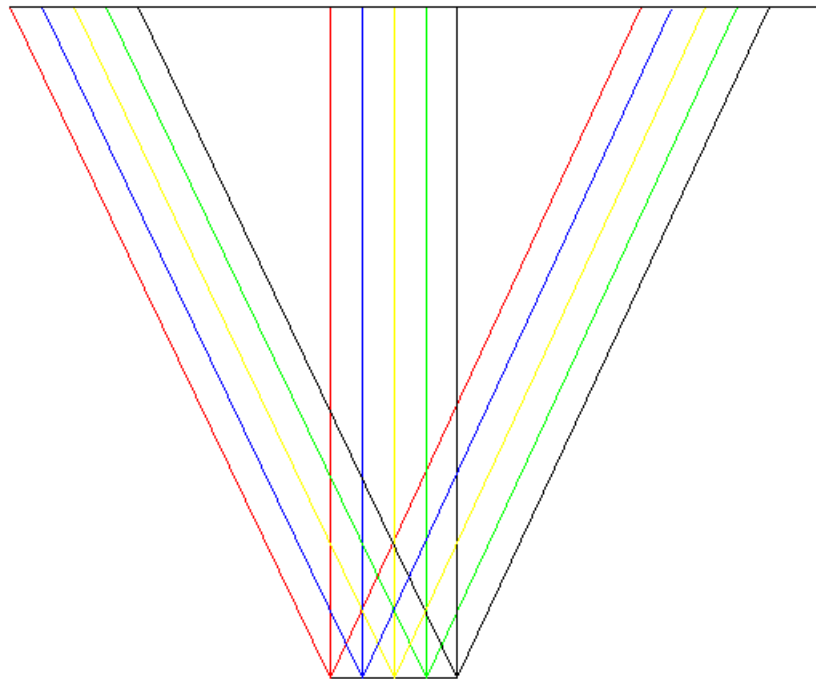


Figur 8.2: Keystoning

## 8.2 Parallele kameraer

Parallele kameraer kan sees på som et grensetilfelle av konvergerende kameraer hvor avstanden til brennpunktet er satt til uendelig. Med parallelle kameraer går siktlinjene fra hvert kamera parallelt, og separasjonen mellom kameraene er dermed konstant lik avstanden mellom kameraene, jamfør figur 8.3. Separasjonen i bildet vil likevel minke som en funksjon av avstanden fra kameraene. Separasjonen i bildet for et objekt utgjør en synkende andel av bildet ettersom avstanden fra kameraene til objektet øker. I motsetning

til ved bruk av konvergerende kameraer, vil separasjonen i bildet endre seg lite med avstanden, og så lenge objekter som filmes befinner seg bak en minimumsavstand bestemt av avstanden mellom kameraene og kameraenes bildevinkel, vil separasjonen i bildet alltid være mindre enn det den autostereoskopiske skjermen kan vise uten feil.



**Figur 8.3:** Fem parallelle kameraer, 50 grader bildevinkel

En autostereoskopisk skjerm kan ikke vise noe fra et kamera som ingen av de andre kameraene ser uten at det blir feil i bildet, jmfør kapittel 7. Bildene fra de parallelle kameraene må derfor også beskjæres. Hvor mye hvert bilde skal beskjæres, bestemmes av avstanden fra kameraene til nærmeste objekt, samt kameraenes bildevinkel. Desto lengre avstand til nærmeste objekt, desto mindre må beskjæres. Det er mulig å beskjære bildene mer enn dette dersom en ønsker å zoom inn på det en filmer, men en bør ikke beskjære mer enn at oppløsningen på det beskjærte bildet er større eller lik oppløsningen den autostereoskopiske skjermen krever at visningene skal ha.

## 8.3 Ulike metoder til ulike scenarier

Det finnes dermed ikke én metode som fungerer i alle scenarier. Hvilken metode som bør benyttes er avhengig av:

- Kameraenes bildevinkel
- Hvor tett kameraene kan plasseres
- Tilgjengelige avstander i rommet
- Avstand fra kameraer til nærmeste objekt
- Avstand fra kameraer til bakerste synlige objekt eller vegg
- Skjermens størrelse
- Optimal seeravstand til skjerm
- Optimal bildeseparasjon for skjerm



# Kapittel 9

## Formler og utregninger

Det er en rekke faktorer som avgjør den opplevde kvaliteten av bildet på en autostereoskopisk skjerm. Ved å uttrykke så mange som mulig av disse faktorene matematisk er det lettere å arbeide systematisk med å finne ut hvilke faktorer som er avgjørende for kvalitetsopplevelsen og sammenhengen mellom dem. Den kanskje viktigste faktoren av de som lar seg uttrykke matematisk er separasjonen i bildet. I alle grafene i dette kapitlet er avstanden mellom sentrum av katedralinsene satt til 4 cm, med mindre annet er oppgitt.

### 9.1 Beskjæring

For å oppfylle kravet om at en visning ikke kan inneholde bildeinformasjon som ikke finnes i en annen visning, må noe beskjæres bort. Hvor mye som må beskjæres er avhengig av hvilket kameraoppsett som er valgt. Felles for begge kameraoppsettene er at alt utenfor innerste bildekant på begge sider må klippes bort. Se figur 8.3 og 8.1. Hvor mye som må klippes bort på hver side av bildet er avhengig av hvilken posisjon kameraet har i kamerastativet. For parallelle kamera er summen av beskjæring for høyre og venstre side lik for alle kamera, mens for konvergerende kamera er beskjæringen mindre på kameraet i midten og større for kameraene ytterst. De påfølgende seksjonene vil vise formlene for nødvendig beskjæring for parallelle og konvergerende kamera.

### 9.1.1 Nødvendig beskjæring for parallelle kamera

$$C_{NH} = \frac{(K_N - 1) - A}{2B \tan(\frac{v}{2})} \quad (9.1.1)$$

$$C_{NV} = \frac{(5 - K_N)A}{2B \tan(\frac{v}{2})} \quad (9.1.2)$$

I formelen for høyre side,(9.1.1), og formelen for venstre side,(9.1.2), gjelder at  $K_N$  er kamera nummer  $N$  og  $1 \leq N \leq 5$ , nummerert når man ser kamerastativet bakfra.

### 9.1.2 Nødvendig beskjæring for konvergerende kamera

Ytterste høyre kamera:

$$C_1 = \frac{C_Y}{D_{YK}}$$

$$C_Y = \left| \frac{H \sin(\frac{v}{2}) 4A \sin(\frac{v}{2})}{\left(\frac{1}{2} \cos(2 \arctan(\frac{2A}{B})) + \cos(v)\right) B \sqrt{\frac{4A^2}{B^2} + 1}} + \frac{4A}{B} \right|$$

$$D_{YK} = \frac{H \sin(\frac{v}{2}) \cos(\frac{v}{2})}{\left(\frac{1}{2} \cos(2 \arctan(\frac{2A}{B})) + \cos(v)\right) \sqrt{\frac{4A^2}{B^2} + 1}}$$

Nest ytterste høyre kamera:

$$C_2 = \left| \frac{A \tan(\frac{v}{2})}{B} + \frac{(\frac{A^3}{B^2} + A) \cos(2 \arctan(\frac{A}{B})) + \cos(v)}{B \sin(v) \sqrt{B^2 + C^2} \sqrt{\frac{A^2}{B^2} + 1} + C} \right| \quad \text{for } C > 0$$

$$C_2 = \left| \frac{A \tan(\frac{v}{2})}{B} + \frac{(\frac{A^3}{B^2} + A) \cos(2 \tan(\frac{A}{B})) + \cos(v)}{B \sin(v)(B + C)} \right| \quad \text{for } C \leq 0$$

Midterste kamera:



$$C_3 = \left| \frac{(B + C) \tan(\frac{v}{2}) + \frac{2AC}{B} - \frac{\left(\sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))}\right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}}{2(B + C) \tan(\frac{v}{2})} \right| \quad \text{for } C > 0$$

$$C_3 = \left| \frac{(B + C) \tan(\frac{v}{2}) + \frac{2AC}{B} - \frac{\frac{(B + C) \sin(\frac{v}{2})}{\cos(\arctan(\frac{2A}{B}))}}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}}{2(B + C) \tan(\frac{v}{2})} \right| \quad \text{for } C \leq 0$$

Nest ytterste venstre kamera:

$$C_4 = \frac{\alpha - \beta + \frac{AC}{B}}{\alpha - \gamma}$$

der

$$\alpha = \frac{\left(\sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{A}{B}))}\right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{A}{B}))}$$

$$\beta = \frac{\left(\sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))}\right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}$$

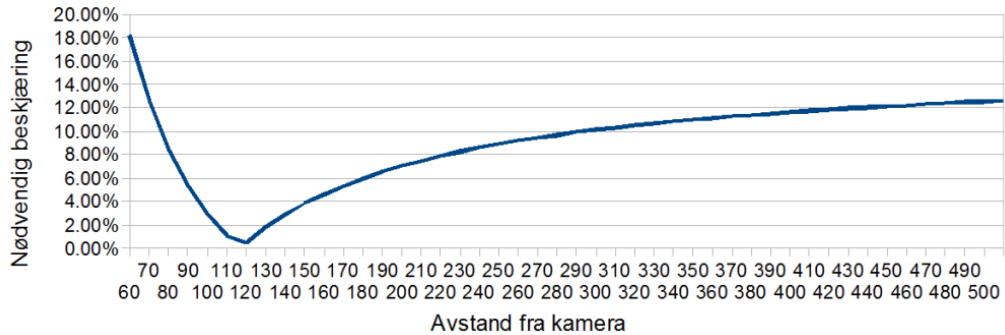
$$\gamma = \frac{\left(\sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))}\right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} + \arctan(\frac{A}{B}))}$$

Ytterste venstre kamera: Skal ikke beskjæres.

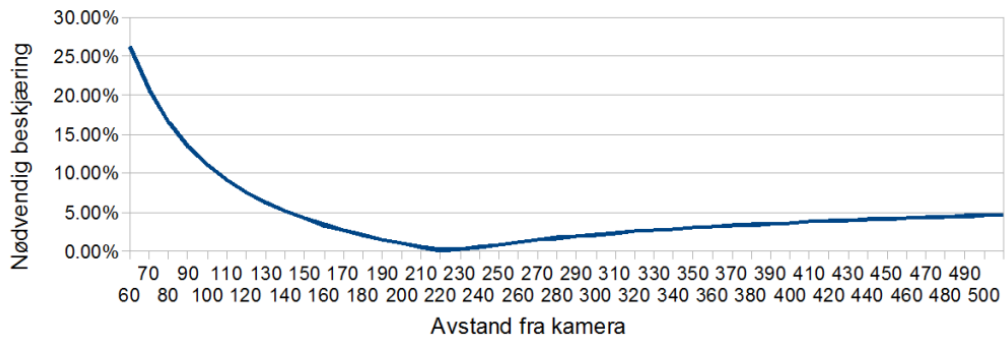
Formlene over gjelder beskjæringen på høyre side av bildet. På venstresiden skal det beskjæres tilsvarende, men i motsatt rekkefølge.

I grafene i figur 9.1 og figur 9.2 for konvergerende kamera er brennpunktet satt til henholdsvis 130 cm og 260 cm. Dette er aktuelle verdier, og ved å

ha med to verdier kan man se for seg hvordan grafen vil endre seg dersom avstanden til brennpunktet økes eller minkes. Grafene for konvergerende kamera viser hvor mye som må beskjæres av bildet fra de ytterste kameraene. For det midterste kameraet og de nest ytterste vil den nødvendige beskjæringen være noe lavere. Figur 9.3 viser gafen for parallelle kameraer.



**Figur 9.1:** Nødvendig beskjæring i prosent for konvergerende kameraer som funksjon av avstanden til objektet. Avstanden til brennpunktet er 130 centimeter.

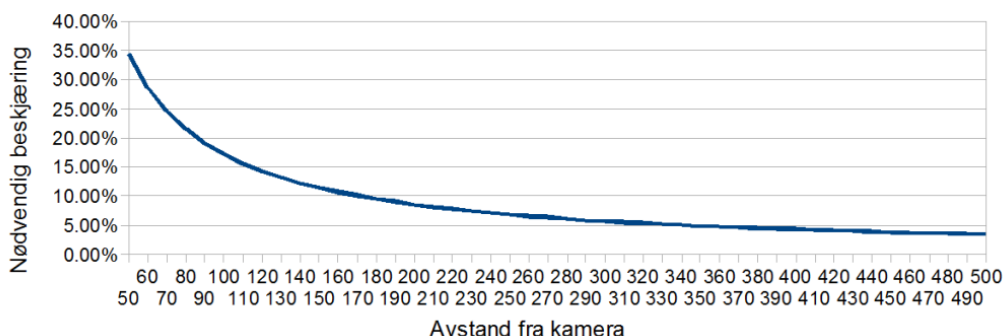


**Figur 9.2:** Nødvendig beskjæring i prosent for konvergerende kameraer som funksjon av avstanden til objektet. Avstanden til brennpunktet er 260 centimeter.

For å finne ut hvor mye et bilde må beskjæres, må man se på avstandene til objektene i bildet. Den avstanden til et objekt i bildet som gir størst nødvendig beskjæring bestemmer hvor mye beskjæring som må gjøres. Det vil si at dersom man bruker konvergerende kamera med 260 cm til brennpunktet og det befinner seg objekter fra 160 cm fra kameraene til en vegg 360 cm fra kameraene, må man sammenligne nødvendig beskjæring for disse to avstandene. I dette eksempelet er nødvendig beskjæring størst 360 cm fra kameraene, 3,27%.

## 9.2. SEPARASJON

---



**Figur 9.3:** Nødvendig beskjæring i prosent for parallelle kameraer som funksjon av avstanden til objektet.

Man bør merke seg en prinsipiell forskjell mellom parallelle og konvergerende kamera: Med parallelle kamera trenger man bare ta hensyn til avstanden til det nærmeste objektet når man skal beskjære, mens med konvergerende må man også ta hensyn til avstanden til det bakerste. Dette gjør parallelle kamera til et mer egnet kameraoppsett dersom dybden der man skal filme er stor.

## 9.2 Separasjon

Separasjonen i bildet uttrykkes som andel av skjermbredden og formlene er forskjellig for de to hovedtypene av kameraoppsett.

### 9.2.1 Parallelle kamera

For parallelle kamera er faktorene som avgjør separasjonen i bildet kameraenes bildevinkel,  $v$ , avstanden mellom kameraene,  $A$ , og avstanden til nærmeste objekt,  $B$ . Dersom ønskelig kan bildet beskjæres mer enn det som er nødvendig, denne beskjæringen er lik for alle bildene og betegnes  $C_f$ .  $C_f$  er andelen av bildet som er igjen etter beskjæring.

Separasjonen i bildet:

$$S_B = \frac{A}{2C_f(B \tan(\frac{v}{2}) - 2A)} \quad (9.2.1)$$

For full utregning, se appendiks A.5.

### 9.2.2 Konvergerende kamera

For konvergerende kamera er det noen flere faktorer. Faktorene som avgjør separasjonen i bildet her er kameraenes bildevinkel,  $v$ , avstanden mellom kameraene,  $A$ , avstanden til brennpunktet,  $B$ , og avstanden fra brennpunktet til objektet,  $C$ .  $V$  er positiv hvis objektet ligger bak brennpunktet sett fra kameraene, og negativ hvis objektet er foran.

Separasjonen i bildet:

$$S_B = \frac{CA}{BD_V \left( 1 - \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(\sqrt{B^2 + C^2} \sqrt{\frac{4A^2}{B^2} + 1 + C})} \right| \right)} \quad \text{for } C > 0$$

$$S_B = \frac{CA}{BD_V \left( 1 - \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(B + C)} \right| \right)} \quad \text{for } C \leq 0$$

For full utregning, se appendiks A.6.

### 9.2.3 Grader

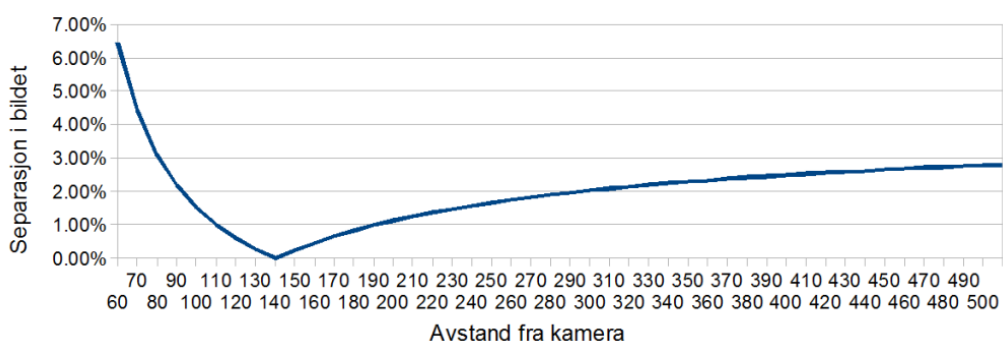
Ved å bruke regelen om at vinkelen i triangelet som dannes av seeren og to visninger av samme objekt i bildet ikke bør overstige 1,5 grader, kan man regne ut maksimal separasjon i bildet. Vi setter vinkelen fast til 1,5 grader og løser for separasjonen i bildet. Her er  $D_s$  bredden av skjermen.

$$S_{b-max} = \frac{2B \tan(0.75)}{D_s}$$

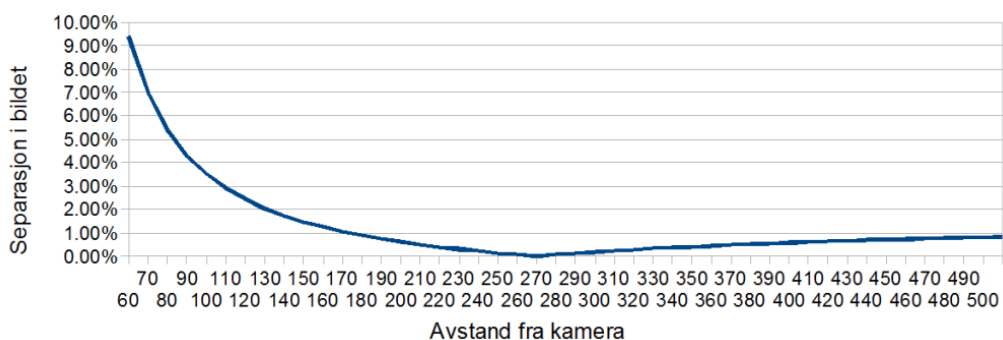
Hvis man ser på vår 24-tommer skjerm, er denne 52 cm bred og har optimal avstand 70 cm. Vi har da  $S_{b-max} = 3,52\%$ . Skjermen på 42 tommer er 93 cm bred og har optimal avstand 335 cm. For denne blir  $S_{b-max} = 9,43\%$ .  $S_{b-max}$  vil øke for begge skjermene når avstanden til skjermene øker.

## 9.3 Avstander

Gitt at man vet intervallet separasjonen i bildet bør ligge innenfor, kan man bruke dette til å regne ut hvor nært og hvor langt unna et objekt kan være kameraene før separasjonen i bildet havner utenfor dette intervallet. Ved å bruke et regneark hvor formelen for separasjonen i bildet er lagt inn og verktøyet “Goal seek”, kan man beregne avstander numerisk og lage grafer som viser separasjonen i bildet som en funksjon av avstanden korrigeret for nødvendig beskjæring [96].

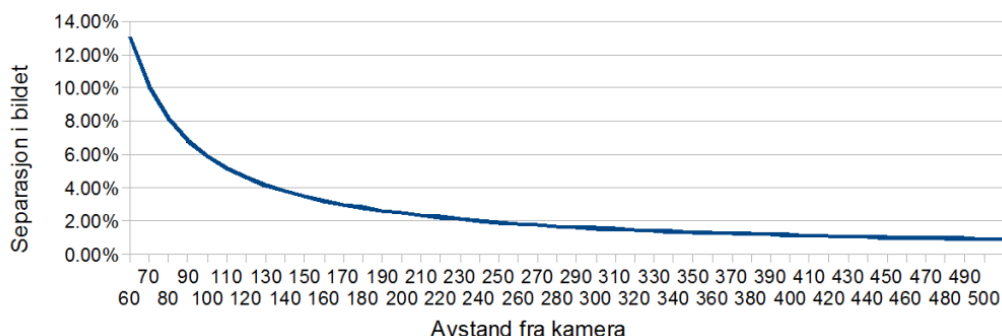


**Figur 9.4:** Separasjonen i bildet som funksjon av avstanden, korrigeret for nødvendig beskjæring. Kameraene er konvergerende og avstanden til brennpunktet er fastsatt til 130 centimeter.



**Figur 9.5:** Separasjonen i bildet som funksjon av avstanden, korrigeret for nødvendig beskjæring. Kameraene er konvergerende og avstanden til brennpunktet er fastsatt til 260 centimeter.

Her ser man av grafene at desto lengre unna kameraene brennpunktet er, desto mer vil funksjonen for separasjonen i bildet for konvergerende kameraer ligne funksjonen for separasjonen i bildet for parallelle kameraer. Det er



**Figur 9.6:** Separasjonen i bildet som funksjon av avstanden, korrigert for nødvendig beskjæring. Parallelle kameraer.

forøvrig viktig å merke seg at grafene som viser separasjonen i bildet er korrigert for nødvendig beskjæring. Det vil si at dersom man bruker parallelle kameraer og nærmeste objekt er 130 cm unna, vil separasjonen i bildet for dette objektet være 3,8%, gitt at man har beskjært de nødvendige 13,2%. Da kan man ikke se på grafen for å finne ut hvordan separasjonen vil være for de andre objektene i bildet. Disse grafene gir med andre ord ikke informasjon om hvordan separasjonen i bildet utvikler seg som en funksjon av avstanden. De gir informasjon om hvor stor separasjonen er i en gitt avstand gitt at bildet er beskjært for denne avstanden.

Hvis man vet hvor stor separasjonen maksimalt bør være, kan slike grafer hjelpe til å finne ut hvor nært eller hvor langt unna kameraene objekter kan være. Som for beskjæring, ser vi også her at for parallelle kameraer trenger man bare å ta hensyn til hvor nært objekter kan være, mens for konvergerende kameraer må man også ta hensyn til hvor langt unna de kan være.

Hvis man ser bort fra separasjonen, vil også skjermenes egenskap med at visningene ikke kan inneholde noe bildeinformasjon ingen andre visninger inneholder uten at det oppstår feil i bildet, begrense avstanden til nærmeste objekt. Dette kan uttrykkes ved formlene (9.3.1) og (9.3.2).

Parallelle kameraer:

$$D_{min} = 2A \tan\left(\frac{\pi}{2} - \frac{v}{2}\right) \quad (9.3.1)$$

### 9.3. AVSTANDER

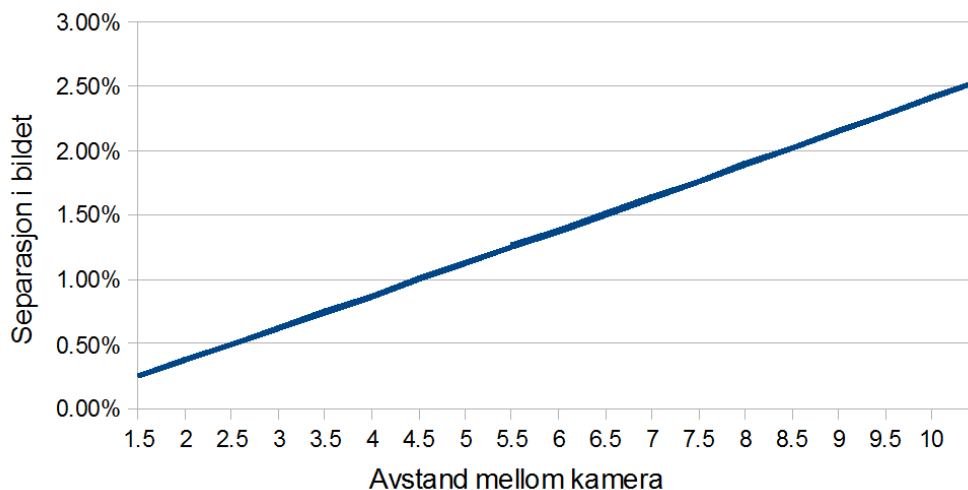
---

Konvergerende kameraer:

$$D_{min} = 2A \tan \left( \arctan \left( \frac{B}{2A} - \frac{v}{2} \right) \right) \quad (9.3.2)$$

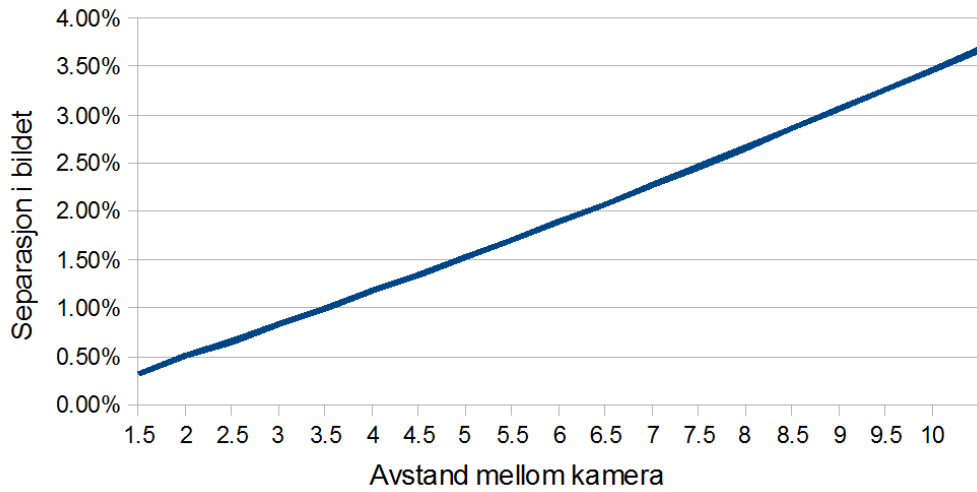
I praksis vil separasjonen i bildet være den faktoren som bestemmer hvor nært et objekt kan være kameraene.

For å finne den optimale avstanden mellom sentrum av kameralinsene, må vi se på hvordan avstanden mellom kameralinsene påvirker separasjonen i bildet. Avstandene som er brukt i grafene i figur 9.7 og figur 9.8 er typiske for parallelle og konvergerende kameraer.



**Figur 9.7:** Separasjonen i bildet som funksjon av avstanden mellom kameralinsene. Kameraene er konvergerende, avstanden til brennpunktet er 130 centimeter og avstanden til objektet er 100 cm.

Som vi ser av grafene, øker separasjonen i bildet når avstanden mellom kameraene øker. For parallelle kameraer er denne funksjonen ganske enkel. Separasjonen i bildet minker som en funksjon av avstanden fra kameraene. Når avstanden mellom kameralinsene øker, vil separasjonen i bildet øke for alle avstander, men graden av hvor mye separasjonen minker med avstanden fra kameraene vil være den samme. Dersom en vet avstanden til nærmeste objekt, kameraenes bildevinkel og hvilket intervall separasjonen i bildet for nærmeste objekt bør ligge innenfor, kan man enkelt beregne intervallet avstanden mellom kameralinsene bør være innenfor.



**Figur 9.8:** Separasjonen i bildet som funksjon av avstanden mellom kameralinsene for parallelle kameraer. Avstanden til objektet er 335 cm.

For konvergerende kameraer er det litt mer komplisert. Her øker separasjonen i bildet som en funksjon av avstanden fra brennpunktet. Når avstanden mellom kameralinsene øker, vil økingen av separasjonen i bildet som en funksjon av avstanden fra brennpunktet forsterkes. En økning i avstanden mellom kameralinsene vil altså begrense det området hvor separasjonen i bildet er lav nok.

Når avstanden mellom kameralinsene økes uten at brennpunktet flyttes konvergerer kameraene mer. Desto mer kameraene konvergerer, desto mer må beskjæres av bildene. Dersom en setter en grense for hvor mye en kan beskjære bildet, enten for å hindre tap av oppløsning eller for å hindre objektene i bildet å blir for store på skjermen, vil denne begrense hvor stor avstanden mellom kameraene kan være. For å regne ut intervallet avstanden mellom sentrum av kameralinsene bør være innenfor for konvergerende kameraer, må man altså ta hensyn til hvor stor separasjonen i bildet kan være, og hvor stor beskjæringen av bildet kan være.

I praksis vil valget av avstand mellom kameralinsene være gitt av den fysiske bredden av kameraene. I de fleste situasjoner vil problemet være å få separasjonen lav nok, og dersom en ønsker å gjøre den høyere, vil det være like enkelt å beskjære bildene som å øke avstanden mellom kameralinsene.



# Kapittel 10

## Forsøk

For å lage en 3D-film i så høy kvalitet at den kan sammenlignes med en 2D-film i full HD, er det viktig å finne ut om verdiene fra utregningene stemmer i praksis. Det er dessuten viktig å gjøre forsøk som finner ut optimal avstand, innsynsvinkel og maksimal separasjon i bildet for den autostereoskopiske skjermen som skal brukes i det sammenlignende forsøket.

### 10.1 Kalibrering av kamera

For å unngå feil i bildet, er det svært viktig å kalibrere kameraene nøye før opptak. De påfølgende seksjonene forklarer hvordan vi har kalibrert i våre forsøk.

#### 10.1.1 Parallele kamera

Kameraene settes parallelt i stativet med jevn avstand mellom seg, så tett som mulig, eller med ønsket avstand. Den totale bredden av kamerarekken måles og avstanden mellom sentrum av hver kameralinse regnes ut. I vårt tilfelle var den totale bredden av kamerarekken 20 cm, og avstanden mellom sentrum av hver kameralinse blir dermed 4 cm. På en papplate har vi tegnet opp fem parallelle streker med 4 cm mellomrom. Denne plata settes noen meter fra kameraene, slik at den midterste streken på plata havner cirka i sentrum av bildet til det midterste kameraet. Opptaksprogramvaren har mulighet til å vise et kryss i sentrum av bildet, og kameraet justeres slik at dette krysset overlapper perfekt med den tilhørende streken på plata. Dette

gjentas for alle kameraene. På denne måten blir kameraene parallelle med relativt god nøyaktighet. Ved å øke avstanden fra kameraene til plata blir kalibreringen vanskeligere, men også mer nøyaktig. Det er forøvrig viktig at papplata og kamerarekka står parallelt i forhold til hverandre, ellers vil ikke avstanden mellom strekene på plata stemme.

Like viktig som å få kameraene helt parallelle er det at høyden på kamerabil-  
dene er like. For å sikre dette brukes først et vater for å se at kamerastativet  
er i vater. Deretter siktes kameraene inn mot en gjenstand som vises i bildet  
på alle kameraene, for eksempel et bord. Denne gjenstanden må også være  
målt for å sikre at den er i vater. Deretter justeres kameraene slik at kryss-  
et i opptaksprogramvaren overlapper med toppen av gjenstanden. Dersom  
bildene kameraene leverer er litt på skrå i forhold til hverandre selv om ka-  
merastativet er i vater, er det viktig at høydejusteringen i gjort i forhold til  
midten av bildet. På den måten vil feilene som oppstår i bildet på grunn  
av ulik høyde på kameraene befinne seg i ytterkantene av bildet. Grunnen  
til at vi heller ønsker å ha feilene i kantene av bildet er først og fremst fordi  
mennesker har fokus i midten av bildet, men også fordi bildet beskjæres etter  
opptak slik at en del av feilene forsvinner [2].

### 10.1.2 Konvergerende kamera

Kameraene settes i stativet med jevn avstand mellom seg, så tett som mulig,  
eller med ønsket avstand. Ønsket avstand til brennpunkt måles langs sikt-  
linjen til det midterste kameraet, og en gjenstand kameraene kan sikte på  
plasseres i brennpunktet. Krysset i opptaksprogrammet brukes deretter til å  
få kameraenes siktlinjer til å gå gjennom gjenstanden plassert i det ønske-  
de brennpunktet. Ettersom det kan være vanskelig å rotere kameraene om  
linsene, er det lurt å gjøre en ny måling av avstanden mellom kameraene  
for å sjekke om denne fortsatt er riktig etter at roteringen er gjort. Høyden  
justeres som for parallelle kamera.

## 10.2 Måling av optimal avstand og vinkel til skjerm

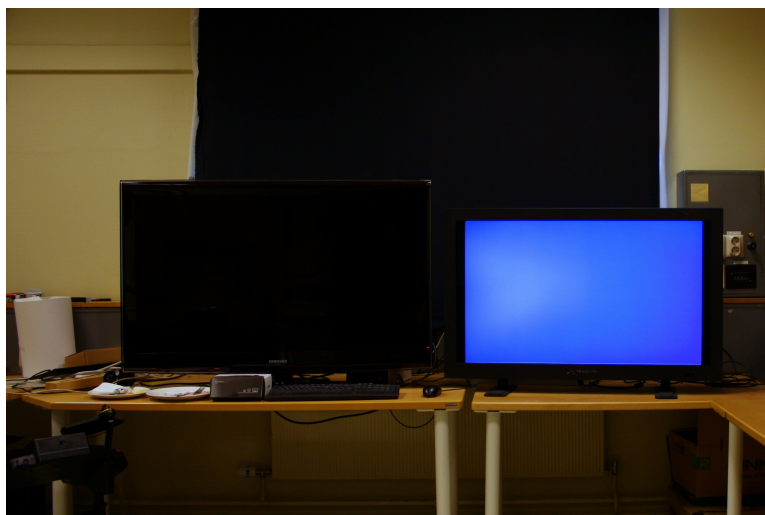
For å finne den optimale avstanden til den største av de to autostereoskopiske  
skjermene brukte vi følgende metode: Skjermen ble satt til å vise én farge  
per visning. Fargene hvit, rød, blå, grønn og gul ble brukt for å enkelt kunne

## 10.2. MÅLING AV OPTIMAL AVSTAND OG VINKEL TIL SKJERM

---

skille visningene fra hverandre. Deretter satte vi oss midt foran skjermen på relativt kort avstand og beveget oss bakover helt til vi så to forskjellige, distinkte farger på hvert øye, jamfør figur 10.1. For å bekrefte at dette noe kunstige forsøket er relevant for den typen bilder skjermen normalt sett viser, gjentok vi forsøket med en video fra et annet forsøk. Vi satt nært, midt foran skjermen og beveget oss bakover til bildet på skjermen ble uten feil. Resultatet ble det samme i begge tilfellene, optimal avstand er ca 360 cm fra skjermen. Bildet er uten feil derfra og fram til ca 460 cm fra skjermen, mens bak denne avstanden blir bildet gradvis dårligere.

For å finne skjermens innsynsvinkel plasserte vi oss på optimal avstand og satte skjermen til å vise én farge per visning. Vi beveget oss deretter sidelengs langs en linje parallelt med skjermen inntil det ikke lenger var mulig å få sett to ulike farger for hvert øye, jamfør figur 10.2. Deretter målte vi avstanden fra dette punktet til en linje som går rett ut fra midten av skjermen. Som i forsøket for å finne optimal avstand gjentok vi også dette forsøket med en video fra et annet forsøk. Vi beveget oss sidelengs til vi ikke lenger klarte å finne et bilde uten feil. Resultatet i begge forsøkene ble også denne gang like, ca 150 cm fra linja fra midten av skjermen.



**Figur 10.1:** Skjermen viser én farge, altså én visning, for hvert øye ved optimal avstand



**Figur 10.2:** Skjermen viser flere farger (visninger) sett litt fra siden

### 10.3 Måling av separasjon i bildet

Tridelity hevder i en epost til oss at separasjonen i bildet ikke kan være større enn 1,5% av bredden av skjermen før det oppstår bildefeil og ghosting. I de tidlige bildeforsøkene fikk vi en mistanke om at dette ikke nødvendigvis var riktig. Dette stemmer heller ikke med regelen om maks 1,5 grader separasjon for de optimale avstandene til skjermen, jamfør kapittel 9. Vi utførte derfor noen forsøk for å finne ut hva som er maksimal separasjon i bildet før det oppstår feil, hvor liten separasjonen i bildet kan være for nærmeste objekt for å fortsatt gi en 3D-effekt og hvilken separasjon i bildet som gir den mest realistiske dybdefølelsen.

Forsøket ble gjennomført ved at vi først filmet en scene med relativt mye ulik dybde med parallelle kamera. Deretter ble denne filmen redigert og beskåret i ulike omganger slik at vi hadde fem utgaver av den samme filmen med ulik separasjon. Ved 312 cm avstand fra kameraet var separasjonen i bildet for de ulike utgavene 1,44%, 1,55%, 1,63%, 1,72% og 1,80%. Ved sammenligning av utgavene ble det klart at alle gir en brukbar 3D-opplevelse. Det er snakk om relativt små forskjeller i dybdefølelse, og for en person med lite erfaring med å se 3D er det vanskelig å se forskjell på videoene. Vi konkluderte likevel med at 1,72% separasjon i bildet er svært nært optimalt og at 1,8% nærmer seg en litt overdreven dybde, mens 1,55% og nedover gir et litt flatt bilde. For parallelle kamera vil separasjonen være størst nært kameraet og avta med avstanden. Det er forskjellene i separasjonen i bildet som funksjon av

avstanden som gir dybdefølelsen, så selv om separasjonen for objekter som befinner seg langt unna er lav, vil man likevel oppfatte avstanden til disse, forutsatt at separasjonen i bildet nært kameraet ikke er like lav.

Ved videre forsøk med separasjoner på 1,9% og 2% ga dette feil i bildet i form av ghosting og uskarpe objekter. Basert på disse forsøkene kan vi ikke se at informasjonen fra Tridality er riktig, verken at separasjonen i bildet ikke kan være større enn 1,5% eller at den kan være så stor som 9,43%, jamfør kapittel 9. Vi vil heller tilstrebe maksimal separasjon i bildet mellom 1,55% og 1,80% for objektene i våre videoer.

## 10.4 Tidlige bildeforsøk

Før vi hadde kamerafestet klart, gjorde vi en serie forsøk med stillbilder. Disse var til stor hjelp i forståelsen av hvor viktig det er å kalibrere nøyaktig, hvor viktig det er å beskjære bildene og fjerne bildeinformasjon ingen andre kamera ser, de matematiske sammenhengene mellom faktorene som avgjør separasjonen i bildet, forskjellene mellom parallelle og konvergerende kamera og hvilke egenskaper en autostereoskopisk skjerm har.

## 10.5 Sammenligning av 2D og 3D

### 10.5.1 Bakgrunn

For å vurdere om man kan få en bedre brukeropplevelse i videosamtaler ved å bruke 3D, er det naturlig å sammenligne en videosamtale i 3D med en i 2D. Det er da viktig at kvaliteten på videosamtalene som skal sammenlignes er så god at ikke en eventuell kvalitetsforskjell ikke blir en feilkilde i sammenligningen.

### 10.5.2 Gjennomføring

Det ble laget totalt fire videoer til sammenligningen, gjort i to opptak: To videoer i 2D og to i 3D. Begge opptakene viser et typisk videosamtaleoppsett, med én eller flere personer bak et bord og en vegg ikke langt bak. Kameraene er parallelle med 4 cm mellom sentrum av hver kameralinse og er satt i en høyde som tilsvarer høyden til en 42-tommer skjerm plassert på et bord. Den

autostereoskopiske skjermen på 42 tommer ble brukt til å vise 3D-videoene, mens 2D-videoene ble vist på en 46-tommer Samsung-skjerm, jamfør 7.2.

3D-videoene ble laget ut fra hva vi har erfart og regnet ut at gir god 3D, samtidig som vi har prøvd å etterligne et realistisk scenario.

I det første opptaket sitter én person bak bordet, og befinner seg 270 cm fra kameraene. I redigering ble ikke bildene beskåret mer enn nødvendig for å sikre lav nok separasjon. Separasjonen i bildet er 1,70% for gjenstandene som befinner seg 270 cm fra kameraene. Dette tilsvarer 1,58 cm av vår 42-tommer skjerm. Veggene befinner seg ca én meter bak personen, og her er separasjonen i bildet ca 1,2%. 2D-kameraet er plassert litt lengre bak enn kamerarekka som blir brukt til å filme 3D, men optisk zoom ble brukt slik at utsnittet ble tilnærmet likt. Det andre opptaket er likt, bortsett fra at det sitter to personer bak bordet, og avstanden til kameraene er økt til 335 cm. I redigeringen av dette opptaket ble det foretatt beskæring slik at separasjonen i bildet for gjenstandene som befinner seg 335 cm unna er 1,61%, eller 1,5 cm av vår 42-tommer skjerm. Separasjonen i bildet i veggens avstand er ca 1,24%.

Selve sammenligningen av 3D-bildet og 2D-bildet ble gjort ved hjelp av en kvalitativ metode, jamfør 2.2.1. Som testperson brukte vi vår veileder og professor ved institutt for telematikk, NTNU, Leif Arne Rønningen. Rønningen har jobbet mye med 3D og har erfaring i å identifisere feil og skille gode 3D-bilder fra dårlige. Rønningen fikk først se 2D-filmen fra det første opptaket, deretter ble han bedt om å kommentere filmens kvalitet og bedt om å vurdere i hvor stor grad bildet virket realistisk. Før Rønningen fikk se 3D-filmen ble han bedt om å ignorere enkelte aspekter ved 3D-filmen som vi mener er resultater av begrensninger i vårt utstyr, men ikke er gjeldende for denne type 3D generelt. Disse er en horisontal strek som befinner seg i midten av bildet i den ene visningen, samt noe ghosting som en følge av at visningene ikke er perfekt synkroniserte. Det ble også presisert at det er viktig å sitte i riktig posisjon i forhold til skjermen, i en sweetspot, jamfør 6.2, slik at han opplever en best mulig 3D-effekt. Etterpå fikk han se 3D-filmen fra det første opptaket og ble spurt de samme spørsmålene om kvalitet og realisme. I tillegg ble han bedt om å kommentere følgende:

- Hvordan dybdefølelsen var
- Om opplevelsen av 3D var god
- Hvilke bildefeil han kunne observere
- Hvor vanskelig det var å finne en hodeposisjon hvor bildet opplevdes

som godt

- Om han kunne tenke seg å bruke et bilde av en slik kvalitet til en videosamtale
- Om størrelsen på personene i bildet føltes naturlig og om den opplevde avstanden til personene på bildet var behagelig
- Om den reduserte oppløsningen på bildet gikk ut over bildekvaliteten
- Om han opplevde samme kvalitet over hele bildet

Det ble også åpnet for andre kommentarer og det hele forløp som en samtale der vi noterte Rønningens innspill og kommentarer. Etterpå ble den samme prosedyren gjentatt for det andre opptaket, og Rønningen fikk også se 2D- og 3D-videoene side ved side og fikk kommentere forskjellene mellom dem.

### 10.5.3 Resultater

Ut fra samtalen med Rønningen kan vi oppsummere følgende:

I 2D-filmene oppfattes bildet som veldig skarpt og fint, og bordet gir bildet en viss dybdefølelse. Personen i bildet oppleves som litt liten i forhold til avstanden. Oppløsningen er god nok for å gi et realistisk bilde og bildet oppleves som 70-80% realistisk. I 3D-filmene er dybdevirkningen av bordet bedre enn på 2D-filmen, og det er generelt mye lettere å se dybden i bildet. Dybdevirkningen gjør bildet mer realistisk og 3D-filmen oppleves som 80-90% realistisk, men det forholdsvis lille utsnittet på skjermen gjør dybden vanskelig å bedømme. Dette ville vært enklere dersom skjermen hadde dekket en større del av synsfeltet.

Ansiktsuttrykkene i 3D-bildet er bedre enn i 2D-bildet og oppleves som mer naturlige og nærmere virkeligheten, selv om bildet ikke er skarpere. Det oppleves som positivt for en videosamtale at ansiktsuttrykkene blir tydeligere. Det er tydelig at kvaliteten på bildet er dårligere langs kantene enn midt i, men dette er et lite problem når det bare er én person i bildet og fokus holdes på denne. Den reduserte oppløsningen går i liten grad ut over opplevelsen av bildekvalitet. Vi har ikke funnet noen kilder som kan bekrefte at bruk av 3D kan kompensere for redusert oppløsning, men ut fra resultatet i forsøket må vi være åpne for at dette kan være tilfelle.

Man må flytte på seg for å få bildet skarpt og den beste posisjonen for å få bildet skarpt er ikke nødvendigvis den samme gjennom hele filmen. Det oppleves ellers ikke som vanskelig å finne posisjonen som gir det beste bildet.

Det kan i noen tilfeller være vanskelig å få hele bildet skarpt på en gang. Det oppleves ikke som slitsomt å holde hodet rolig i en god posisjon, men dersom dette hadde vart like lenge som en spillefilm kan det tenkes å bli slitsomt. Utover de feilene som Rønningen ble bedt om å ignorere er det lite feil i bildet, utover at bordet i bildet har noen feil i de fremste hjørnene som vises delvis dobbelt. Bildet slik det framstår ellers kan fungere fint for en videosamtale, gitt at feilene som Rønningen ble bedt om å ignorere lar seg fjerne.

På spørsmål om forholdet mellom den opplevde avstanden til personen i bildet og størrelsen av personen, svarte Rønningen at størrelsen på personen ble opplevd som naturlig i forhold til avstanden, men at denne avstanden er litt lang. Han vil foretrekke at personen oppleves i naturtro størrelse, og vil heller ha kortere avstand til skjermen enn å gjøre personen på skjermen større og beholde avstanden.

Det er i liten grad forskjeller på de to 3D-filmene, men dybden virker noe kortere på den andre filmen og personene i bildet virker å være lengre unna. Kvaliteten på bildet i den andre filmen er også bedre langs kantene, og bordet i bildet har mindre feil.

#### **10.5.4 Feilkilder**

Det er en klar ulempe at vi bare har hatt én testperson. Det har rett og slett vært vanskelig å finne testpersoner som har erfaring med ulike 3D-teknikker, erfaring med å sammenligne opplevd kvalitet og som er i stand til å ignorere de feilene som er et resultat av det konkrete utstyret vi har brukt. Vi føler likevel at samtalen med Rønningen bidro positivt til vår forståelse av opplevelsen av autostereoskopisk 3D og gjør oss i bedre stand til å vurdere hvorvidt autostereoskopisk 3D er egnet for videosamtaler.

#### **10.5.5 Konklusjoner**

Rønningens innspill og kommentarer samsvarer godt med de tankene vi selv har gjort oss om bildene. At den første 3D-filmen ble oppfattet som å ha større dybdevirkning enn den andre samsvarer godt med at den har noe større forskjell på største og minste separasjon i bildet. Det er derimot litt mer overraskende at personene i den andre filmen blir vurdert til å være lengre unna enn i den første. De er riktignok filmet på lengre avstand, men bildet er beskåret i etterkant slik at de skal være omtrent like store som i



den første. Denne beskjæringen er også trolig årsaken til at den andre filmen har mindre feil i kantene av bildet. Kameraene er kalibrert slik at feilene som følger av at kameraene ikke er helt like blir verre desto lengre unna kalibreringspunktet midt i bildet man kommer. Dermed vil et bilde som er mye beskjært ha færre synlige feil enn et som er lite beskjært.

Det er forøvrig interessant å merke seg at den, i forhold til 2D-videoen, begrensede oppløsningen i 3D-videoene ikke ser ut til å gå ut over opplevelsen av kvalitet og realisme. Oppløsningen i 3D-videoen er en femtedel av oppløsningen i 2D-videoen, jmfør 6.4.

Når det gjelder avstanden til skjermen i forhold til størrelsene på personene på skjermen, er det Tridality-skjermen som er den begrensende faktoren. Å sitte 360 cm unna en skjerm på 42-tommer gjør at en ganske liten del av synsfeltet dekkes. Å ha en samtale med noen på en avstand som oppfattes som omtrent seks meter er også noe merkelig. Det hadde med andre ord vært en fordel å hatt en skjerm med noe kortere optimal seeravstand.

## 10.6 3D-i-3D

I et forsøk litt på siden av de andre forsøkte vi å finne ut hva som skjer dersom man spiller av en 3D-film på den autostereoskopiske skjermen, filmer dette i 3D og spiller av filmen på den autostereoskopiske skjermen. Blir 3D-effekten bevart? Ettersom kameraene er ment å fange verden slik øynene våre gjør det og skjermen viser bilder som hjernen skal tolke til å være i 3D, er det naturlig å anta at 3D-effekten blir bevart. Forsøket blir dermed en test på om utstyret fungerer etter hensikten. I forsøket brukte vi en film med god 3D-effekt og plasserte kameraene i optimal avstand fra skjermen.

Resultatet ble som forventet. I videoen ser man en autostereoskopisk skjerm i 3D hvor det som foregår på skjermen også foregår i 3D. 3D-effekten blir altså bevart. Videre forskning på dette området bør se på hva som skjer dersom en slik film blir filmet i 3D og vist på en autostereoskopisk skjerm<sup>1</sup>.

---

<sup>1</sup>Et forsøk på humor



# Kapittel 11

## Diskusjon

Hvorvidt det er mulig å lage god autostereoskopisk 3D til bruk i videosamtaler er avhengig av en del valg som gjøres. Disse valgene handler om man skal bruke konvergerende eller parallelle kamera og hvilken avstand man eventuelt skal ha til brennpunktet. Videre må man velge hvor stor skal avstanden være mellom kameraene og hvor mye skal bildet beskjæres ekstra. Det finnes ingen fasitsvar til disse valgene, og valgene man tar vil alltid være avhengige av hvilke andre faktorer som er til stede. Det er likevel mulig basert på våre utregninger og erfaringer med å lage autostereoskopisk 3D å komme med noen generelle betraktninger. I tillegg vil vi i dette kapitlet vise hvilke utfordringer vi har kommet over i arbeidet med å lage 3D. Til slutt vurderer vi hvilke egenskaper ved autostereoskopi som er relevante for videosamtaler.

### 11.1 Generelle betraktninger

Dersom man har mulighet til å benytte seg av parallelle kamera bør man gjøre det. Selv om konvergerende kamera i noen tilfeller kan by på bedre separasjonsverdier og mindre behov for beskjæring, vil konvergerende kamera alltid ha feil i bildet som en følge av at perspektivet er ulikt for kameraene, jamfør 8.1. Disse feilene blir større desto mer kameraene konvergerer. De autostereoskopiske skjermene vi har brukt virker å være mer følsomme for feil i bildet enn annen måte å vise 3D på, så denne typen feil bør man unngå så langt som det er mulig. Ettersom avstanden mellom kameraene sjeldent kan reduseres, er den eneste måten å minke kamerakonvergensens på å øke avstanden fra kameraene til brennpunktet. Det optimale blir dermed en uendelig avstand til brennpunktet og følgelig parallelle kamera.

En egenskap med konvergerende kamera som vi erfarte i arbeidet med prosjektoppgaven, er at objekter helst bør befinne seg i den positive parallaksen, området bak brennpunktet [36]. Dette er fordi objekter i den negative parallaksen, området foran brennpunktet, har en tendens til å bli oppfattet som at de befinner seg foran skjermen. Dette kan gi en flott 3D-effekt, men oppleves som slitsomt i lengden. Når konvergerende kamera brukes i autostereoskopisk 3D, ønsker man å holde kamerakonvergensens lav for å minke keystone-feil og gjøre øke størrelsen av området hvor separasjon i bildet er under 1,8%. Det fører ofte til at brennpunktet blir satt bak nærmeste objekt, slik at dette befinner seg i den negative parallaksen. Effekten av at objekter befinner seg foran skjermen er ikke like slitsom på autostereoskopiske skjermer som på stereoskopisk, polarisert 3D, men det er likevel en effekt som bør brukes med varsomhet.

## 11.2 utfordringer

Det har vært en del utfordringer med å lage god autostereoskopisk 3D, noen av disse er relatert til utstyret som er brukt, mens andre er relatert til teknologien i seg selv.

### 11.2.1 Programvare og maskinvare

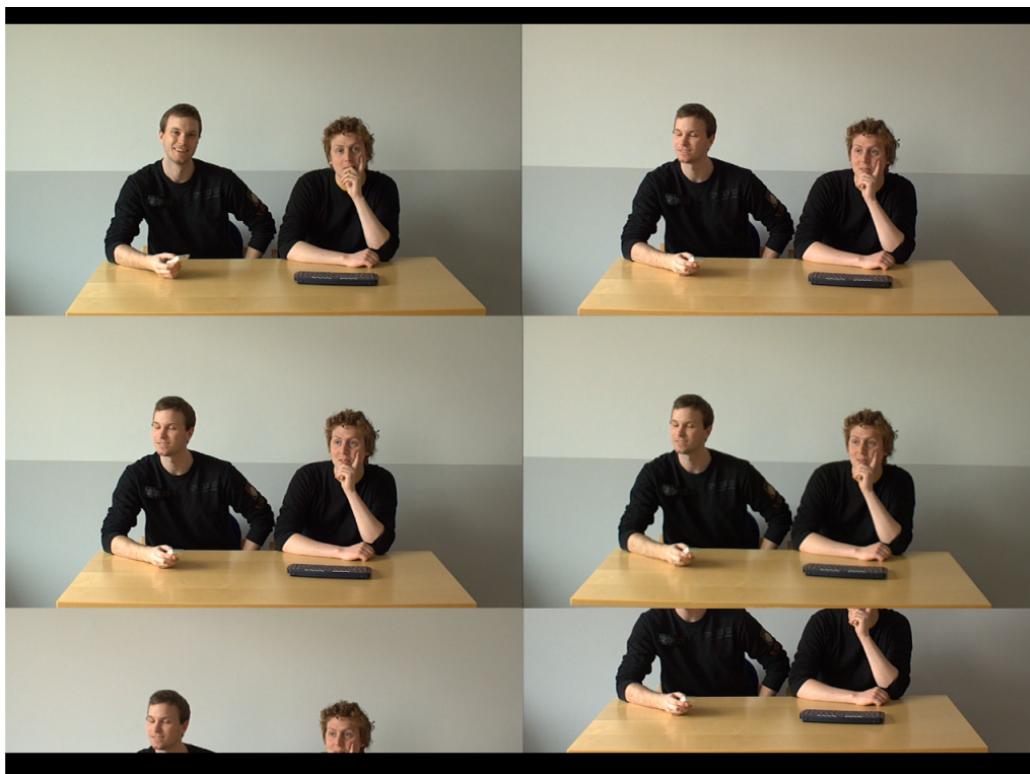
#### Hacking

Det har vist seg å være vanskelig å gjøre opptak som ikke har forstyrrelser i form av hacking eller korte perioder med redusert bilderate, fra spillverden kjent som “frame drop”. Denne hackingen er ofte ikke synlig dersom en ser på videoen fra ett kamera alene, men blir veldig synlig på den autostereoskopiske skjermen hvor en enkelt ser at den ene visningen ikke er synkronisert med de andre. Dette vises også godt ved avspilling på en vanlig skjerm som vist i figur 11.1. Her ser man tydelig at bildet øverst til venstre ikke er synkronisert med de andre.

Det har ikke vært enkelt å fastslå årsaken til disse feilene. Vi trodde først at det trolig skyldes at datamengden kameraene sender ut er større enn det selv våre relativt kraftige datamaskiner klarer å håndtere. Dette skulle vise seg å være delvis riktig.

## 11.2. UTFORDRINGER

---

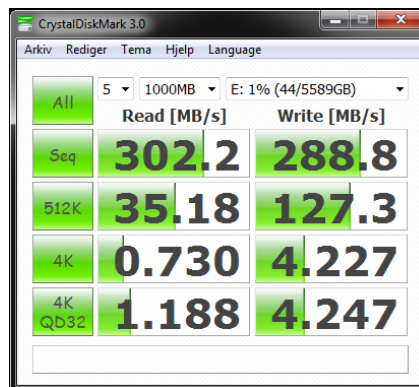


**Figur 11.1:** Skjermskudd som viser synkroniseringsfeil, samt høyde- og vinkelforskjeller

Flere tiltak ble gjort, det første var å dele lasten på to datamaskiner der den første datamaskinen tar i mot videostrøm fra to kameraer og den andre fra tre. Dette reduserte problemet noe, men ikke tilstrekkelig. Det neste var å redusere oppløsningen kameraene tar opp i for å redusere datamengden. I utgangspunktet ønsker vi å ha så høy oppløsning på bildene som mulig, men i behandlingen av bildene etter opptaket må bildestørrelsen uansett redigeres til  $968 \times 544$  bildepunkter. Det ble gjort opptak i oppløsningen  $1024 \times 576$  og dette løste problemet med hakkingen så langt vi kunne se. Dessverre gir denne oppløsningen oss svært liten frihet når det kommer til beskjæring for å øke separasjonen eller den beskjæringen som uansett må gjøres for å ta bort delene av bildet ingen andre kamera ser. Å redusere oppløsningen er dermed en dårlig løsning.

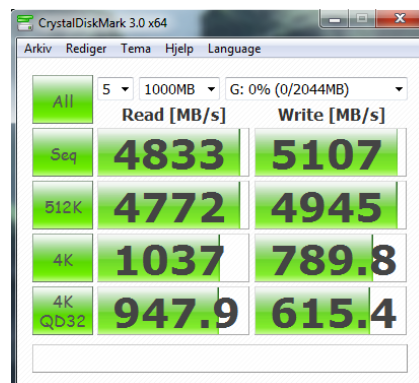
Ettersom datamaskinene helt fint klarer å vise bilder fra alle fem kameraene samtidig uten noen bildefeil når det ikke blir gjort opptak, er det rimelig å anta at det er i lagringen feilene oppstår. Ved hjelp av Windows' innebygde program Resource Monitor ser vi at datamengden opptaksprogrammet skri-

ver til disk på det meste er 30 MB per sekund, jamfør figur 11.5. I følge en test av harddiskene vi bruker skal de være i stand til å skrive 143 MB per sekund sekvensielt [97]. Selv datamaskinen som tar i mot video fra tre kameraer samtidig skal kunne klare å ta unna datamengden kameraene leverer. I et forsøk på å øke skrivekapasiteten satte vi i den ene datamaskinen opp tre harddisker i RAID 0<sup>1</sup>. Dette økte skrivekapasiteten til nesten 300 MB per sekund, jamfør figur 11.2.



**Figur 11.2:** Skjermskudd av programmet CrystalDiskMark

På den andre datamaskinen opprettet vi en såkalt RAMDisk som reserverer en del av minnet til lagring og gjør det skrivbart som en vanlig harddisk [99]. Tester av denne viste skrivehastigheter opp mot 5000 MB per sekund, se figur 11.3.



**Figur 11.3:** Skjermskudd av programmet CrystalDiskMark

<sup>1</sup>I RAID0, også kjent som striping, slås flere fysiske diskere sammen til en logisk enhet der datastrømmen fra en skriveprosess fordeles på diskene. Dette øker skrivekapasiteten [98].

## 11.2. UTFORDRINGER

Selv med dette oppsettet oppstod det hakking med ujevne mellomrom. Det var derfor grunn til å anta at prosessoren var flaskehalsen. Windows' Resource Monitor viste at prosessorbruken på PC-en med tre videokameraer tilkoblet var oppe i 96%. Vi bestemte oss derfor for å finne en ny kodek. Ulike kodeker er ofte optimalisert for ulike parametre. Vi ønsket å finne en som minimerte prosessorbelastningen og lot oss beholde så mye videokvalitet som mulig. Valget falt på HuffYUV som er en tapsfri kodek med svært lav prosessorbelastning [100]. Denne reduserte prosessorbelastningen til ca 40%, som vist i figur 11.4, men hakkingen er fortsatt ikke helt borte. Minnet er heller ikke fullt, så dette kan heller ikke være årsaken, jamfør figur 11.6.

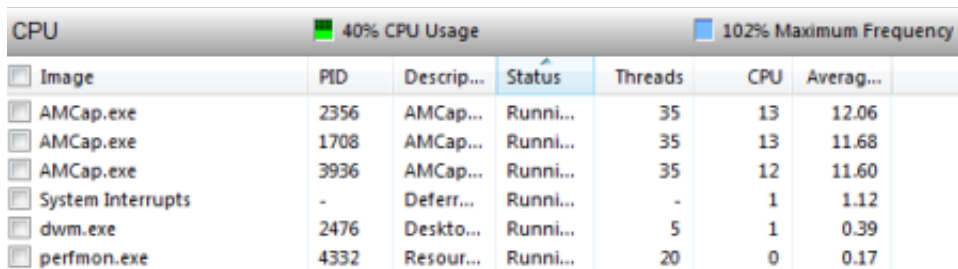


Image	PID	Descrip...	Status	Threads	CPU	Averag...
AMCap.exe	2356	AMCap...	Runni...	35	13	12.06
AMCap.exe	1708	AMCap...	Runni...	35	13	11.68
AMCap.exe	3936	AMCap...	Runni...	35	12	11.60
System Interrupts	-	Deferr...	Runni...	-	1	1.12
dwm.exe	2476	Deskto...	Runni...	5	1	0.39
perfmon.exe	4332	Resour...	Runni...	20	0	0.17

**Figur 11.4:** Skjermsskudd av programmet Resource Monitor, prosessorbelastning

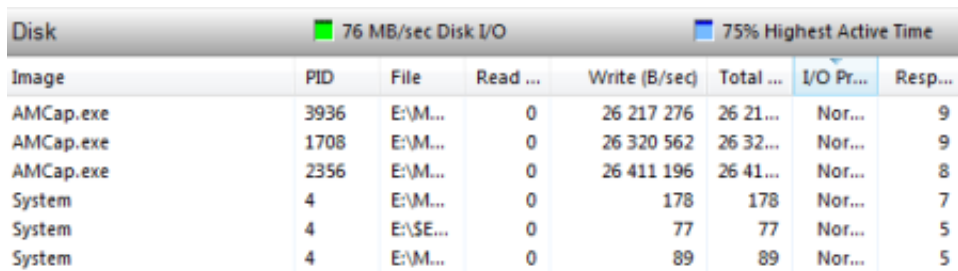


Image	PID	File	Read ...	Write (B/sec)	Total ...	I/O Pr...	Resp...
AMCap.exe	3936	E:\M...	0	26 217 276	26 21...	Nor...	9
AMCap.exe	1708	E:\M...	0	26 320 562	26 32...	Nor...	9
AMCap.exe	2356	E:\M...	0	26 411 196	26 41...	Nor...	8
System	4	E:\M...	0	178	178	Nor...	7
System	4	E:\SE...	0	77	77	Nor...	5
System	4	E:\M...	0	89	89	Nor...	5

**Figur 11.5:** Skjermsskudd av programmet Resource Monitor, diskbruk

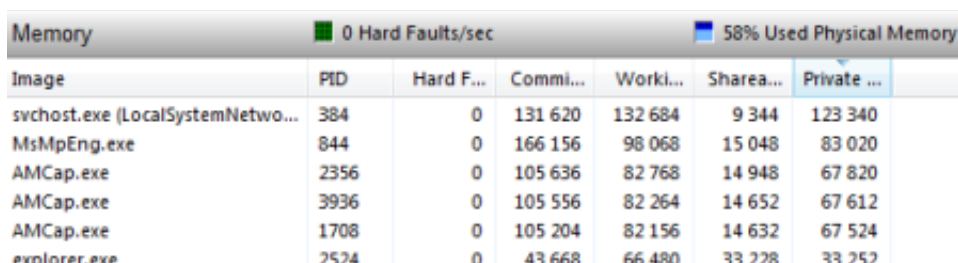


Image	PID	Hard F...	Commi...	Worki...	Sharea...	Private ...
svchost.exe (LocalSystemNetwo...	384	0	131 620	132 684	9 344	123 340
MsMpEng.exe	844	0	166 156	98 068	15 048	83 020
AMCap.exe	2356	0	105 636	82 768	14 948	67 820
AMCap.exe	3936	0	105 556	82 264	14 652	67 612
AMCap.exe	1708	0	105 204	82 156	14 632	67 524
explorer.exe	2574	0	43 668	66 480	33 228	33 252

**Figur 11.6:** Skjermsskudd av programmet Resource Monitor, minnebruk

Videofilene som lagres ved bruk av HuffYUV er svært store, men dette går bra så lenge de ikke skal sendes over et nettverk. Dersom en skal bruke autostereoskopisk 3D i videosamtaler over et nettverk bør videostrømmene fra kameraene komprimeres i maskinvare som er spesielt designet for dette.

### Redigering og enkoding

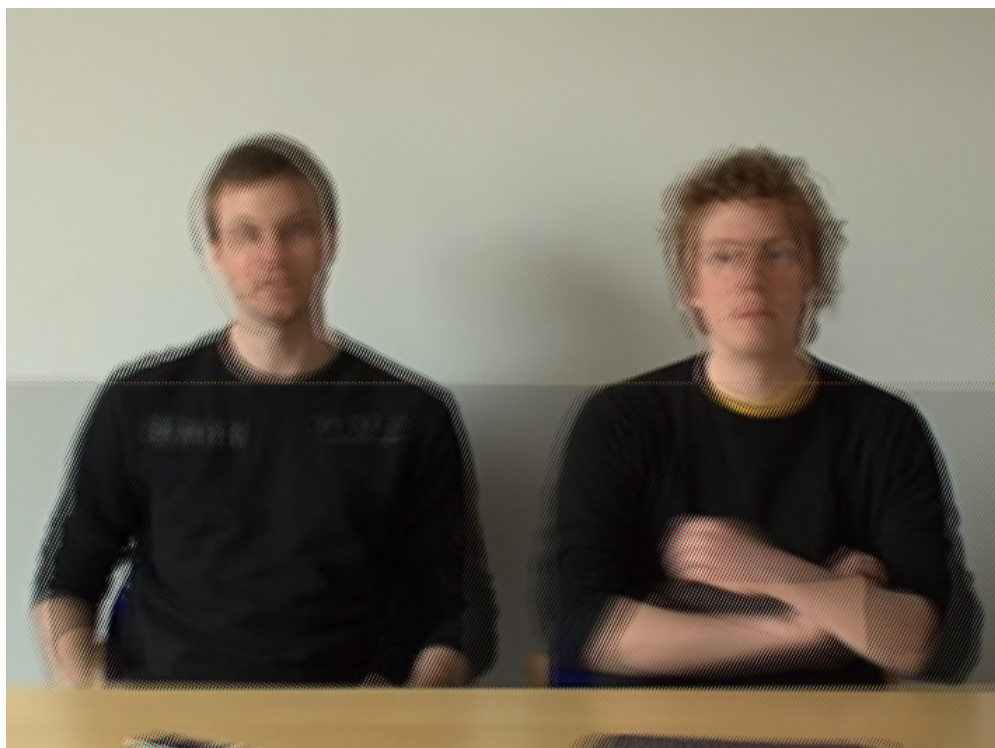
Video som skal spilles av på multiview-skjermen skal i følge Tridality ha en oppløsning på  $1936 \times 1360$  [101]. Denne oppløsningen er unormal, og vi har ikke fått til å eksportere video i den oppløsningen som HuffYUV fra Adobe Premiere. Vi ble derfor nødt til å velge en ny kodek da den ferdige videoen bestående av alle fem visningene skulle eksporteres fra Adobe Premiere. Den kodeken som ga det beste resultatet var Microsofts WMV. I innstillingene satte vi bitraten til 10 000 kb/s for å unngå å tape bildekvalitet som en følge av kompresjon. Ved å sammenligne den opprinnelige HuffYUV-videoen med den endelige WMV-videoen side ved side ser man ved første øyekast liten forskjell. Ved nærmere sammenligning ser man imidlertid at transkodingen har introdusert noe digital støy [102]. Dette vil trolig føre til noe redusert opplevd bildekvalitet.

Et større problem er at WMV-kodeken tilsynelatende sliter med de skarpe kontrastene mellom visningene i det ferdige bildet. Disse kontrastene prøver den å jevne ut, og lager dermed streker øverst og nederst på noen av visningene. Visningen fra kamera fem er horisontalt delt i to i bildematriksen og har en kant opp mot en annen visning. Når den ferdige videoen spilles av på den autostereoskopiske skjermen vises det dermed en strek horisontalt midt over bildet i visning fem, se figur 11.7. Det vises også noen streker øverst og nederst i bildet, men disse er i liten grad forstyrrende sammenlignet med streket midt over bildet. Denne streken er altså et produkt av kodeken som er brukt og ikke en feil som gjelder autostereoskopisk 3D generelt. For å bli kvitt denne streken bør man enten finne en kodek som takler kontrastene mellom visningene bedre, eller et avspillingsprogram som ikke er avhengig av at noen visninger deles i to i en matrise.

### 11.2.2 Oppløsning og beskjerping

I noen tilfeller ved bruk av konvergerende kamera, blir behovet for nødvendig beskjerping stort, jamfør 9.1. For at bildet skal beholde sin aspektrate, må bildet beskjerpes like mye i topp og bunn som på sidekantene. Høy beskjerping





**Figur 11.7:** Skjermkudd som viser en horisontal strek tvers over bildet. Skjermkuddet er tatt ved at TriView er brukt på en vanlig skjerm, og viser dermed også hvordan TriView fordeler visningene på den autostereoskopiske skjermen.

har mange ulemper. For det første mister man mange detaljer som befinner seg i kantene av bildet. Dette er ille nok i seg selv, men det fører også til at det blir vanskeligere å sørge for at objekter i bildet ikke blir overlappet av kanten av skjermen, jamfør 4.2. Når mye av bildet er beskåret bort, fører det også til at det resterende bildet blir blåst opp. Dette fører til at objekter i bildet blir større, og det kan bli vanskelig å få objektene til å framstå i naturtro størrelse. Man kan også risikere ved mye beskjæring at oppløsningen i det ferdig beskjærte bildet blir lavere enn det den autostereoskopiske filmen krever, i vårt tilfelle  $968 \times 544$  [101]. Da må bildet forstørres for å passe til skjermen, og dette kan føre til en alvorlig reduksjon i bildekvalitet.

## Ulike kamera

Et problem vi kom over etter å ha fått kamerastativet, var at kameraene er noe ulike. Ved hjelp av stativet får vi satt opp kameraene slik at de alle står helt flatt mot stativet og med samme høyde. Likevel er det enkelt å observere høyde- og vinkelforskjeller mellom kameraene når vi ser på bildene fra dem, jamfør figur 11.1. Vi har kommet fram til at disse forskjellene trolig skyldes at kameraene ikke er eksakt like. Mer presist tror vi det er plasseringen av bildebrikken bak optikken i kameraet som ikke er helt lik på alle kameraene. Denne forskjellen internt i kameraet er trolig svært liten, men forskjellen slik vi observerer den er avvik på 1-2 cm over to meter. Avvik i høyden kan i stor grad minimeres ved å justere kameraene fysisk, men forskjell i vinkler er det vanskeligere å gjøre noe med.

Dette ble svært synlig i de første testvideoene våre som inneholdt store bildeforstyrrelser. Ved å vise bildet som sendes til den autostereoskopiske skjermen på en vanlig skjerm får vi sett hvor godt kameraene er stilt inn, og det er lett å se at kameraene har filmet med litt ulike vinkler, med høydeforskjell og med ujevn avstand mellom seg. Ved å kalibrere kameraene slik at de er like i et punkt midt i bildet sørger vi for at ulikhetene mellom kameraene skaper størst feil i ytterkanten av bildet. Når bildet beskjæres forsvinner områdene hvor disse feilene er mest markante, men den samme beskjæringen sørger også for at de gjenværende feilene i bildet blir større når bildet forstørres igjen etter beskjæring. En bør derfor sørge for at kameraene en bruker er så like som overhode mulig.

Kameraene som blir brukt har automatisk styring av farger, lyssensitivitet, og fokus. Denne styringen er individuell for hvert kamera og er styrt av elektronikk inne i selve kameraene. Selv om kameraene har nesten helt identiske lysforhold og avstander til omgivelsene når de står ved siden av hverandre, blir videoene fra kameraene litt ulike. Dette blir ganske synlig på den autostereoskopiske skjermen, spesielt hvis en eller flere av visningene er ufokuserte. Forskjellene mellom visningene når det gjelder farger og lysstyrke er også ganske synlige og kan oppfattes som ubehagelige. For å løse dette må en bruke kamera der styringen av farger, lyssensitivitet og fokus kan synkroniseres.

### 11.2.3 Visninger

I kapittel 4 nevnes det at autostereoskopiske skjermer har muligheten til å vise bilder som gjenskaper bevegelsesparallaksen. Dette har vist seg å være

noe vanskelig å få til i praksis. Først og fremst er det overgangene mellom visningsvinduene som er problemet. Når man ser på skjermen og flytter seg, vil man i en periode være mellom to visningsvindu. Mellom visningsvinduene er bildet fullt av feil, og 3D-illusjonen brytes, i tillegg til at bildet oppleves som ubehagelig å se på. Dette problemet kan reduseres ved å øke antallet visninger, men vil ikke forsvinne fullstendig før antallet visninger kommer opp mot 30 per øyeavstand, jamfør 6.5. En mulig løsning vil være å interpolere optisk mellom visningsvinduene, for eksempel ved å skråstille parallaksebarrieren [103].

En opplever heller ikke en troverdig sammenheng mellom hvor mye man beveger seg til siden og hvor mye objektet man ser på relativt sett roterer. For å løse dette må kameraene plasseres i samme bredde som seeren kan bevege seg i. For å holde separasjonen lav, bør ikke avstanden mellom hvert kamera økes. Dermed må antallet kamera økes betraktelig. Det er mulig å generere virtuelle visninger mellom kameraene ved hjelp av dybdekart, jamfør 6.4. Dette kan redusere antallet kameraer.

I noen tilfeller vil bildet også bli feil selv om øynene ser to distinkt forskjellige visninger. Dette kommer av at visningene repeterer seg. Dersom man ser både ytterste venstre visning og ytterste høyre visning samtidig, vil separasjonen i bildet bli alt for stor og bildet blir ubehagelig å se på. Dette kan løses ved å ha så mange visninger at man ikke trenger å repetere dem.

Det kan virke som at den autostereoskopiske skjermen gir en del cross talk, jamfør 6.7. Dette ser man tydelig når man setter skjermen til å vise én farge per visning. Det er rett og slett ikke mulig å finne en posisjon hvor man på ett øye ser én helt ren farge. På sitt beste ser man fortsatt en overgang til fargen til neste visning i kanten av bildet. Dette oppleves ikke som et problem når videoene er laget riktig, men gjør skjermen mer følsom for feil i bildet som for eksempel vertikale forskyvninger.

## 11.3 Relevans for videosamtaler

### 11.3.1 3D som neste steg i telepresence

Telepresence har kommet et stykke på vei mot å kunne tilby naturlige og realistiske samtaler mellom personer som fysisk befinner seg forskjellige steder, jamfør 3.1. Samtidig er det fortsatt rom for forbedringer for å gjøre opplevelsen enda mer virkelighetsnær, jamfør 3.4. Et naturlig neste steg med

tanke på hvor aktuelt 3D er, jamfør kapittel 5, vil være å introdusere dybde-dimensjonen vi opplever i virkeligheten. Akkurat som at overgangene til farge-TV og HDTV kan sees på som generasjonsskifter i sin tid, er det en overensstemmelse blant mange teknologer i dag at det neste generasjonsskiftet vil være overgangen til 3DTV [80]. Dessverre finnes det ingen magisk formel for å gi skarpe og levende 3D-bilder som kan måle seg med de høyoppløselige 2D-bildene vi er vant med. Det er mange utfordringer knyttet til bruk av 3D i videokonferanser, jamfør 3.4, men med riktige valg, riktig utstyr og riktig teknologi kan resultat bli tilfredsstillende og også overgå 2D-systemene. For at opplevelsen av 3D skal bli så bra at det oppleves som en forbedring av et tilsvarende 2D-bilde, er det viktig å kjenne den aktuelle 3D-teknologiens styrker og svakheter.

### 11.3.2 Autostereoskopi i telepresence

Våre erfaringer med autostereoskopisk 3D er at på sitt beste gir denne 3D-teknikken minst like gode bilder som vanlig stereoskopisk 3D basert på polariserte filter. Det er ingenting i veien med dybdefølelsen, bildet oppleves som skarpt og fint, og verken lysnivå eller farger forringes. Selv om oppløsningen reduseres med antallet visninger, gir ikke dette, iallefall med fem visninger, en betydelig reduksjon i opplevd bildekvalitet. En viktig bemerkning er likevel at autostereoskopisk 3D er mye mer følsom for feil i bildet enn annen type 3D. Dette skyldes trolig at sammenlignet med andre 3D-teknologier har de autostereoskopiske skjermene mer cross talk, jamfør 6.2. Når dette settes i sammenheng med at det er vanskelig å kalibrere kamera, og vanskelig å finne optimale kameraoppsett, ser man at det kan være vanskelig å produsere innhold til autostereoskopiske skjermer. Dette passer godt med våre erfaringer, og kan også være en forklaring på hvorfor vi har funnet veldig lite kildemateriale til de autostereoskopiske skjermene vi har brukt. Et positivt aspekt ved at skjermene er mer sensitive for feil er at bildet kan oppleves som mindre slitsomt over tid. Egenskaper ved bildet som på en autostereoskopisk skjerm vil se ut som feil, vil når andre 3D-teknikker brukes i noen tilfeller vises tilsynelatende uten feil, men vil oppleves som ubehagelig over tid. Typiske slike feil er høydeforskjeller mellom visningene, ulike farger, for stor separasjon og høy divergens [36].

Autostereoskopisk 3D er preget av å være en umoden teknologi. Forsøkene på å lage gode autostereoskopiske skjermer spriker i flere retninger med hensyn på hva som er det beste kompromisset mellom antall visninger og oppløsning per visning, hvilket kameraoppsett som er best egnet, hva som er best av pa-

rallaksebarriere og lentikulære linser, hvordan skjermen også skal kunne brukes til 2D-video og om man skal bruke tilleggsteknikker som head-tracking, jamfør kapittel 6 og 7. Våre erfaringer er med et ganske lite utvalg av de mulighetene som finnes, men kan likevel si oss mye om hva som er styrkene og svakhetene til autostereoskopisk 3D.

De styrkene autostereoskopi har som er mest relevante for telepresence er først og fremst friheten fra å måtte bruke 3D-briller. Dette gjør at man kan berike bildene med en dybde dimensjon uten at 3D-brillene vil forstyrre den kritiske øyenkontakten. Samtidig har vi erfart, både gjennom det sammenlignende forsøket, men også i arbeidet ellers, at i en video av en person i autostereoskopisk 3D er det lettere å forstå ansiktsuttrykk og kroppsspråk enn i tilsvarende 2D-video. Dette er gunstig når autostereoskopi skal brukes i telepresence.

En egenskap ved autostereoskopisk 3D som ikke er like kritisk for telepresence, er gjengivelsen av bevegelsesparallaksen. I et telepresence-system vil deltakerne hovedsaklig sitte rolig og i liten grad ha behov for å bevege seg for å se samtalepartneren fra et nytt perspektiv. Våre erfaringer er at det er svært vanskelig i praksis å gjenskape bevegelsesaksen på en god måte. Det er derfor ikke negativt at det ikke er kritisk å gjengi denne på en god måte, jamfør 6.2.

En betydelig svakhet ved autostereoskopisk 3D med utstyr tilsvarende som vi har brukt, er at man ikke kan sitte hvor man vil. For å kunne se bildet uten feil må brukeren sitte slik at to visninger havner på hvert sitt øye. Dersom øynene havner mellom visningsvinduer, vil bildet bli fullt av feil. I optimal avstand fra skjermen er visningene slik at man kan sitte relativt naturlig avslappet uten å måtte holde hodet helt rolig, men det er likevel veldig begrenset hvor mye man kan bevege seg sidelengs. Dette kan virke noe hemmende for mennesker som naturlig beveger seg mye i en samtale, og det begrenser utformingen av rommet hvor videosamtalen skal foregå en del.

Utstyret vi har brukt har hatt noen svakheter som begrenser bruken av autostereoskopisk 3D til videokonferanser noe. Først og fremst handler dette om skjermenes innsynsvinkel og optimal avstand. Innsynsvinkelen er lav og begrenser dermed hvor mye på skrå skjermen kan stå i forhold til deltakerne. Den optimale avstanden er for høy i forhold til skjermens størrelse og gjør dermed at skjermen dekker en litt liten del av synsfeltet. Når deltakerne må sitte langt unna skjermen, blir de også sittende langt unna kameraene. Disse avstandene gjør at samtalen foregår på en opplevd avstand som ikke er naturlig for en samtale. Det er ikke usannsynlig at ettersom teknologien modnes vil disse egenskapene forbedres.



# Kapittel 12

## Scenarier

I dette kapitlet vil vi gå gjennom noen aktuelle og realistiske scenarier hvor autostereoskopisk 3D kan brukes i en videosamtale. For de ulike scenariene presenterer vi hvilke parameterverdier vi ut fra utregninger og erfaring tror vil gi best 3D med tilsvarende utstyr som vi har hatt tilgjengelig. Vi vil også presentere et idealscenario basert på det som teoretisk sett gir den beste autostereoskopiske 3D-opplevelsen. Idealscenariet vil også gi et innblikk i hva vi kan vente oss av 3D-teknologi i framtiden.

### 12.1 Videotelefon

For dette scenarioet tenker vi oss et typisk kontor med en videotelefon lik Tandberg EX90 [104]. Denne har en 24-tommer skjerm og brukeren sitter typisk 100 til 150 cm unna skjermen. I mange tilfeller er kontorer små slik at det ikke er veldig lang avstand fra brukeren til veggen bak, men for brukere som sitter i åpne kontorlandskap vil denne avstanden kunne være relativt lang.

På grunn av den korte avstanden fra skjermen og kameraene til brukeren i dette scenarioet, er ikke parallelle kamera noe alternativ. På 100 cm avstand og 4 cm mellom sentrum av katedralinsene gir parallelle kamera en separasjon på hele 5,18%, og en må opp i avstander på over 250 cm før separasjonen er lav nok til å ikke gi feil i bildet. Et bedre alternativ er å bruke konvergerende kamera med brennpunktet satt til 130 cm. Da vil separasjonen være under 1,8% mellom 85 cm og 255 cm fra kameraene. Den nødvendige beskjæringen som må gjøres hvis dybden i bildet er 255 cm er 9,31% av bildet for de

ytterste kameraene. Dersom dybden i bildet er større enn 255 cm, kan en sette brennpunktet 160 cm fra kameraene. Da vil separasjonen være under 1,8% mellom 100 cm og 410 cm. I dette tilfellet blir den nødvendige beskjæringen ved 410 cm avstand 8,85% av bildet for de ytterste kameraene.

Hvilken av disse posisjonene for brennpunktet som velges vil være avhengig av hvor dypt rommet videotelefonen står i er. Generelt ønsker man at kameraene konvergerer så lite som mulig for å minimere keystone-feil, jamfør 11. Samtidig bør en passe på at ikke brukeren befinner seg på grensen av det avstandsintervallet som gir best separasjon i bildet. Det er uheldig om det skulle oppstå feil i bildet hos brukerens samtalepartner fordi brukeren eller et objekt befinner seg for nært skjermen. Det bør med andre ord legges inn en margin slik at brukeren ikke trenger å passe nøye på hvor han eller hun sitter.

De kameraene vi har brukt i våre forsøk vil være egnet i et slikt scenario. Selv om større kamera kan gi bedre bildekvalitet, vil den økte avstanden mellom sentrum av kameralinsene gjøre at et allerede smalt avstandsintervall for god separasjon blir enda smalere. Med 160 cm fra kameraene til brennpunktet vil bare én cm økning i avstanden mellom kameraene minke intervallet hvor separasjonen i bildet er under 1,8% til 105 cm til 310 cm. Mindre kamera med tilsvarende bildekvalitet er det få av, og selv om det er mulig å filme bilder med større dybde med mindre avstand mellom kameraene, er det vanskelig å argumentere for at dette veier opp for dårligere bildekvalitet og et bilde med mindre dybdevirkning.

## 12.2 Møterom

For dette scenarioet ser vi for oss et ordinært middels stort møterom med plass til cirka ti deltakere. Et slikt møterom vil typisk være cirka åtte meter langt og fem meter bredt, ha et bord i midten og en skjerm plassert i den ene kortsiden av rommet. I et slikt scenario er det svært utfordrende å få til god autostereoskopisk 3D for alle brukerne med tilsvarende utstyr som vi har hatt tilgjengelig. For det første må skjermen være stor for at også de som sitter lengst unna skjermen, opptil 6-7 meter, skal se godt. Autostereoskopiske skjermer av stor størrelse har gjerne stor optimal avstand. Tridexterity, produsenten av skjermene vi har brukt, har en 57-tommer skjerm på markedet. Denne har oppgitt optimal avstand 4,25 meter [105]. For at det i det hele tatt skal være mulig å bruke autostereoskopiske skjermer i slike møterom må den optimale avstanden reduseres betraktelig.



En annen utfordring man vil støte på er at deltakerne ikke bare vil sitte i ulik avstand til skjermen, men også ha ulik innsynsvinkel. Skjermene vi har brukt har vært relativt følsomme på at seerens innsynsvinkel må være riktig for at bildet på skjermen skal kunne sees uten feil. Det vil dermed være veldig vanskelig å sikre at alle deltakerne i møterommet har et optimalt bilde.

Det vil også være vanskelig å finne optimale kamerainnstillinger for møterommet. Separasjonen må være under 1,8% over en ganske lang avstand, dette betyr normalt at man enten må bruke parallelle kameraer, eller sette brennpunktet langt unna kameraene. Ulempen med dette igjen er at hvis man bruker parallelle kameraer med 4 cm mellom sentrum av kameralinsene, må skjermen stå minst 255 cm unna nærmeste deltaker eller objekt, jamfør 9. Denne avstanden er kurant i forhold til de som sitter nærmest, men menneskene på andre siden av bordet kan risikere å bli veldig små i bildet som kameraene fanger. Denne avstanden gjør det også vanskeligere å plassere skjermen, bordet og deltakerne i rommet. Ved å sette brennpunktet i veggen på andre siden av rommet, 800 cm unna, vil nærmeste objekt kunne være 195 cm unna. Ulempen med dette igjen er at hele rommet da vil befinne seg i den negative parallaksen, jamfør 11.1. Et alternativ er å bruke mindre kameraer for å korte ned avstanden mellom kameralinsene. Med parallelle kameraer med 3 cm mellom sentrum av kameralinsene kan skjermen være 195 cm unna nærmeste objekt. Som nevnt i 12.1 er det vanskelig å få tak i så små kameraer og bildekvaliteten vil bli redusert. Etersom skjermen i dette scenarioet må være stor, er redusert bildekvalitet ekstra kritisk.

## 12.3 Videosamtalerom

Med videokonferanserom tenker vi oss et rom spesialdesignet for videokonferanser av høy kvalitet, lik Tandberg T3 [106]. Dette rommet har typisk seks deltakere som sitter på rekke langs et bord og ser på tre store skjermer, der de to ytterste skjermene er vinklet litt innover. Avstanden fra skjermene til bordet er omtrent 250 cm, bordet er omtrent en meter dypt og avstanden fra brukerne til veggen bak er omtrent en meter.

I likhet med møterommet, vil det for dette rommet være vanskelig å finne en autostereoskopisk skjerm som dekker stor nok del av synsfeltet. Skjermene som eksisterer i dag har for lang optimal seeravstand i forhold til skjermstørrelsen. Gitt at man finner en stor autostereoskopisk skjerm med optimal seeravstand rundt 250 cm, vil det være mulig å bruke autostereoskopisk 3D i et slikt rom. En fordel med et spesialdesignet rom i forhold til

et vanlig møterom er at stolene kan plasseres og skjermene kan vinkles slik at brukerne sitter i perfekt posisjon. I et videomøte av denne typen vil deltakerne typisk ikke bevege seg så mye at det vil være behov for å gjenskape bevegelsesparallaksen. Når det gjelder kameraene, vil parallelle kameraer over hver skjerm være den beste løsningen. Ved å vinkle kameraenes høydevinkel slik at kanten av bordet ikke vises, vil separasjonen være under 1,8% for hele bildet når avstanden mellom sentrum av kameraninsene er 4 cm.

I et slikt scenario er det i teorien mulig å bruke større kameraer for å forbedre bildekvaliteten, og konvergere kameraene for å få ned separasjonen igjen. I dette tilfellet vil likevel det ikke være en god idé, ettersom det vil være vanskelig å finne kameraer som forbedrer bildekvaliteten så mye at det veier opp for den reduserte kvaliteten konvergering av kameraene innebærer.

## 12.4 Optimalt oppsett

Et optimalt oppsett vil være et oppsett hvor man ikke kan skille videosamtalen fra en virkelig samtale. Det vil si at skjermen viser video i sanntid som innehar alle karakteristikkene til den opprinnelige scenen. Dette er kun mulig med holografiske teknikker. Holografi sees på som den ultimate 3D-teknologien [107]. Holografi rekonstruerer bølgefronten til den opprinnelige scenen og gjengir på denne måten alle dybdehint man opplever i virkeligheten, jamfør 4.1 [108]. Prinsipielt vil rekonstruksjonen av et objekt fra et perfekt hologram være umulig å skille fra det korresponderende objektet i den virkelige verden. Holografi har alle egenskaper som er nødvendige i et ideelt 3D-system [39]. Mange brukere kan benytte et slikt system samtidig og de kan bevege seg fritt over et stort område. Til og med problemet med fokusering og konvergering, jamfør 6.7 [109], unngås.

Holografi er fortsatt på prototypenivå og står foran mange teknologiske utfordringer. Det er lite trolig at en god holografisk løsning vil være klar for markedet de nærmeste årene, selv om blant andre SeeReal<sup>1</sup> er optimistiske i forhold til at produkter kan være klare så tidlig som 2010 [111]. Dette vil i så fall være små systemer med begrensede visningsområder. Det er sannsynlig at virkelig gode hologrammer blir realiserbare en stund inn i framtiden etter hvert som teknologien avanserer.

Generelt kan man si at det ultimate oppsett må:

- Støtte mange brukere

---

<sup>1</sup>SeeReal utvikler autostereoskopiske og holografiske skjermer [110]

#### 12.4. OPTIMALT OPPSETT

---

- La brukerne bevege seg fritt
- Tilby jevn bevegelsesparallakse
- Gjengi alle dybdehint man opplever i virkeligheten
- Levere en så høy oppløsning at bildet oppleves som helt realistisk
- Usynliggjøre all mellomliggende teknologi, jamfør 7.1



# Kapittel 13

## Konklusjon

Autostereoskopisk 3D er en spennende måte å øke den opplevde kvaliteten av en videosamtale på. Å innføre en dybde dimensjon i bildet gjør at dette oppleves som mer naturlig og virkelighetsnært. Når dette kan gjøres uten at deltakerne trenger å bruke spesielle briller, har man fjernet det kanskje viktigste ankepunktet mot innføring av 3D i videosamtaler. Autostereoskopisk 3D har likevel noen begrensninger som gjør at det ikke i alle typer videosamtaler vil være veldig vellykket å bruke det. Først og fremst handler disse om at brukeren ikke kan plasseres fritt i forhold til skjermen. I dag har autostereoskopiske skjermer høy optimal seeravstand i forhold til størrelsen, innsynsvinkelen er lav, visningsvinduene er smale og overgangene mellom dem ødelegger 3D-illusjonen. Dette er noe som vil bli bedre ettersom skjermteknologien utvikler seg.

På kamerasiden er det også utfordringer, de fleste av disse handler om å måtte tilpasse kameraoppsettet til begrensninger satt av skjermene. Den viktigste av disse er hvor mye separasjon i bildet skjermene tåler før det oppstår ghosting. Det er også viktig å begrense separasjonen i bildet for at hjernen skal kunne klare å sammenføre bildene fra øynene til et 3D-bilde uten at dette gir ubehag over tid. I noen tilfeller vil også kameraenes fysiske størrelse begrense friheten i kameraoppsettet.

Vi har likevel tro på at når autostereoskopi modner som teknologi, vil den finne veien inn i telepresence-løsninger. 3D kan tilby virkelighetsnære bilder på en måte som 2D aldri kan måle seg med. Hvis man er ute etter å tilby en kvalitet som i så høy grad som mulig ligner virkeligheten, kommer man rett og slett ikke utenom 3D, og for videosamtaler tror vi autostereoskopisk 3D er fremtiden.



# Bibliografi

- [1] METODEFORELSNINGER, GRUNNFAG PSYKOLOGI. <http://www.svt.ntnu.no/psy/Mons.Bendixen/>.
- [2] S. Salater. Videobjektors oppløsning kontrollert av blikkpunkt, 2005.
- [3] Dag Ingvar Jacobsen. *Hvordan gjennomføre undersøkelser?* Høyskoleforlaget AS, 2005.
- [4] The Assumptions of Qualitative Designs. <http://www.gifted.uconn.edu/siegle/research/Qualitative/qualquan.htm>.
- [5] Kvalitative og kvantitative forskningsmetoder - likheter og forskjeller. <http://www.etikkom.no/no/Forskningsetikk/Etiske-retningslinjer/Medisin-og-helse/Kvalitativ-forskning/1-Kvalitative-og-kvantitative-forskningsmetoder--likheter-og-forskjeller/>, 2010.
- [6] Qualitative versus Quantitative Research: Key Points in a Classic Debate. <http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html>.
- [7] Forskningsetiske retningslinjer for naturvitenskap og teknologi. <http://www.etikkom.no/Forskningsetikk/Etiske-retningslinjer/Naturvitenskap-og-teknologi/>.
- [8] Reidun Anita Gussiås. Militært lederskap - nye utfordringer for den militære sjef? Master's thesis, Forsvarets stabsskole, 2009.
- [9] Lynn Marotta. How does video conferencing work? A look at some of the basics of setting up a video conference. <http://www.video-conferencing-guide.com/how-does-video-conferencing-work.html>, 2006. Sist besøkt 20.05.10.

- 
- [10] Video for Cisco Unified Videoconferencing. <http://www.radvision.com/Unified-Communications/Video-for-Cisco-Unified-Video-Conferencing/cisco.htm>.
- [11] full hdtv. [http://whatis.techtarget.com/definition/0,,sid9\\_gci1071898,00.html](http://whatis.techtarget.com/definition/0,,sid9_gci1071898,00.html).
- [12] John Jackman. *Lighting for Digital Video and Television*. CMP Books, 2004.
- [13] Tandberg. Produkter. <http://www.tandberg.no/telepresence-produkter/index.jsp>, 2010. Sist besøkt 20.05.10.
- [14] Tandberg Telepresence T1/T3. [http://www.tandberg.com/collateral/product\\_brochures/TANDBERG%20Telepresence%20T3%20and%20T1%20Product%20Sheet.pdf](http://www.tandberg.com/collateral/product_brochures/TANDBERG%20Telepresence%20T3%20and%20T1%20Product%20Sheet.pdf).
- [15] Image and video compression techniques. <http://www.autosophy.com/videcomp.htm>.
- [16] Computer glossary. <http://www.meetingtomorrow.com/cms-category/computer-glossary>.
- [17] Overview of h.264/avc. <http://www.vcodex.com/h264overview.html>.
- [18] Cisco. Cisco TelePresence Solutions. [http://www.cisco.com/en/US/solutions/ns669/networking\\_solutions\\_products\\_genericcontent0900aec80546cd0.html](http://www.cisco.com/en/US/solutions/ns669/networking_solutions_products_genericcontent0900aec80546cd0.html), 2010. Sist besøkt 20.05.10.
- [19] Lossless. <http://www.techterms.com/definition/lossless>.
- [20] Axis. Compression Formats. [http://www.axis.com/products/video/about\\_networkvideo/compression\\_formats.htm](http://www.axis.com/products/video/about_networkvideo/compression_formats.htm). Sist besøkt 20.05.10.
- [21] Nefsis. The “Signing” of The Times - Video Conferencing Extends the Reach of American Sign Language Interpreters. <http://www.nefsis.com/Best-Video-Conferencing-Software/applications-asl.html>, 2010. Sist besøkt 20.05.10.
- [22] Cable and Wireless Worldwide. VIDEO CONFERENCING HAS ARRIVED. <http://www.cw.com/the-future-of-the-face-to-face-business-meeting-video-conferencing-has-> 2010. Sist besøkt 21.05.10.



## BIBLIOGRAFI

---

- [23] Gilda Bonanno. Non-verbal communications matter. Available at <http://ezinearticles.com/?Non-Verbal-Communications-Matter&id=1213104>, 2008.
- [24] Don Acquisition. Myth-Information: Communication is 93% Non-verbal. <http://www.wifcon.com/discussion/index.php?autocom=blog&blogid=6&showentry=322>, 2009. Sist besøkt 21.05.10.
- [25] Bjørn Bergslien. Spar heftig på videokonferanse. <http://www.ergogroup.no/default.aspx?path={B8AC4BF6-6C08-4404-A14D-190%41D48D0EB}>. Sist besøkt 21.05.10.
- [26] Lee Hopkins. Nonverbal communication in business. Available at <http://ezinearticles.com/?Nonverbal-Communication-in-Business&id=5406>.
- [27] S. McNelley. Immersive Group Telepresence and the Perception of Eye Contact. 2005.
- [28] About dve. <http://www.dvetelepresence.com/about/>.
- [29] Stephen Lawson. DVE brings 3-D look to meeting system. <http://www.networkworld.com/news/2010/030210-dve-brings-3-d-look-to.html>.
- [30] About telcordia. <http://www.telcordia.com/aboutus/index.html>.
- [31] D.A. Braun, T.J. Nelson, and L.S. Smoot. Eye contact video-conferencing system and screen, March 21 1995. US Patent 5,400,069.
- [32] beam splitter. <http://dictionary.reference.com/browse/beam+splitter>.
- [33] Sony. Sony 3D TelePresence. The new dimension. [http://www.video-audio-conference.com/brouchers/sony\\_3d\\_telepresence.pdf](http://www.video-audio-conference.com/brouchers/sony_3d_telepresence.pdf), 2008.
- [34] Tandberg. TANDBERG Total Telepresence. The Ultimate Meeting Experience. [http://www.tandberg.com/collateral/product\\_brochures/total-telepresence-solution-sheet.pdf](http://www.tandberg.com/collateral/product_brochures/total-telepresence-solution-sheet.pdf).
- [35] HOW TO VIDEO CONFERENCE. A Guide to Video Conferencing. <http://www.howtovideoconference.com/>.
- [36] S. Grønningen H. Smeplass. Nær-naturlige virtuelle læringsmiljø i 3d. 2009.
- [37] S.Ilstad. *Generell psykologi*. Tapir akademiske forlag, 2007.

- 
- [38] Tridexterity “Multiview” 3D Monitors. [http://www.pixelution.co.uk/Products/Computers/tridexterity\\_multiview\\_autostereoscopic\\_3D\\_LCD\\_monitors.html](http://www.pixelution.co.uk/Products/Computers/tridexterity_multiview_autostereoscopic_3D_LCD_monitors.html).
- [39] P. Surman, K. Hopf, I. Sexton, W.K. Lee, and R. Bates. Solving the 3D Problem - The History and Development of Viable Domestic. *Three-Dimensional Television*, pages 471–503.
- [40] M. Lombard, T.B. Ditton, M.E. Grabe, and R.D. Reich. The role of screen size in viewer responses to television fare. *Communication Reports*, 10(1):95–106, 1997.
- [41] Real d 3d system. <http://reald.com/Content/Files/REALD3Dsystems.pdf>.
- [42] Types of stereoscopic delivery. <http://usa.autodesk.com/adsk/servlet/item?siteID=123112&id=11906946>.
- [43] Trond Solem. På innsiden av planene. <http://www.forskning.no/artikler/2007/oktober/1192713234.52>, 2010.
- [44] Flat-TV. <http://www.elprice.no/tv-dvd/flat-tv>. Sist besøkt 29.05.10.
- [45] Abhinav Lal. Sony to release 3D TVs by June, amidst heavy Japanese competition [update]. [http://www.thinkdigit.com/TVs/Sony-to-release-3D-TVs-by-June\\_4435.html](http://www.thinkdigit.com/TVs/Sony-to-release-3D-TVs-by-June_4435.html), 2010.
- [46] Jack Burden. 3D TV Technology Guide for 2010 and Beyond. <http://www.lcdtvbuyingguide.com/hdtv/3d-tv-explained.html>, 2010.
- [47] About CES. <http://www.cesweb.org/aboutces.asp>. Sist besøkt 29.05.10.
- [48] Shardendu Gautam. CES 2010: The winning technologies and the tech future. [http://www.thinkdigit.com/Features/CES-2010-The-winning-technologies-and-the\\_3903.html](http://www.thinkdigit.com/Features/CES-2010-The-winning-technologies-and-the_3903.html), 2010.
- [49] Matthew DeCarlo. Blu-ray 3D video standard finalized, product launches imminent . <http://www.techspot.com/news/37373-bluray-3d-video-standard-finalized-product-launches-imminent.html>, 2009.
- [50] Scott Stein. Summer 2010: When the PS3 goes 3D. [http://news.cnet.com/8301-17938\\_105-10450842-1.html](http://news.cnet.com/8301-17938_105-10450842-1.html), 2010.

## BIBLIOGRAFI

---

- [51] Kshitij Sobti. YouTube adds a dimension, 3D goggles not included. [http://www.thinkdigit.com/Internet/YouTube-adds-a-dimension-3D-goggles-not\\_3219.html](http://www.thinkdigit.com/Internet/YouTube-adds-a-dimension-3D-goggles-not_3219.html), 2009.
- [52] TriDef Promotions Google Earth in Stereoscopic 3D! <http://www.tridef.com/promotions/google-earth.html>.
- [53] K. Muller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand. View synthesis for advanced 3D video systems. *EURASIP Journal on Image and Video Processing*, 2008.
- [54] About Google Trends. <http://www.google.com/intl/en/trends/about.html>.
- [55] autostereoscopic. <http://www.google.com/trends?q=autostereoscopic>.
- [56] Avansert søkestatistikk. <http://www.google.com/insights/search/?hl=no#>.
- [57] Nettsøk Interesse: autostereoscopic, stereoscopic. <http://www.google.com/insights/search/?hl=no#q=autostereoscopic%2C%20stereoscopic&cmpt=q>.
- [58] Jose Feroso. Philips 3D HDTV Might Destroy Space-Time Continuum, Wallets. <http://www.wired.com/gadgetlab/2008/10/philips-3d-hdtv/>, 2008.
- [59] 3D Eye Solutions. <http://www.3deyesolutions.com/>.
- [60] xyz 3D Display: The true 3D experience. <http://www.xyz3d.tv/>.
- [61] Japan Mobile Phone Goes 3D with Special Screen. <http://www.mobiledia.com/news/69201.html>, 2009.
- [62] Nintendo Plans New 3-D Hand-Held To Beat Apple. <http://thetechjournal.com/electronics/gaming-electronics/nintendo-plans-new-3-d-hand-held-to-beat-apple.xhtml>, 2010.
- [63] Re: Launch of New Portable Game Machine. <http://www.nintendo.co.jp/ir/pdf/2010/100323e.pdf>, 2010.
- [64] Chris Davies.
- [65] D.E. Roberts. History of lenticular and related autostereoscopic methods. *Hillsboro, WI: Leap Technologies, LLC*, 2003.

- 
- [66] Lenticular lens. [http://en.wikipedia.org/wiki/Lenticular\\_lens](http://en.wikipedia.org/wiki/Lenticular_lens).
- [67] A. Boev, M. Georgiev, A. Gotchev, and K. Egiazarian. OPTIMIZED SINGLE-VIEWER MODE OF MULTIVIEW AUTOSTEREOSCOPIC DISPLAY. *eye*, 1:2.
- [68] N. Holliman. 3D display systems. *to appear*, pages 0–7503.
- [69] Tridexterity. FAQ. <http://www.tridexterity.com/FAQ.267.0.html#c1082>.
- [70] N.A. Dodgson. Variation and extrema of human interpupillary distance. *Stereoscopic Displays and Virtual Reality Systems XI*, 5291:36–46, 2004.
- [71] T. Peterka. Dynallax: Dynamic parallax barrier autostereoscopic display, 2007.
- [72] H. Stolle, J.C. Olaya, S. Buschbeck, H. Sahm, and A. Schwerdtner. Technical solutions for a full-resolution auto-stereoscopic 2D/3D display technology. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6803, page 24, 2008.
- [73] A. Dodgson Neil. Autostereoscopic 3D displays. *Computer*, 38(8):31–36, 2005.
- [74] Wide VGA. [http://en.wikipedia.org/wiki/Wide\\_VGA](http://en.wikipedia.org/wiki/Wide_VGA).
- [75] mobile 4YOU NEWS. <http://www.mobilenews4u.com/tag/wvga/>. Sist besøkt 27.05.10.
- [76] O. Schreer, I. Feldmann, N. Atzpadin, P. Eisert, P. Kauff, and HJW Belt. 3DPresence<sub>i</sub>: A System Concept for Multi-User and Multi-Party Immersive 3D Videoconferencing. In *Visual Media Production (CVMP 2008), 5th European Conference on*, pages 1–8, 2008.
- [77] G.J. Woodgate, D. Ezra, J. Harrold, N.S. Holliman, G.R. Jones, and R.R. Moseley. Observer tracking autostereoscopic 3D display systems. In *PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, pages 187–198. SPIE INTERNATIONAL SOCIETY FOR OPTICAL, 1997.
- [78] Ian. Autostereoscopic Displays . 2009.
- [79] I. Sexton, R. Bates, WK Lee, P. Surman, S. Date, and M. Exports. A roadmap for autostereoscopic multi-viewer domestic tv displays. 2006.

## BIBLIOGRAFI

---

- [80] S. Pastoor. Human factors of 3dtv: An overview of current research at heinrich-hertz-institut berlin. Available at <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=193706&userType=&tag=1>.
- [81] K. Mashitani, H. Takahashi, and T. Aida. Multi-View Glass-Less 3-D Display by Parallax Barrier of Step Structure. *MEMOIRS-FACULTY OF ENGINEERING OSAKA CITY UNIVERSITY*, 48:1, 2007.
- [82] Corning Incorporated. Display Technologies . <http://www.corning.com/displaytechnologies/en/index.aspx>.
- [83] T. Annen, W. Matusik, H. Pfister, HP Seidel, and M. Zwicker. Distributed rendering for multiview parallax displays. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6055, pages 231–240. Citeseer, 2006.
- [84] J. Konrad and P. Agniel. Subsampling models and anti-alias filters for 3-D automultiscopic displays. *IEEE Transactions on Image Processing*, 15(1):128–140, 2006.
- [85] T. AkutsuTPF. Stereoscopic display which shows 3D natural scenes without contradiction of accommodation and convergence.
- [86] Reviews and awards. [http://www.adobe.com/products/premiere/reviews\\_awards/](http://www.adobe.com/products/premiere/reviews_awards/). Sist besøkt 01.06.10.
- [87] AMCap 9.20 review. <http://rbytes.net/software/amcap-review/>. Sist besøkt 10.06.10.
- [88] Tridality Display Solutions. <http://tridality.com/Displays.179.0.html>. Sist besøkt 11.06.10.
- [89] Tandberg PrecisionHD USB Camera. <http://www.tandberg.com/video-conferencing-telepresence-peripherals/hd-usb-camera.jsp>. Sist besøkt 07.04.10.
- [90] Test av Panasonic TM300. [http://www.akam.no/artikler/en\\_ny\\_legende/70430](http://www.akam.no/artikler/en_ny_legende/70430). Sist besøkt 15.06.10.
- [91] AMCap. <http://www.noeld.com/programs.asp?cat=video#AMCap>. Sist besøkt 05.04.10.
- [92] FFDSHOW. <http://sourceforge.net/projects/ffdshow/>. Sist besøkt 12.03.10.
- [93] Adobe Premiere Pro CS5. <http://www.adobe.com/no/products/premiere/>. Sist besøkt 12.06.10.

- 
- [94] How 3-D movie is made. <http://www.the3drevolution.com/3dscreen.html>.
- [95] How a 3D movie is made - from Home to Hollywood 3D. <http://www.the3drevolution.com/3dscreen.html>.
- [96] About Goal Seek. <http://office.microsoft.com/en-us/excel-help/about-goal-seek-HP005203894.aspx>. Sist besøkt 10.06.10.
- [97] Seagate Barracuda XT 2TB SATA 6G HD Review. <http://hothardware.com/Articles/Seagate-Barracuda-XT-2TB-SATA-6G-HD-Review/?page=6>, 2010. Sist besøkt 05.04.10.
- [98] RAID - Redundant Array of Inexpensive Disks. <http://www.interscapetech.com/blog/?cat=19>, 2008. Sist besøkt 03.04.10.
- [99] What is RAM disk? [http://www.webopedia.com/TERM/R/RAM\\_disk.html](http://www.webopedia.com/TERM/R/RAM_disk.html), 2004. Sist besøkt 13.06.10.
- [100] Huffuyuv. <http://neuron2.net/www.math.berkeley.edu/benrg/huffuyuv.html>. Sist besøkt 16.04.10.
- [101] Tridality tech note series 1 -resolutions, 2008.
- [102] JPEG and JPEG2 Artifacts. <http://www.stat.columbia.edu/~jakulin/jpeg/artifacts.htm>, 2004. Sist besøkt 13.06.10.
- [103] Autostereoscopic displays. <http://sp.cs.tut.fi/mobile3dtv/technology/displays.shtml>, 2009. Sist besøkt 05.06.10.
- [104] TANDBERG EX90. <http://www.tandberg.com/telepresence-products/personal-telepresence-ex90.jsp>. Sist besøkt 05.06.10.
- [105] autostereoscopic Multi-View 3D-display: MV5700. [http://usa.tridality.net/fileadmin/user\\_upload/downloads/datasheets/MV5700\\_datasheet.pdf](http://usa.tridality.net/fileadmin/user_upload/downloads/datasheets/MV5700_datasheet.pdf). Sist besøkt 05.06.10.
- [106] TANDBERG T3. <http://www.tandberg.com/telepresence-products/telepresence-t3-t1.jsp>. Sist besøkt 15.06.10.
- [107] S. Reichelt, R. H  
"aussler, G. F  
"utterer, and N. Leister. Depth cues in human visual perception and

## BIBLIOGRAFI

---

- their realization in 3D displays. In *Proceedings of SPIE*, volume 7690, page 76900B, 2010.
- [108] A. Schwerdtner, R. Hussler, and N. Leister. Large holographic displays for real-time applications. *Practical Holography XXII: Materials and Applications, Proceedings of the SPIE*, 6912.
- [109] R. Hussler, A. Schwerdtner, and N. Leister. Large holographic displays as an alternative to stereoscopic displays. In *Proc. SPIE*, volume 6803.
- [110] Company Group. <http://www.seereal.com/en/company/index.php>. Sist besøkt 05.06.10.
- [111] Like being there. Holography. <http://www.seereal.com/download/docs/Prospekt%20HoloTech%20SeeReal.pdf>. Sist besøkt 05.06.10.





# Tillegg A

## Utregninger

### A.1 Konstanter og variabler

$A$ :	Avstand mellom kamera
$B$ :	Avstand fra kamera til brennpunkt
$C$ :	Avstand fra brennpunkt til objekt
$C_1$ :	Beskjæring av bildet fra ytterste kamera
$C_2$ :	Beskjæring av bildet fra nest ytterste kamera på andre siden av midten for $C_4$
$C_3$ :	Beskjæring av bildet fra midterste kamera
$C_4$ :	Beskjæring av bildet fra nest ytterste kamera på andre siden av midten for $C_2$
$C_F$ :	Beskjæring utover det nødvendige
$C_H$ :	Beskjæring på høyre side av bildet
$C_N$ :	Nødvendig beskjæring
$C_T$ :	Total beskjæring
$C_V$ :	Beskjæring på venstre side av bildet
$D_{min}$ :	Minimumsavstand med tanke på at ingen kameraer skal se noe ingen andre ser
$D_S$ :	Skjermbredde
$D_V$ :	Bredden av bildet i objektets avstand
$D_{YK}$ :	Bredden av bildet av ytterste kamera
$H$ :	Siktlinje fra kamera til linje parallelt med kamerarekka i avstand $B+C$
$S_B$ :	Separasjonen i bildet
$S_V$ :	Separasjonen i virkeligheten

- $V$ : Kameraenes bildevinkel  
 $W$ : Vinkelen mellom visninger, sett fra brennpunkt

## A.2 Absolutt minimumsavstand

Minimumsavstand med tanke på at ingen kamera skal kunne se noe de andre kameraene ikke ser Parallele kamera:

$$D_{min} = 2A \tan\left(\frac{\pi}{2} - \frac{v}{2}\right)$$

Konvergerende kamera:

$$D_{min} = 2A \tan\left(\arctan\left(\frac{B}{2A} - \frac{v}{2}\right)\right)$$

## A.3 Nødvendig beskjæring parallelle kamera

I formlene for høyre og venstre side gjelder at  $K_N$  er kamera nummer  $N$  og  $1 \leq N \leq 5$ , nummerert bakfra.

Høyre side:

$$C_{NH} = \frac{(K_N - 1) - A}{2B \tan(\frac{v}{2})}$$

Venstre side:

$$C_{NV} = \frac{(5 - K_N)A}{2B \tan(\frac{v}{2})}$$

## A.4 Nødvendig beskjæring for konvergerende kamera

Ytterste høyre kamera:

$$C_1 = \frac{C_Y}{D_{YK}}$$

$$C_Y = \left| \frac{H \sin(\frac{v}{2}) 4A \sin(\frac{v}{2})}{\left(\frac{1}{2} \cos(2 \arctan(\frac{2A}{B})) + \cos(v)\right) B \sqrt{\frac{4A^2}{B^2} + 1}} + \frac{4A}{B} \right|$$

$$D_{YK} = \frac{H \sin(\frac{v}{2}) \cos(\frac{v}{2})}{\left(\frac{1}{2} \cos(2 \arctan(\frac{2A}{B})) + \cos(v)\right) \sqrt{\frac{4A^2}{B^2} + 1}}$$

Nest ytterste høyre kamera:

$$C_2 = \left| \frac{A \tan(\frac{v}{2})}{B} + \frac{(\frac{A^3}{B^2} + A) \cos(2 \arctan(\frac{A}{B})) + \cos(v)}{B \sin(v) \sqrt{B^2 + C^2} \sqrt{\frac{A^2}{B^2} + 1} + C} \right| \quad \text{for } C > 0$$

$$C_2 = \left| \frac{A \tan(\frac{v}{2})}{B} + \frac{(\frac{A^3}{B^2} + A) \cos(2 \arctan(\frac{A}{B})) + \cos(v)}{B \sin(v)(B + C)} \right| \quad \text{for } C \leq 0$$

Midterste kamera:

$$C_3 = \left| c \frac{(B + C) \tan(\frac{v}{2}) + \frac{2AC}{B} - \frac{\left(\sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))}\right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}}{2(B + C) \tan(\frac{v}{2})} \right| \quad \text{for } C > 0$$

$$C_3 = \left| \frac{(B + C) \tan(\frac{v}{2}) + \frac{2AC}{B} - \frac{\frac{(B + C) \sin(\frac{v}{2})}{\cos(\arctan(\frac{2A}{B}))}}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}}{2(B + C) \tan(\frac{v}{2})} \right| \quad \text{for } C \leq 0$$

Nest ytterste venstre kamera:

$$C_4 = \frac{x - y + \frac{AC}{B}}{x - z}$$

der

$$x = \frac{\left( \sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{A}{B}))} \right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{A}{B}))}$$

$$y = \frac{\left( \sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))} \right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} - \arctan(\frac{2A}{B}))}$$

$$z = \frac{\left( \sqrt{B^2 + C^2} + \frac{C}{\cos(\arctan(\frac{2A}{B}))} \right) \sin(\frac{v}{2})}{\cos(\frac{v}{2} + \arctan(\frac{A}{B}))}$$

## A.5 Separasjon i bildet for parallelle kamera

$$C_T = C_V + C_H$$

$$C_V = 1 - \frac{1 - C_F}{2} - C_F(1 - C_N)$$

$$C_H = 1 - \frac{1 - C_F}{2} - C_F$$

$$C_T = 1 - C_F(1 - C_N)$$

$$C_N = \frac{4A}{D_V}$$

$$S_V = \frac{A}{1 - C_T}$$

$$S_V = \frac{A}{C_F(1 - C_N)}$$

$$D_V = 2B \tan\left(\frac{v}{2}\right)$$

$$S_B = \frac{S_V}{D_V}$$

$$S_B = \frac{A}{2C_F(B \tan(\frac{v}{2}) - 2A)}$$

## A.6 Separasjon i bildet for konvergerende kamera

$$S_V = \frac{C \tan(w)}{1 - C_T}$$

$$w = \arctan\left(\frac{A}{B}\right)$$

$$S_V = \frac{CA}{B(1 - C_T)}$$

$$S_B = \frac{S_V}{D_V}$$

$$S_B = \frac{CA}{BD_V(1 - C_T)}$$

$$H = \sqrt{B^2 + C^2} + \frac{C}{\sqrt{\frac{4A^2}{B^2} + 1}} \quad \text{for } C > 0$$

$$H = \frac{B + C}{\sqrt{\frac{4A^2}{B^2} + 1}} \quad \text{for } C \leq 0$$

$$C_T = \frac{C_Y}{D_{YK}}$$

$$C_Y = \left| \frac{H \sin\left(\frac{v}{2}\right) 4A \sin\left(\frac{v}{2}\right)}{\left(\frac{1}{2} \cos\left(2 \arctan\left(\frac{2A}{B}\right)\right) + \cos(v)\right) B \sqrt{\frac{4A^2}{B^2} + 1}} + \frac{4A}{B} \right|$$

$$D_{YK} = \frac{H \sin\left(\frac{v}{2}\right) \cos\left(\frac{v}{2}\right)}{\left(\frac{1}{2} \cos\left(2 \arctan\left(\frac{2A}{B}\right)\right) + \cos(v)\right) \sqrt{\frac{4A^2}{B^2} + 1}}$$

$$C_T = \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(\sqrt{B^2 + C^2} \sqrt{\frac{4A^2}{B^2} + 1 + C})} \right| \quad \text{for } C > 0$$

$$C_T = \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(B + C)} \right| \quad \text{for } C \leq 0$$

$$S_B = \frac{CA}{BD_V \left( 1 - \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(\sqrt{B^2 + C^2} \sqrt{\frac{4A^2}{B^2} + 1 + C})} \right| \right)} \quad \text{for } C > 0$$

$$S_B = \frac{CA}{BD_V \left( 1 - \left| \frac{2A \tan(\frac{v}{2})}{B} + \frac{(\frac{8A^3}{B^2} + 2A) \cos(2 \arctan(\frac{2A}{B} + \cos(v)))}{B \sin(v)(B + C)} \right| \right)} \quad \text{for } C \leq 0$$