

Blikk på skjermen

Bruk av eyetracking for evaluering av visuelt
komplekse TV-programmer

Amund Berg



Masteroppgave
Master i medieteknikk
30 ECTS
Avdeling for informatikk og medieteknikk
Høgskolen i Gjøvik, 2007

Avdeling for
informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Faculty of Computer Science
and Media Technology
Gjøvik University College
Box 191
N-2802 Gjøvik
Norway

Sammendrag

Etter hvert som TV går inn i en ny digital hverdag får produsentene nye verktøy. Det gir muligheter for å produsere langt mer komplekse TV-bilder enn tidligere. Det er lite kunnskap om hvordan seeren påvirkes av denne typen komplekse bilder. Formålet med denne oppgave er å finne ut om eyetracking kan måle seeratferd på TV og i hvilken grad ulik kompleksitet i grafikken på TV påvirker seeratferden.

Et eksperiment med eyetracking har blitt gjennomført med 21 deltakere som så klipp fra tre ulike TV-programmer med varierende grad av dynamisk grafikk. Resultatene fra dette eksperimentet har blitt kontrollert for ulike parametere og analysert for å finne en metode for å måle seeratferd ved hjelp av eyetracking.

En vektor som beskriver distansen øyet beveger seg i en sakkade ble funnet som mål på seeratferd. Denne vektoren har en endring mellom sett med ulik mengde grafikk.

Abstract

As new digital world emerges, new tools become available for the producers of television. This opens up for new possibilities to create more complex TV images. There is little knowledge of how these kinds of complex images affect the viewer. The goal of this thesis is to see if eyetracking is suitable for measuring viewer behavior in television, and to what extent the complexity of TV graphics affects viewer behavior.

An experiment with eyetracking has been conducted with 21 participants. They were asked to watch three different TV programs with varying degree of dynamic graphics. The results of this experiment has been checked for different parameters and analyzed to find a method for measuring viewer behavior using eyetracking.

A vector was used to measure the distance the eye moves during a saccade. This vector is found as a measurement for viewer behavior. There is a change in this vector for the different sets whit different amount of graphics.

Forord

Denne masteroppgaven er utført ved Institutt for informatikk og medieteknikk ved Høgskolen i Gjøvik. Formålet har vært å kartlegge om eyetracking kan brukes for å måle seeratferd og om seeratferd påvirkes av ulik grafikk på TV.

I løpet av prosjektperioden har jeg fått hjelp fra flere ansatte og medstudenter ved Høgskolen i Gjøvik. Jeg vil rette en takk til Frode Volden som har vært veileder for prosjektet. Rune Hjelsvold fortjener en stor takk for verdifulle innspill. Jeg ønsker også å takke Roel Puijk og resten av TiDE-miljøet for støtten de har gitt til prosjektet, og spesielt Yngvar Kjus for hjelpen med å skaffe til veie Idol-opptak til stimulumaterialet. En takk går også til Thorsten Ahlrichs hos SMI for raske, forståelige og oppklarende svar på tekniske spørsmål.

Den viktigste støttespilleren for prosjektet har vært Marius Pedersen. Han har hatt gode innspill med stor faglig dybde gjennom hele prosjektet, og tatt seg tid til å sette seg inn i og komme med forslag til løsninger på små og store problemer.

Medstudentene på A220 fortjener en takk for å ha vært til betydelig faglig hjelp og minst like betydelig sosial atspredelse.

En stor takk går til min samboer Vivian som har hatt en uovertruffen tålmodighet med meg.

Amund Berg, 30. juni 2007

Innhold

Sammendrag	iii
Abstract	v
Forord	vii
Innhold	ix
Figurer	xi
Tabeller	xiii
1 Innledning	1
1.1 Bakgrunn	1
1.2 Forskningsspørsmål	1
1.3 Motivasjon og nytteverdi	2
1.4 Oppbygning av rapporten	2
2 Teori og relaterte arbeid	3
2.1 Bruksområder for eyetracking	3
2.2 Synsfeltet	4
2.3 Hendelsesdeteksjon	5
2.4 Målevariable	5
2.5 Målemetoder for sakkade	6
2.6 Utvelgelse av deltakere	7
2.7 Begreper	7
3 Eksperimentoppsett	9
3.1 Stimulimateriale	9
3.1.1 Sett 1 - Cindy	9
3.1.2 Sett 2 - Aleksander	9
3.1.3 Sett 3 - World Exchange	10
3.1.4 Lengde	10
3.1.5 Format	10
3.1.6 Rekkefølge	10
3.2 Område av interesse (AOI)	10
3.3 Fysisk oppsett	11
3.4 Javaprogrammet ETmessage	12
3.5 Dominant øye	13
3.6 Kalibrering	13
3.7 Instruksjoner	14
3.8 Datafilene fra iView X	14
3.9 BeGaze	15
3.10 Innlesing av dataene i MATLAB	15
3.11 Behandling av data i SPSS	15
3.12 Forkasting av deltakere	16
4 Resultat	17
4.1 Generelt	17

4.2	Sammensetning av deltakere	17
4.3	Resultater fra spørreskjema	17
4.4	Antall fikseringer, sakkader og blunk	18
4.5	Fiksasjonstid, sakkadetid og blunktid	18
4.6	Sakkadeamplitude	20
4.7	Sakkadevektor	20
4.8	Fikseringsfrekvens	22
5	Diskusjon	25
5.1	Forkasting av deltakere	25
5.2	Tid og antall	26
5.3	Amplitude og vektorlengde	26
5.4	Fikseringsfrekvens	26
5.5	Scanpath og heatmap	27
6	Konklusjon	29
7	Videre arbeid	31
	Bibliografi	33
A	Vedlagte filer	37
B	Spørreskjema	39
C	Resultater fra spørreskjema	43
D	Sammendrag av hendelser i dataene	47
D.1	Forklaring	49
D.2	Hendelser i sett 1	50
D.3	Hendelser i sett 2	52
D.4	Hendelser i sett 3	56

Figurer

1	Synsfeltet	5
2	Ulike mål på øyebevegelse	6
3	Eksempel fra stimulumaterialet	9
4	Fordeling mellom grafikk og program	11
5	Eksperimentoppsett	12
6	SMI RED eyetracker kamera	12
7	Histogram – aldersfordeling i deltakergruppen	18
8	Kjønnfordeling i deltakergruppen	18
9	Estimert gjennomsnitt for sakkadevektoren	21
10	Estimert gjennomsnitt for frekvensfordelingen	23
11	Vektor og amplitude for deltaker 15 sett 1	27
12	Scanpath og heatmap	28
13	Spesielle hendelser i Sett 1	50
14	Spesielle hendelser i Sett 2	53
15	Spesielle hendelser i Sett 3	57

Tabeller

1	Beskrivelse av kolonnene i datafilene	15
2	Inkluderte og ekskluderte deltakere	17
3	Antall fikseringer, sakkader og blunk	19
4	Gjennomsnittlig tid per fiksering, sakkade og blunk i millisekund	19
5	Fikserings- sakkade- og blunktid relativt til total tid i prosent	19
6	Amplitude i grader	20
7	Gjennomsnittlig vektorlengde for sakkader	20
8	Parret t-test	21
9	Gjennomsnitt fikseringer per AOI	22
10	Parret t-test	22
11	Gjennomsnittlig pikselverdi for settet	25

1 Innledning

1.1 Bakgrunn

Prosjektet inngår som en del av paraplyprosjektet «Fjernsyn i digitale omgivelser», forkortet TiDE etter det engelske «Television in digital enviroment» [1]. Hovedformålet for TiDE er å ta for seg konsekvenser av digitaliseringen som foregår innen fjernsyn. Bidraget fra «Blikk på skjermen» i dette prosjektet er å analysere hvordan seeren ser visuelt komplekse bilder i fjernsynsproduksjon.

TV har tradisjonelt sett vært produsert som en enkel strøm av bilder, med noe statisk grafikk som overlegg. Etter hvert som fjernsynet går inn i en digital hverdag får produsentene helt nye verktøy for å kunne lage mer komplekse bilder. Det er enkelt å lage dynamisk grafikk som kan brukes alene eller sammen med tradisjonelle fjernsynsbilder. Vi ser denne formen for TV i en rekke ulike former allerede. For eksempel har vi fått nyhetskanaler hvor det i tillegg til den tradisjonelle nyhetsoppleseren kontinuerlig ruller nyheter som tekst over skjermen.

Et annet aspekt for fjernsynsprodusenten i den digitale hverdagen er nye måter å kommunisere med seeren, eller snarere for seeren å kommunisere med fjernsynsprodusenten og andre seere. Nye kommunikasjonskanaler som internett og SMS, gjør at tilbakemelding fra seeren raskt og enkelt kan brukes som et virkemiddel i produksjonen.

Dette ser vi for eksempel i såkalt «SMS-TV» hvor seeren er en aktiv deltaker i programmet ved å sende meldinger til en chat som ruller over skjermen. Det er gjerne flere ulike typer felter seeren kan sende meldinger til, en vanlig chat i en del av skjermen, et mer eksklusivt felt hvor meldingen blir på skjermen lenger, og et eget felt for MMS-bilder. Sammenlagt fører dette til at bildet blir sammensatt og komplekst å se på.

I tillegg til at nye digitale verktøy har gjort det mulig å legge grafikk over og ved siden av tradisjonelle bilder har det også blitt mulig å bygge opp hele studio med dynamisk grafikk. Dette har vi sett eksempel på i nyhetssendinger og valgsendinger hvor programlederen omgis av digitalt fremstilte dynamiske elementer.

Alle disse ulike grafiske elementene gjør at TV-bildet blir mer komplekst enn i tradisjonell TV. Dette gjør at seeren må forholde seg til programmet på en annen måte enn tidligere. Med tradisjonelle TV-bilder er det i hovedsak ett budskap som kommuniseres, selve programinnholdet. Med komplekse bilder er det ikke lenger bare programinnholdet seeren må forholde seg til, men også informasjonen som kommer gjennom de grafiske elementene. Det er gjort lite forskning på hvordan atferden til seeren endrer seg.

Eyetracking har etter hvert vist seg å være et godt verktøy for forskere innen mange ulike områder. Det har etter hvert blitt vanlig å bruke blant annet innen brukervennlighetsstudier for web. Likevel er det få som har brukt eyetracking for studier innenfor TV, og enda færre som har sett på komplekse TV-bilder.

1.2 Forskningsspørsmål

Gjennom denne masteroppgaven er det i hovedsak to forhold det er ønskelig å finne svar på:

Spørsmål 1: *Er eyetracking egnet for studier av seeratferd?*

Seeratferd kan være vanskelig å tallfeste. Gjennom eksperiment er det ønskelig å finne ut om eyetracking er egnet til å gjøre studier av seeratferd. Det er ønskelig å se om det kan dannes en målev variabel på seeratferd ut fra resultatene fra bruk av eyetracking.

Spørsmål 2: *Påvirkes seeratferden av mengden grafikk på TV?*

Dersom det lar seg gjøre å finne en målev variabel tilsvarende den i forskningsspørsmål 1, er det ønskelig å bruke denne variabelen for å se om det er endring i seeratferd hos deltakerne i eksperimentet. Det er spesielt interessant å se om det er endring i seeratferden dersom mengden grafikk i bildet øker.

Med seeratferd forstås i denne sammenheng hvordan seeren søker informasjon i skjermen. Det kan være hvordan blikket forflyttes eller hvor det festes. Med endring i seeratferd menes hvordan seeren endrer måten han ser på skjermen for å finne informasjon.

1.3 Motivasjon og nytteverdi

Motivasjonen for denne masteroppgaven er at den kan være med på å kartlegge et nytt bruksområde for eyetracking. Eyetracking er et nyttig verktøy i mange sammenhenger, og det er interessant å se om denne oppgaven kan finne en måte å bruke det. Seeratferd er vanskelig å måle og denne oppgaven vil forsøke å finne en metode for å måle dette.

Kartlegging av seeratferd anses som viktig kunnskap for fjernsynsbransjen. Området er ikke i særlig grad utforsket tidligere. Betydningen av å undersøke forholdet vil derfor være av stor hjelp for bransjen. Det vil gjøre det enklere å planlegge oppbygningen av grafiske profiler for TV-produksjoner og gjøre brukervennlighetstester for fjernsyn.

1.4 Oppbygning av rapporten

Kapittel 1 starter med innledning og tar for seg bakgrunn for oppgaven og de problemstillingene oppgaven skal besvare. Deretter følger kapittel 2 som presenterer nødvendige teorier for oppgaven og andres relaterte arbeid.

I kapittel 3 presenteres gjennomføringen av eksperimentet fra forberedelse av stimulumateriale i avsnitt 3.1 – 3.2, gjennom det fysiske oppsettet i avsnitt 3.3 – 3.4 til gjennomføringen av eksperimentet i avsnitt 3.5 – 3.7. Metodikken for håndteringen av data i etterkant er omtalt i avsnitt 3.8 – 3.12.

Kapittel 4 presenterer resultatene for de ulike målev variablene som blir diskutert videre i kapittel 5. Kapittel 6 oppsummerer og gir en felles konklusjon. Forslag til videre arbeid er presentert i kapittel 7.

Vedleggene inneholder et bredere grunnlag for resultatene som har blitt presentert.

2 Teori og relaterte arbeid

2.1 Bruksområder for eyetracking

Eyetracking har en rekke svært ulike anvendelsesområder. Spennvidden er stor fra Gullberg og Holmqvist [2] sine studier at persepsjon av gestikulering i samtaler, via Ware og Mikaelian [3] sine eksperimenter for å bruke eyetracking til å styre datamaskinen, til Babcock m. fl. [4] sine studier av hvordan mennesker ser på scener og bilder fra de er i ferd med å ta bildet, for å nevne noe.

Eyetracking har også blitt brukt til en rekke studier av hva som trekker oppmerksomhet. Lundqvist og Holmqvist [5] gjennomførte eksperimenter hvor de så på hvilken betydning størelsen på annonser i avis har for hvorvidt leseren stopper opp ved de og husker de. De fant at deltakerne stoppet opp lenger og husket mer av større annonser. Det samme gjelder fargeannonser mot sort/hvitt annonser. Stark og Garcia [6] fant at deltakerne bruker mer tid på fargeannonser enn sort/hvite annonser

Radach m.fl. [7] gjorde eksperimenter hvor de så på forskjellen mellom annonser med direkte budskap og indirekte budskap. De benyttet eyetracking for å se om det var noen forskjell i måten deltakerne så på annonsene med direkte kontra inndirekte budskap. De fant at det var en endring i hvor lenge deltakeren så på de ulike annonsene, men det var ingen signifikant endring i antallet fikseringer, tiden per fiksering eller lengden på sakkadene. Et annet spennende funn de gjorde var at deltakerne spesielt i annonsene med inndirekte budskap ofte «parkerte» øynene på et område tilsynelatende uten særlig interessant informasjon mens de prosesserte informasjonen de hadde sett.

Av mer relevante eksperimenter er det gjort to studier på teksting (subtitling) på TV. Det ene studiet er utført av Jonas Borell i hans masteroppgave fra 2000 [8]. Han viser to forskjellige versjoner av samme film, en med original lyd på fransk og teksting på deltakerens eget språk, svensk, og en annen versjon uten teksting men med lyd på svensk. Deltakerne i dette eksperimentet var tredelt; en kontrollgruppe som så den svenske versjonen, en gruppe som så den tekstede versjonen og ikke kunne fransk og en gruppe som så den tekstede versjonen og kunne fransk. Borell fant at de som kunne fransk og derved forsto lydsporet hadde 57% færre fikseringer på tekstingen.

Bruyker og d'Ydewalle [9] gjorde et studie på lesing av teksting hvor de har både barn på 5-6 år og voksne som deltakere. Deltakerne fikk se programmer tekstet både på sitt eget (tysk) og et fremmed språk (svensk). De brukte antallet fikseringer, fiksasjonstid og sakkadetid i området med teksting kontra resten av bildet for å måle forskjellen mellom deltakerne. De målte også hvor lang tid det tok fra teksten kom inn i bildet til deltakeren flyttet blikket dit. De gjorde også forsøk med enkelt linje teksting mot dobbelt linje teksting. De fant at deltakerne brukte mindre tid på å lese fremmedspråklig teksting enn teksting på sitt eget språk. De fant også at færre droppet å lese tekstinger som går over to linjer i forhold til tekstinger på en linje. De fant ingen forskjell på barna og de voksne deltakerne, bortsett fra at barna gjerne fikserte lengre og hadde kortere sakkader. Det stemmer også med det blant annet Rayner [10] har rapportert angående barns lesemønstre.

Aoki og Itoh har gjort flere eksperimenter med eyetracking og reklame på tv hvor de forsøker å avdekke en sammenheng mellom øyebevegelser og holdninger til reklame. I ett eksperiment fra 2000 [11] hadde de fire kvinnelige deltakere som fikk se et halvtimes talkshow med to reklamepauser med til sammen 26 reklamer. De registrerer total tid brukt på sakkader og fikseringer i områder med informasjon av betydning for reklamen og danner ut fra dette et forholdstall for hvor interessant reklamen er. De bruker dette forholdstallet videre og sammenligner med det deltakeren selv har sagt de synes er interessant. De fant at det var en sammenheng mellom hva deltakeren selv sa og hva eyetrackingen målte. De fant også at en kjendis i reklamefilmen har liten effekt, utenom at deltakerne kanskje husket filmen bedre i ettertid.

Et annet eksperimentet Aoki og Itoh gjennomførte i 2004 [12] har et veldig lignende oppsett ved at de har to reklamer de ønsker å teste effekten av. Disse to reklamene vises blant 11 reklamer i to reklamepauser i et halvtimes talkshow. De hadde elleve deltakere i eksperimentet som alle så reklamefilmene fire ganger. I dette eksperimentet fant de en sammenheng mellom total tid brukt på sakkader og fikseringer på reklamen og hvor mange spørsmål deltakerne kunne svare korrekt på i etterkant. Dette stemmer godt med det Jarbus [13] konkluderte med i sine studier at øyet både frivillig og ufrivillig fokuserer på de elementene i et objekt som kan gi viktig og nyttig informasjon. Desto mer informasjon som finnes i elementet, desto lenger holdes blikket der.

Vi gjør en del øyebevegelser som ikke er beviste. Godjin og Theeuwes [14] nevner studier hvor deltakerne fikk se en skjerm hvor det dukket opp et objekt. Instruksjonene til deltakerne var å se til motsatt side av skjermen i forhold til hvor dette objektet dukket opp. Deltakeren ble spurt om de selv trodde de hadde gjort en feilaktig sakkade mot objektet, og 20% mente de hadde gjort det. I 57% av tilfellene viste det seg at deltakerne hadde flyttet blikket uten å selv være klar over det.

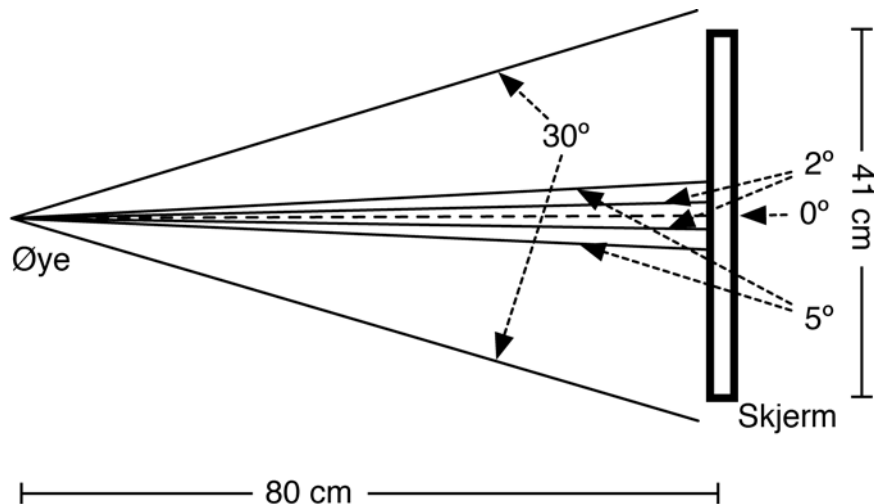
Vi tiltrekkes også automatisk av ansikter. Dette har blitt rapportert gjennom en rekke studier slik som Daly m. fl. [15], Bradley og Stentiford [16] og Endo m. fl. [17]. Disse har vist at vi for eksempel når vi snakker i videotelefon fester blikket på ansiktet og aller helst leppene til personen vi snakker med.

2.2 Synsfeltet

Det normale synsfeltet for et menneske er innenfor 180° horisontalt og 130° vertikalt. Øyet er skarpest innenfor 2° av synsfeltet. Utover dette faller skarpheten raskt. Den parafoveale sonen strekker seg til 5° [18]. Ved 5° reduseres skarpheten med omlag 50% [19]. Det brukbare synsfeltet blir regnet å være opp til 30°, og alt utover det regnes å kun være egnet til å registrere bevegelser (sidesynet) [18].

I forbindelse med eyetracking er dette viktig for å forstå hvor nøyaktig eyetracking egentlig er. Eksempelvis utgjør 5° på 80 cm avstand ca 7 cm. Det vil si at deltakeren i et eyetrackingeksperiment kan oppfatte objekter inntil 7 cm rundt det som registreres som fokuseringspunkt. 2°, altså området som er skarpt under en fiksering, utgjør på samme avstand ca 2,8 cm [20]. Dette er illustrert i Figur 1

Skjermen som brukes i dette prosjektet er cirka 27° × 21° på 80 cm avstand (se avsnitt 3.3)



Figur 1: Synsfeltet

2.3 Hendelsesdeteksjon

Det er tre typer hendelser som er interessant i eyetracking; fikseringer, blunk og sakkader. En fiksering er punktet hvor deltakeren fester blikket og holder det for en periode. Sakkade er det som er mellom fikseringene, altså bevegelser øyet gjør mellom fikseringer. Blunk anses som et særtilfelle av en fiksering, hvor den horisontale og vertikale blikkposisjonen er lik 0 [21].

Med 50Hz/60Hz eyetrackingsystem, slik som det som brukes i dette eksperimentet, er det anbefalt å bruke fikseringsdeteksjon som hovedhendelse [21]. Det betyr i praksis at ved hendelsesdeteksjon vil programvaren først lete etter fikseringer og tolke data mellom disse som sakkader.

Algoritmen som benyttes i programvaren brukt i dette eksperimentet kalles Dispersion-Threshold Identification (I-DT). To parametere settes for identifiseringen av en fiksering; varighet (duration) og terskelverdi (threshold). Algoritmen utnytter den egenskapen at datapunktene i en fiksering grupperes tett sammen. I-DT algoritmen bruker et vindu som dekker en mengde datapunkter for å lete etter fikseringer innen disse punktene. Vinduet starter først i datasettet og dekker et minimum av datapunkter, begrenset av varighetsparameteren og samplingsfrekvensen. I-DT sjekker avstanden mellom punktene (dispersion) ved å summere differansen mellom punktenes maksimum og minimum x- og y-verdier, med andre ord $D = [\max(x) - \min(x)] + [\max(y) - \min(y)]$. Hvis avstanden er over terskelverdien representerer ikke vinduet en fiksering og vinduet flytter ett punkt videre. Hvis avstanden er under terskelverdien representerer vinduet en fiksering og vinduet utvides til avstanden i datapunktene innenfor vinduet overskrider terskelverdien. Det endelige vinduet registreres som en fiksering med medianverdien som senterpunkt og en maksimal avstand (dispersion) og tid [20, 22, 23].

2.4 Målevariable

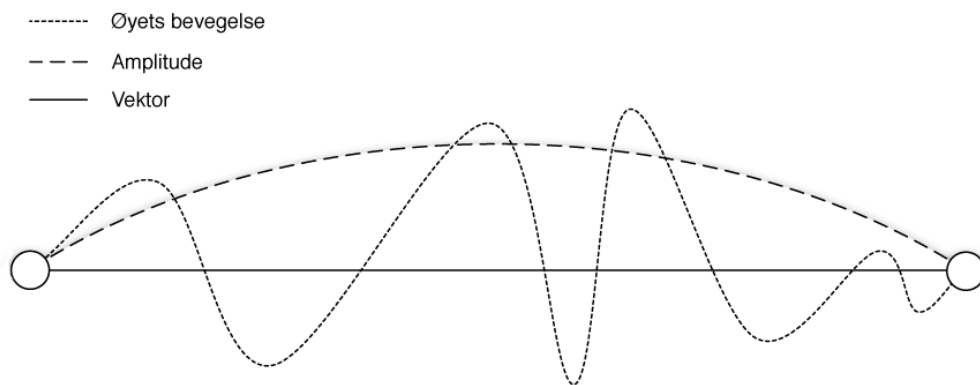
Som nevnt over er det i hovedsak tre typer hendelser i eyetracking, fiksering, sakkader og blunk. Det er disse tre ulike typene hendelser som benyttes i rapportering av funn i eyetracking studier, men måten de blir benyttet på varierer. Noen eksempler på hvordan

andre har valgt å bruke eyetracking følger.

- Bruyker og d'Ydewalle [9] og Borell [8] deler skjermen opp i ulike områder og ser på hvor lang tid som er brukt på fikseringer innen de ulike områdene.
- Aoki og Itoh [11, 12] brukte tid brukt på fikseringer og sakkader innen et område i forhold til den totale lengden på den viste videoen
- Zangemeister m.fl. [24] bruker blunkefrekvensen som en måte å måle aktivitet. Visuelt mer krevende bilder gjør at blunkefrekvensen blir lavere.
- Radach m. fl. [7] bruker amplituden på sakkaden målt i antall piksler
- Henderson m. fl. [25] bruker amplituden på sakkaden målt i grader

2.5 Målemetoder for sakkade

De fleste målemetodene er relativt selvforklarende, slik som antall, frekvens og tid. Metodene benyttet for å måle sakkader er kanskje ikke så intuitivt å forstå. Figur 2 er en illustrasjon av de to målemetodene. I hver ende er det en fiksering.



Figur 2: Ulike mål på øyebegivelse

Øyets bevegelse mellom disse punktene går ikke i en rett linje men beveger seg i ulike retninger før det ender opp på neste fikseringspunkt. Det er derfor to ulike måter å representere en sakkade. Den første er amplituden. Den beregnes på følgende måte [26]:

Amplitude = gjennomsnittlig vinkelhastighet \times varighet

Vinkelhastigheten er oppgitt i grader per sekund [27] og varigheten i millisekund. Resultatet fra dette blir oppgitt i antall grader, det vil si hvor mange grader ut av en hel sirkel blikket har forflyttet seg. Denne buen beskriver hvor langt utenfor en rett linje blikket har forflyttet seg under sakkaden.

Den andre måten å representere en sakkade på er å beskrive den rette linjen mellom fikseringspunktene. Ved å beregne en euklidsk avstand slik som i formel 2.1 mellom de to fikseringspunktene får man en vektor som beskriver distansen mellom de to fikseringspunktene uten å ta hensyn til retning. Denne vektoren tar ikke hensyn til de «omveiene» øyet tar mellom to fikseringspunkter. Synet er undertrykt under sakkader [28] slik at det oppfattes ikke ny informasjon under bevegelsen.

$$v = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2.1)$$

2.6 Utvelgelse av deltakere

I de fleste eyetrackingeksperiment er det en viss andel av deltakerne som ikke gir gode nok resultater til å kunne brukes.

- Widdel [23] forkaster 2-14% av de individuelle dataene på grunn av tekniske feil, blunk eller andre spesifikke problemer
- Underwood og Foulsham [29] forkastet 2 av 16 deltakere (12,5%) på grunn av manglende data for halve settet eller feil fiksering i starten av eksperimentet
- Aoki og Itoh [12] hadde 16 deltakere og endte opp med 11 (forkastet 31,25%) på grunn av dårlig kalibrering
- Crowe og Narayanan [30] gjennomførte eksperimenter med to ulike deltakergrupper for menneske-maskin interaksjon. Ene gruppen brukte mus og tastatur mens andre gruppen brukte stemmestyring for å styre maskinen. Blant gruppen som brukte mus og tastatur mistet de i gjennomsnitt 18% av dataene mens for gruppen med stemmestyring mistet de 35%.

Goldberg og Wichansky [31] sier at opp til 20% av de rekrutterte deltakerne vil ha problemer med at kalibreringen blir tapt. De nevner noen ulike metoder for å unngå problemer.

- Minimer mulighetene for å flytte på hodet, spesielt i kalibreringsprosessen. Om nødvendig kan hodestøtte eller bitestykke brukes.
- Dersom deltakeren har veldig mørk iris, prøv med både lys og mørk pupil tracking dersom det er tilgjengelig.
- Pass på å plassere kameraet på en slik måte at øyelokket ikke kommer i veien for kameraet.
- Prøv å unngå deltakere med briller eller linser. Personer med svak korrigering kan være lettere å kalibrere enn deltakere med sterk korrigering (sterke briller eller linser).
- Rekrutter 20% flere kandidater enn det som er nødvendig for eksperimentet.

Nguyen m.fl. [32] fant at om deltakeren har veldig mørk iris kan kontrasten mellom iris og pupill bli for liten til at systemet greier å finne pupillen, og det blir vanskelig å følge øyets bevegelse. De forklarer også en årsak til at man bør unngå deltakere med synsfeil. Noe av årsaken til langsynthet eller nærsynthet er at lyset ikke fokuseres korrekt på retina. Det samme gjelder for IR-lyset fra eyetrackeren slik at refleksjonen fra øyet får feil vinkel.

2.7 Begreper

Innen eyetracking er det en rekke uttrykk og begrep, og få av disse har blitt oversatt til norsk på en tilfredsstillende måte. Der hvor det ikke finnes gode norske begreper har de engelske blitt benyttet i stedet for å oversette. Det er foreslått noen oversettelser slik som øyebevegelsessporing eller blikkvandringsregistrering for eyetracking, søkebane for scahpath og hetekart for heatmap. Dette er begreper som ikke er innarbeidet og det er

derfor valgt å benytte de engelske begrepene i denne oppgaven.

3 Eksperimentoppsett

3.1 Stimulimateriale

Det har blitt benyttet tre sett stimulimateriale i eksperimentet. De tre settene er valgt ut fra en økende grad av dynamisk grafikk i bildet. De to første settene er hentet fra «Idol – Jakten på en superstjerne» som gikk på TV2 våren 2006.

Siden prosjektet er en del av TiDE var det ønskelig at stimulimaterialet hadde en knytting mot andre deler av TiDE. Idol har tidligere blitt brukt av Yngvar Kjus inn i TiDE [33], så det var derfor naturlig å se på om det også kunne brukes i dette prosjektet.

Alle sendingene fra Idol ble sendt to ganger. Først live fredag kveld med seeravstemning på sms. Reprisen ble sendt lørdag morgen og i denne sendingen var informasjon om stemmegivning erstattet med en sms-chat hvor seerne kunne sende inn spørsmål og kommentarer.

Valget falt på to ulike klipp fra forskjellige tidspunkter i konkurransen. Sett 1 med Cindy Selven Røe er fra sendingen 24/2-06, som er relativt tidlig i konkurransen. Sett 2 med Aleksander Denstad With er fra reprisen sendt 30/4-06.

CNBC World Exchange er en form for nyhetsprogram som omhandler trender i verdensøkonomien, rettet mot forretningsfolk som jobber på et internasjonalt marked. Settet som har blitt brukt er hentet fra sendingen 21/02-07 og er et tilfeldig klipp av en typisk sending.



Figur 3: Eksempel fra stimulimaterialet

3.1.1 Sett 1 - Cindy

Første sett er fra direktesendingen av Idol. Den grafiske informasjonen er begrenset til en boks med navnet på artisten og informasjon om hvordan man kan stemme på sin favoritt. Grafikken kommer inn etter 27 sekunder og ligger fast og uforandret til den forsvinner ut etter 2 minutter og 10 sekunder. Hele settet er 2 minutter og 19 sekunder.

3.1.2 Sett 2 - Aleksander

Andre sett er hentet fra reprisesendingen av Idol. Plasseringen og størrelsen på grafikken er den samme som i sett 1, men i stedet for informasjon om stemmegivning er det en sms-chat. I løpet av det 6 minutter og 5 sekunder lange settet vises det 31 sms-meldinger.

De fleste består bare av hvit tekst, mens 7 meldinger er fra moderator «Sølvi». Sammen med disse meldingene vises et bilde av moderatoren (som vist i figur 3(b)). Tre ganger i løpet av settet vises informasjon og regler for moderering med gul tekst.

3.1.3 Sett 3 - World Exchange

Siste settet er hentet fra CNBC World Exchange. Dette er det mest komplekse settet. Øverst i bildet er en linje som til en hver tid inneholder tre ulike børsverdier av større betydning, slik som DAX-indeksen og valutakurs for britiske pund kontra sveitsiske franc. Disse skifter med 3 sekunders mellomrom. Nederst i bildet er det to ulike rullende felt. Begge viser aksjekurs for ulike internasjonale selskaper. Disse tre grafiske feltene er synlige gjennom hele det 6 minutter og 46 sekunder lange settet, uavhengig av hva som vises av programinnhold. Hoveddelen av bildet består store deler av tiden av et delt bilde hvor venstre del er en grafisk fremstilling av historisk verdi for f.eks. en valutakurs og høyre del er et bilde av personen som snakker om den aktuelle grafen (figur 3(c)). De to rullende feltene ruller over skjermen i ulikt tempo. Den øverste bruker ca. 5 sekunder over skjermen og den nederste 7 sekunder. Det tilsvarer en hastighet på henholdsvis 5,4° per sekund og 3,9° per sekund. Kitamura m.fl. [34] fant i sine studier at deltakerne i deres eksperiment best oppfattet teksten når den rullet over skjermen med 5° per sekund og at de måtte opp i 10° per sekund før deltakeren ikke fikk med seg innholdet i teksten. Det skulle tilsi at disse to feltene skal være lett å lese for deltakeren.

3.1.4 Lengde

Nøyaktig lengde for de tre settene i millisekund er

1. 140776 ms
2. 364569 ms
3. 407421 ms

De tre settene er altså av ulik lengde. Resultatene gjengitt i kapittel 4 er på forskjellig vis kompensert for dette, slik at tallene og resultatene er sammenlignbare for de tre settene.

3.1.5 Format

Alle de tre programmene ble opprinnelig sendt som widescreen (16:9) format. Skjermen som brukes i dette eksperimentet er en dataskjerm med forhold 4:3 (se avsnitt 3.3). Stimulimaterialet ble derfor vist til deltakerne i letterboxing, det vil si med sort øverst og nederst i bildet.

3.1.6 Rekkefølge

For å kunne korrigere for rekkefølgeeffekter ble de tre settene vist i ulik rekkefølge til de ulike deltakerne.

3.2 Område av interesse (AOI)

For å kunne avgjøre om et fikseringspunkt er innenfor grafikk eller handlingen i bildet er det nødvendig å dele inn bildet. I alle tre settene er det en klar grense mellom hva som er grafikk og hva som er program, og grafikken dekker hele bildebredde. Det gjør at det kan defineres ulike områder av interesse (AOI etter det engelske Area Of Interest).

Inndelingen for sett 3 er litt annerledes enn for de to første. Den faste grafikken i

sett 1 og 2 er nederst i bildet og kan derfor merkes med ett AOI. I sett 3 er det to felter med fast grafikk, ett øverst og ett nederst. Det gjør at sett 3 må deles i tre AOI. Under behandlingen av dataene i senere kapitler blir AOI 2 og 3 i sett 3 omtalt som om det var ett felt.

For alle tre settene er hovedmotiv et 400 linjer og grafikken 176 linjer. For sett 1 og 2 er bildet delt i to områder slik at de øverste 400 linjene av bildet er hovedmotiv og nederste 176 er grafikk. For sett tre er linjene fra 0 til 75 og 476 til 576 regnet som grafikk, og 76 til 475 er regnet som hovedbildet. Sagt på en annen måte er det for alle settene 69% bilde og 31% grafikk. Inndelingen er illustrert i figur 4.



Figur 4: Fordeling mellom grafikk og program

Viktig å bemerke her er at for sett 3 er det tildels mye grafikk i selve programmet. Grafikk i bildet er derfor i denne sammenhengen betegnet som grafikk som er i bildet gjennom hele sendingen. Det vil si at grafikk av typen til venstre i bildet i figur 3(c) regnes som programinnhold, ikke grafikk i denne sammenhengen.

3.3 Fysisk oppsett

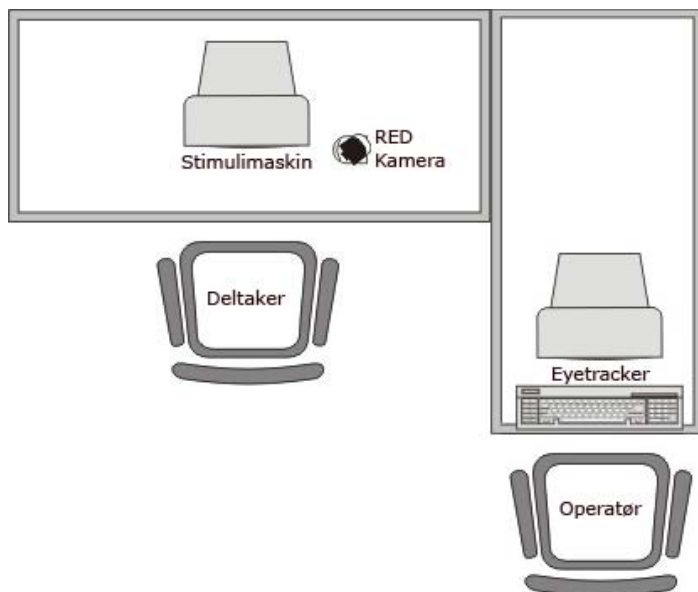
Eyetrackeren som har blitt brukt i dette eksperimentet er levert av tyske SensoMotoric Instruments (SMI) [35] og er en «Remote Eyetracking Device» (RED). Systemet består av et kamera med motorisert panorering, tilt, zoom og fokus, og alt styres via programvaren iView X fra SMI. Oppå kameraet, parallelt med linsen sitter en infrarød lyskilde. Se Figur 6. Kameraet plasseres på bordet foran deltakeren slik at det har mest mulig klar sikt til øyet. Under forsøkene ble kameraet plassert på høyre side av skjermen for de med dominant høyre øye og på venstre side for de med dominant venstre øye (se avsnitt 3.5).

Stimulimaskinen var en PC med Intel Pentium III 731MHz prosessor og 384MB Ram. Til denne sto en «LaCIE Electron 22 blue II» 21" CRT monitor med en bildeflate på 41x31 cm (51 cm eller 20" målt diagonalt) og tilnærmet flatt bilderør. Skjermen ble kalibrert på forhånd slik at fargegjengivelsen var mest mulig korrekt under eksperimentet. Oppløsningen på skjermen var satt til 1280x1024 piksler. Musen og tastaturet ble flyttet bort til eyetrackermaskinen av to årsaker:

1. Deltakeren trengte ikke å bruke det og derfor ville det kun distrahere
2. Det var nødvendig å betjene stimulimaskinen fra eyetracker maskinen

Deltakeren fikk lyden fra stimulumaterialet gjennom AKG K141 hodetelefoner. Deltakeren satt i en stol med armlener, ryggstøtte og fire ben uten hjul. Denne ble valgt for å begrense hvor mye deltakeren beveger seg i løpet av eksperimentet. Deltakeren satt med en avstand på cirka 80 cm til skjermen.

I følge manualen fra leverandøren fungerer eyetrackeren best med dempet og jevn



Figur 5: Eksperimentoppsett



Figur 6: SMI RED eye-tracker kamera

belysning [36]. En Skonic Flashmate L-308s lysmåler [37] ble brukt for å måle lyset. Dette gir lyset målt i EV (Exposure Value) Forholdet mellom EV og lux er gitt av Formel 3.1 [38, 39]. Lyset i rommet ble målt til 2.1 EV på veggen bak skjermen og 2.8 EV foran skjermen. Det tilsvare 11 lux på veggen og 17 lux foran skjermen. Til tross for at lyset i rommet er jevnt vil lyset fra skjermen variere noe etter hvor lyst bildet er på stimulimaterialet.

$$E = 2.5 \cdot 2^{EV} \quad (3.1)$$

3.4 Javaprogrammet ETmessage

Det var ønskelig å automatisere så store deler av eksperimentet som mulig. Hovedargumentet for å automatisere er å redusere muligheten for forskjell i resultatene på grunn av manuell gjennomføring. Det er for eksempel tilnærmet umulig å starte avspillingen av videoen på stimulimaskinen og datainnsamlingen på eyetracker maskinen eksakt likt hver gang om man skulle gjøre det manuelt. Derfor ble java-programmet ETmessage utviklet. Kildekoden til programmet ligger vedlagt i vedlegg A.

Eyetracker programvaren har en mulighet for «remote control» via nettverk (UDP). Det gjør at ETmessage kan bruke en socket-connection mellom stimulimaskinen og eyetrackermaskinen. Over denne tilkoblingen kan et sett «remote commands», slik som å starte og avslutte opptak, benyttes for å styre eyetracker-programvaren.

ETmessage har tre hovedfunksjoner: Starte avspillingen av stimuli-videoen på stimulimaskinen, starte logging av data og opptak av scenevideo på eyetracker maskinen, og logge hendelser i stimulimaterialet sammen med eyetracker dataene.

Hendelsene i stimulimaterialet var på forhånd lagret i en XML-fil som ETmessage kan lese. XML-filen har formatet vist i listing 3.1. hr, min og sec markerer avstanden i tid

fra forrige hendelse i stimulimaterialet slik at i eksemplet skjer hendelsen «Stemmeinfo inn» 27 sekunder etter «Start Cindy». Settnummeret blir økt og meldingene blir lagret i datafilene fra eyetrackeren.

Listing 3.1: Eksempel på xml-fil

```

1 <?xml version="1.0" encoding="ISO-8859-1" ?>
2 <timestamps>
3   <event>
4     <hr>0</hr>
5     <min>0</min>
6     <sec>2</sec>
7     <msg>Start Cindy</msg>
8   </event>
9   <event>
10    <hr>0</hr>
11    <min>0</min>
12    <sec>27</sec>
13    <msg>Stemmeinfo inn</msg>
14  </event>
15 </timestamps>

```

3.5 Dominant øye

Før deltakeren satte seg med eksperimentet ble en test for å finne dominant øye gjennomført. Omtrent 97% av befolkningen har ett øye som er dominant over det andre. For 65% er det høyre øye som er dominant, og for 32% er det venstre dominant. De resterende 3% har ingen signifikant preferanse [40]. Når vi fokuserer på et punkt bruker vi ubevisst det ene øyet i større grad enn det andre. Metoden som ble brukt for å finne det dominante øyet ble først beskrevet av Porta i boken «De Refractione» i 1593 og kalles Porta-metoden [40, 41]. Deltakeren fikk instruksjoner om å holde opp en finger på strak arm og peke på et punkt et stykke unna. Deltakeren lukket så ett og ett øye. Når det dominante øyet er åpent vil fingeren fortsatt peke på punktet, men for det andre øyet er det en forskyvning slik at fingeren bommer på punktet.

Det er uenighet om det er et poeng å følge det dominante øyet. Nguyen m. fl. [32] fant ingen bevis for at det er forskjell mellom det dominante og det andre øyet, mens Goldberg [31] sier at mesteparten av den «brukbare» visuelle informasjonen blir fanget opp av det dominante øyet. Det ble gjort et valg på å følge det dominante øyet i dette eksperimentet.

3.6 Kalibrering

Alle har forskjellig form på øyet. Det gjør at refleksjonen av IR-lys i øyet også er forskjellig fra person til person. Derfor er det nødvendig å kalibrere eyetrackeren for hver deltaker i eksperimentet [23]. En god kalibrering er kritisk for å få gode resultater fra eyetrackingen. Dårlig kalibrering kan føre til en forskyvning mellom det registrerte punktet og det faktiske punktet deltakeren har sett på. Leverandøren anbefaler å bruke 9-punkts kalibrering [36]. 9-punkts kalibrering er også det som har blitt brukt i andre prosjekter (f.eks. Babcock [42] og Vuori m.fl. [43]), derfor ble det brukt i dette prosjektet.

Kalibreringen gjennomføres ved at deltakeren får opp et rødt punkt på grå bakgrunn på skjermen. Deltakeren bes om å se på dette punktet. Punktet aksepteres av operatøren av eyetracker maskinen og neste punkt kommer opp på deltakerens skjerm. Dette gir en mapping mellom punktet på skjermen og øyebevegelsen til deltakeren.

3.7 Instruksjoner

Under rekrutteringen av deltakerne ble de informert om at de skulle se tre klipp fra forskjellige TV-program mens eyetrackeren fulgte bevegelsene til øyet. Når deltakeren hadde satt seg til rette ble personen bedt om å unngå å bevege hodet samt å se fremover mot skjermen under hele eksperimentet for at eyetrackeren ikke skulle miste kontakten med øyet.

Etter gjennomført eksperiment og før deltakeren fylte ut spørreskjemaet ble det informert om det egentlige formålet med forsøket. Dette ble gjort for å gjøre deltakeren i stand til å svare korrekt på spørsmålet om de kjente til formålet med eksperimentet (spørsmål 6 på spørreskjemaet, vedlegg B).

3.8 Datafilene fra iView X

Resultatet fra eyetrackingen er tre datafiler for hver deltaker. I tillegg ble en scenevideo tatt opp som en form for sikkerhets kopi. Dersom datafilene ikke gav mening hadde man videoene å gå tilbake til for å visuelt sjekke resultatene.

Datafilene som blir lagret fra iView X er i et binært format. IdfConvert er et konverteringsprogram som kan gjøre om disse binære filene til tekstfiler som er mulig å behandle videre. I listing 3.2 og 3.3 er det vist eksempler på hvordan datafilen kan se ut.

I konverteringen fra idf til tekst er det to parametere som kan settes:

- Min. Fix. Duration : 80
- Dispersion Threshold : 100.00

Det vil si, det blir ikke regnet som en fiksering før den varer 80ms og maksimal spredning i piksler fra senterpunktet i fikseringen er satt til 100 (se også avsnitt 2.3).

Det er fire ulike typer hendelser, Blink, Fixation, Saccade og UserEvent. Betydningen for de ulike feltene er vist i tabell 1. I listing 3.2 er det først to blunk, deretter en fiksering, etterfulgt av en sakkade over i en ny fiksering. Dette er hentet fra starten av filen, slik at første blunk starter 0 ms ut i filen, går til 180 ms og varer i 180 ms. Neste hendelse er et blunk som starter 160 ms ut i filen, varer til 380 ms som gir en varighet på 220 ms. Tredje hendelse er en fiksering fra 560 ms til 640 ms med varighet på 80 ms. Den starter på piksel 623.173 i X-retning og 335.087 i Y-retning. For denne fikseringen er det en spredning på 22 piksler i X-retning og 72 i Y-retning.

Videre er det fra denne fikseringen en sakkade som starter etter 640 ms og varer til 1020 ms og derved har en lengde på 380 ms. Den starter i pikselposisjon $X=630.854$ $Y=335.271$, altså innenfor foregående fiksering ($X \approx 623 \pm 22$, $Y \approx 335 \pm 72$), og ender opp i $X=585.103$ $Y=251.383$ som er innenfor neste fiksering ($X \approx 613 \pm 72$, $Y \approx 266 \pm 23$).

Listing 3.2: Eksempel på datafil

1	Type	Set	Start	End	Duration					
2	Blink R	1	0	180	180					
3	Blink R	1	160	380	220					
4	Fixation R	1	560	640	80	623.173	335.087	22	72	
5	Saccade R	1	640	1020	380	630.854	335.271	585.103	251.383	41.335
			159.048	0.421	108.762	2442.2	-4229.4	1835.7		
6	Fixation R	1	1020	1160	140	613.953	266.634	72	23	

Legg merke til at en fiksering eller sakkade har starttidspunkt likt endetidspunkt på foregående sakkade eller fiksering, mens et blunk kan komme midt i en sakkade eller fiksjon.

Listing 3.3 viser et eksempel på UserEvent. Ved spesielle hendelser i stimulamateri-

alet sender javaprogrammet ETmessage en melding til iView X (se avsnitt 3.4). Denne meldingen blir logget i datafilen som en UserEvent med en melding, og sett-nummeret blir økt. Som vist i listing 3.3 kan en slik hendelse logges midt i en sakkade eller fiksering.

Listing 3.3: Eksempel på datafil

```

1 Fixation R      1      4600    4700    100      442.204 312.348 79      17
2 Saccade R      1      4700    4760    60      451.256 315.481 476.876 309.855 5.329
   122.970 1.000 88.869 2496.8 -2110.4 1770.0
3 UserEvent     2      4744    # Message: Melding
4 Fixation R      2      4760    4940    180      468.098 318.586 43      49

```

	Fixations:	Saccades:	Blinks:	User Events:
1	Event Type	Event Type	Event Type	Event Type
2	Set	Set	Set	Set
3	Start	Start	Start	Start
4	End	End	End	Description
5	Duration	Duration	Duration	
6	Location X	Start Loc.X		
7	Location Y	Start Loc.Y		
8	Dispersion X	End Loc.X		
9	Dispersion Y	End Loc.Y		
10		Amplitude		
11		Peak Speed		
12		Peak Speed At		
13		Average Speed		
14		Peak Accel.		
15		Peak Decel.		
16		Average Accel.		

Tabell 1: Beskrivelse av kolonnene i datafilene

3.9 BeGaze

SMI har utviklet et analyseverktøy for eyetracker data som kalles BeGaze. I BeGaze er det relativt enkelt å hente ut heatmaps, scanpaths og andre former for visuelle bedømningsverktøy. BeGaze gir dessverre ikke resultater i form av talldata i særlig stor grad, så for å gjøre analyser på datamaterialet var det nødvendig å håndtere datafilene i andre verktøy, først og fremst MATLAB og SPSS

3.10 Innlesing av dataene i MATLAB

Den første håndteringen av dataene ble foretatt i MathWorks MATLAB. MATLAB er et godt egnet verktøy for å gjøre beregninger på et stort antall data slik som er tilfelle for datafilene fra eyetrackeren [44]. Datafilene beskrevet i avsnitt 3.8 ble lest inn i MATLAB ved hjelp av funksjonen ReadFiles i scriptet ETanalyze som kan sees i vedlegg A.

I MATLAB ble dataene lagt inn i en struct som videre kunne behandles på ulike måter for å hente ut igjen de resultatene som er interessante. For eksempel kunne antallet fikseringer i hvert AOI enkelt hentes ut fra MATLAB til Microsoft Excel eller SPSS for videre analyse.

3.11 Behandling av data i SPSS

Fra MATLAB ble det hentet ut data for videre analyse i SPSS. SPSS er et statistikk og datahåndteringsverktøy med de analyseverktøyene som var nødvendig for dette prosjek-

tet [45]. Dataene ble filtrert etter reglene i 3.12. Andre variable nødvendig for analysene ble beregnet. Alle disse omregningene finnes i syntaks-filen i vedlegg A.

3.12 Forkasting av deltakere

Av ulike årsaker er det ikke alle datasettene som gir korrekte eller tilstrekkelige data. Dette kan skyldes for eksempel dårlig kalibrering. I dette prosjektet er følgende satt som kriterie for at datasettet skal aksepteres

1. Tid medgått til fikseringer må være over 10% av total tid
2. Antall fikseringer må være over 5
3. Tid medgått til sakkader må være over 1% av total tid
4. Antall sakkader må være over 5

Alle kriteriene må være tilfredsstillt for at datasettet regnes som brukbart. I tillegg må lengden på sakkadevektoren overstige 5 for at datasettet skal kunne tas med i betraktningen i beregninger hvor sakkadevektoren blir brukt. Terskelverdien er satt ut fra at det antas at under normale omstendigheter har en person mer enn 5 fikseringer og 5 sakkader i løpet av to minutter (det korteste settet). På samme måte bør det være mulig å finne fikseringer for mer enn 14 sekunder eller sakkader for mer enn 1,4 sekunder i en sekvens på to minutter. Terskelverdier på henholdsvis 10% og 1% er satt ut fra en visuell bedømmning av materialet. Tersklene kan virke noe lave, men de datasettene som er over terskelverdiene er relativt mye over. Terskelverdiene er satt som et minimum.

4 Resultat

4.1 Generelt

Dersom det ikke er oppgitt noe annet så er det brukt et statistisk signifikansnivå på 95% ($\rho = 0.050$) for alle resultatene. For det meste er statistikkprogrammet SPSS benyttet for å beregne resultatene.

For alle settene er det samme gruppe deltakere. Det gjør at det i stor grad er innengruppe variansanalyse (Whitin-subjects ANOVA) som har blitt benyttet for å hente ut resultatene. Etter modell beskrevet i Langdridge [46] har først en Mauchlys test blitt gjennomført for å sjekke for korrelasjon mellom settene. Mauchlys test har vist seg ikke å være signifikant, slik at signifikansen oppgitt for innengruppe testene baserer seg på «Greenhouse-Geisser»-resultatene. Videre er signifikansen for de tre vilkårene funnet gjennom en paret t-test med Bonferroni-korreksjon.

4.2 Sammensetning av deltakere

Eksperimentet ble gjennomført med 21 deltakere. Disse ble rekruttert i hovedsak blant studenter ved Høgskolen i Gjøvik. Dette reflekteres i alder og kjønnsfordelingen i gruppen. Samtlige deltakere deltok frivillig og uten betaling. En del av deltakerne ble forkastet etter vilkårene gitt i avsnitt 3.12. Som man kan se av tabell 2 må nærmere 50% av deltakerne forkastes fra sett 1.

	Inkludert		Ekskludert	
	Antall	Prosent	Antall	Prosent
Sett 1	11	52,4	10	47,6
Sett 2	13	61,9	8	38,1
Sett 3	14	66,7	7	33,3
Ett sett	16	76,2	5	23,8
Alle sett	10	47,6	11	52,4

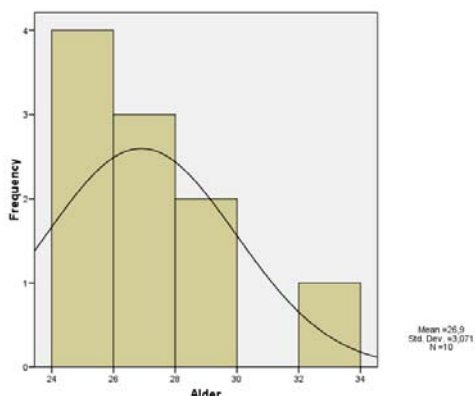
Tabell 2: Inkluderte og ekskluderte deltakere

Det er de ti deltakerne som hadde tilfredsstillende data for alle settene som har blitt brukt videre i analysen. Denne gruppen har en aldersfordeling mellom 24 og 34 år hvor mennene er mellom 24 og 34 og kvinnene mellom 26 og 29. Med 7 menn og 3 kvinner blir gjennomsnittsalderen for gruppen 26.9 år, 26.71 for menn og 27.33 for kvinner (se figur 7 og 8).

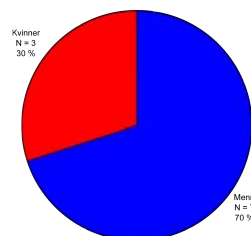
4.3 Resultater fra spørreskjema

Hensikten med spørreskjemaet var å kunne dele inn deltakerne i ulike grupper og kontrollere dataene for ulike variabler. Med resultater fra bare 10 deltakere vil gruppene bli for små til at man kan si noe sikkert om man deler de inn i enda mindre grupper. Alle 21 deltakerne svarte på spørreskjemaet. Under er noen av resultatene

- 10 har eller er i ferd med å ta en bachelorgrad, 11 en mastergrad



Figur 7: Histogram – aldersfordeling i deltakergruppen



Figur 8: Kjønnfordeling i deltakergruppen

- 7 hadde tidligere deltatt i eksperiment med eyetracking mens 14 var ukjent med eyetracking
- 11 viste formålet med eksperimentet før de deltok
- 8 kjente programkonseptet World Exchange fra før mens alle kjente til Idol.
- 2 svarte at de aldri så på TV, 10 så 1-2 timer mens de resterende 9 så 3-5 timer hver dag.
- Til tross for at 5 svarte at de så på sms-tv var det ingen som svarte at de pleide å delta på sms-tv.
- 12 deltakere synes at World Exchange var interessant eller i liten grad interessant mens 9 synes ikke det var interessant.

Ett resultat å legge merke til er at det til tross for at det var over ett år siden programmene ble sendt så var det 6 stykker som kunne huske å ha sett ett eller flere av Idol-programmene.

Spørreskjemaet ligger vedlagt som vedleggB og alle resultatene kan ses i vedlegg C

4.4 Antall fikseringer, sakkader og blunk

For å finne et sammenlignbart forholdstall har antallet fikseringer, sakkader og blunk blitt delt på den totale tiden for settet omregnet i sekund (avsnitt 3.1.4). Det gir et forholdstall som gir antall per sekund for de ulike målevariablene. Tabell 3 viser gjennomsnitt, standardavvik, minimum og maksimum for alle tre settene (10 deltakere).

Det synes å være individuelt hvor mange fikseringer, sakkader og blunk hver deltaker har, og en parett-test viser ingen signifikant sammenheng mellom endret grafikk og antall ($p > 0,5$).

4.5 Fiksasjonstid, sakkadetid og blunktid

To mål for tid har blitt brukt:

- Gjennomsnittlig tid per fiksering, sakkade og blunk (fikseringstid delt på antall fikseringer)

		Gjennomsnitt	Std. Avvik	Minimum	Maksimum
Fikseringer	Sett 1	2,226	0,910	0,490	3,680
	Sett 2	2,302	0,691	1,585	3,607
	Sett 3	2,291	0,682	1,384	3,564
Sakkader	Sett 1	1,794	0,960	0,313	3,133
	Sett 2	1,789	0,773	1,020	3,069
	Sett 3	1,748	0,767	0,498	2,953
Blunk	Sett 1	1,083	0,877	0,135	2,983
	Sett 2	0,968	0,524	0,156	1,926
	Sett 3	0,977	0,567	0,292	1,767

Tabell 3: Antall fikseringer, sakkader og blunk

- Fikserings-, sakkade- og blunktid relativt til total tid

Dette gir to forskjellige målevariabler på henholdsvis hvor lang hver fiksering, sakkade og blunk er og hvor stor del av den totale tiden deltakeren har brukt til å fikse, flytte blikket eller blunke. Tabellene 4 og 5 viser gjennomsnitt, standardavvik, minimum og maksimum av de ulike målevariablene for de tre settene (antall deltakere = 10).

Det synes å være individuelt hvor lange fikseringer, sakkader og blunk hver deltaker har, og en paret t-test viser ingen signifikant sammenheng mellom endret grafikk og lengden i tid ($\rho > 0,5$).

		Gjennomsnitt	Std. Avvik	Minimum	Maksimum
Fikseringer	Sett 1	265,38	114,48	151,05	501,60
	Sett 2	251,70	97,41	157,69	423,51
	Sett 3	278,43	96,92	178,36	477,38
Sakkader	Sett 1	55,14	30,11	24,91	110,96
	Sett 2	59,83	38,67	25,78	145,37
	Sett 3	51,54	31,24	27,29	133,57
Blunk	Sett 1	272,07	30,23	225,00	329,01
	Sett 2	259,71	36,52	210,25	324,44
	Sett 3	262,30	33,79	218,05	304,39

Tabell 4: Gjennomsnittlig tid per fiksering, sakkade og blunk i millisekund

		Gjennomsnitt	Std. Avvik	Minimum	Maksimum
Fikseringer	Sett 1	57 %	26 %	10 %	91 %
	Sett 2	57 %	22 %	27 %	91 %
	Sett 3	61 %	20 %	32 %	86 %
Sakkader	Sett 1	9 %	6 %	2 %	18 %
	Sett 2	10 %	5 %	3 %	17 %
	Sett 3	9 %	5 %	1 %	15 %
Blunk	Sett 1	30 %	24 %	3 %	81 %
	Sett 2	25 %	14 %	4 %	49 %
	Sett 3	26 %	16 %	6 %	54 %

Tabell 5: Fikserings- sakkade- og blunktid relativt til total tid i prosent

4.6 Sakkadeamplitude

En måte å representere lengden på en sakkade er amplituden. Den er oppgitt i antall grader og representerer den gjennomsnittlige buelengden øyet har beveget seg i fra ett fikseringspunkt til det neste. Amplituden kan regnes ut på følgende måte [26]

Amplitude = gjennomsnittlig vinkelhastighet \times varighet

Se avsnitt 2.5 for en detaljert beskrivelse. Tabell 6 viser gjennomsnitt, standardavvik, minimum og maksimum for alle tre settene (10 deltakere). Legg merke til at standardavviket er tilnærmet like stort som gjennomsnittet og at spennet mellom minste og største er stort. Det synes å være individuelt hvor lange sakkader hver deltaker har, og en parret t-test viser ingen signifikant sammenheng mellom endret grafikk og antall ($\rho > 0,5$).

	Gjennomsnitt	Std. avvik	Minimum	Makismum
Sett 1	5,41	3,86	1,41	11,74
Sett 2	6,13	5,41	1,76	18,28
Sett 3	5,25	4,76	2,07	18,13

Tabell 6: Amplitude i grader

4.7 Sakkadevektor

For hver sakkade er det et startpunkt og et endepunkt. Ved å beregne euklidsk avstand (formel 4.1) mellom disse punktene får vi en vektor som beskriver hvor lang denne sakkaden er. En totalvektor kan beregnes ved å summere disse vektorene. Videre har gjennomsnittsvektoren for hver deltaker og hvert sett blitt beregnet ved å ta denne totalvektoren og dele på antallet sakkader. Formålet med denne vektoren er å danne et mål på hvor langt deltakeren flytter blikket under forskjellige vilkår.

$$v = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \quad (4.1)$$

Se avsnitt 2.5 for en detaljert beskrivelse. Tabell 7 viser gjennomsnitt, standardavvik, maks og min for de tre settene (10 deltakere). Vi kan se at den gjennomsnittlige lengden øker med mer grafikk. Forskjellen mellom settene er signifikant ($F = 7.2$, $df = 1.6$, $\rho = 0.007$). Effekten er også grafisk fremstilt i figur 9.

	Gjennomsnitt	Std. avvik	Minimum	Makismum
Sett 1	70,68	13,25	51,09	97,66
Sett 2	82,73	13,75	62,71	108,65
Sett 3	94,70	18,69	72,60	137,05

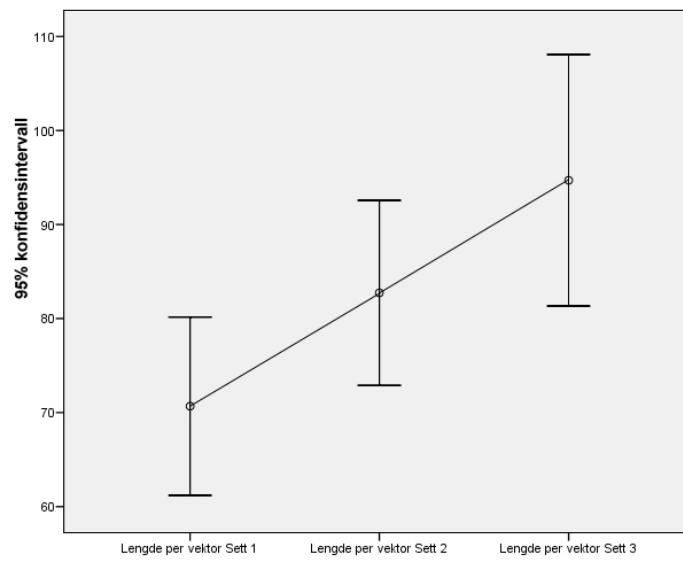
Tabell 7: Gjennomsnittlig vektorlengde for sakkader

Tre parrede t-testere ble utført med Bonferroni-korreksjon, og de viser en signifikant forskjell mellom to av de tre vilkårene, sett 1 og 3 og sett 2 og 3. Det er ingen signifikant forskjell mellom sett 1 og 2.

Det sier oss at deltakerne flytter blikket lenger per sakkade når det er mer grafikk i bildet. Det er en forskjell mellom CNBC-vilkåret og de to Idol-vilkårene, men ikke mellom de ulike Idol-vilkårene.

	t	df	Signifikans	Bonferroni-korrigert signifikans
Sett 1 – Sett 2	-0,755	9	0,465	1,395
Sett 1 – Sett 3	-3,074	9	0,010	0,029
Sett 2 – Sett 3	-3,806	9	0,003	0,008

Tabell 8: Parret t-test



Figur 9: Estimert gjennomsnitt for sakkadevektoren

4.8 Fikseringsfrekvens

Bildet ble delt opp i ulike områder av interesse, AOIer, etter modellen beskrevet i avsnitt 3.2 på side 10. Disse områdene gir muligheten til å telle hvor mange fikseringer det har vært på henholdsvis grafikken og hovedbildet. For å jevne ut forskjellen i lengde på settene er antall fikseringer på ett område delt på totalantallet fikseringer for å gi en prosentverdi innen hvert område. Det er denne prosentverdien som har blitt brukt i sammenligningene.

Tabell 9 viser gjennomsnitt, standardavvik, minimum og maksimum for de tre settene og de to områdene bilde og grafikk. Antall deltakere her er 10. Vi ser at antallet fikseringer som er innenfor området grafikk minker med økende grafikk. Forskjellen mellom settene er signifikant ($F = 19.2$, $df = 1.7$, $\rho = 0.001$).

Tabell 10 viser resultatene fra tre parrede t-tester på de ulike vilkårene. Vi ser at det er ingen signifikant sammenheng mellom sett 1 og 2, men mellom sett 1 og 3 og sett 2 og 3 er forskjellen signifikant.

Det sier oss at deltakerne har flere fokuseringer på grafikken når det er mindre grafikk i bildet. Det er en forskjell mellom CNBC-vilkåret og de to Idol-vilkårene, men ikke mellom de ulike Idol-vilkårene.

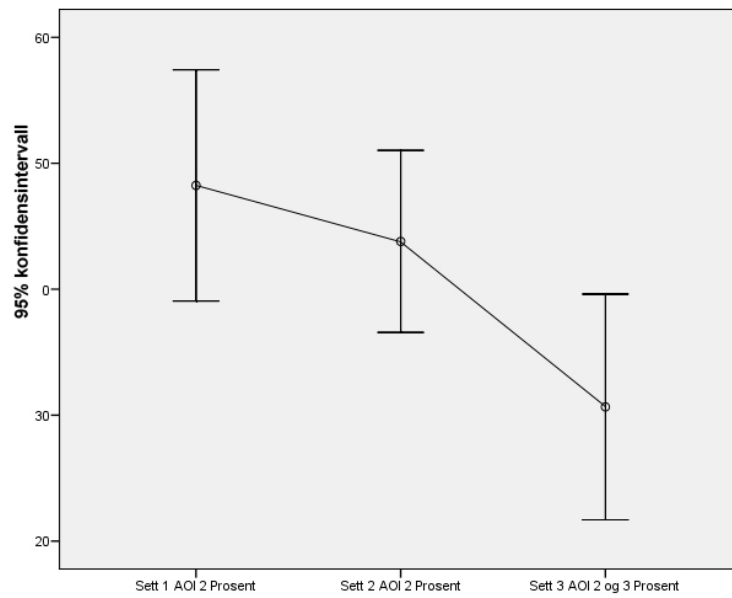
Legg merke til at dette er rent tallmessige resultater. I vedlegg D er en visuell analyse av mulige årsaker til fordelingen av fikseringer presentert.

		Gjennomsnitt	Std. avvik	Minimum	Makismum
Bilde (AOI1)	Sett 1	51,8 %	12,83	35,2 %	76,9 %
	Sett 2	56,2 %	10,10	42,4 %	72,1 %
	Sett 3	69,3 %	12,53	42,6 %	83,5 %
Grafikk (AOI2+3)	Sett 1	48,2 %	12,83	23,1 %	64,8 %
	Sett 2	43,8 %	10,10	27,9 %	57,6 %
	Sett 3	30,7 %	12,53	16,5 %	57,4 %

Tabell 9: Gjennomsnitt fikseringer per AOI

	t	df	Signifikans	Bonferroni-korrigert signifikans
Sett 1 – Sett 2	1,585	9	0,147	0,442
Sett 1 – Sett 3	4,965	9	0,001	0,002
Sett 2 – Sett 3	5,498	9	0,001	0,001

Tabell 10: Parret t-test



Figur 10: Estimert gjennomsnitt for frekvensfordelingen

5 Diskusjon

5.1 Forkasting av deltakere

Det er en relativt stor prosentandel av deltakerne i eksperimentet som har blitt forkastet. Det har ikke lyktes å kunne forklare nøyaktig hvorfor. De faktorene som er gitt i utvelgelse av deltakere når det gjelder briller, linser, mørke pupiller osv (se avsnitt 2.6 side 7) ser ikke ut til å ha noen innvirkning på hvilke deltakere som har dårlig kalibrering eller lav mengde data. Blant annet er en deltaker med linser blant de som har gitt mest eksakte og sammenhengende data. Det som derimot har vist helt tydelig å ha en innvirkning er øyesminke. Det virker til at øyenskygge, mascara og lignende har IR-reflekterende egenskaper slik at eyetrackeren tolker refleksjonen fra sminken som refleksjon fra øyet.

De tre settene er vist i forskjellig rekkefølge for de ulike deltakerne. Dataene er kontrollert for rekkefølgeeffekter uten at det kunne påvises noen effekt for de ulike målvariablene. Det er derfor ingen grunn til å tro at det skulle være det i forkastingen av deltakere.

En mulig forklaring på at det må forkastes flere deltakere fra første enn fra andre settet og flere fra andre enn tredje kan være hvor lyst settet er. Det første settet er veldig mørkt, det er mørk bakgrunn og dempet lys gjennom store deler av settet. Det andre settet har en lysere bakgrunn enn det første. Det går i stor grad i mørke blåfarger, men deler er filmet ute blant publikum hvor det er dempet belysning. Det siste settet er filmet i studio med god og jevn belysning over hele skjermen til en hver tid. For å danne et mål på hvor lyst bildet er har et gjennomsnitt av pikselverdiene gjennom hele settet blitt beregnet som vist i formel 5.1. R, G og B tilsvarer de tre fargekanalene. D er dimensjonen på bildet og F er frame (ramme). Resultatet blir en tallverdi som representerer gjennomsnittlig pikselverdi mellom 0 og 255 for hele settet. Lavere tall betyr mørkere bilde. Tabell 11 viser verdien for de tre settene. Det er en tydelig forskjell i verdien for de tre settene.

Sett	Verdi
Sett 1	37,33
Sett 2	54,96
Sett 3	68,25

Tabell 11: Gjennomsnittlig pikselverdi for settet

$$V = \frac{\sum_{F=0}^F \frac{\sum_{D=0}^D \frac{R + G + B}{3}}{D}}{F} \quad (5.1)$$

Grunnen til at dette kan være av betydning er at eyetrackeren har behov for en viss mengde lys for å kunne følge øyet og fungere på en tilfredsstillende måte. Dersom skjermen avgir for lite lys er det mulig at den dempede belysningen i rommet var for liten til at eyetrackeren greide å følge øyet. Som nevnt i avsnitt 3.3 ble lyset i rommet og lyset fra skjermen målt før eksperimentet, men da med et statisk bilde på skjermen og ikke et gjennomsnittlig lysmå. Det har ikke vært tilgjengelig teknisk utstyr for å måle det fak-

tiske lyset avgitt fra skjermen under avspilling av settene. Dette blir derfor en spekulasjon på en mulig årsak til problemet.

5.2 Tid og antall

Antallet sakkader, fikseringer og blunk og den gjennomsnittlige tiden av disse er som forventet. De tilsvarer det andre har funnet for disse variablene i andre forsøk. Det er heller ingen forskjell mellom de tre settene.

En typisk fiksering skal være 200-400 millisekund [20, 10]. Resultatene funnet i dette eksperimentet er innenfor disse verdiene, 265 ms i gjennomsnitt.

Vanligvis gjør et menneske 2-4 sakkader og 2-3 fikseringer per sekund [28] [47]. I dette eksperimentet er gjennomsnittet 1,7 sakkader og 2,3 fikseringer per sekund, altså noe under det andre har funnet. Årsaken til dette kan være at det ikke er kompensert for manglende data. Antallet er beregnet ut fra totaltid

Zangemeister m.fl. [24] bruker blunkfrekvensen som en måte å måle aktivitet. Visuelt mer krevende bilder gjør at blunkfrekvensen blir lavere. Denne effekten har ikke vært synlig i dette eksperimentet.

5.3 Amplitude og vektorlengde

Amplituden og sakkadevektoren er to metoder for å måle det samme, lengden på en sakkade. Samtidig måler de to metodene forskjellige ting. Amplituden forsøker å beskrive hvor langt blikket har beveget seg totalt sett med alle avvik fra en rett linje, mens vektoren kun er ment for å beskrive distansen mellom to punkter og derved hvor langt blikket har flyttet seg på materialet i løpet av sakkaden.

Figur 11 viser en graf for amplituden og vektoren hvor vektoren er skalert ned slik at de kan sammenlignes visuelt. Man kan se at disse to kurvene følger hverandre. Der hvor sakkadevektoren er lengre er også amplituden tilsynelatende større. Dette kan virke logisk ut fra at øyet vil ha en lengre bane utenom den rette linjen når blikket forflyttes lenger. Denne sammenhengen er synlig i et plott slik som i figur 11, men det holder ikke i en korrelasjonstest. Sammenhengen er ikke signifikant.

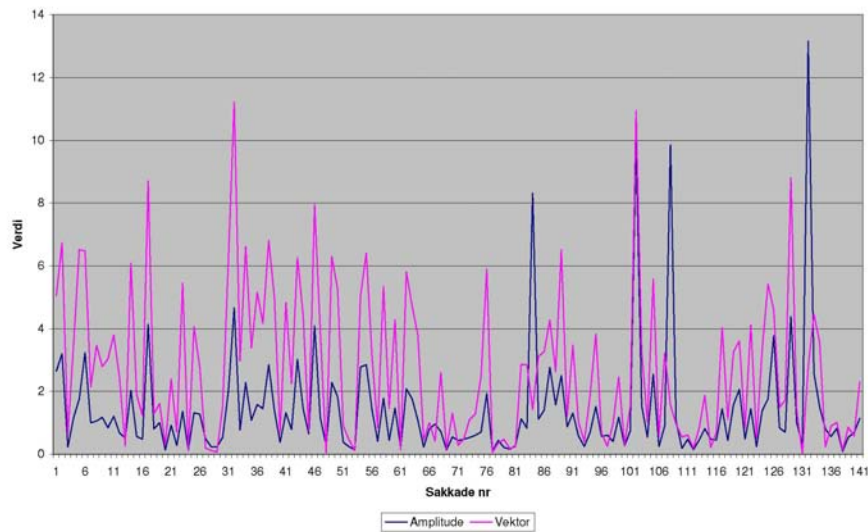
5.4 Fikseringsfrekvens

Som avsnitt 4.8 viser så er det en endring i hvor mange fikseringer deltakerne i gjennomsnitt gjør innenfor områdene med grafikk i bildet. Det som er noe overraskende er at det er færre fikseringer på grafikken i de mer kompliserte settene. Første antakelse var at det ville være flere fikseringer på grafikken når det ble mer kompleks grafikk, altså at deltakeren ville feste blikket på grafikken for å lese informasjonen der.

I vedlegg D er det gjort en visuell analyse av stimulumaterialet opp mot fikseringsfrekvens for å forsøke å avdekke en årsakssammenheng. Denne subjektive analysen peker i retning av to hypoteser

- I settene med mer grafikk er det også mer som foregår i selve programmet. Derfor er det også mer som trekker oppmerksomheten bort fra grafikken.
- Informasjonen i området med grafikk er mindre interessant å følge med på i det mest komplekse settet.

Det er tendenser som tyder på at det kan være en uheldig produksjon som gjør at første settet fikk stort utslag i fikseringer på området med grafikk. Det er deler av det vesentlige



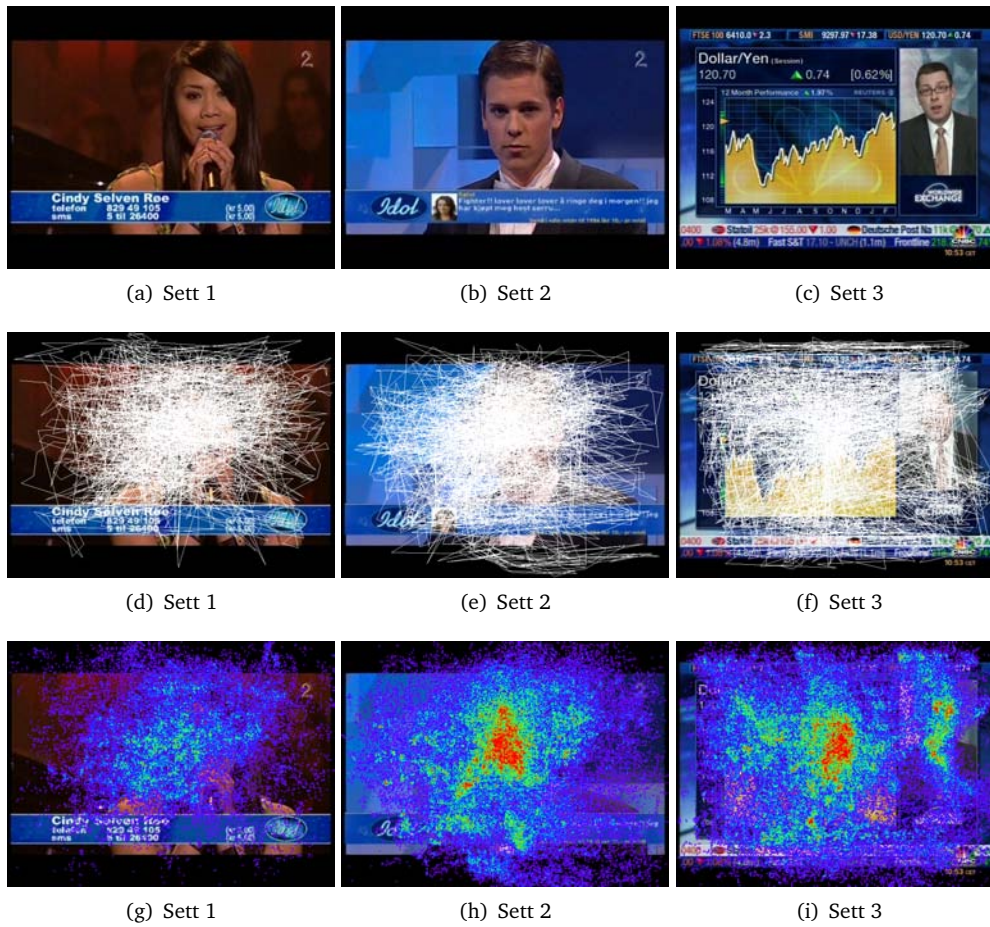
Figur 11: Vektor og amplitude for deltaker 15 sett 1

innholdet i programmet som havner bak grafikken og derfor i en tallmessig analyse oppfattes som fikseringer på grafikken til tross for at det deltakeren forsøker å gjøre er å se «gjennom» grafikken. Det samme kan være tilfelle for sett 2 som også er en «live» produksjon og derved ikke er helt og holdent kontrollert for hva som er hvor i bildet til en hver tid. Sett 3 er en studioproduksjon og er mer strukturert, det er derfor ikke noe som kan ende opp bak grafikken.

5.5 Scanpath og heatmap

Figur 12 viser scanpath for de første 2 minuttene og 20 sekundene for alle tre settene, alle 10 deltakere sammenslått. Disse illustrerer samme effekt som ble vist i avsnitt 4.7. I sett 1 er de aller fleste sakkadene samlet rundt senter i bildet. Det er få sakkader som beveger seg ut i hjørnene. For sett 3 er det sakkader som strekker seg ut i hjørnene. Blikket har beveget seg over et større område av skjermen.

Dette underbygges også av heatmapene i figur 12(g)-12(i). I sett 1 er fokuseringspunktene i stor grad samlet rundt sentrum i bildet. Dette er også tydelig i sett 2 (figur 12(h)) hvor det er en tydelig konsentrasjon midt i bildet, men også ned mot grafikken. I sett 3 er det også en konsentrasjon midt i bildet, men det er en betydelig større spredning ut mot kantene. Blikket har beveget seg over et større område av bildet. Deler av denne spredningen har sin naturlige forklaring i at under store deler av settet er ansiktet til personen som snakker plassert ute på høyre side slik som vist i 12(c). Som omtalt i avsnitt 2.1 tiltrekkes vi naturlig av ansikter. Det samme gjelder for de to andre settene, men her er ansiktene som regel plassert nokså sentrert i bildet.



Figur 12: Eksempel fra stimulumaterialet(a-c) med scanpath(d-f) og heatmap(g-i)

6 Konklusjon

Det ble presentert to forskningsspørsmål i avsnitt 1.2:

Spørsmål 1: Er eyetracking egnet for studier av seeratferd?

Spørsmål 2: Påvirkes seeratferden av mengden grafikk på TV?

Oppgaven har undersøkt hvordan seeratferd kan måles med eyetracking ved å gjennomføre et eksperiment hvor 21 deltakere så tre ulike klipp fra forskjellige TV-program. Det resulterte i 30 valide datafiler, 10 for hvert sett. Resultatene fra eyetrackingen har blitt kontrollert for ulike parametere for å avdekke hva som gir et godt mål på hvordan deltakerne søker informasjon i bildet.

En vektor som beskriver lengden på sakkaden har blitt avdekket som et mål på seeratferd. I settet med mest dynamisk grafikk flyttet deltakerne blikket lenger for hver sakkade enn de gjorde for settet med statisk grafikk. Det indikerer at ved mer kompleksitet i bildet søker deltakerne å flytte blikket lenger for hver sakkade.

Eyetracking har vist seg som et godt verktøy for dette prosjektet hvor det var ønskelig med en objektiv måling av persepsjon. Det har vært et stort frafall av deltakere og det er en viktig faktor i praktisk bruk. Dette prosjektet har vist at det er flere faktorer som er viktig for å gjennomføre et vellykket eksperiment. Et godt forarbeide med tilpassing av miljøet rundt, riktig utvelgelse av stimulumateriale og selektiv rekruttering av deltakere er viktige faktorer for å få pålitelige resultater.

7 Videre arbeid

Gjennom arbeidet med oppgaven har det fremkommet flere områder som egner seg for videre forskning.

Dette prosjektet startet som et eksperiment for å avdekke effekter av ulik mengde og type grafikk i TV-produksjoner, med eyetracking som et verktøy for å undersøke effekten. Gjennom prosjektperioden har det dreid mer over i retning av å bli en evaluering av eyetracking i bruk på forskning innen TV. Det gjør at det fortsatt gjenstår åpne spørsmål angående hvordan seeren påvirkes av ulike typer grafikk.

Et nytt eksperiment på de samme effektene kan med fordel utføres. En idé ville være å bruke en mer rendyrket produksjon med tydelige skifter mellom ulike effekter. Det at det ble brukt «live»-produksjoner som stimulimateriale i dette prosjektet har gitt en ekstra utfordring og usikkerhet i analysen.

Alle settene som har vært brukt i dette eksperimentet har hatt originale lydspor. Det har ikke blitt gjort noen analyse av hvordan endringer i lydbildet påvirker hvordan seeren forflytter blikket. Et stimulimateriale med kontrollerte effekter i lyden vil kunne gi en indikasjon på om blikket styres ut fra hvilken type lyd deltakeren hører.

Bibliografi

- [1] Television in digital enviroment. Mai 2007. <http://tide.hil.no>.
- [2] Holmqvist, K. & Gullberg, M. 1998. Focus on gesture. *Lund University Cognitive Studies 70*.
- [3] Ware, C. & Mikaelian, H. H. 1987. An evaluation of an eye tracker as a device for computer input2. In *CHI '87: Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, 183–188, New York, NY, USA. ACM Press.
- [4] Babcock, J. S., Lipps, M., & Pelz, J. B. 2002. How people look at pictures before, during and after scene capture: Buswell revisited. In *Human Vision and Electronic Imaging V, SPIE Proceedings*, B.E.Rogowitz & Pappas, T. N., eds, 34–47.
- [5] Lundqvist, D. & Holmqvist, K. 2004. Bigger is better: How size of newspaper advertisement and reader attitude relate to attention and memory. *Submitted*.
- [6] Stark, P. & Garcia, M. *Poynter Eye Tracking Study*, chapter Eye on the news. Poynter, 1990.
- [7] Radach, R., Lemmer, S., Vorstius, C., Heller, D., & Radach, K. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, chapter Eye movements in the processing of print advertisements, 609–632. Elsevier, Amsterdam, 2003.
- [8] Borell, J. Subtitling or dubbing? an investigation of the effects from reading subtitles on understanding audiovisual material. Master's thesis, Lund University, 2000.
- [9] Bruycker, W. D. & d'Ydewalle, G. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, chapter Reading Native and Foreign Language Television Subtitles in Children and Adults, 671–684. Elsevier, Amsterdam, 2003.
- [10] Rayner, K. November 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- [11] Aoki, H. & Itoh, K. 2000. Analysis of cognitive attitudes to commercial films on basis of eye tracking data. *Proceedings of the human factors and ergonomics society annual meeting*, 1, 38–41.
- [12] Aoki, H. & Itoh, K., eds. *Eye tracking analysis for unconscious process of information gathering from television commercials*. 33rd EMAC Conference, Murcia, Spain, May 2004.
- [13] Jarbus, A. L. 1967. *Eye movements and vision*. Plenum Press, New York.
- [14] Godijn, R. & Theeuwes, J. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, chapter The Relationship Between Exogenous and Endogenous Saccades and Attention, 3–26. Elsevier, Amsterdam, 2003.

- [15] Daly, S., Matthews, K., & Ribas-Corbera, J. 1998. Face-based visually-optimized image sequence coding. 443–447 vol.3.
- [16] Bradley, A. P. & Stentiford, F. W. 2003. Visual attention for region of interest coding in jpeg 2000. *Journal of Visual Communication & Image representation*, 14, 232 – 250.
- [17] Endo, C., Asada, T., Haneishi, H., & Miyake, Y. 1994. Analysis of the eye movements and its applications to image evaluation. In *IS&T and SID's 2nd Color Imaging Conference: Color Science, Systems and Applications*, 153–155.
- [18] Duchowski, A. T. 2003. *Eye tracking methodology: theory and practice*. Springer, London, 1 edition.
- [19] Irwin, D. E. *Eye movements and visual cognition: scene perception and reading*, chapter Visual memory within and across fixations, 146–165. Springer Series in Neuropsychology. Springer, New York, 1992.
- [20] Salvucci, D. D. & Goldberg, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the Eye Tracking Research and Applications Symposium*, 71–78. New York: ACM Press.
- [21] SMI SensoMotoric Instruments. 2005. BeGaze 1.2 user manual, document version 1.06.03.
- [22] Byrne, M. D., Anderson, J. R., Douglass, S., & Matessa, M. 1999. Eye tracking the visual search of click-down menus. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, 402–409, New York, NY, USA. ACM Press.
- [23] Widdel, H. *Theoretical and Applied Aspects of Eye Movement Research*, chapter Operational problems in analysing eye movements, 22–29. Elsevier, New York, 1984.
- [24] Zangemeister, W. H., Sherman, K., & Stark, L. W. 1995. Evidence for a global scan-path strategy in viewing abstract compared with realistic images. *Neuropsychologia*, 33(8), 1009–1025.
- [25] Henderson, J. M., Williams, C. C., Castelhana, M. S., & Falk, R. J. 2003. Eye movements and picture processing during recognition. *Perception & Psychophysics*, 65, 725–734(10).
- [26] SMI BeGaze event detection ver 1.0. Technical report, SMI SensoMotoric Instruments, March 2007.
- [27] Wikipedia.org - the free encyclopedia. Juni 2007. http://en.wikipedia.org/wiki/angular_velocity.
- [28] Privitera, C. M. & Stark, L. W. 2000. Algorithms for defining visual regions-of-interest: comparison with eye fixations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(9), 970–982.
- [29] Underwood, G. & Foulsham, T. 2006. Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *The Quarterly Journal of Experimental Psychology*, 59(11), 1931–1949.

- [30] Crowe, E. C. & Narayanan, N. H. 2000. Comparing interfaces based on what users watch and do. In *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, 29–36, New York, NY, USA. ACM Press.
- [31] Goldberg, J. & Wichansky, A. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, chapter Eye tracking in Usability Evaluation: A Practitioner's Guide, 493–517. Elsevier, Amsterdam, 2003.
- [32] Nguyen, K., Wagner, C., Koons, D., & Flickner, M. 2002. Differences in the infrared bright pupil response of human eyes. In *ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications*, 133–138, New York, NY, USA. ACM Press.
- [33] Kjus, Y. 2006. Når publikum blir produsenter. Deltakelse i det kommersielle underholdningsformatet Idol. *Norsk Medietidsskrift*, (3).
- [34] Kitamura, Y., Horii, K., Takeuchi, O., Kotani, K., & d'Ydewalle, G. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, chapter Determining the parameters for the scrolling text display technique, 645–656. Elsevier, Amsterdam, 2003.
- [35] SMI SensoMotoric Instruments. Mai 2007. <http://www.smi.de/>.
- [36] SMI SensoMotoric Instruments. 2005. iView X system manual, document version 1.05.09.
- [37] Sekonic Light Meters. Juni 2007. <http://www.sekonic.com>.
- [38] Conrad, J. August 2003. Exposure metering relating subject lighting to film exposure. 16/3/2007: <http://www.largeformatphotography.info/articles/conrad-meter-cal.pdf>.
- [39] Kerr, D. A. Juli 2006. Apex - the additive system of photographic exposure. <http://doug.kerr.home.att.net/pumpkin/APEX.pdf>. Issue 5.
- [40] Porac, C. & Coren, S. 1976. The dominant eye. *Psychological Bulletin*, 83(5), 880–897.
- [41] Roth, H. L., Lora, A. N., & Heilman, K. M. September 2002. Effects of monocular viewing and eye dominance on spatial attention. *Brain*, 125(9), 2023–2035.
- [42] Babcock, J. S. Eye tracking observers during color image evaluation tasks. Master's thesis, Rochester Institute of Technology, 2000.
- [43] Vuori, T., Olkkonen, M., Pölonen, M., Siren, A., & Häkkinen, J. 2004. Can eye movements be quantitatively applied to image quality studies? In *NordiCHI '04: Proceedings of the third Nordic conference on Human-computer interaction*, 335–338, New York, NY, USA. ACM Press.
- [44] The Mathworks. Mai 2007. <http://www.mathworks.com>.
- [45] SPSS inc. Mai 2007. <http://www.spss.com>.

- [46] Langdridge, D. & Tvedt, S. D. 2006. *Psykologisk forskningsmetode: en innføring i kvalitative og kvantitative tilnærminger*. Tapir, Trondheim.
- [47] Jaimes, A., Pelz, J. B., Grabowski, T., Babcock, J. S., & Chang, S.-F. 2001. Using human observer eye movements in automatic image classifiers. In *Proceedings of SPIE: Human Vision and Electronic Imaging VI*, volume 4299, 373–384.

A Vedlagte filer

De vedlagte filene gjør det mulig å gjennskape alle resultatene som er funnet i dette prosjektet. Følgende filer ligger på <http://www.utsnitt.com/masteroppgave>

- Datafiler – Datafilene etter konvertering i idfConvert
 - 0101.txt
 - 0102.txt
 - 0103.txt
 - ...
 - 2103.txt
- Java – Kildekode til javaprogrammet ETmessage.java
 - ETCommunicate.java
 - ETmessage.java
 - ExecCommand.java
 - ReadXMLFile.java
- MATLAB – Script for behandling av datafilene i matlab
 - ETanalyze.m
- Rapport – Denne rapporten i pdf-format
 - Thesis.pdf
- SPSS – Syntax-fil for behandling av data i SPSS
 - Syntax.sps
- Spørreskjema – Resultatene fra spørreskjemaet
 - sporreskjema.xls

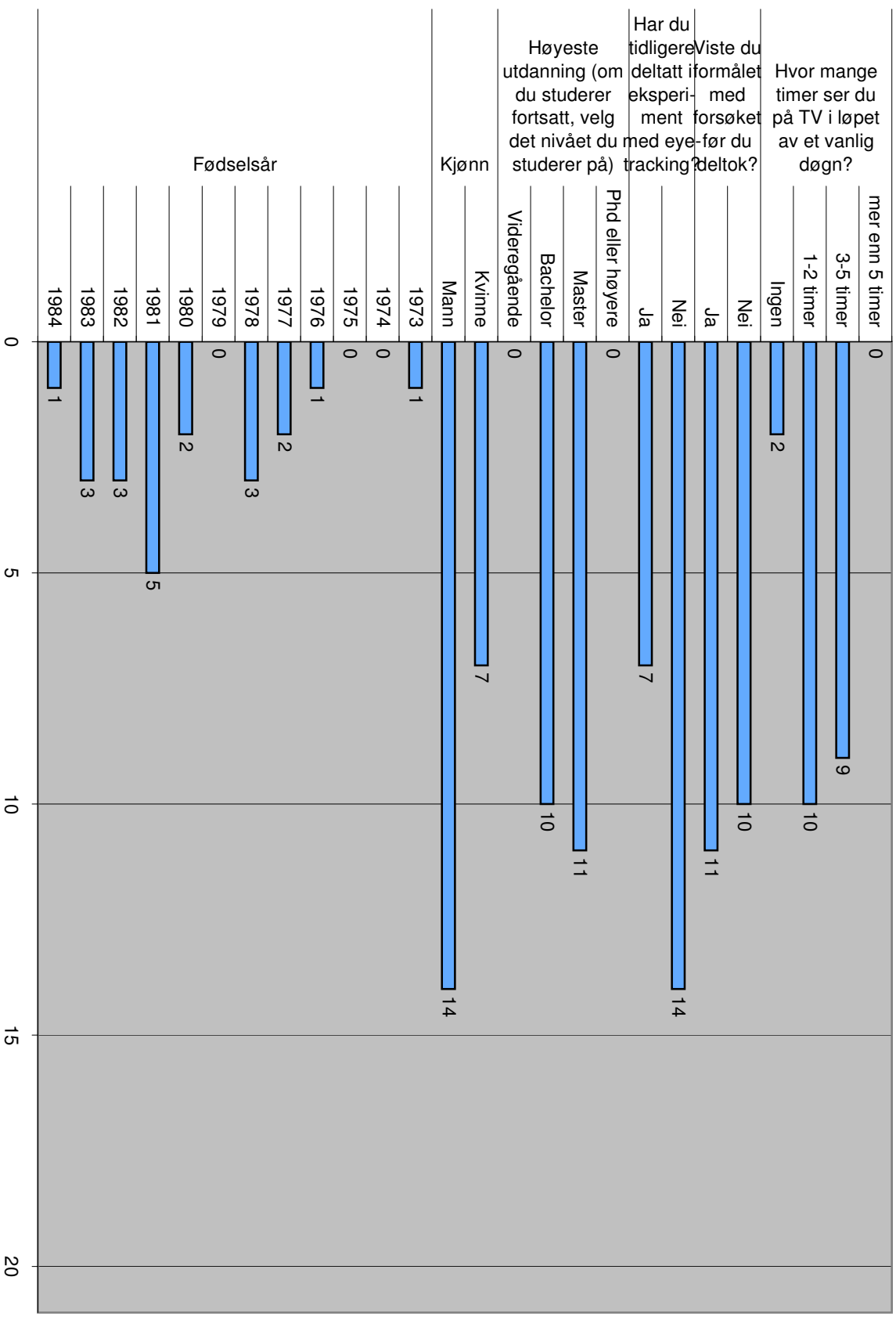
B Spørreskjema

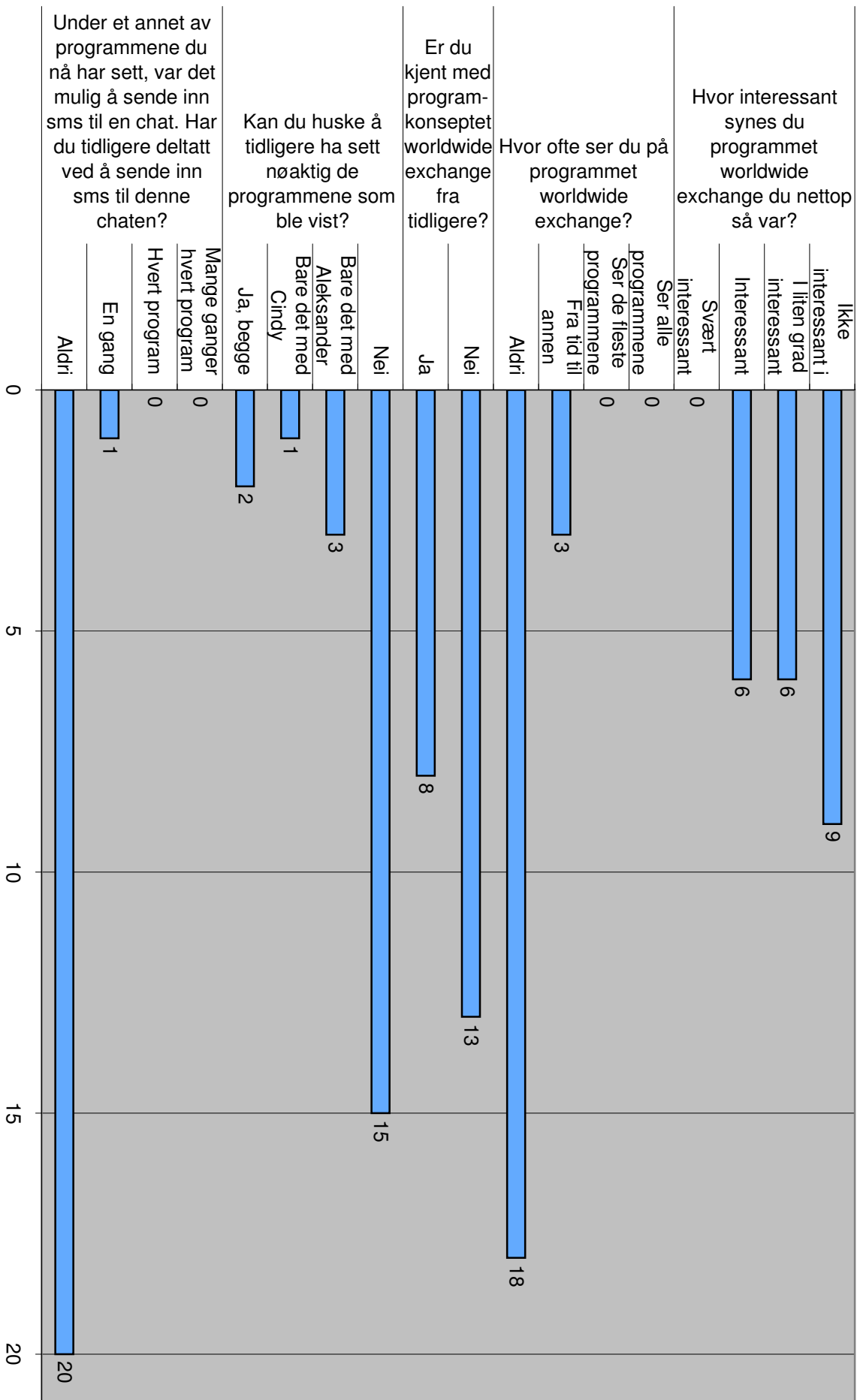
Spørreskjema

- 1 Deltaker nummer _____
- 2 Fødselsår _____
- 3 Kjønn Mann
 Kvinne
- 4 Høyeste utdanning (om du studerer fortsatt, velg det nivået du studerer på) Videregående
 Bachelor
 Master
 Phd eller høyere
- 5 Har du tidligere deltatt i eksperiment med eyetracking? Ja
 Nei
- 6 Viste du formålet med forsøket før du deltok? Ja
 Nei
- 7 Hvor mange timer ser du på TV i løpet av et vanlig døgn? Ingen
 1-2 timer
 3-5 timer
 mer enn 5 timer
- 8 Hvor ofte *ser* du på sms-tv (MessTV, Svisj o.l.)? Aldri
 1-2 ganger per uke
 3-5 ganger per uke
 mer enn 5 ganger per uke
- 9 Hvor ofte *deltar* du på sms-tv (MessTV, Svisj o.l.)? Aldri
 1-2 ganger per uke
 3-5 ganger per uke
 mer enn 5 ganger per uke
- Du har du nettopp sett tre TV-programmer. To av delene er hentet fra programserien «Idol» som ble vist på TV2 våren 2006. Den tredje er hentet fra CNBC «worldwide exchange»**
- 10 Er du kjent med programkonseptet Idol fra tidligere? Ja
 Nei
- 11 Fulgte du med på Idol i sesongen 2006? Nei
 Fra tid til annen
 Så de fleste programmene
 Så alle programmene
- 12 Under ett av programmene du nå har sett, var det mulig å stemme frem sin favoritt. Har du tidligere deltatt i avstemningen ved å ringe/sende sms? Aldri
 En gang
 Hvert program
 Mange ganger hvert program

- 13 Under et annet av programmene du nå har sett, var det mulig å sende inn sms til en chat. Har du tidligere deltatt ved å sende inn sms til denne chaten? Aldri
 En gang
 Hvert program
 Mange ganger hvert program
- 14 Kan du huske å tidligere ha sett nøyaktig de programmene som ble vist? Ja, begge
 Bare det med Cindy
 Bare det med Aleksander
 Nei
- 15 Er du kjent med programkonseptet «worldwide exchange» fra tidligere? Ja
 Nei
- 16 Hvor ofte ser du på programmet «worldwide exchange»? Aldri
 Fra tid til annen
 Ser de fleste programmene
 Ser alle programmene
- 17 Hvor interessant synes du programmet «worldwide exchange» du nettopp så var? Svært interessant
 Interessant
 I liten grad interessant
 Ikke interessant i det hele tatt

C Resultater fra spørreskjema





D Sammendrag av hendelser i dataene

D.1 Forklaring

For hver av settene er det et binding-diagram hvor y-aksen er relativ fikseringstid i prosent og x-aksen er tid i sekunder. Søylene representerer ett sekund. Den blå delen av søylen viser fikseringer innenfor grafikken mens den røde delen er fikseringer innenfor programinnholdet, inndelt etter modellen beskrevet i avsnitt 3.2.

Det er gjort en visuell analyse av videomaterialet hvor det er forsøkt å identifisere mulige årsaker til de segmentene hvor det er mer enn 15% fikseringer på grafikken i bildet. 15% er markert med en linje i binding-diagrammet. Det ble valgt å identifisere hendelser med mer enn 15% på grunn av at det er en høy nok terskelverdi til at årsaken kan identifiseres med en viss sansynlighet.

I tillegg til tekstlig beskrivelse av mulige årsaker til utslagene er også skjermdump fra hendelsen lagt ved som støtte.

D.2 Hendelser i sett 1

12 sek Som figur 13(a) viser er det ikke grafikk i bildet på dette tidspunktet, men hovedpersonen er i denne delen av bildet.

29 sek Grafikken kommer inn i bildet (fig. 13(b))

1 min 17 sek Gitaristen dukker opp bak ryggen på artisten og gitaren er delvis skjult bak grafikken (fig. 13(c)). Deltakerne ser derfor kanskje ikke på grafikken men forsøker å se gitaristen bak.

2 min 3 sek Som over er hovedmotivet i bildet gitaristen som havner bak grafikken (fig. 13(c)).



(a) 12 sek



(b) 29 sek

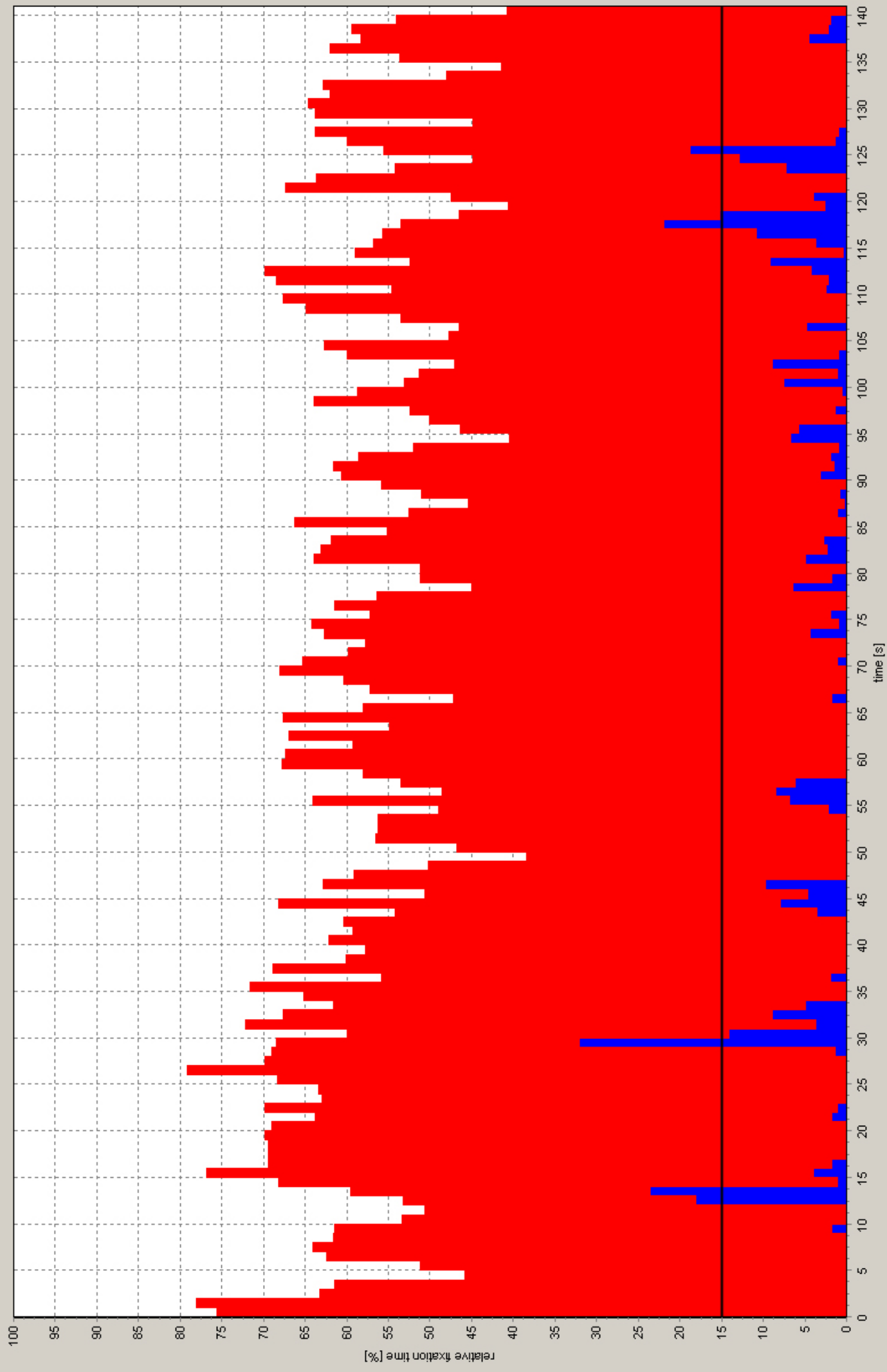


(c) 1 min 17 sek



(d) 2 min 3 sek

Figur 13: Spesielle hendelser i Sett 1



D.3 Hendelser i sett 2

- 17 sek** Etter 17 sekunder kommer den første infomeldingen med gul tekst inn (fig 14(a)).
- 1 min 12 sek** Første melding fra moderator, altså første melding med bilde (fig 14(b)). Sammenlignet med de foregående meldingene så er det mange som leser denne. De meldingene som følger rett etter denne blir også lest i større grad.
- 1 min 40 sek** Artisten løper for å sette seg ved pianoet (fig. 14(c)). Etter dette ligger grafikken over hendene til artisten mens han spiller, slik at det kanskje er et forsøk på å følge hendene snarere enn å lese grafikken.
- 2 min 9 sek** Igjen er det melding fra moderator med bilde (fig. 14(d)).
- 2 min 16 sek** En gul informasjonmelding dukker opp. Den trekker oppmerksomhet et kort øyeblikk (fig. 14(e)).
- 2 min 27 sek** Det skjer veldig lite i bildet og det har ikke skjedd noe i grafikken på en stund. Når det da kommer inn en ny melding trekker denne oppmerksomhet (fig. 14(f)).
- 2 min 49 sek** Musikken er slutt, applausen har lagt seg og dommerne begynner å prate. Når det da kommer inn en ny tekstmelding trekkes oppmerksomheten til denne (fig. 14(g)). Det samme skjer ved 2 min 54 sek og 3 min 2 sek .
- 3 min 13 sek** Det har ikke skjedd noe i grafikken på en stund og det virker til at deltakerne sveiper blikket innom grafikken for å bekrefte at de ikke har gått glipp av noe. Det samme skjer ved 3 min 26 og 3 min 33 sek
- 3 min 40 sek** En ny melding kommer inn og blikket flyttes (fig. 14(h))
- 3 min 48 sek** Moderator svarer. (fig. 14(i)).
- 4 min 11 sek** Ny melding trekker oppmerksomhet (fig. 14(j))
- 4 min 17 sek** En relativt lang melding som deltakeren bruker mer tid på å lese (fig. 14(k))
- 4 min 25 sek** Dommerenes tilbakemeldinger fortsetter og nye meldinger som kommer inn trekker umiddelbart oppmerksomhet (fig. 14(l)). Det samme etter 4 min 34 sek (fig. 14(m)) og 4 min 41 sek (fig. 14(n)).
- 4 min 49 sek** Gul informasjonmelding trekker oppmerksomhet et kort øyeblikk (fig. 14(o))
- 5 min 6 sek** Moderator sin melding med bilde trekker oppmerksomhet et kort øyeblikk. Det foregår mye i bildet og deltakerne fikserer ikke lenge på grafikken. (fig. 14(p))
- 5 min 46 sek** Den første nye melding på en stund og i tillegg starter den med to spørsmålsteget. Dette får deltakerne til å lese (fig. 14(q)).
- 5 min 52 sek** Melding fra moderator som svar til foregående melding (fig. 14(r)). Dette er en lengre melding, deltakerne bruker noe tid på denne.



(a) 17 sek



(b) 1 min 12 sek



(c) 1 min 40 sek



(d) 2 min 9 sek



(e) 2 min 16 sek



(f) 2 min 27 sek



(g) 2 min 49 sek



(h) 3 min 40 sek



(i) 3 min 48 sek



(j) 4 min 11 sek



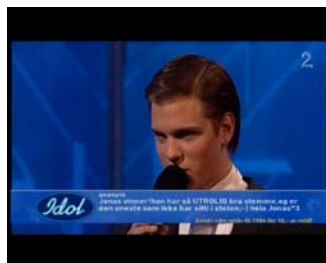
(k) 4 min 17 sek



(l) 4 min 25 sek



(m) 4 min 34 sek



(n) 4 min 41 sek



(o) 4 min 49 sek



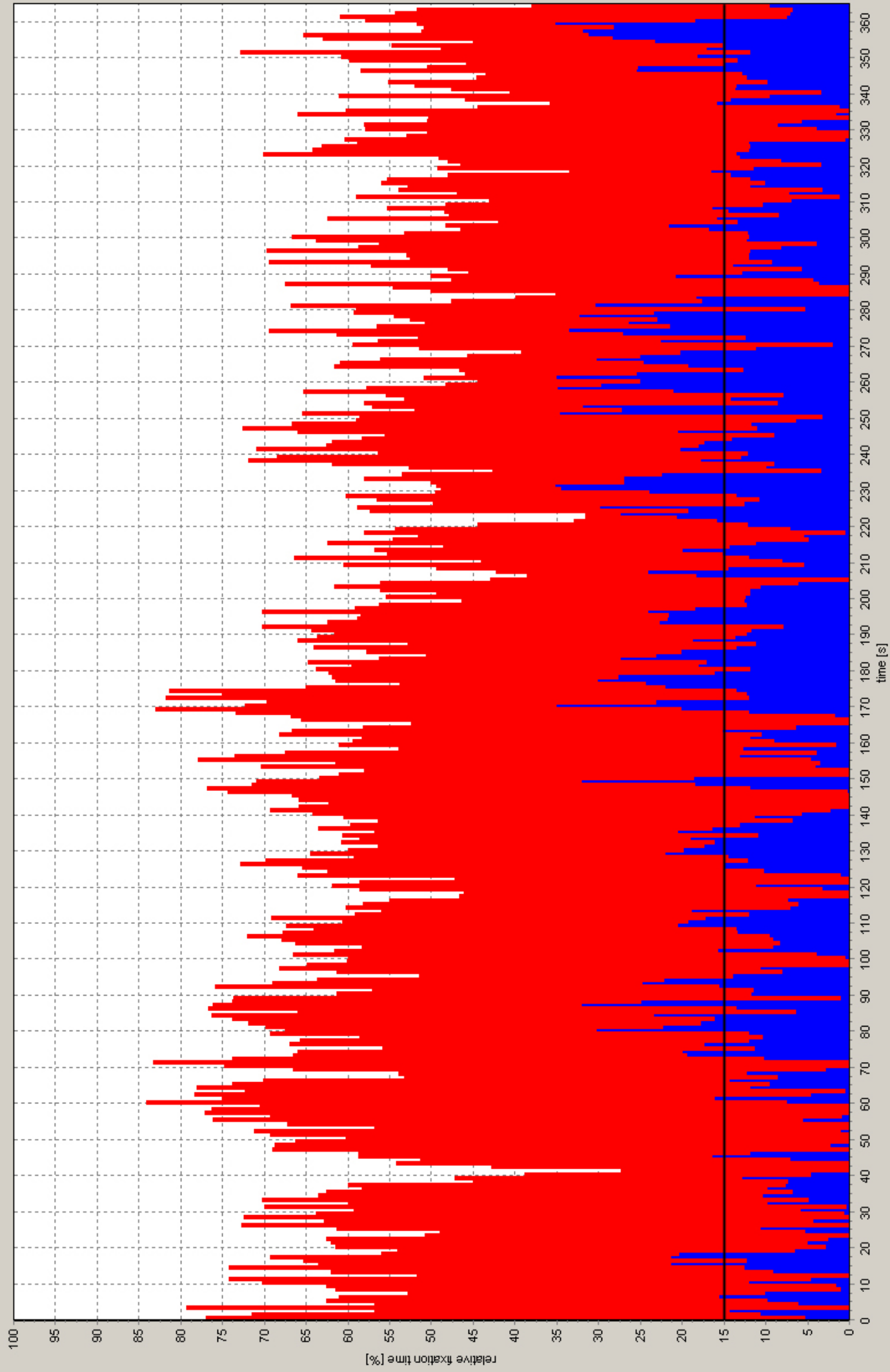
(p) 5 min 06 sek



(q) 5 min 46 sek



(r) 5 min 52 sek



D.4 Hendelser i sett 3

- 5 sek** Deltakeren forsøker sansynlig vis å orientere seg i bildet og er innom delene med grafikk for å se hva de inneholder (fig. 15(a)).
- 1 min 11 sek** Det er et skifte fra en graf til en annen i hovedbildet (fig. 15(a)), noe som skulle gjøre at blikket ble trukket bort fra grafikken og ikke mot grafikken.
- 1 min 33 sek** Bildet har nettopp skiftet fra å være graf til å bli nærbilde av Leuchtmann ((fig. 15(c)). Helt frem til 1 min 56 sek virker det som deltakerne benytter anledningen til å studere grafikken rundt om i bildet mens innholdet i programmet er et nærbilde
- 3 min 0 sek** Nærbilde av Leuchtmann igjen, deltakeren ser på grafikken
- 3 min 17 sek** Fra 3 min 17 sek er deltakerne hyppig innom grafikken. Bildet består i denne perioden stort sett av nærbilde av Leuchtmann helt frem til 3 min 30 sek hvor et anker kommer inn, takker Leuchtmann for deltakelsen og henter inn en ny ekspert (Pedrazzi) på 3 min 35 (fig. 15(d))
- 3 min 41 sek** 6 sekunder etter Pedrazzi kommer inn i bildet begynner deltakeren igjen å se på grafikken. Det er nå et nærbilde av Pedrazzi i hovedbildet (fig. 15(e)).
- 3 min 45 sek** Grafene kommer igjen inn i bildet (fig. 15(f)) og deltakerne ser ikke på grafikken i bildet før rundt 3 min 56 sek.
- 3 min 56 sek** Deltakerne begynner igjen å se på grafikken i stor grad. Samme graf har vært på skjermen i 11 sekunder. Deltakerne bruker en del tid på grafikk helt frem til 4 min 2 sek
- 4 min 2 sek** I perioden fra 4 min 2 sek frem til 5 min 17 sek er det mer aktivitet i hovedbildet. Blant annet vises det et klipp fra innsiden av en fabrikk (fig. 15(g)).
- 5 min 6 sek** Det har vært en del ulike grafer og deltakerne tar igjen en titt på grafikken.
- 5 min 27 sek** Fra 5 min 17 sek har Pedrazzi vært i nærbilde. Deltakerne viser mer interesse for grafikken i en periode.
- 5 min 45 sek** Pedrazzi takker for seg. Frem til dette punktet har deltakerne brukt tid på å se på grafikken, men trekkes mot hovedbildet når han takker for seg og setter tilbake til studio (fig. 15(h)).
- 6 min 17 sek** Fra 6 min 13 sek vises et oversiktsbilde av en storby (fig. 15(i)). Deltakerne bruker tid på dette bildet i noen sekunder men går igjen til grafikken.
- 6 min 26 sek** Programmet går over i reklamepause. En jingel markerer overgangen (fig. 15(j)). Fra denne jingelen starter er det mye aktivitet i bildet og deltakerne er ikke innom grafikken for resten av settet.



Figur 15: Spesielle hendelser i Sett 3

