

Publikumsinteraktivitet i konserter

Trond Christian Frøhaug



Masteroppgave
Master i Teknologi - Medieteknikk
30 ECTS
Institutt for informatikk og medieteknikk
Høgskolen i Gjøvik, 2005



Masterprogrammet i medieteknikk
har blitt kjørt i samarbeid med
Kunliga Tekniska högskolan (KTH),
Stockholm, Sverige

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

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

Sammendrag

Målet med prosjektet var å komme frem til en løsning hvor publikum kunne være aktive deltakere under konserten og påvirke stemningen. Prosjektet har sett på en teknisk løsning for hvordan slik interaktivitet kan tilbys publikum ved at de i dette tilfellet kan påvirke lyd og bilde. Som kontrollgrensesnitt er det benyttet videokamera med påfølgende bevegelsesanalyse for at systemet skulle være transparent og ikke legge begrensninger på publikums fysiske utfoldelse.

Et problem ved bruk av video som kontrollgrensesnitt er at forandringer i lysforhold gjør det vanskelig å gjennomføre bevegelsesanalyse. For å unngå dette ble det gjennomført eksperimenter hvor synlig lys ble filtrert vekk og hvor bevegelsesanalysen ble gjennomført på video fra det infrarøde området. Resultatene viste at det under bestemte forhold var mulig å fjerne forstyrrelser fra varierende lyskilder ved å benytte et IR-filter. Siden systemet skal brukes i sanntid ble det også eksperimentert med forskjellige typer kamera og innstillinger for å redusere kravet til maskinkraft. Det viste seg at ramme-frekvensen kunne reduseres mye uten vesentlig reduksjon av nøyaktigheten i bevegelsesanalysen. Eksperimentene viste også at kvaliteten på et ordinert webkamera var god nok, slik at det ikke var nødvendig å benytte et DV-kamera.

Det ble gjennomført en brukertest med 50 ungdommer. På grunn av problemer under gjennomføringen ga denne ingen klare resultater. Likevel viste testen en tendens til at de publikummerne som mente at interaktiviteten fikk de til å danse mer, også svarte mer positivt på alle de andre spørsmålene. DJ-en som deltok under brukertesten opplevde at systemet var nyttig fordi han fikk en bekreftelse på når publikum "tok av". Bekreftelsen førte til at han selv ble oppfordret til å yte mer.

Abstract

The aim of this project was to find a technical solution for offering interactivity to a concert audience. The objective was to have the audience becoming active participators in the concert and affect the concert's mood. This was accomplished by giving them some control over sound and graphic. As a controller for this interaction, a video camera combined with motion analysis was used. This solution made the system transparent and did not restrict the physical expression of the audience.

Motion analysis is sensitive to variations in the light conditions. To avoid this it was experimented with a filter that removed the visible light, and the motion analysis was done in the infrared area. The results show that it was possible to reduce the influence from varying light with an IR-filter. The system is a real time system and therefore it was experimented with different types of camera and settings to reduce the need of CPU power. The experiment show it possible to reduce the frame rate significantly, without reducing the motion analysis accuracy. Experiments indicate also that the quality from a webcam is sufficient and that it is not necessary to use a DV camera.

A user study with 50 young people was conducted. The study is somewhat inconclusive due problems faced during this experiment. Still it show that the audience who stated that they danced more because of the interactivity, also responded more positively to the other questions. The DJ who played music during the experiment, experienced that the system was useful because it gave him a confirmation when the audience moved and danced more expressively. This confirmation encouraged him to improve his performance.

Forord

Prosjektet har blitt gjennomført som en del av studiet Master i teknologi, retning medieteknikk ved Høgskolen i Gjøvik. Ideen til oppgaven ble lagt frem og gjennomført av undertegnede. Hovedveileder har vært Rune Hjelsvold og undertegnede vil takke for alle konstruktive innspill under prosjektperioden. I forbindelse med planlegging og analyse av brukertesten har Frode Volden også fungert som veileder.

Undertegnede ønsker i tillegg å takke Alexander Refsum Jensenius for samtaler og veiledning. Jensenius er ansatt som stipendiat ved Institutt for musikkvitenskap ved Universitetet i Oslo, og har fungert som ekstern veileder for prosjektet.

- 1. juli 2005, Trond Christian Frøhaug -

Innhold

Sammendrag	iii
Abstract	v
Forord	vii
Innhold	ix
1 Innledning	1
1.1 Problemstilling	2
1.2 Metode	3
2 Bakgrunn	5
2.1 Tilsvarende arbeid	5
2.1.1 Interaktiv messe	5
2.1.2 Cinematrix interactive entertainment system	5
2.1.3 Galvactivator	6
2.1.4 Large group interaction using disposable wireless motion sensors	6
2.1.5 The interactive dance club	6
2.2 Bevegelsessensorer	7
2.2.1 Tråkkematter	7
2.2.2 Joysticker og gamepads	8
2.2.3 Trykknapper	8
2.2.4 Laser/IR-matrise	8
2.2.5 Videokamera	9
2.3 Introduksjon til bevegelsesanalyse	9
2.3.1 Fargerom	9
2.3.2 Det infrarøde området	10
2.3.3 Termisk stråling	11
2.3.4 Bevegelsesdeteksjon	12
2.3.5 Glatting	13
2.3.6 Aktivitetsverdi	14
2.4 Software	15
2.4.1 Pure Data	15
2.4.2 EyesWeb	16
2.4.3 OpenSound Control	16
3 Teknologiske avklaringer med videokamera som sensor	17
3.1 Rammefrekvens	17
3.2 Kamerastøy	20
3.3 IR-filter	22
3.3.1 Kamasammenligning uten IR-filter	24
3.3.2 Kamasammenligning med IR-filter	25
3.4 IR-filter og lysscanner	27
4 Systemoppsett og brukertest	31
4.1 Forberedelser	31

4.1.1	Bevegelsesintervall og patcher	31
4.1.2	Plassering av kamerasensor	31
4.1.3	Valg av musikk	31
4.1.4	Publikumsinteraktiviteten	32
4.1.5	Maskinoppsett	32
4.1.6	Programoversikt	33
4.2	Pre-test Garvern	34
4.3	Brukertest Mesnali	35
5	Konklusjon	41
6	Videre arbeid	43
	Bibliografi	45
A	Spørreskjema	47
B	Rådata fra rammefrekvenseksperimentet	49
C	Programkode fra brukertesten	51
D	Diagram fra brukertesten	55

1 Innledning

Temaet for oppgaven kommer fra ønsket om å involvere publikum på en konsert i selve konserten. Det snakkes altså om interaktivitet mellom publikum og selve konserten som helhet. I dag finnes det få kommersielle løsninger som gjør dette mulig, og det er en utfordring å finne ut hvordan dette kan løses i praksis.

Før dette utdypes ytterligere, kan det være nyttig å se på begrepene interaksjon og interaktivitet. Begrepene brukes ofte om hverandre, og det kan være nyttig å operasjonalisere dem. Geir Haugsbakk skriver i boka “Interaktivitet, teknologi og læring” [1], at det er begrepet interaksjon som har lengst tradisjon av de to. Begrepet er knyttet til mellommenneskelig kommunikasjon, til gjensidighet, samspill og vekselvirkning. Det er først i senere tid man har begynt å bruke begrepet interaktivitet. Haugsbakk sier videre at interaktivitetsbegrepet i første rekke har vært knyttet til teknologibruk. Det har sin hovedforankring i IKT-feltet og har blitt brukt til å beskrive relasjonen mellom menneske og maskin.

Begrepet interaksjon har også blitt brukt om en persons kontakt med datamaskinen når personen får kontinuerlig tilbakemelding¹. På bakgrunn av dette kan altså begrepet interaksjon brukes om publikums påvirkning av musikken hvis man sier at “publikum er i interaksjon med konserten”, og det er fristende å bruke begrepet når publikum ved hjelp av sin aktivitet kan påvirke konserten. Denne oppgaven har derimot holdt seg til Haugsbakk sine definisjoner, men begrepet *påvirkning* er i tillegg benyttet som erstatning for begrepene.

Når begrepet interaktivitet brukes i denne oppgaven, henviser det til den muligheten publikum har for å påvirke elementer under konserten, slik som avspilling av lydssamples og kontroll av scene-grafikk. For å kunne realisere denne interaktiviteten må det finnes et system som tilbyr dette.

Opgaven tar for seg teknologi og metoder som kan realisere interaktiviteten mellom publikum i salen og det som blir fremført på scenen i form av lyd og bilde. Resultatet av publikums påvirkning kan muligens føre til en ny form for interaksjon mellom musiker(e) og publikum. Slik konserter er i dag kan de fra publikums ståsted oppfattes som en “monolog” fra scenen. Det er musikerne sammen med lyd- og lysteknikere som former konserten ved hjelp av musikalsk innhold og “stemning” i lokalet. Publikum har derfor lite direkte innvirkning på sceneeffekter eller det musikalske uttrykket. Den eneste muligheten de har er å formidle sin begeistring eller misnøye ved hjelp av klapping, roping og lignende. Uttrykksformene som publikum benytter kan videre påvirke musikerne eller teknisk personell. Disse kan så påvirke musikk eller sceneeffekter, men denne “signalkjeden” gir ikke publikum direkte kontroll. Det gir derfor en liten grad av interaksjon mellom publikum i salen og utøverne på scenen.

¹Det er en vanlig oppfatning at interaktivitetsbegrepet først ble tatt i bruk på 1960-tallet i forbindelse med overgangen til ulike “online”-løsninger for programmering og databehandling. “Online”-programmering - som raskt ble kalt “interaktiv” programmering - innebar at programmereren hele tiden kunne stå i kontakt med datamaskinen og dermed få kontinuerlig tilbakemelding om eventuelle feil eller alternative løsninger. Dette ble av mange opplevd som å føre en dialog, eller med andre ord; å være i interaksjon med maskinen [2].

Det er ønskelig å komme frem til en teknisk løsning som gir publikum mulighet til å påvirke stemningen i konserten, derav tittelen “Publikumsinteraktivitet i konserter”. Dette kan oppnås ved at publikum får mulighet til å gjøre forandringer i musikk/lyd, bilde og lys. Hensikten er at publikum skal føle seg mer som deltakere i stedet for tilhørere under konserten. Musikerne på scenen vil også i større grad merke publikums engasjement, eller fravær av dette siden det gjenspeiles i lyd, lys og bilde.

Som et eksempel kan publikum få mulighet til å kontrollere noen elementer i musikken. Det kan være styring av lydfiltere på trommelopper, volum på synthpadder (lydtepper) og lignende. De kan også være med å styre visuelle elementer som lyssetting i salen, på scenen, eller styring av animasjoner. På store konserter eller arrangement vises det noen ganger video eller animasjon på en storskjerm. Avspilling og manipulasjon av elementene på denne skjermen kan for eksempel styres av publikum.

Gjennom eksperimenter skal noen av disse mulighetene prøves ut, og det er ønskelig å komme frem til en løsning der publikums engasjement påvirker konserten både lyd-messig og visuelt. I neste omgang vil dette kanskje påvirke publikums opptreden slik at resultatet blir en tilbakemeldingssløyfe. Publikum påvirker “stemningen”, stemningen “svarer” og publikum påvirkes på nytt. Kanskje blir det til og med en positiv tilbakemeldingssløyfe?

1.1 Problemstilling

Oppgaven ønsker å finne en løsning på følgende scenario: *Du er på en konsert og synes at bandet spiller så bra at du har lyst til å være med selv!*

Det er vanskelig for utøverne å vite hva du og de andre publikummerne tenker og føler, men dette kommer ofte til uttrykk i fysiske handlinger. Graden av fysisk aktivitet blant publikum kan derfor fungere som et mål på engasjementet. Utfordringen blir å registrere denne aktiviteten som for eksempel kan være klapping, dansing og hopping. Med dette som bakgrunn har oppgaven tatt opp flere problemstillinger. Disse kan deles inn i to hovedgrupper som er kalt *teknologiske avklaringer med videokamera som sensor* (kapittel 3) og *systemoppsett og brukertest* (kapittel 4). Nedenfor forklares de problemstillingene som oppgaven har sett på. Det overordnede spørsmålet for oppgaven har vært:

- hvilke komponenter og metoder kan benyttes for å realisere et rimelig system for publikumsinteraktivitet og hvordan skal de konfigureres?

Kapittel 3 beskriver eksperimenter når videokamera er benyttet for å registrere den fysiske aktiviteten. Det er bare denne sensortypen som er testet i praksis, men andre sensortyper er vurdert i kapittel 2.2. Hovedgrunnen til dette er et ønske om at teknologien som benyttes skal oppleves transparent for publikum. Det betyr at teknologien ikke skal være et forstyrrende element, og avlede publikums oppmerksomhet fra konserten ved at de må forandre sin fysiske adferd. Samtidig skal publikum få mulighet til å påvirke konserten bevisst eller ubevisst, noe som blir ivaretatt ved å bruke videokamera. Selv om videokamera trolig er et godt valg er følgende problemstilling interessant:

- Hvilke typer sensorer er det aktuelt å benytte?

Siden videokamera har et fortrinn ved at det er en transparent sensor, bør det vurderes hva slags kvalitet som kreves av kameraet. Det er i utgangspunktet ønskelig med et så rimelig system som mulig. Kostnadene på systemet kan holdes nede ved å benytte

et webkamera i stedet for et kamera med DV-kvalitet hvis dette ikke er nødvendig. Ulempen er at webkamera kan ha mye egenstøy og det bør tas hensyn til dette ved valg av kamera. For at systemet skal fungere best mulig i sanntid er det også en fordel å redusere datastrømmen som skal analyseres. Det er interessant å finne ut hvor mye denne kan reduseres samtidig som det er mulig å gjennomføre en tilfredsstillende bevegelsesanalyse. En reduksjon av datastrømmen vil også senke kravene til maskinvare og gjøre systemet rimeligere. Dette gir følgende punkter:

- Sammenligning av kvalitet på tilgjengelige webkamera. Nødvendig rammefrekvens, egenstøy og kvalitet på kamera. Webkamera vs. DV-kamera.

Konserter og lignende arrangement gjennomføres ofte i svakt opplyste lokaler eller lokaler med skiftende lysstyrke. Kameraet som skal benyttes som sensor må derfor testes under disse forholdene. Muligens kan noen av problemene som oppstår fjernes ved å filme i IR-området. Det er derfor interessant å finne ut:

- Hvordan kan bevegelse detekteres i et halvmørkt rom med videokamera?
- Vil feilaktig deteksjon av bevegelse reduseres ved filming i IR-området?

Kapittel 4 dokumenterer hvordan systemet er bygget opp og hvordan det er konfigurert for bruk på konsert. Prototypen av systemet er formet på bakgrunn av litteraturstudie, ønskede egenskaper og de ressursene som har vært tilgjengelig under prosjektet. Et poeng er også at selve systemet skal være så rimelig som mulig slik at flere har mulighet til å bruke det. Økonomiske og tidsmessige begrensninger har ført til at det ikke har vært mulig å teste alle løsninger som virket interessante. Dette har påvirket den endelige prototypen og har lagt premisser som fikk konsekvenser for utfallet av brukertesten.

Kapittelet dokumenterer også testingen av systemet i praksis og resultatene fra dette arbeidet. Det ble gjennomført en større brukertest sammen med 50 deltagere, og en musiker som i dette tilfellet var representert ved en DJ. Hensikten var å finne ut hvordan systemet fungerte i praksis og om publikum opplevde at de kunne påvirke lyd og bilde. Testen skulle også forsøke å finne ut om systemet bidro til at konserten ble en mer positiv opplevelse både for publikum og musikeren. Dette ga følgende spørsmål:

- Fungerer systemet slik at publikum opplever interaktivitet?
- Opplever musikeren at systemet har verdi?

1.2 Metode

Problemstillingene i kapittel 1.1 ble løst ved litteraturstudie, eksperiment og brukerstudium. Gjennom litteraturstudiet kom det frem flere prosjekter som hadde jobbet med interaktivitet for mange personer samtidig. Noe av dette arbeidet er referert til i kapittel 2.1. Litteraturstudiet pågikk gjennom hele prosjektperioden siden nye problemstillinger og løsninger dukket opp etterhvert som prosjektet utviklet seg. Det ble hovedsakelig benyttet fagdatabaser på Internett for å finne relevant litteratur. Disse inneholder store mengder informasjon og det er hurtig å gjennomføre søk. I tillegg ble bibliotekene på HiG og UiO benyttet, og fagmiljøet på begge steder var også behjelpelig med noe litteratur. Det er vanskelig å vite om den mest relevante litteraturen ble funnet men den er godt forankret i tidligere forskning. Litteraturen inneholdt derfor også referanser til annet arbeid og dette har gjort det mulig å få en oversikt over området. Det skal også nevnes

at siden oppgaven benytter mye forskjellig teknologi, har den også mange områder som ikke nødvendigvis refererer til hverandre. Dette har gjort litteratursøket vanskeligere siden det var viktig å finne riktige søkeord og sammensetninger av disse.

Eksperimentene som ble gjennomført kan karakteriseres som en blanding av ingeniørmetode og kvalitativ metode. De kom som en naturlig følge av at videokamera ble valgt som sensor, og sier i første rekke noe om egenskapene til de kameraene som ble testet. Allikevel kan resultatene til en viss grad generaliseres, eller ihvertfall gi en pekepinn på hvilke innstillinger som kan benyttes som utgangspunkt for bruk av systemet.

Brukerstudiet ble gjennomført sammen med en live-testing av systemet. Systemet ble også testet før denne live-testen. Verdien av brukerstudiet ble likevel redusert fordi det ikke var tid eller ressurser til å gjennomføre store tester av systemet før live-testingen. Dette førte til at man hadde lite erfaring med gjennomføringen av en slik live-brukertest. Noen av problemene som oppsto skyldtes faktorer som ikke hadde direkte tilknytning til systemet. Disse kunne vært unngått i senere tester og på den måten gjort resultatene fra senere live-tester mer verdifulle. Den kvantitative delen av oppgaven ble svekket, og dette ble det tatt hensyn til under analysen av datamaterialet fra spørreundersøkelsen. Spørreskjemaet ble utformet med bakgrunn i problemstillingen og ved å studere en tilsvarende undersøkelse gjennomført ved MIT [3].

2 Bakgrunn

Dette kapittelet ser på tidligere arbeid på området og problemstillinger som er aktuelle for å realisere en prototype av systemet. Det begrunner også forskjellige teknologivalg, slik som valg av sensor og tilgjengelig programvare.

2.1 Tilsvarende arbeid

Det har tidligere blitt gjennomført prosjekter hvor mange mennesker samtidig har fått en bestemt type kontroll. Hva de kunne kontrollere har vært forskjellig, men en fellesnevner har vært at utfallet har gjenspeilet hele gruppens mening eller bevegelse. For å klare dette har både passive og aktive enheter/sensorer blitt benyttet. Eksempelene nedenfor viser noen av prosjektene og metodene der deltakerne benytter slike enheter for påvirkning.

2.1.1 Interaktiv messe

Høsten 2004 ble det arrangert Interaktiv messe [4] ved Østria (Høgskolen i Agder) og Betong (Oslo). Dette var en “ikke-lineær multimedia-gudstjeneste” i regi av Norges Kristelige Studentforbund. Følgende tekst er hentet fra messens informasjonsside [5] og forklarer begrepet og interaktiviteten i messen.

Rommet er fylt av forskjellige type sensorer som til enhver tid følger menneskemassens bevegelser i rommet og enkeltpersoners handlinger. Dette danner grunnlaget for messens struktur og innhold. Materialet som er forberedt på forhånd er lagret slik at det kan settes sammen på et utall forskjellige måter, basert på signaler fra de interaktive elementene.

Sensorene som det refereres til var tråkkematter, webkamera, IR-detektorer og “kunstobjekter” som hang ned fra taket. Kunstobjektene var festet til joysticker, som sammen med de andre sensorene registrerte verdier som ble analysert og bestemte messens struktur og innhold.

2.1.2 Cinematrix interactive entertainment system

Cinematrix interactive entertainment system [6] [7] har blitt testet med 4000 deltakere. Systemet tillot deltakerne å konkurrere i et elektronisk spill og ble første gang demonstrert i 1991 på ACM SIGGRAPH¹, som er en konferanse om datagrafikk og interaktive teknikker.

Systemet har blitt brukt i konkurranser der publikum er delt i to lag som spiller mot hverandre. Det er utviklet flere spill for dette formålet og et av dem ble brukt under åpningsseremonien til ARS Electronica [8] i 1994. Festivalen avholdes i Tyskland hvert år og har overskriften kunst, teknologi og samfunn. Systemet fungerte ved at alle deltakerne fikk utdelt hver sin passive enhet i form av en paddel (flat pinne). Paddelen ble holdt i hodehøyde og var grønn på den ene siden og rød på den andre. Ved hjelp av en storskjerm kunne deltakerne stilles spørsmål eller være medspillere i et dataspill. Deltakerne svarte ja eller nei på spørsmålet ved å snu paddelen. I dataspillet kunne de for eksempel styre en avatar ved å snu på paddelen. Deltakerne som testet Cinematrix var i alle aldre og med

¹Special Interest Group on Computer Graphics of the Association for Computing Machinery

forskjellig grad av teknisk erfaring. Verten som åpnet arrangementet og ga instruksjoner snakket tysk, men dette var det ikke alle deltakerne som forstod siden de kom fra mange forskjellige land. På tross av dette moret deltakerne seg og klarte å samarbeide som en felles gruppe [6]. En kvinnelig deltaker uttrykte etter åpningen at hun hadde følt gruppetilhørighet samtidig som hun opplevde å beholde individualiteten.

Cinematrix interactive entertainment system ble så suksessfylt at Cinematrix Incorporated ble etablert.

2.1.3 Galvactivator

Handlingene til publikum kan også måles ved å gi dem en aktiv enhet. En aktiv enhet er utformet slik at brukeren må betjene den eller at den selv kan registrere forandringer. Galvactivator [9] er en slik aktiv enhet som selv registrerer forandringer. Enheten er utformet som en hanske som brukeren tar på seg og som måler ledningsevnen på brukerens håndoverflate. Hvis brukeren blir stresset stiger fuktigheten på håndoverflaten. Dette fører til økt ledningsevne som igjen gjør at LED-en (lysdioden) som er montert på hansken begynner å lyse kraftigere. Lyset som emitteres fra alle brukernes hansker blir målt med et kamera og gjør det mulig å kvantisere publikums engasjement eller stress.

Ved SENS*BLES symposiumet på MIT i 1999 ble det delt ut 1200 galvactivatorer. Under et foredrag på det samme symposiumet ble systemet testet med rundt 200 hansker. Når deltakerne ble utsatt for ytre påvirkning, ble reaksjonen registrert ved måling av lyset som hansken sendte ut. Det viste seg at man klarte å registrere følelsesmessige topper og bunner når deltakerne lo, applauderte, slappet av, og når ballonger ble blåst opp og siden eksploderte.

2.1.4 Large group interaction using disposable wireless motion sensors

En annen type aktiv enhet som selv registrerer brukerens bevegelser er et akselerometer [3]. Ved hjelp av trådløse akselerometere som ble delt ut til deltakerne på en danseklubb, var det deltakerne selv som bygget opp musikken. Deltakerne holdt det trådløse akselerometeret i hånden og når bevegelsen passerte en forhåndsdefinert grenseverdi, sendte den aktive sensoren fra seg en puls. To mottagerstasjoner registrerte pulsene fra akselerometerne som ble analysert og bygget opp rytmiske strukturer. Når pulsene oversteg bestemte energinivå, ble det introdusert nye elementer i musikken. Ved lav aktivitet var det for eksempel bare trommer og basslinjer som ble trigget. Ved høyere energinivå ble melodiske elementer og lyder lagt til. Resultatet var at intensiteten i musikken økte i takt med hvor aktive deltakerne var. Dette systemet ble testet flere ganger og var initiert av MIT. Det største antall deltakere systemet har blitt testet med er 200 personer.

Den største fordelen ved å benytte denne typen sensor er at den ikke krever fri sikt til brukerne, slik som blant annet en kamerasensor gjør. Akselerometeret blir heller ikke påvirket av variasjoner i lysforhold og fungerer like godt om det er helt mørkt i lokalet. Ulempen er at hver person trenger ett akselerometer hver. Kostnadene blir derfor vesentlig høyere enn ved bruk av kamera.

2.1.5 The interactive dance club

I motsetning til de foregående eksemplene finnes det også sensorer som er aktive, men ikke krever at deltakerne må bære eller betjene en enhet. Noe av dette ble brukt på "The interactive dance club" [10] under 25-års jubileet til SIGGRAPH. Målet var å lage et møtested der mennesker kunne få mulighet til å bli medspillere i et stort sammenkoblet

interaktivt musikalsk og visuelt miljø. For å gjennomføre dette ble det blant annet brukt sensorer som registrerte deltakernes bevegelser uten at fysisk kontakt med deltakerne var nødvendig. For eksempel ble personene på dansgulvet filmet med et infrarødt kamera. Bildet med deltakernes kroppsvarme ble vist på en storskjerm som en effekt.

Det ble også satt opp soner bestående av parallelle lysstråler over deltakernes hode. Brudd på lysstrålen ble detektert og deltakerne kunne på den måten trigge samplede musikalske fraser. I en annen sone ble deltakerne oppfordret til å bryte lyset med den hensikt å kaste forskjellige skygger på sensorene som var plassert i en ring på gulvet. Det ble også brukt 10 tråkkematter i installasjonen. Her kunne deltakerne tråkke på mattene og på den måten spille av korte musikalske fraser og i tillegg styre projisert datagrafikk.

2.2 Bevegelsessensorer

Som det ble nevnt i kapittel 1.1 er det et ønske at sensoren ikke virker forstyrrende på publikum. Med dette menes at sensoren ikke må “stjele” publikums oppmerksomhet ved at de fokuserer på betjeningen av denne. I tillegg må publikums frihet ivaretas slik at de ikke trenger å forandre sin fysiske oppførsel for at sensoren skal kunne registrere den fysiske aktiviteten.

Muligens vil resultatet som fremkommer ved bruk av sensoren forandre publikums oppførsel, men poenget er at sensoren i seg selv ikke skal bidra til dette. Sensoren skal oppleves transparent for publikum og skal registrere publikums engasjement uten å forstyrre publikums “normale” fysiske aktivitet. Dette kan oppsummeres i noen punkter som sensorene har blitt vurdert i forhold til. Sensoren må:

- oppleves transparent for publikum og fungerer uten fysisk berøring
- være relativt rimelig i innkjøp
- være mobil, hurtig og enkel å sette opp
- være robust og fungere under forskjellige lysforhold
- ha lav reaksjonstid slik at publikum opplever interaksjon

2.2.1 Tråkkematter

Tråkkematter (jfr. figur 1) har form omtrent som dørmatter og aktiviseres når en person tråkker på den. Kort sagt fungerer tråkkematten som en bryter som enten er av eller på. En mulig løsning kan være å dekke hele gulvet med slike matter. Dette vil gjøre det mulig å registrere bevegelse blant publikum med en “oppløsning” som tilsvarer antall matter. Ulempen er blant annet at mattene vil få stor fysisk påkjenning når publikum hopper og danser på dem. Dette kan redusere levetiden på mattene og samtidig skape en usikkerhetsfaktor rundt stabiliteten til systemet.

Tråkkemattene som har vært aktuelle for prosjektet, har en størrelse på 58x17cm og 73x38cm, og koster henholdsvis 45 og 64 kroner stykket [11]. Andre tråkkematter er Tap Tile [12] og Z-Tiles [13]. Sistnevnte er et modulbasert system som gjør det enklere å montere “gulvet” på forskjellige måter.

På grunn av størrelsen på konsertlokalet trengs det mange tråkkematter for å dekke hele gulvet. Løsningen er derfor ikke testet fordi den regnes som upraktisk og dyr. En annen faktor er at tråkkemattene ikke vil oppleves som en transparent sensor, siden det gir en annen følelse å tråkke på mattene, enn å tråkke på et vanlig gulv.



Figur 1: Tråkkematter i to forskjellige størrelser [11]

2.2.2 Joysticker og gamepads

Vanlige joysticker (jfr. fig 2) og gamepads som brukes i data og TV-spill har ofte mange knapper og et x/y-aksesystemet. Hvis man knytter knappene og aksesystemet til forskjellige variable gir det store muligheter for kontroll [14]. Prisen på slike kontrollenheter er fra 150 kroner og oppover [15].

Ulempen ved å bruke joysticker og gamepads som sensorer er at publikum trenger en kontroller hver. Eventuelt kan et bestemt antall kontrollere plasseres rundt i konsertlokalet. Publikum må da dele på disse og bare et fåtall av deltakerne kan interagere med systemet på samme tid.

Begge løsningene fører til at publikum må endre sitt vanlige aktivitetsmønster og fysisk betjene sensorene. Sensoren vil derfor ikke oppleves som transparent og er ikke vurdert som aktuell i denne sammenhengen.



Figur 2: Joystick som kan brukes som kontrollenhet [15]

2.2.3 Trykknapper

Ved bruk av trykknapper er det mulig å lage en løsning hvor publikum uttrykker sitt engasjement ved å trykke på en eller flere knapper. På samme måte som joystick trenger denne typen sensor fysisk berøring og må utplasseres blant publikum. Resultatet er liten transparenss og trykknapper ansees derfor ikke som en aktuell sensortype.

2.2.4 Laser/IR-matrise

Laser/IR-matrise vil si at flere parallelle lysstråler “skyter” på kryss og tvers i konsertlokalet. Lysstrålene lager en matrise hvor oppløsningen/nøyaktigheten er avhengig av antall stråler. Hvis strålen blir brutt kan dette tolkes som aktivitet på det bestemte punktet. Ved å legge matrisen litt over hodehøyde kan man fange opp publikums aktivitet i form av hopp og hender som beveges over hodehøyde. Samtidig kan man bestemme hvor i rommet denne bevegelsen er.

Ulempen er at en slik installasjon krever mye utstyr og mye tid til montasje av systemet. Det er derfor en lite mobil løsning og ansees som lite aktuelt i denne sammenhengen.

2.2.5 Videokamera

Et eller flere videokamera som filmer publikum vil kunne gi overblikk over aktiviteten i lokalet. Kameraet kan monteres på scenen slik at det filmer ovenfra og ned mot publikum. Det vil da være mulig å fange opp bevegelse fra publikum som befinner seg bak i lokalet og ikke bare de på første rad. En slik løsning vil være transparent for publikum og forholdsvis hurtig å sette opp. En av ulempene er at konsertlokalet ofte har varierende lysforhold noe som kan påvirke filmingen og bevegelsesanalysen. På tross av dette ansees videokamera som den beste implementasjonen av en transparent sensorløsning. Sensoren krever ingen fysisk betjening av publikum og vil trolig være relativt mobil og hurtig å montere. Som tidligere nevnt har oppgaven derfor benyttet videokamera som sensor.

Siden kamera er valgt som sensor er det aktuelt å vurdere bruk av webkamera i forhold til DV-kamera. Et DV-kamera gir høyere oppløsning og bedre kvalitet på bildene, men er også dyrere enn et webkamera. Fordelen med webkameraet i tillegg til lav pris er at det er enkelt å redusere oppløsning og rammefrekvens. Dette gir lavere datastrøm, som igjen reduserer kravet til maskinkraft. Det er en fordel med lav datastrøm siden systemet skal benyttes i sanntid.

For å fange opp bevegelse i store lokaler kan det være aktuelt å benytte flere kamera eller et kamera med vidvinkel slik det er gjort i oppgaven. Kamera finnes i mange prisklasser og et av de rimeligste er webkameraet som er brukt i oppgaven. Dette er vist i figur 3 og koster rundt 300 kroner [16].



Figur 3: Sandberg NightCam [16] som brukes senere i prosjektet

2.3 Introduksjon til bevegelsesanalyse

Dette kapittelet tar for seg problemstillinger ved realisering av systemet som skal tilby interaktivitet til publikum. Det ser på muligheten for å bruke kamera som sensor i et mørkt rom med varierende belysning, og teorien som ligger bak bevegelsesdeteksjon. I tillegg blir programmene som er benyttet i oppgaven beskrevet.

2.3.1 Fargerom

Det er mulig å gjennomføre bevegelsesanalyse i forskjellige fargerom slik som RGB, HSI og YCbCr [17]. I dette tilfellet skal bevegelsesanalysen gjennomføres i et miljø hvor både farger og lysintensitet er i forandring. Dette gir en ekstra utfordring siden systemet bare

skal detektere bevegelse blant publikum og ikke variasjon i lysintensiteten. Hvis bare intensiteten hadde vært i forandring hadde det vært mulig å holde intensiteten utenfor analysen ved å benytte et fargeområde hvor denne komponenten var skilt ut. Bevegelse blant publikum kunne da blitt detektert på bakgrunn av fargeforandringer uten at intensitetsforandringer i lyset “forstyrret” dette. Disse problemene kan kanskje også unngås ved å gjøre analysen i det infrarøde området.

2.3.2 Det infrarøde området

Ved bruk av vanlig videokamera er det en fordel med mye lys og stabile lysforhold for at opptaket skal bli best mulig. Dette er nødvendig for å unngå at kameraet må justere lysfølsomheten hele tiden, og for at objektene som filmes skal reflektere nok lys til at de blir synlige på opptaket. På konserter kan dette være en utfordring siden det ofte bare er scenen som er opplyst. Området hvor publikum står er som regel bare delvis opplyst eller opplyses med lys som for eksempel varierer i takt med musikken. Særlig på klubber og diskotek brukes det i tillegg lyseffekter som scannere og strober. Det er derfor interessant å komme frem til en metode som kan gjøre det mulig å foreta bevegelsesdeteksjon under slike vanskelige lysforhold.

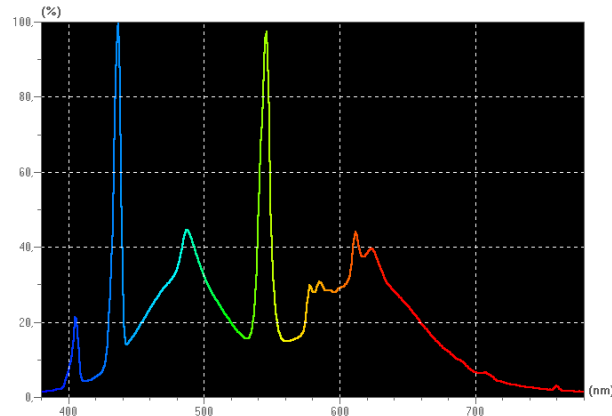
En fordel i denne sammenhengen er at datastrømmen som kameraet leverer, ikke skal betraktes visuelt av mennesker slik som et vanlig videoopptak blir. Det kan derfor være aktuelt å benytte utstyr som registrerer bølgelengder utenfor det synlige området. Ved bruk av et kamera med nightshot-funksjon blir det infrarøde området transponert ned i området som er synlig for det menneskelige øyet. Bilder fra dette området er i gråtoner siden det ikke finnes farger i det infrarøde området. Muligens kan noe av det forstyrrende lyset fjernes slik at man kan få en bedre bevegelsesdeteksjon ved å analysere bildet som er laget ut fra det infrarøde området.

Figur 4 viser en oversikt over ultrafiolett, synlig og infrarødt lys. Lyset vi mennesker oppfatter som synlig, befinner seg i området 430nm - 790nm [17]. Selv om mennesker ikke kan se bølgelengder over 790nm, registrerer de fleste videokamera beregnet på forbrukermarkedet bølgelengder opp til ca 1100nm. I dette området finnes infrarødt lys, ofte forkortet som IR-lys.

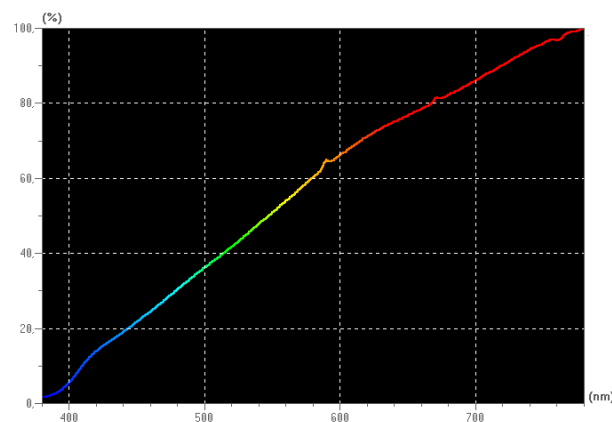


Figur 4: Bølgelengder for synlig og usynlig lys [18]

For å se hvilke bølgelengder som emitteres fra et standard lysstoffrør, ble det foretatt en måling med et radiospektrometer. Resultatet er vist i figur 5 og viser relativ energi-stråling ved forskjellige bølgelengder. Figur 6 viser “signaturen” til en standard glødelampe som brukes som filmbelysning. Signaturene viser at det er stor forskjell på de to lyskildene. Radiospektrometeret som målingene ble utført med kunne ikke måle bølgelengder over 800 nm. På tross av dette, antyder likevel kurvene at glødelampen emitterer mye energi i IR-området. Dette bekreftes også i oppslagsverk [19] som opplyser at bare 3% av den elektromagnetiske energien fra en glødelampe er synlig. Resten emitteres i form av infrarød varmeenergi. Hvis det er ønskelig å opplyse et rom med IR-lys bør



Figur 5: Bølgespekter ved måling av lysstoffør på hvit bakgrunn



Figur 6: Bølgespekter ved måling av glødelampe på hvit bakgrunn

det derfor benyttes en glødelampe. For å fjerne bølgelengdene som gir synlig lys, kan man benytte et høypassfilter som bare slipper igjennom bølgelengder over 790nm. Ved å montere dette på lampen blir rommet bare belyst med IR-lys.

På samme måte kan man også bruke et høypassfilter på videokameraet slik at bare refleksjonene av IR-lyset blir registrert. På denne måten er det mulig å filme publikum selv om det er helt mørkt i rommet. Fordelen med dette er at “stemningen” i lokalet ikke blir ødelagt ved å opplyse det med synlig lys.

2.3.3 Termisk stråling

En annen mulighet er å registrere den varmestrålingen som publikum emitterer. Varmestrålingen er som tidligere nevnt det samme som termisk stråling og ligger høyt i IR-spekteret. Ved lite aktivitet blant publikum kan man anta at de stråler ut lite termisk energi. Etterhvert som publikum blir mer og mer aktive vil strålingen øke og vil fungere som et akkumulert mål for aktivitet hos publikum. Muligens kan den også være et godt mål på deres fysiske aktivitet.

Den termiske strålingen har en bølgelengde fra 10.400nm til 12.500nm. Vanlige CCD-brikker som brukes i digitalkamera registrerer som tidligere nevnt bølgelengder i IR-området, men ikke så høye verdier. For å kunne registrere disse bølgelengdene trenger

man andre typer brikker. Prisen for et kamera med slik teknologi er dessverre høy [20] og prosjektet har ikke hatt tilgang på et slikt kamera. På grunn av prisen er heller ikke løsningen aktuell hvis systemet skal holde en akseptabel pris.

I tillegg til høy pris på utstyret kan en annen ulempe ved å bruke termisk stråling som aktivitetsmål være at det trolig er en treg måleverdi. Det vil være en forsinkelse fra personen er fysisk aktiv, til personen blir varm og emitterer ut varmestråling. Det samme vil skje motsatt vei, ved at personen stråler varme en stund etter avsluttet fysisk aktivitet. Siden systemet skal reflektere aktiviteten blant publikum har det krav til kort forsinkelse. Forsinkelsen som finnes i den termiske strålingen kan være for lang og derfor være en ulempe.

2.3.4 Bevegelsesdeteksjon

Det finnes flere metoder for å detektere bevegelse blant objekter i en videofilm. I MPEG benyttes det en blokkorientert metode. Det vil si at hvert bilde deles opp i et bestemt antall blokker. Den forandringen som oppstår innenfor hver blokk, kodes hver for seg sammen med eventuell forflytning. Ved hjelp av bevegelsesvektorene som er en del av MPEG-kodingen, er det senere mulig å detektere objektenes bevegelse i videoen [21] [22].

Det er også mulig å se på hjørner eller kanter i bildet. Ved å registrere hvor mye disse flytter på seg får man et mål på bevegelsen i bildet. Hjørne- og kantdeteksjon bruker forskjellige matriser som i praksis fungerer som filtre og filtrerer vekk alt annet en hjørnene eller kantene i bildet. Moravec og Harris er eksempler på slike matriser [23]. Figur 7 viser eksempel på matrisen til en Sobel kantdetektor. Matrisene finner henholdsvis loddrette og vannrette kanter.

-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Figur 7: Matrisen til Sobel kantdetektor

En annen metode går ut på å sammenlikne to etterfølgende bilder i videoen [24] [25]. Hvis forskjellen på to etterfølgende bilder er stor kan det tyde på bevegelse i scenen. Forskjellen mellom de to bildene kan uttrykkes med formelen:

$$\Delta_n = |I_n - I_{n-1}|$$

I_n er bildet med rammenummer n og Δ_n er forskjellen mellom de to bildene ved dette rammenummeret. Hvis Δ_n overstiger en bestemt verdi, antar man at det har vært signifikant bevegelse i scenen. Denne verdien, som ofte blir kalt grenseverdi eller terskelverdi, kan bestemmes på forhånd eller kontinuerlig justeres på bakgrunn av innholdet i videoen.

Ulempen med histogrammetoden er at den ikke fungerer godt for bevegelsesdeteksjon hvis kameraet står stille eller hvis objektene beveger seg “innenfor” det området som filmes. Objektene må for eksempel forlate scenen, forandre størrelse eller farge for at det skal bli en forandring i histogrammet.

I programmet EyesWeb (omtales i kapittel 2.4.2), blir bevegelsesdeteksjon utført ved å se på silhuetten til objektene. På bakgrunn av forflytningen til disse silhuettenes regnes det ut hvor mye objektene beveger seg. EyesWeb har en egen “blokk” som utfører denne bevegelsesdeteksjonen. Blokken har forkortelsen SMI som står for *Silhouette Motion Image*. SMI-blokken har en inngang som blir matet med en tersklet videostrøm i sort/hvitt og to utganger. Fra den ene utgangen sendes en videostrøm som viser silhuettbildet. Bildet i figur 8 er en sammensetning av dette silhuettbildet og bildet fra den originale videoen. De røde feltene viser hvor i bildet det er detektert bevegelse. Den andre utgangen til SMI-blokka representerer bevegelsen i bildet med en tallverdi. Formelen nedenfor viser hvordan tallverdien blir regnet ut.

$$\text{Bevegelse} = \text{Område}(\text{SMI}[t, n]) / \text{Område}(\text{Silhuett}[t])$$

Tidspunktet i bildestrømmen blir angitt av variabel t . Variabel n angir hvor mange bilder som brukes for å bygge opp silhuetten. *Bevegelse* er tallverdien som angir hvor mye bevegelse det er i bildet. Ved å dividere området som beveger seg på hele objektets masse, blir denne verdien skalert i forhold til størrelsen på objektet. Dette gjør at størrelsen på objektet ikke påvirker verdien som angir bevegelsen, og gir forskjellige verdier for eksempel når objektet nært og fjernt fra kameralinsen.



Figur 8: Bilde fra videoen etter bevegelsesanalysen. De røde feltene viser hvor i bildet det er detektert bevegelse.

2.3.5 Glatting

Aktivitetsverdien skal gjenspeile publikums aktivitet som en gruppe. Det vil si at aktiviteten som registreres må være til stede over en viss tid, før det er interessant å “reflektere” den tilbake til hele publikumsmassen i form av forandring i lyd eller bilde.

Aktivitetsverdien som blir kalkulert av bevegelsesanalysen kan variere mye på grunn av støy, feilmålinger og objekter som kommer nært og delvis dekker hele sensoren. Det er ønskelig å "glatte" denne verdien for å fjerne ekstremverdiene som oppstår innenfor et kort tidsintervall og derfor ikke er representative for alle publikummerne.

Glattingen kan gjennomføres ved at aktivitetsverdien plottes på en tidsskala, sammen med et tidsvindu som flytter seg i forhold til tidsskalaen. Innenfor tidsvinduet kan det regnes ut en verdi som er representativ for hele tidsvinduet. En medianverdi vil for eksempel fjerne ekstremt lave og høye verdier, og på den måten glatte aktivitetskurven. Under prosjektet ved bruk av akselerometer som sensor (jfr. kap. 2.1.4), viste det seg at 10 sekunders glatting fungerte bra for å eliminere virkningen av sporadiske bevegelser. Andre glatteverdier ble benyttet for å kunne forutsi om aktiviteten var økende eller synkende [3].

I tillegg til glatting kan det være aktuelt å benytte relative terskler i forhold til triggering av hendelser og lignende. En mulighet er at disse justeres i forhold til medianverdien innenfor bestemte tidsintervall. På den måten blir systemet mer dynamisk og krever mer eller mindre av publikum i forhold til hvor aktive de er i utgangspunktet.

2.3.6 Aktivitetsverdi

Det er også vanskelig å vite hva som er korrekt aktivitetsverdi i forhold til den aktiviteten som finnes i rommet. For å kunne sammenligne forskjellige verdier bør det benyttes en metrikk. Dessverre er det ikke funnet en god metrikk som kan benyttes i denne sammenhengen, men en løsning kan være å bruke en referansefilm. Ut fra denne filmen kan det lages en fasit for hvordan aktiviteten i denne videoen bør være. Ved hjelp av eksperimenter og analyse er det mulig å komme frem til det oppsettet hvor aktivitetsverdien ligner mest på fasiten. Denne kan lages ved å la en brukergruppe angi aktiviteten de oppfatter på referansefilmen. Det kan for eksempel brukes en skala fra 0-1 der ingen aktivitet tilsvarer 0 og maks aktivitet tilsvarer 1.

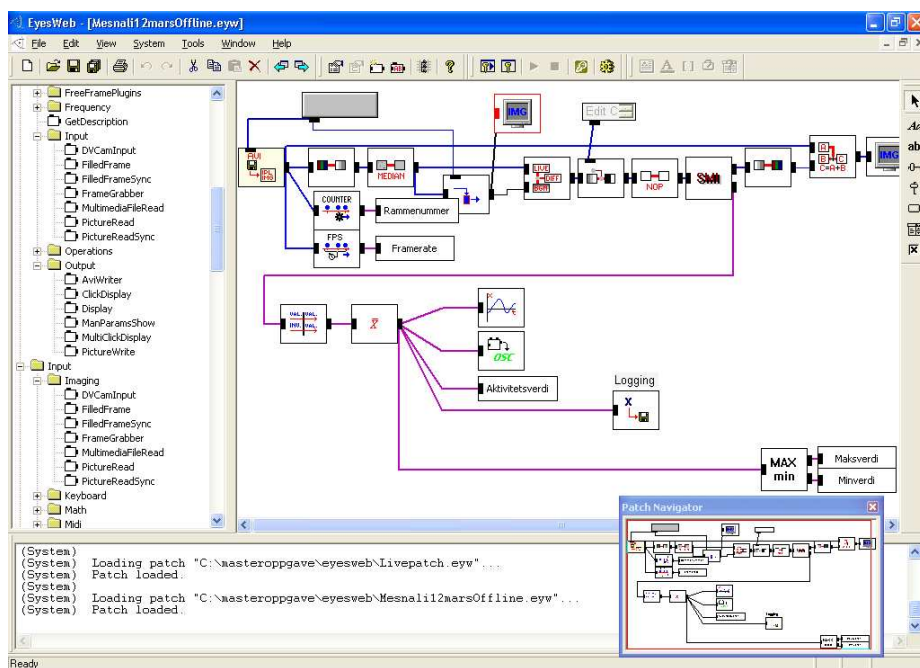
Samtidig er det vanskelig å vite om resultatet fra en slik evaluering kan brukes på annet enn denne ene referansefilmen. Aktivitetsverdien som brukergruppa angir vil være farget av personene i brukergruppa og av innholdet i filmen. Faktorer som antall personer, musikkstil, danseferdigheter og kulturell bakgrunn vil være avgjørende for hvor stor aktiviteten i referansefilmen er. Derfor er det også vanskelig å vite om den fasiten brukergruppen kommer frem til er representativ slik at den kan brukes for å kalibrere systemet.

For å kunne gjennomføre eksperimenter uten en slik referansefilm eller metrikk, har aktivitetsverdien fra en video med en bestemt rammefrekvens blitt brukt som referanse. Dette har først og fremst vært en utfordring i eksperimentet som skulle fastsette rammefrekvens (kapittel 3.1). I dette eksperimentet ble det brukt en video med rammefrekvens på 25 rps (rammer per sekund) som referanse. Dette er valgt fordi 25 rps er den høyeste oppløsningen på kameraet og trolig gir et godt bilde av virkeligheten siden det er brukt som standard ved videoopptak (PAL).

Når det gjelder hva som er "korrekt" aktivitetsverdi er det vanskelig å lage generelle regler fordi dette også er en kunstnerisk avgjørelse.

2.4.2 EyesWeb

EyesWeb [29] er også et grafisk programmeringsmiljø. Programmet muliggjør analyse av video og lyd, og analysen kan utføres på eksisterende videomateriale eller i sanntid. Programmeringsmiljøet er også her basert på blokker som kobles sammen ved hjelp av signalgater (se figur 10). Resultatet fra analysen kan vises på skjerm eller logges til fil. Ved å installere ekstra bibliotek legges det til blokker som utfører bevegelsesanalyse slik som SMI-blokka som er benyttet i prosjektet. EyesWeb er utviklet ved InfoMus Lab ved universitetet i Genova i Italia. Prosjektet har brukt EyesWeb versjon 3.3.0.



Figur 10: Programmeringsmiljøet i EyesWeb

2.4.3 OpenSound Control

OpenSound Control (OSC) er en kommunikasjonsprotokoll som er utviklet ved CNMAT³ ved University of California, Berkeley [30]. Protokollen brukes for kommunikasjon mellom datamaskiner, synthesizere og andre multimediaenheter. Den er optimalisert for moderne nettverksteknologi ved at den blant annet benytter URL-syntaks for adressering [31]. OSC kan brukes i datamaskininterface, wide-area og local-area nettverksbaserte distribuerte musikkssystemer, interprosess-kommunikasjon og i enkeltstående applikasjoner. Spesifikasjoner på OSC finnes på nettsidene til CNMAT [32].

³Center for New Music and Audio Technologies

3 Teknologiske avklaringer med videokamera som sensor

Dette kapittelet tar for seg utfordringer og tekniske løsninger ved bruk av videokamera som sensor. Det er gjennomført flere eksperimenter for å komme frem til et best mulig systemoppsett på bakgrunn av de ressurser som var tilgjengelig under prosjektet. Oversikten viser gjennomførte eksperimenter og motivasjonen for disse:

- 3.1 Rammefrekvens. Motivasjon: redusere datamengden som skal analyseres
- 3.2 Kamerastøy. Motivasjon: redusere mulig feildeteksjon av bevegelse på grunn av egenstøy i kameraet
- 3.3 og 3.4 IR-filter. Motivasjon: redusere feildeteksjon av bevegelse på grunn av lys-scannere

3.1 Rammefrekvens

For å redusere belastningen av systemet og sørge for best mulig ytelse i sanntid er det en fordel å redusere datamengden som skal prosesseres. En måte å gjøre dette på er å redusere antall bilder som skal sammenlignes per sekund. I praksis kan dette gjøres ved å redusere rammefrekvensen på videostreamen fra kameraet.

Ved reduksjon av rammefrekvensen vil videoen oppleves mer hakkete. Dette skjer fordi reduksjonen av antall bilder fjerner noe av informasjonen om objektenes bevegelse i bildet. Resultatet er at objekter i bevegelse vil kunne gi et annet utslag når man foretar bevegelsesanalyse av videoen enn det som er ønskelig. Det ble derfor gjennomført et eksperiment for å kunne foreta et kvalifisert valg av hvilken rammefrekvens som skulle brukes under resten i prosjektet.

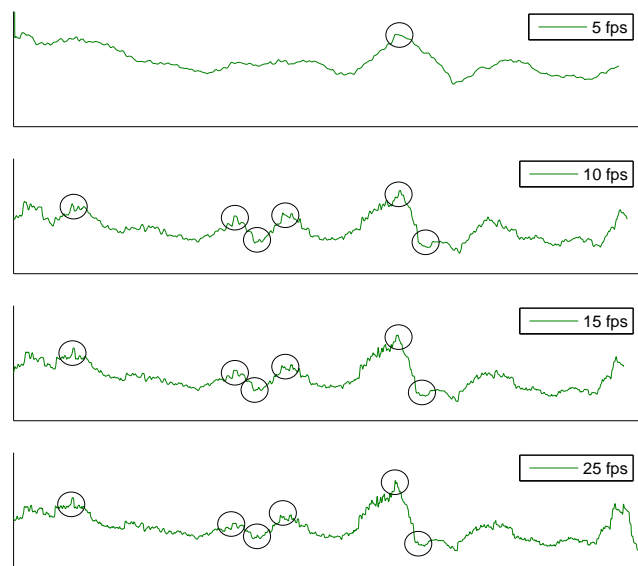
Videoen som ble brukt for analysen ble tatt opp under en konsert ved Viken folkehøgskole og viser publikum i bevegelse (jfr. figur 11). Videokameraet ble plassert på et stativ på den ene siden av scenen og var rettet ut mot publikum.



Figur 11: Bilde fra analysevideoen

Det ble benyttet et Sony DV-kamera med nightshot-funksjon uten ekstra belysning. Etter opptaket ble det laget nedskalerte kopier av videoen med oppløsning på 320x240

For å gjøre det lettere å sammenligne kurvene ved de forskjellige rammefrekvensene ble det brukt et tidsmedianfilter. Tidsmedianfilteret foretar en glatting av kurven over de tre siste sekundene. Resultatet er at ekstremverdier blir filtrert bort og kurvene blir jevnere (jfr. kapittel 2.3.5). Figur 14 viser kurven for aktiviteten ved henholdsvis 5, 10, 15 og 25 rps. Kurvene er linjert og skalert for å gjøre de lettere å sammenligne. De er derfor uten benevning på aksene, men rådataene kan studeres i vedlegg B. Figuren viser at detaljrikdommen endres ved forskjellig rammefrekvens. Siden objektene rekker å flytte seg lenger ved lav rammefrekvens forandres også forholdet mellom laveste og høyeste aktivitetsverdi slik det er vist i tabell 1. Dette kan skaleres med en multiplikasjonsfaktor, men detaljrikdommen er tapt på grunn av nedsamplingen som rammereduksjonen innebærer.



Figur 14: Aktivitetskurven med 3 sekunders glatting

Frame rate	5	10	15	25
Min	0.111	0.036	0.025	0.012
Max	0.245	0.149	0.117	0.063

Tabell 1: Max/min verdier med 3 sekunders glatting

Alle kurvene viser den samme tendensen men de tre nederste kurvene har flere detaljer enn den øverste. Dette er illustrert ved hjelp av sirkene. Det er altså større likhet mellom kurvene til rammefrekvensene 10, 15 og 25, enn kurven for rammefrekvens 5. Siden målet er å beholde flest mulig detaljer men samtidig få lavest mulig datastrøm, vil en rammefrekvens på 10 rammer per sekund være et godt kompromiss. Ved 10 rps beholder man de fleste detaljene og får samtidig redusert datamengden til under det halve.

I følge samplingsteoremet må samplingfrekvensen være minst det dobbelte av frekvensen på forandringen det er ønskelig å registrere. Ved filming av personer i bevegelse vil man ved 10 rps maksimalt kunne registrere 5 bevegelser per sekund. Dette forutsetter

at bildene blir tatt på rett tidspunkt i forhold til utslaget på bevegelsen. Hvis personen beveger seg hurtigere enn dette kan det oppstå en aliasingeffekt. Ved 10 rps vil dette oppstå hvis aktiviteten har en frekvens høyere enn 5 Hz. Resultatet er at systemet registrerer lav aktivitet i stedet for høy. I praksis er det lite sannsynlig at publikum klarer å bevege seg så hurtig og begrensingen er derfor ikke problematisk i denne sammenhengen.

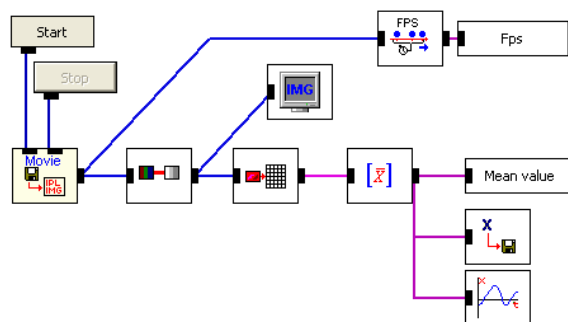
3.2 Kamerastøy

Eksperimentet testet de to webkameraene som var tilgjengelig under prosjektet for å finne ut hvilket som har lavest egenstøy. Dette er interessant fordi bevegelsesanalysen sammenligner to etterfølgende bilder. Egenstøy som fører til forandringer i pikselverdien ved filming av en uniform bakgrunn, kan derfor gi utslag av detektert bevegelse uten at dette er tilfelle. Muligheten for slik feildeteksjon kan reduseres ved å bruke kameraet som har lavest egenstøy. Kameraene som ble testet var to ordinære webkamera med betegnelsen *Sandberg NightCam* [16] (jfr. figur 3) og *Logitech® QuickCam® Messenger™* som er avbildet i figur 15. Høyeste oppløsning på begge kameraene var 640x480 piksler.



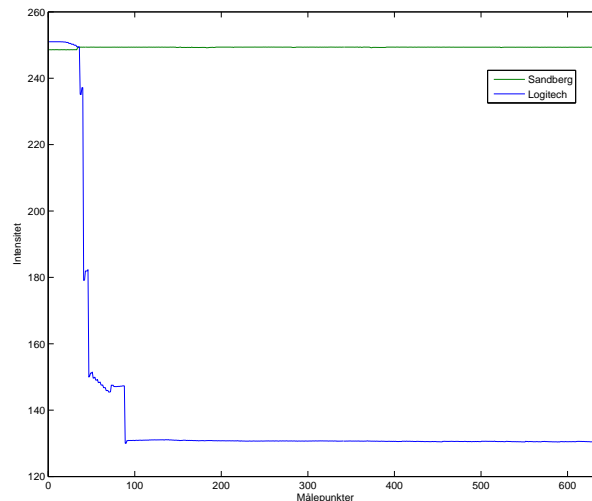
Figur 15: Logitechkameraet som ble testet for egenstøy

Eksperimentet ble gjennomført ved at kameraene gjorde opptak av et hvitt lerret i 1. minutt. Rammefrekvensen var 15 rps og opptaket ble gjennomført med begge kamera samtidig, slik at eventuelle ytre påvirkninger skulle ha lik innvirkning på begge kameraene. Eksperimentet ble gjennomført på HiG sitt fargelaboratorium [33] med kontrollert belysning. Temperaturen på lysstoffrørmaturet var 5000K (D50) og innfallende lys på lerretet ble målt til EV 8,7. For å bestemme graden av støy i videoen ble den først konvertert til gråtoner. På bakgrunn av pikselmatrisen for hvert bilde ble det regnet ut en gjennomsnittsverdi som ble logget til fil. Dette ble programmert i EyesWeb og patchen er vist i figur 16.



Figur 16: Patch for kalkulering av egenstøy

Ved å se på variasjonen i gjennomsnittsverdien er det mulig å få et inntrykk av hvor mye egenstøy det er i kameraet. Siden opptaket viser et uniformt hvitt lerret med konstant lysintensitet skulle kameraene teoretisk sett registrere samme verdi hele tiden. Dette er ikke tilfelle i praksis og figur 17 gir en grafisk fremstilling av variasjonen i intensitetsverdien på de to kameraene.



Figur 17: Intensitetsverdier for Logitech- og Sandbergkamera

Den store variasjonen i intensitetsverdien på de første målepunktene skyldes antakelig kameraets egen driver som foretar automatisk kalibrering. Verdier før målepunkt 100 er derfor ikke tatt med i analysen. For å lettere kunne sammenligne kurvene, er området etter autojusteringen zoomet inn og vist i figur 18.

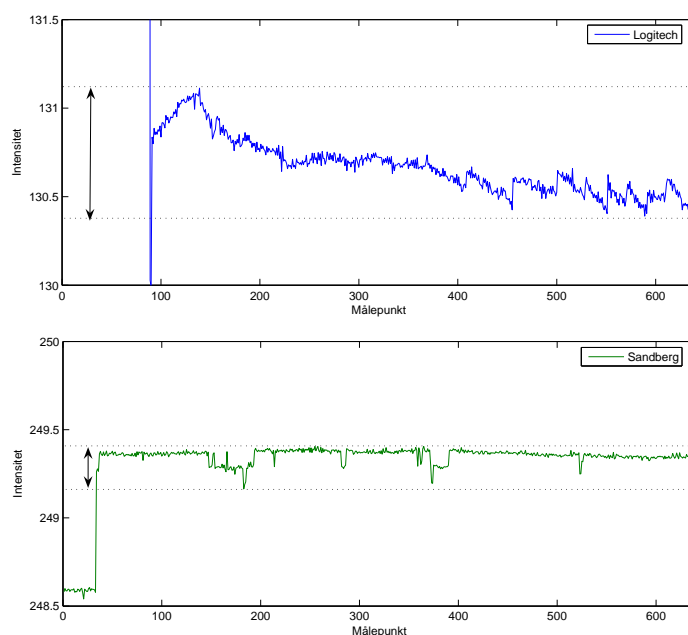
Det kommer tydelig frem at differansen mellom høyeste og laveste intensitetsverdi er forskjellig for de to kameraene:

$$\text{Differanse på Logitech} = 131.113 - 130.390 = 0.723$$

$$\text{Differanse på Sandberg} = 249.405 - 249.164 = 0.241$$

På bakgrunn av dette regnes webkameraet fra Sandberg som det kameraet med lavest egenstøy. Det er derfor brukt videre i oppgaven, mens Logitech-kameraet ikke er testet videre. DV-kameraet fra Sony er ikke med i denne sammenligningen fordi innledende tester viste at det var mulig å gjennomføre bevegelsesanalyse på bakgrunn av datastrømmen fra webkameraene. Dette bekreftes i kapittel 3.3.1 og fordelen er blant annet et rimeligere system.

For å få en mer nøyaktig test kunne man sammenligne etterfølgende bilder piksel for piksel. En slik sammenligning ville ta hensyn til hvor i bildet det var støy og for eksempel om øverste og nederste del av bildet skiftet mellom sort og hvitt mellom hvert bilde. Eksperimentet som ble gjennomført tok ikke hensyn til dette, men ble ansett som god nok i forhold til den bevegelsesdeteksjonen som ble benyttet.



Figur 18: Zoomet og transponert intensitetskurve

3.3 IR-filer

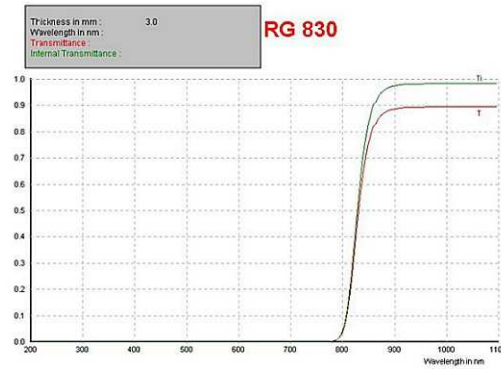
Som det ble nevnt i teorikapittelet brukes det ofte lys for å skape bestemte stemninger under en konsert. Det kan være små og store lyskastere som lyser opp både scenen og området der publikum står. I tillegg brukes det ofte lysscannere som tegner mønster på gulv og tak.

Innledende tester viste at variasjon i slike lyskilder gir stort utslag på bevegelsesdeteksjonen. Dette skjedde fordi bevegelsesanalysen sammenligner lysforholdene i to etterfølgende bilder (jfr. kap. 2.3.4). En forandring av belysningen i rommet gir den samme forandringen i pikselverdi som skapes av objekter i bevegelse. Da systemet ikke er i stand til å skille disse to intensitetsvariasjonene blir begge tolket som bevegelse blant publikum.

Det ble derfor eksperimentert med å benytte et IR-filer for å filtrere bort endringer i synlig lys som kommer fra lysscannere og annet lysutstyr som typisk befinner seg på et utested. Hensikten var å basere bevegelsesanalysen på lysforandringer i IR-området. IR-fileret [34] som ble benyttet til dette formålet hadde betegnelsen B+W 093 og en transmisjonskurve vist i figur 19. Kurven viser at filteret blokkerer alt synlig lys og slipper gjennom bølgelengder i IR-området.

For at kameraet skal kunne "se" i IR-området, må det finnes nok IR-lys i rommet slik at dette kan reflekteres fra objektene. Som beskrevet i kapittel 2.3.2 stråler vanlige glødelamper ut IR-lys. Hvis dette lyset er konstant vil menneskene på dansegulvet bli belyst med et jevnt nivå av IR-lys. Antar man at lysscannere stråler ut lite IR-lys, er det teoretisk sett mulig å filtrere vekk lyset fra lysscanneren som tegner mønster på dansegulvet og forstyrrer bevegelsesdeteksjonen.

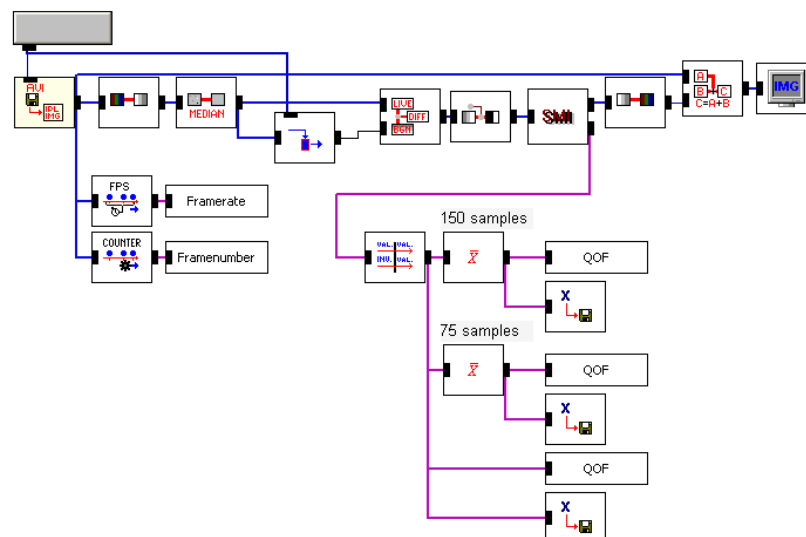
Eksperimentet ble gjennomført ved å benytte et Sandberg NightCam webkamera og et Sony DV-kamera med nightshot-funksjon. Opptakene ble gjort på studenthuset, men det var bare ett IR-filer tilgjengelig under eksperimentet. Første video ble derfor tatt



Figur 19: Transmisjonskurve for IR-filteeret [34]

opp med begge kamera uten IR-filteer for å kunne sammenligne kameraene. Andre video ble tatt opp med IR-filteeret montert på Sony-kameraet og tredje video med IR-filteeret montert på Sandberg-kameraet. Det ble ikke benyttet ekstern IR-belysning men Sony-kameraet hadde innebygd en IR-diode og Sandberg-kameraet hadde 6 IR-dioder. Disse diodene var aktive under alle opptakene. Utover dette ble ikke rommet opplyst mer en det som var naturlig på diskoteket.

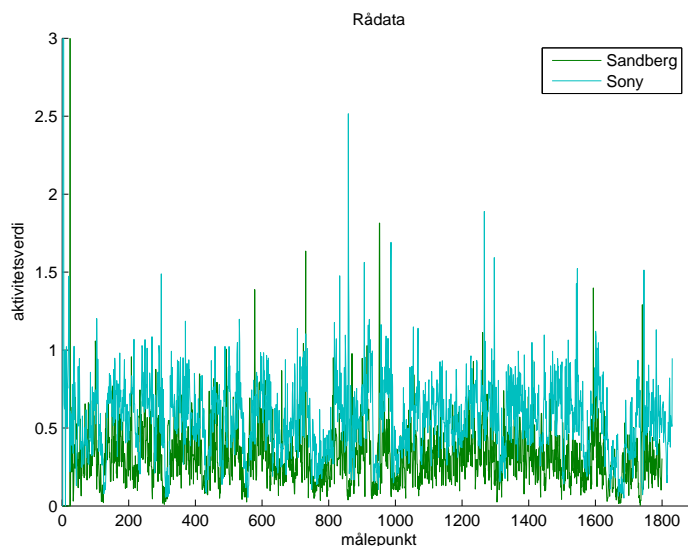
Siden opptakene fra diskoteket var forholdsvis mørke, ble det foretatt en beregning av gjennomsnittlig gråtoneverdi for hver film. Dette ble gjort ved å først konvertere videoen til gråtoner og siden regne ut gjennomsnittlig pikselverdi for hvert bilde. Gjennomsnittet av alle disse verdien ble så regnet ut og brukt som terskelverdi når filmene ble analysert i EyesWeb. Figur 20 viser analysepatchen som prosesserte filmene og laget kurver med forskjellige glatteverdiier. For hver video ble det laget 3 forskjellige kurver som var glattet over 0, 75 og 150 målepunkter. Ved rammefrekvens på 15 rps tilsvarer dette en glatting på 0, 5 og 10 sekunder.



Figur 20: Analysepatch i EyesWeb for IR-eksperimentet

3.3.1 Kamerasammenligning uten IR-filter

Figur 21 viser aktivitetskurvene fra begge kameraene uten bruk av IR-filter. Kurvene til aktivitetsverdien er vanskelig å sammenligne siden de varierer med høy frekvens. Analysen er derfor foretatt etter 5 og 10 sekunders glatting. Glatting i dette tidsområdet ble også brukt ved testing av systemet og ble valgt på bakgrunn av litteraturstudiet og egne eksperimenter.

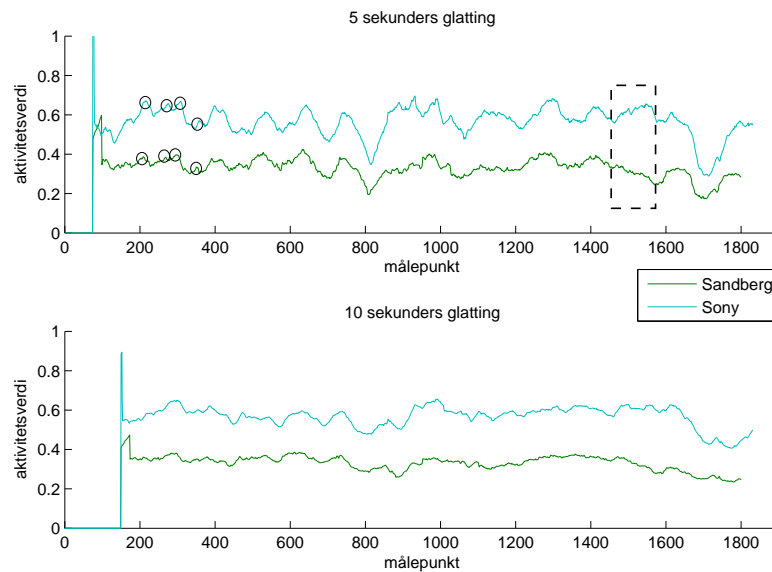


Figur 21: Aktivitetskurvene slik de ser ut uten glatting

Aktivitetskurvene fra begge kameraene etter glatting er vist i figur 22. Her er det også gjort markeringer for noen av sammenligningspunktene. Figuren viser at kurvene i stor grad har samme kontur, bortsett fra helt i starten og innenfor det avmerkede rektangelet. Forskjellen i starten kommer antakelig av at analysepatchen og de to kameraene bruker litt tid for å justere seg første gang de får signal på inngangen. Avviket markert med rektangelet er vanskelig å forklare, men det kan skyldes målefeil eller at bare dette kameraet har registrert økt aktivitet. Avstanden mellom kameraene var ca 20 cm så dette var fysisk mulig, men kan ikke bekreftes ved visuell gjennomgang av videoen.

Sandberg-kameraet har vidvinkel og dette betyr i praksis at det klarer å fange opp større deler av dansegulvet. Prosentvis andel av områder som er i bevegelse blir derfor mindre enn ved filming med Sony-kameraet uten vidvinkel. Resultatet er at det registrerer mindre aktivitet enn Sony-kameraet selv om begge kameraene filmer den samme menneskemassen. Trolig er dette forklaringen på kurvenes dynamikkforskjell. Valget om man skal bruke et kamera med eller uten vidvinkel, bør derfor gjøres i forhold til rommets størrelse og antall personer som deltar. Best mulig "oppløsning" ved registrering av aktiviteten får man hvis kameraet bare filmer det området hvor publikum danser. Teoretisk vil da høyest utslag oppstå når alle personene er i bevegelse og lavest mulig utslag oppstå når alle står stille.

Sony-kameraet som er brukt i eksperimentet koster rundt 5000 kroner og Sandberg-kameraet koster rundt 300 kroner. På tross av denne prisforskjellen tyder kurvene på at Sandberg-kameraet klarer å fange opp aktiviteten like godt som Sony-kameraet. Det



Figur 22: Aktivitetskurvene med glatting

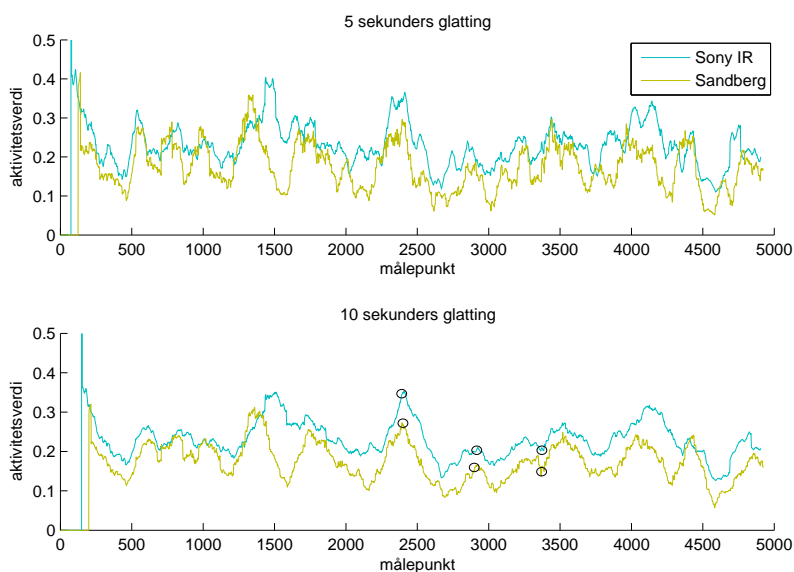
ser derfor ut til at systemet fungerer like bra med et rimelig webkamera som et dyrt DV-kamera. Dette kan forklares ved at DV-kameraet består av mer enn bare CCD-brikke og linse. Ved kjøp av DV-kamera betales det mye for opptaksmulighet på tape, LCD-skjerm, menyer, batteri, lader og lignende. Det bør allikevel testes videre siden opptaket som ble brukt for å analysere kameraene var fra et tomt lokale med varierende lysmønster på dansegulvet. Mønsteret på dansegulvet blir tegnet opp av en lysscanner. Prinsippet blir det samme, men testen bør også gjennomføres med personer i bevegelse for å underbygge teorien.

3.3.2 Kamasammenligning med IR-filter

Etter kamasammenligningen ble IR-filteret montert på Sony-kameraet. Figur 23 viser aktivitetskurvene for begge kameraene med henholdsvis 5 og 10 sekunders glatting. Opptaket ble gjort med Sandberg-kameraet og startet litt senere enn opptaket med Sony-kameraet. Kurven fra Sandberg-kameraet er derfor flyttet 50 målepunkter mot høyre for å kunne sammenligne kurvene på samme tidspunkt. Ved 15 rps tilsvarer dette ca. 3 sekunders forskyvning, noe som resulterer i at starttidspunktet på kurvene er forskjellig. Ved å sammenligne flere punkter, slik sirklene angir, ser man at kurvene i stor grad har samme kontur. Det er riktignok et stort avvik rundt målepunkt 1500, men ellers er kurvene ganske like. Hva dette avviket skyldes er vanskelig å si men det kan være målefeil. Videoen på dette punktet viser ingen vesentlig forskjell som skulle tilsi et slikt fall i kurven.

Det ble nevnt i kapittel 2.3.6 at det er vanskelig å lage en fasit for aktivitetskurven og det finnes derfor heller ikke noen "fasitkurve" å forholde seg til. Dette gjør det vanskelig å sammenligne de forskjellige kurvene fra kameraene og bestemme hvilken som gjengir virkeligheten best. På tross av dette er det mulig å sammenligne i hvor stor grad kurvene følger hverandre.

Det ble også fremsatt en hypotese om at IR-filteret skulle fjerne bølgelengdene som



Figur 23: Aktivitetskurve med IR-filter på Sony-kameraet

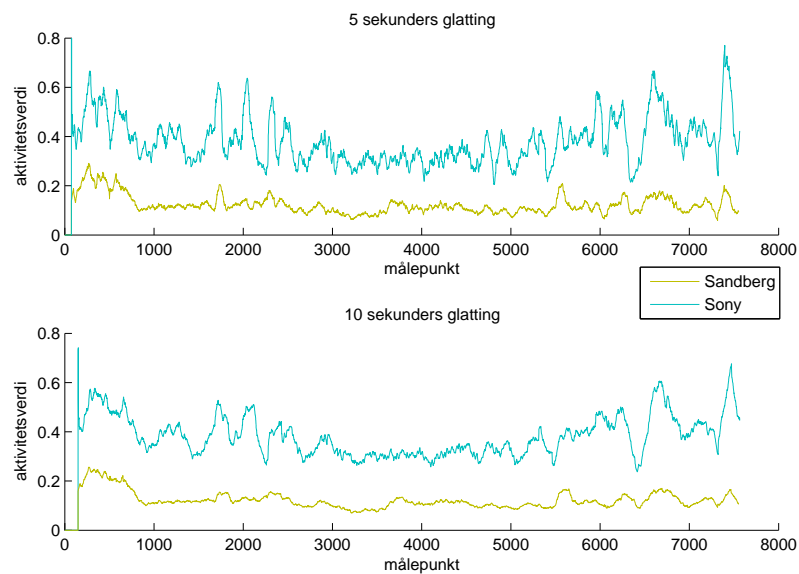
lys-scanneren sendte ut og på den måten filtrere vekk feilaktig bevegelsesdeteksjon som denne innførte. Hvis dette hadde vært vellykket, ville effekten trolig kommet frem i form av en “roligere” aktivitetskurve fra kameraet som filmet gjennom IR-filteret. Det ser imidlertid ikke ut til at det er noen nevneverdig forskjell på kurven som representerer aktivitetsverdi med IR-filter og den uten.

En forklaring på dette kan være at IR-filteret ikke har noen hensikt eller at det er andre faktorer som har spilt inn. Ut fra visuell inspeksjon av videoen ser det ut til at lyset som varierer, i tillegg til lys i det synlige spekteret, sender ut lys i IR-området slik at analysen blir lurt av dette.

Det er tidligere nevnt at det må finnes tilstrekkelig med stabilt IR-lys i rommet. Hvis ikke vil det heller ikke være mulig å registrere den bevegelsen som er ønskelig å detektere. For å løse dette kan det være en mulighet å flombelyse hele rommet med IR-lys. Dette vil også redusere innvirkningen fra IR-lys som lyskasterne stråler ut. Eventuelle forandringer i lysintensiteten på dette IR-lyset, vil kunne “viskes ut” av sterke IR-lamper som brukes til flombelysning. Ulempen er at IR-belysning er kostbart og det var derfor ikke tilgjengelig for testing under prosjektperioden.

Figur 24 viser aktivitetskurvene fra filmen der IR-filteret er montert på Sandberg-kameraet. Aktivitetskurven til Sandberg-kameraet med IR-filter ser betraktelig roligere ut enn den fra Sony-kameraet. Muligens kan dette forklares med “vidvinkel-teorien” som er nevnt ved analysen av figur 21, det kan også tenkes at filteret klarte å fjerne noe av “støyen” fra lysscanneren.

På bakgrunn av de refererte eksperimentene er det altså ikke mulig å se en vesentlig forbedring ved bruk av et IR-filter. Det skal likevel nevnes at ved visuell inspeksjon av videoene kan se at mange av lysstrålene som sendes ut fra scanneren ikke er synlig når det filmes gjennom IR-filteret. Når lysstrålene treffer mennesker eller andre objekter reflekteres de godt, men den tiden de er i lufta oppfattes de sterkt redusert i forhold til opptaket uten IR-filter. Dette er vist i figur 25.



Figur 24: Aktivitetskurvene med IR-filter på Sandberg-kameraet



Figur 25: Utsnitt av videoen fra diskoteket uten og med IR-filter

Det kunne vært interessant å gjøre tester med flere lyskastere som lyste opp rommet. På et diskotek benyttes det ofte mindre statisk belysning siden det er ønskelig at mønstrene fra lysscannerne skal komme tydelig frem. Det hadde allikevel vært interessant å se om glødelampene fra mange lyskaster hadde vært nok, slik at det ikke var nødvendig med ekstern IR-belysning.

3.4 IR-filter og lysscanner

Resultatene i forrige kapittel klarte ikke tallmessig å påvise at IR-filteret reduserte feildeteksjon av aktivitet som følge av lysscanneren. Visuell betraktning av videoene antydte likevel at noe av lyset fra scanneren ble filtrert vekk. På bakgrunn av denne observasjonen ble det foretatt nye eksperimenter under mer kontrollerte omgivelser.

Resultatet fra disse eksperimentene viste at IR-filteret klarte å fjerne lysstrålene fra scannere som sender ut lys med en fargetemperatur på 6000 kelvin. Figur 26 viser opptak av lyskjeglen til en Roboscan Pro 518 med HSD250 pære. Lysscanneren belyser gulvet på ett bestemt sted men skifter mellom forskjellige farger. Bilde *a* er fra opptaket uten IR-filter og bilde *b* er fra opptaket med IR-filter.

Det samme eksperimentet ble utført med en lysscanner som hadde pære med far-



Figur 26: Bilde a er fra videoopptaket uten IR-filer og bilde b er fra opptaket med IR-filer. Pæren i scanneren hadde en fargetemperatur på 6000 kelvin på begge opptakene.

getemperatur på 3200 kelvin. Rommet ble i tillegg delvis opplyst med vanlige lyskanner. Pærene i disse kannene hadde en fargetemperatur på 3100 kelvin. Figur 27 viser et bilde fra videoopptaket gjort med IR-filer. Her var det mulig å se lyset fra scanneren og lyskannene på gulvet. Disse bølgelengdene slipper altså igjennom filteret og vil påvirke bevegelsesdeteksjonen.



Figur 27: Bilde fra videoopptaket med IR-filer. Pæren i lyskannen hadde en fargetemperatur på 3100 kelvin og scanneren hadde en fargetemperatur på 3200 kelvin.

Eksperimentet viser at det er mulig å blokkere synlig lys ved hjelp av IR-filer under bestemte lysforhold. I tillegg til dette må det også bevises at IR-fileret ikke fjerner eventuell bevegelse i rommet hvis resultatet skal kunne benyttes til bevegelsesanalyse. Det ble derfor gjort opptak av en person som beveget seg i rommet. Figur 28 viser et bilde fra videoen hvor personen beveger seg, samtidig som lysscanneren med pæren på 6000 kelvin belyser et punkt på gulvet. Bildet viser at personen fortsatt er synlig selv om lyskjeglen på gulvet er filtrert vekk. Det ser ut til at det er mulig å foreta en bevegelsesanalyse som bare tar hensyn til personens bevegelser, selv om rommet belyses av en lysscanner.



Figur 28: Bilde fra videoopptaket med IR-filter og bevegelse. Pæren i lysscanneren har en fargetemperatur på 6000 kelvin og blir filtrert vekk av IR-filteret.

4 Systemoppsett og brukertest

Kapittelet forklarer i korte trekk hvordan den endelige prototypen ble bygget. Det vil si hva de forskjellige programmene gjør og hvordan de kommuniserer.

Hensikten med brukertestingen var å benytte systemet i en reell situasjon og se hvordan det fungerte. Testene skulle finne ut om systemet lyktes med å tilby deltakerne interaktivitet slik at de opplevde at de kunne påvirke lyd og bilde og hva de syntes om denne interaktiviteten. Det var også interessant å finne ut om utøveren, som i dette tilfellet var en DJ, opplevde systemet som positivt.

4.1 Forberedelser

I kapittel 1 ble det nevnt eksempler på forskjellige elementer det var aktuelt at publikum kunne påvirke under konserten. Noen av disse mulighetene ble testet i praksis, og dette er dokumentert i underkapitlene som følger.

4.1.1 Bevegelsesintervall og patcher

Med bevegelsesintervall menes det verdiområdet som befinner seg mellom laveste og høyeste aktivitetsverdi. Dette intervallet skal mappes til et verdiområde hvor de forskjellige triggerverdiene for animasjonen og samplene befinner seg. Det er vanskelig å forutsi hvor stor aktivitet det vil være blant publikum. Aktivitetsverdien som kommer fra Eyes-Web blir derfor tatt imot og skalert i PD. Patchen i PD gjør det mulig å skalere verdien opp eller ned avhengig av hvor godt sensoren registrerer bevegelse og hvor stor publikumsmassen er.

4.1.2 Plassering av kamerasensor

Hensikten med kameraet er å fange opp den samlede aktiviteten blant alle deltakerne. Det må derfor plasseres slik at det fanger opp flest mulig personer. Hvis man ikke klarer å fange opp et stort nok antall personer, kan man risikere at noen få personer som er veldig aktive får uforholdsmessig stor innflytelse på resultatet.

Under brukertestene som er gjengitt senere i dette kapittelet, er kamerastativet med kameraet plassert på et bord. På denne måten får man fanget opp bevegelse lenger bak i lokalet og ikke bare de personene som står nærmest kameraet. Kameraet er vinklet på skrå nedover mot publikum slik at det filmer hele dansegulvet.

Hvis lokalet er stort kan det benyttes flere kamera. Aktivitetsverdien kan da beregnes på flere steder i lokalet og verdiene kan for eksempel kombineres ved å regne ut et gjennomsnitt, eller at de benyttes hver for seg for å kontrollere forskjellige elementer.

4.1.3 Valg av musikk

Bevegelsesmønsteret til publikum har mye å si for registreringen av aktiviteten og hvor stort verdiområde aktivitetsverdien strekker seg over. Publikum beveger seg forskjellig til ulike typer musikk. Systemet er avhengig av variasjon i bevegelse mens musikken spilles, og ikke bare etter hvert musikkstykke slik som ved applaus i en konsertsal. For å få testet systemet er det best å benytte en type musikk som gjør at publikum får lyst til å danse

eller være i aktivitet under hele konserten. Det ble derfor innledet et samarbeid med en DJ som spiller house-gospel. En av fordelene med dette var at DJ-en ordnet musikk. Det var derfor ikke nødvendig å lage musikk spesielt for anledningen og man kunne begynne testingen med en gang.

4.1.4 Publikumsinteraktiviteten

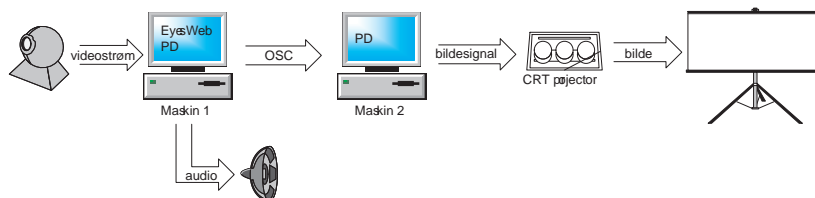
For å gjøre interaktiviteten multimedial ble det bestemt at publikum skulle få muligheten til å påvirke både lyd og bilde. Dette ble i praksis ordnet ved at de kunne påvirke en abstrakt animasjon som ble projisert opp på et lerret. I tillegg kunne de styre avspillingen av lydsamples.

Både animasjonen og lydsamplene reflekterte publikums aktivitetsnivå. Animasjonen bestod av en ball som vokste og minket, skiftet farge og roterte avhengig av aktivitetsverdien. Bakgrunnen bak ballen forandret også lysintensitet og farge. Hvis aktivitetsverdien passerte et bestemt nivå “sprengte” ballen og gikk inn i neste “energinivå”.

Avspillingen av lydsamplene ble også styrt av publikum. Det ble forhåndsdefinert 3 aktivitetsnivå som trigget 3 forskjellige samples når deltakerne passerte disse nivåene. Samplene som ble brukt hadde ikke et harmonisk sentrum. Fordelen med dette var at de kunne spilles av når som helst uten at de brøt med resten av musikken. Samplene ble avspilt på samme PA-anlegg som resten av musikken.

4.1.5 Maskinoppsett

Under brukertestene ble det to datamaskiner. Figur 29 viser hvordan disse maskinene og de andre enhetene var koblet sammen.



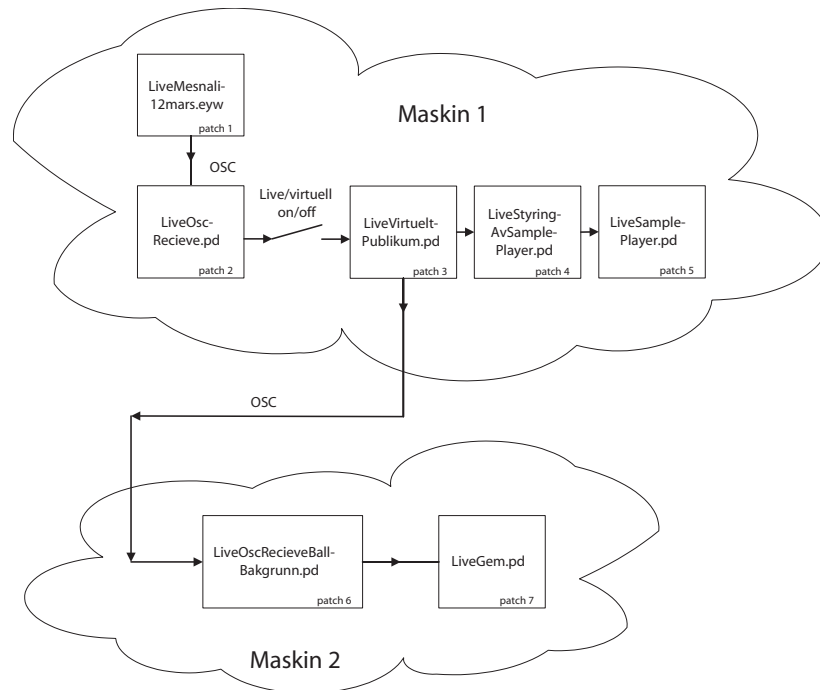
Figur 29: Maskinoppsett

Som **maskin 1** ble det brukt en Fujitsu Siemens Lifebook med 1GHz Pentium III CPU og 256MB RAM. Denne fungerte som analysemaskin og kjørte programmene EyesWeb og PD på Windows XP plattform. Sensoren som var tilkoblet var et Sandberg NightCam webkamera (jfr. fig. 3). Beregningen av aktivitetsverdien ble foretatt i EyesWeb og overført til PD internt i maskinen ved hjelp av OSC. Ved hjelp av forskjellige patcher i PD ble en sampleavspiller trigget. Denne er også programmert i PD og spiller av lydsamplene på samme PA-anlegg som resten av lyden under konserten.

Ved hjelp av OSC blir aktivitetsverdien også sendt over nettverket til **maskin 2**. Denne maskinen rendrer grafikk på bakgrunn av aktivitetsverdien. Maskinen og rendringen blir delvis fjernstyrt fra **maskin 1** over OSC-protokollen. Skjermutgangen til denne maskinen er koblet direkte til prosjektoren som viser bildet på lerretet. **Maskin 2** hadde en 500 MHz Pentium CPU, 256 MB RAM og kjørte Windows XP og PD.

4.1.6 Programoversikt

Figur 30 gir en oversikt over de forskjellige programpatchene i systemet. Navnet på blokkene tilsvarer patchene de representerer. Nedenfor gis en kort forklaring for hver enkelt blokk og programkoden finnes i vedlegg C.



Figur 30: Oversikt over hvilke patcher som kjøres på maskin 1 og 2

Patch 1 er den eneste patchen som er programmert i EyesWeb. Den tar imot datatrømmen fra webkameraet, foretar bevegelsesdeteksjon og regner ut en aktivitetsverdi. Aktivitetsverdien blir glattet over et forhåndsdefinert tidsintervall (jfr. kap. 2.3.5). Både video og aktivitetsverdi blir logget til fil. Patchen sender også aktivitetsverdien videre til PD. Verdien blir sendt over OSC-protokollen internt på maskinen.

Patch 2 er første patch i PD. Patchen mottar og skalerer aktivitetsverdien. Det er brukt lineær skalering som justeres manuelt etter antall personer som deltar, og hvor aktive de er. Etter skaleringen sendes verdien videre til patch 3 internt i PD.

Patch 3 inneholder en kurve som representerer et "virtuelt publikum". Aktivitetskurven til det virtuelle publikumet kan kobles inn og tegnes fritt etter ønske. Dette ble brukt under deler av brukertesten for å se om deltakerne oppdaget at de selv ikke hadde kontroll over lyd og bilde. Patchen sender også aktivitetsverdien videre for styring av sampleavspilleren i patch 4. Start og stopp av grafikkrendringen på maskin 2 kontrolleres også fra denne patchen. Dette gjøres ved å sende kommandoer over OSC-protokollen. Aktivitetsverdien blir også sendt til maskin 2 fra denne patchen.

Patch 4 mottar aktivitetsverdien og trigger sampleavspilleren når verdien overstiger fastsatte nivå. Det er fastsatt tre forskjellige nivå som sender triggerimpulser til patch 5.

Patch 5 er selve sampleavspilleren og har på forhånd lest inn 3 forskjellige lydsamples fra disk slik at de ligger klare. Samplene blir spilt av gjennom maskinens lydkort og sendt

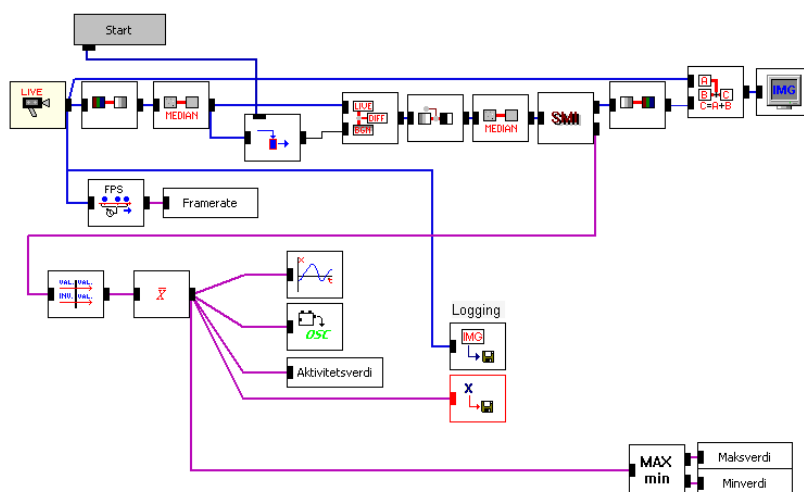
til PA-anlegget.

Patch 6 er første patch på maskin 2 og mottar kontrolldata fra patch 3 fra maskin 1 over OSC-protokollen på nettverket. Kontrolldataene blir hentet ut av datastrømmen og sendt videre til patch 7.

Patch 7 rendrer grafikken som blir vist på lerretet. Grafikken består av en bakgrunn og en ball som forandrer størrelse, farge og rotasjon. Dette styres av aktivitetsverdien som kommer fra patch 6.

4.2 Pre-test Garvern

For å teste systemet og få litt erfaring før brukertesten på Mesnali, ble det gjort en uformell test på utestedet Garvern. Sammen med noen velvillige kamerater danset vi foran kameraet til en CD med musikken DJ-en skulle spille live på Mesnali. Aktivitetsverdien ble kalkulert i EyesWeb og patchen som ble brukt er vist i figur 31. Figur 32 viser verdiene som ble logget til fil.

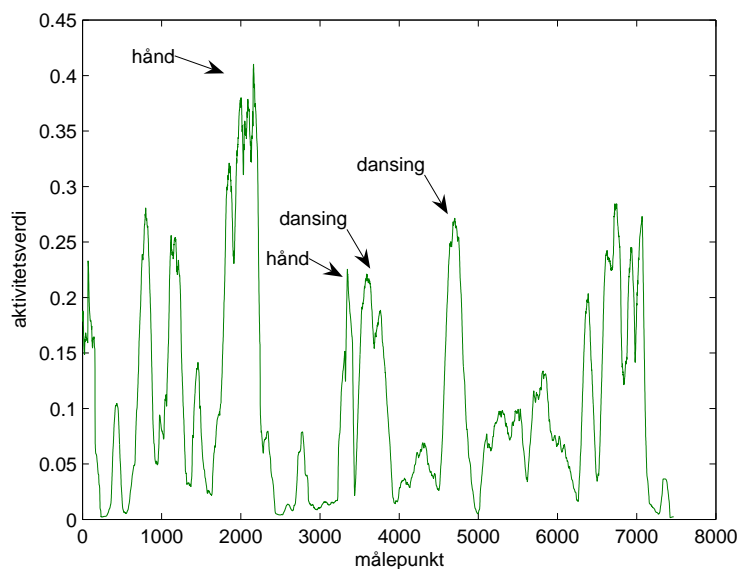


Figur 31: EyesWeb patch brukt på Garvern

Det var bare den visuelle delen som ble testet og det ble avdekket noen svakheter på systemet. Blant annet vokste ikke animasjonen til maksimal størrelse. Det viste seg at en programmeringsfeil gjorde at verdien som kontrollerte animasjonenens størrelse ble skalert ned.

Testen viste videre at webkameraet klarte å fange opp vår bevegelse og at det var relativt lite støy i bildet. En fordel med Sandberg-kameraet som ble brukt var det hadde vidvinkel. Kameraet klarte derfor å fange opp hele dansegulvet. Ulempen med vidvinkel kan være at objekter som befinner seg i ytterkant av kameraobjektivet blir forstørret. Dette kan bety at objektene får forskjellige innvirkning etter hvor i bildet de befinner seg, men SMI-blokka som er brukt i prosjektet tar hensyn til dette (jfr. kap. 2.3.4).

En annen observasjon var at det var vanskelig å få aktivitetsverdien til å overstige 0.3. Dette på tross av store fysiske anstrengelser fra deltakerne. Det ble derfor innført en justerbar multiplikasjonsfaktor slik at systemet kan kalibreres i forhold til hvor mange personer som er i bevegelse. Dette kan videreutvikles slik at systemet automatisk beregner multiplikasjonsfaktoren på bakgrunn av aktivitetsnivået i lokalet.



Figur 32: Kurven for aktivitetsverdiene som ble målt på Garvern

4.3 Brukertest Mesnali

Brukertesten på Mesnali ble arrangert som en del av en konfirmantleir. Ved hjelp av velvillige ledere var det mulig å arrangere diskotek med interaktiv house-gospel på lørdag kveld. Dette fungerte som en del av det faste programmet denne kvelden. Før diskoteket begynte ble deltakerne fortalt at de kunne påvirke lyd og bilde. Samtidig fikk de beskjed om at dette ikke var tilfelle hele tiden. Følgende informasjon ble gitt til deltakerne:

Velkommen til interaktiv house-gospel! Dere skal nå få være med på et eksperiment. Som dere ser er det satt opp en skjerm som viser animasjoner. Det som skjer på skjermen avhenger av hvor aktive dere er. Det vil si at jo mer aktive dere er jo mer skjer på skjermen. På samme måte finnes det 3 forskjellige lyder som spilles av når dere passerer bestemte aktivitetsnivå. Eksperimentet varer i 30 minutter og jeg kommer til å justere noen parametre underveis. Det betyr at aktiviteten deres ikke vil ha innvirkning på skjermen og lyden hele tiden. For at systemet skal fungere best mulig må noe av lyset i rommet være skrudd på.

Informasjonen skulle veilede deltakerne men samtidig ikke fortelle hele sannheten. Hensikten var å informere om hva som skulle skje, uten å avsløre at publikum bare hadde påvirkningsmulighet de siste 20 minuttene. Testen ble gjennomført slik at systemet selv bestemte aktivitetsverdien de første 10 minuttene. Denne verdien ble avlest fra en forhåndsprogramert kurve (jfr. fig. 41 i vedlegg C). De neste 20 minuttene ble aktivitetsverdien bestemt av publikums bevegelser. Hensikten med denne oppdelingen var å finne ut om publikum oppdaget når de selv kunne påvirket lyd og bilde og når systemet styrte dette.

Rundt 50 personer deltok på diskoteket, men siden eksperimentet ble gjennomført på en leir var ikke alle deltakerne tilstede hele tiden. Etter eksperimentet svarte deltakerne på et spørreskjema (jfr. vedlegg A). Spørsmål fra en lignende brukertest [3] ble brukt som referanse under utviklingen av spørreskjemaet.

Hele eksperimentet ble tatt opp på video for analyse i ettertid. Til dette ble det brukt et vanlig DV-kamera som filmet dansegulvet og skjermen. Figur 33 viser et bilde fra denne videoen hvor man kan se publikum, DJ-en og lerretet med animasjonen. I tillegg ble videostrømmen fra kameraet og den kalkulererte aktivitetsverdien logget til fil.

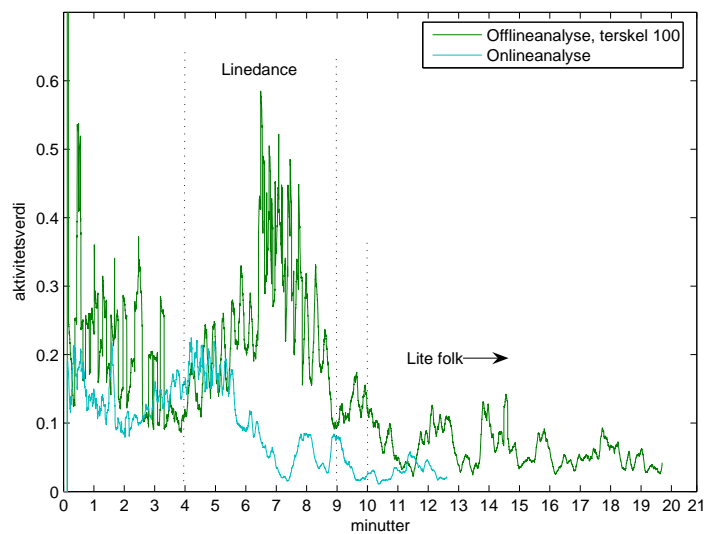


Figur 33: Bilde fra dokumentasjonsvideoen som viser publikum, DJ og storskjerm fra eksperimentet på Mesnali

Under initieringen av systemet før deltakerne ankom ble systemet testet og det fungerte etter hensikten. På dette tidspunktet var rommet normalt opplyst (vanlig taklys) og terskelen hadde en verdi på 127. Under eksperimentet ble denne lyssettingen forandret fordi deltakerne ikke ville danse uten at alt lys var skrudd av. Resultatet var at deltakerne selv skrudde av lyset mens eksperimentet pågikk. Lyset ble forsøkt tent igjen av noen av lederne, men deltakerne skrudde det av igjen. Dette skjedde flere ganger, men det ble tilslutt inngått et kompromiss hvor noe av lyset bak i salen var tent.

Systemet var ikke kalibrert for denne belysningen og det ble derfor nødvendig å justere terskelverdien mens eksperimentet pågikk, for å prøve å fange opp mer av aktiviteten. Ulempen var at bakgrunnsbildet da ble oppdatert mens det var mange mennesker i rommet. Dette er ugunstig fordi bakgrunnsbildet brukes som en referanse og blir trukket fra scenen for å lettere skille statisk bakgrunn fra bevegelser i bildet. Når referansebildet ble tatt med mennesker i rommet, førte det til at bakgrunnen som ble trukket fra også inneholdt personer. Sensoren er også følsom for forandringer i lysforholdene (jfr. kap. 3.3) og forandringene i lysforholdene har hatt innvirkning på systemet og redusert interaktiviteten.

Figur 34 viser aktivitetskurven når terskelen hadde verdi 100, sammen med aktivitetskurven som ble logget til fil under selve eksperimentet. Kurven viser at loggingen under eksperimentet stoppet etter ca 12,5 minutt. For å finne ut hvor mye kameraet registrerte, ble videologgen fra kameraet brukt for å rekonstruere aktivitetskurven. Denne er altså laget i ettertid, mens det er kurven som er merket "Onlineanalyse" som påvirket lyd og bilde under eksperimentet. Som figuren viser er det forskjell på de to kurvene. Dette kommer antakelig av at terskelverdien ble forsøkt justert mens eksperimentet pågikk.



Figur 34: Aktivitetskurvene fra live-testen på Mesnali

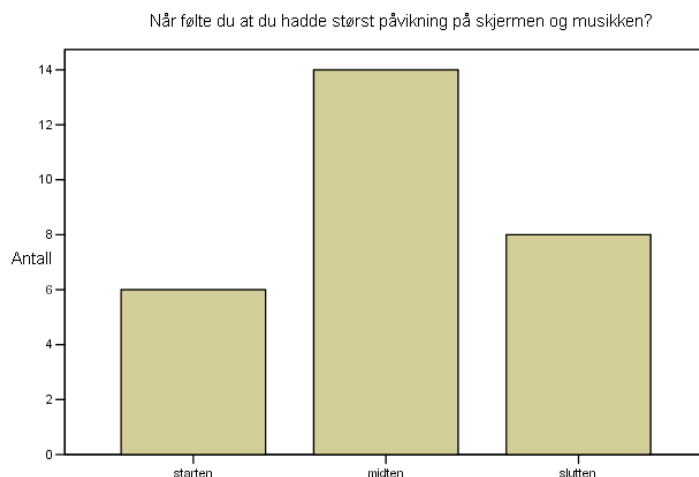
På spørreskjemaet ble deltakerne spurt om de likte å danse (jfr. sp. 3). Tabell 2 viser at de fleste personene likte å danse. Man kan derfor regne med at de i utgangspunktet er positive til diskoteket og ønsker å danse. Fordelingen mellom gutter og jenter var 18 mot 29 og 95,8 % av disse var i aldersgruppen 14-16 år. Det var altså liten spredning i alder blant deltakerne. Den geografiske spredningen var heller ikke stor, siden alle leirdeltakerne var konfirmanter fra Gjøvik-området. Dette førte til at testpersonene i eksperimentet utgjorde en relativt homogen gruppe, noe som trolig også har gjort utslag på resultatene.

Svaralternativ	ja	litt	nei
Antall	20	16	11

Tabell 2: Tabellen viser deltakernes svar på spørsmålet om de liker å danse (sp.3)

På spørsmål 4 om deltakerne hørte på elektronisk musikk var fordelingen ganske annerledes enn på spørsmål 3. Her svarte bare 3 personer *ja*, 11 personer *litt* og 32 personer *nei*, noe som tyder på at ca 70 % av deltakerne ikke var vant til denne type musikk. Dette kan bekreftes ved at 38 personer stilte seg mer eller mindre *uenige* eller *nøytrale* på spørsmålet om variasjonen i musikken var stor nok til å holde på interessen (sp. 13). Også spørsmål 14 bekrefter dette ved at det kom mange kreative forslag på andre musikkstiler som kunne vært brukt for å gjøre opplevelsen bedre.

For å få tak i den mest aktive deltakergruppen ble de som oppga at de ikke likte å danse filtrert vekk. Filtreringen plukket ut 36 personer som hadde svart *ja* eller *litt* på spørsmål 3: *“jeg liker å danse”*. Svarene til disse 36 personene ble sett i sammenheng med spørsmål 15: *“når følte du at du hadde størst påvirkning på skjermen og musikken?”*. Diagrammet i figur 35 viser at 28 stykker har svart et av de gyldige alternativene som var *starten*, *midten* eller *slutten*. Resten markerte flere svaralternativer eller unnlot å svare.



Figur 35: Diagrammet viser svaret på spørsmålet om når i konserten deltakerne følte at de kunne påvirke skjermen og musikken. Det er bare svaret til deltakerne som svarte at de likte å danse som er tatt med.

Eksperimentet ble utført slik at deltakerne bare hadde påvirkningsmulighet i midten og på slutten av diskoteket. Ved sammenligning av de 28 deltakernes svar med denne fasiten, har 22 stykker svart “rett” og 6 svart “feil”. De fleste deltakerne har altså svart rett, men svarfordelingen ser ut som en standard normalfordeling med vilkårlige svar. Det kan være tilfeldig hva de har svart og for å undersøke dette ble svarene på spørsmål 9 og 10 studert nærmere. På disse spørsmålene skulle deltakerne svare på om de følte de kunne påvirke musikken og skjermen. Alle de 36 danseglade deltakerne svart et av alternativene, men svarene er mer eller mindre negative. 23 av deltakerne svarte at de ikke følte at de kunne påvirke musikken og 21 svarte at de ikke følte at de kunne påvirke skjermen. En så stor andel av negative svar kan tyde på at deltakerne ikke har opplevd interaktiviteten og derfor svart vilkårlig på spørsmål 15.

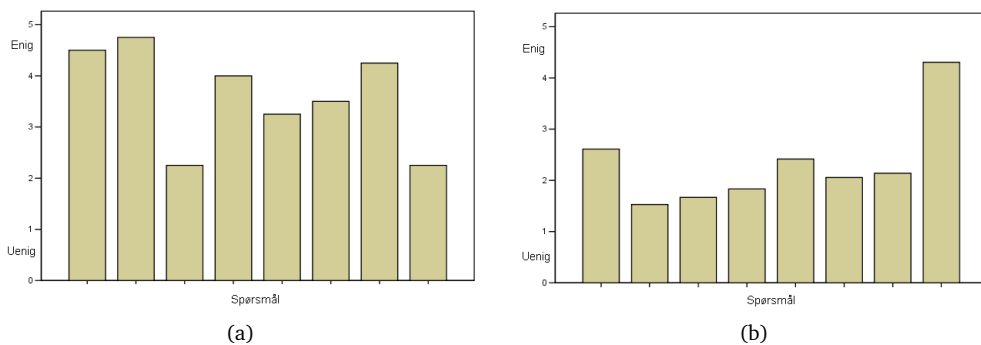
Dette kan skyldes at interaktiviteten i systemet ble redusert og på den måten påvirket resultatene fra brukertesten. En annen mulighet er at deltakerne ikke helt forstod hva interaktiviteten innebar og at de kunne påvirke lyd og bilde. Det ble gitt en forklaring på dette før eksperimentet startet (jfr. sitatet tidligere i kapitlet), men deltakernes oppførsel og avkryssing på spørreskjema kan likevel tyde på at de ikke har forstått hva som skulle skje. I tillegg til avkryssingen på skjemaet, kan dette underbygges ved at deltakerne selv forandret lyssettingen i rommet flere ganger under eksperimentet. Dette ble gjort på tross av informasjonen om at lyset måtte være tent for at systemet skulle fungere.

Muligens kunne dette vært løst ved å demonstrere systemet med en liten gruppe (4-5 deltakere) mens resten så på. Man kunne da oppnådd større forståelse blant deltakerne for hva de kunne påvirke i lyd og bilde. Kanskje ville de på denne måten også blitt mer bevisst på forandringene de selv provoserte frem. Under eksperimentet på Mesnali ble en slik demonstrasjon ikke gjennomført siden diskoteket var en del av konfirmantleiren. Deltakerne var derfor i utgangspunktet forberedt på et diskotek med “interaktiv house-gospel”, men ikke at dette samtidig skulle være et eksperiment hvor noe av lyset måtte

være tent hele tiden.

Flertallet har videre svart at de ikke ønsker å kunne påvirke musikken eller skjermen mer enn de kunne. Hvis man antar at publikum ikke har oppfattet hva de kan påvirke, er det også forståelig at de svarer negativt på disse spørsmålene. Uansett årsak tyder svarene på at deltakerne ikke opplevde større deltakelse under diskoteket på denne testen.

Det er også interessant å se på svarene til de deltakerne som svarte at interaktiviteten fikk dem til å danse mer. Filtrering av svarene på denne betingelsen gir 5 deltagere. Diagram *a* i figur 36 viser hvordan disse 5 deltakerne har svart på de andre spørsmålene, og diagram *b* viser hvordan resten av deltakerne har svart (store diagram med spørsmål-sreferanse finnes i vedlegg D).



Figur 36: Diagram a viser svarene til de 5 deltakerne som svarte positivt på spørsmålet om det som skjedde på skjermen og i lyden fikk dem til å danse mer. Diagram b viser svarene til de 43 andre som svarte negativt på dette spørsmålet.

Diagrammene antyder at det er en positiv tendens blant de 5 personene som svarer at de danset mer på grunn av det som skjedde på skjermen og i lyden. Mulige forklaringer på dette kan være at personene i utgangspunktet er mer positive til ny teknologi eller at de faktisk har forstått konseptet med interaktivitet. Hvis grunnen til deres positive innstilling er at de har forstått konseptet, er det interessant å se hva de har svart på spørsmål 15 om hvor i konserten de følte at de hadde størst påvirkning på skjermen og musikken. Det viser seg at de har den samme normalfordelingen på svarene som de deltakerne som svarte at de ikke danset mer på grunn av interaktiviteten. De har altså ikke truffet bedre i forhold til fasiten enn de andre deltakerne. På tross av dette er de likevel mer positive enn de andre. Muligens er dette psykisk betinget fordi deltakerne i utgangspunktet tror de har påvirkningskraft. På grunn av denne innstillingen leter de i større grad etter bekreftelse på denne kontrollen i både lyd og bilde. Når de får en bekreftelse på dette, enten det tilfeldig eller reell tilbakemelding fra systemet, blir de enda mer positivt innstilt.

Det kan derfor tyde på at innstillingen til deltakerne under testen var vel så viktig som funksjonaliteten til selve systemet. Det er mulig at man her kan ane resultater av en positiv tilbakemeldingssløyfe. Muligens kunne denne sløyfen også blitt startet ved en vilkårlig manipulasjon av lyd og bilde, slik det ble gjort de 10 første minuttene av eksperimentet. Uansett tyder det på at deltakernes innstilling og forventning er viktig. Kanskje er det også mulig å trekke en parallell til TV kontra Internett. Ved TV-titting er man i en mottaks- og underholdningsposisjon, sitter tilbakelent og regner ikke med at

man skal yte noe selv.

Surfing på Internett arter seg derimot annerledes ved at man sitter lent fremover og er mer aktiv i form av å selv lete frem informasjon. Det kan diskuteres om personer på diskotek kommer med en innstilling om å bli underholdt eller å selv underholde. Trolig regner de fleste med at de ikke selv behøver å yte noe for at det skal oppstå forandringer i musikk eller animasjoner. Parallellen på dette området blir derfor til TV. Samtidig ønsker mange å delta på dansegulvet, så til en viss grad har man også den aktive "fremoverlente Internett-innstillingen". Siden denne innstillingen er i forhold til dansing, vil utfordringen være å videreføre denne til å omfatte interaktiviteten også.

Under gjennomføringen av eksperimentet kom det frem noen personer som var nysgjerrige på hensikten til de forskjellige datamaskinene og kameraet. Etter en kort forklaring skjønnte de fort konseptet og var merkbart mer aktive og fokusert på animasjonen enn de andre deltakerne under resten av konserten. Dette kan tyde på at deltakerne ikke oppfattet den felles informasjonen som ble gitt om systemets funksjon før eksperimentet startet.

Det ble tidlig i oppgaven satt som et krav at sensoren skulle oppleves transparent for deltakerne. Dette målet ble oppnådd ved å bruke webkameraet, men det er mulig at hele systemet av denne grunn også ble for transparent. Hvis ikke deltakerne er oppmerksomme på webkameraet og ikke fikk med seg informasjonen før diskoteket begynte, er det forståelig at de svarte negativt på spørreundersøkelsen. Det kan også hende at systemets transparens førte til at deltakerne glemte interaktiviteten og ikke hadde fokus på dette under selve diskoteket.

Uansett grunn viste svarene fra spørreundersøkelsen at publikum ikke opplevde at de kunne påvirke lyd og bilde under dette eksperimentet. Som et resultat av dette følte de heller ikke at de kunne påvirke konserten. Deltakerne oppdaget heller ikke når i konserten det var de som påvirket lyd og bilde og når dette ble styrt av systemet selv. Disse forandringene kunne derfor vært trigget vilkårlig under dette eksperimentet.

Det var også interessant å se på hvordan musikeren, som i dette tilfellet var representert ved en DJ, opplevde å opptre når systemet var i bruk. DJ-en ble derfor spurt om han som musiker opplevde/følte at menneskene i salen påvirket han på en annen måte når de kunne trigge samplene og animasjonene? På dette svarte han:

...ja, de hadde påvirkning på meg fordi jeg hørte samplene, og da visste jeg at publikum var i mer aktivitet. Det er også lett å se ved å følge med på publikum, om de har det bra eller ikke, men en bekreftelse i form av denne samplene ga meg et ekstra smil, og kick til å yte enda mer fra scenen, som igjen førte til at publikum tok mer av. God dialog, og gøy var det.

Det ser altså ut til at selv om det er vanskelig å dokumentere en positiv virkning på publikum, opplevde DJ-en at systemet ga han mer konkret tilbakemelding fra publikum. Siden han har denne holdningen ble han også spurt om han ville brukt systemet på vanlige spillejobber? På dette svarte han "*...jeg ville brukt det med glede, hvis det var greit å sette opp, og ingen knotter å skru på for meg underveis.*"

Kanskje nytten av systemet er at det gir DJ-en tilbakemelding og påvirker lyd og bilde uten at publikum er bevisst på dette. Publikum kan delta lykkelig bevisst eller ubevisst i forhold til at de påvirker stemningen i lokalet og dermed sin egen situasjon.

5 Konklusjon

Prosjektet kom frem til et rimelig system for publikumsinteraktivitet i konserter. Systemet ble også testet i praksis for å se hvordan det fungerte og hvordan det ble mottatt av både publikum og utøver. All programmering ble gjort i programmene EyesWeb og Pure Data som finnes gratis på Internett. OpenSound Control-protokollen ble brukt for å kommunisere mellom programmene.

Det ble i starten av prosjektet bestemt at publikums fysiske bevegelse skulle brukes som et mål på engasjementet. Dette ble gjort fordi det er vanskelig å registrere hva publikum føler og synes på andre måter. Registreringen av engasjementet kunne gjøres ved å benytte forskjellige sensortyper slik som tråkkematte, joystick, gamepad, lasermatrise, videokamera og lignende. Valget falt på videokamera blant annet fordi det oppleves transparent for publikum og ikke forstyrrer deres fysiske adferd. Det er i tillegg enklere og hurtigere å koble opp enn de andre sensortypene.

Det ble gjennomført flere eksperiment med kamera som bevegelsessensor. Disse viste blant annet at det ikke var nødvendig å benytte et kamera med DV-kvalitet. Bevegelsesdeteksjonen som ble utført i denne sammenhengen fungerte godt nok ved bruk av et rimelig webkamera. Eksperimentene viste også at rammefrekvensen på videostrømmen kunne reduseres fra 25 rps til 10 rps, samtidig som tilstrekkelig informasjon om bevegelse i bildet ble bevart.

Støy i videostrømmen kan redusere bevegelsesdeteksjonens pålitelighet. Det ble derfor gjort en sammenligning av de to webkameraene som var tilgjengelig under prosjektet. Sandberg-kameraet hadde lavest egenstøy og ble derfor benyttet videre i prosjektet.

En annen utfordring var å registrere bevegelse i et halvmørkt rom hvor belysningen varierer. Dette er ofte tilfelle på konserter og diskotek. Systemet klarer ikke å skille mellom bevegelse hos publikum, og bevegelsen i lyset fra lysscannerne som tegner mønster på gulv og vegger. Det ble fremsatt en hypotese om at problemet kunne løses ved å filme i IR-området. Denne ble bekreftet og løsningen var å benytte et webkamera som registrerte bølgelengder i IR-området, i tillegg til synlig lys. Ved å benytte et IR-filter som blokkerte for alt synlig lys, var det mulig å foreta en bevegelsesanalyse som bare tok hensyn til bevegelse i IR-området. Resultatene viser at det er mulig å fjerne lyset fra noen lysscannere. Lyset fra lysscannerne som har pære med fargetemperatur på 6000 kelvin ble filtrert vekk. Det ble også testet med pærer som hadde fargetemperatur på 3200 kelvin, men lyset fra disse ble ikke helt fjernet. Trolig vil en løsning med å flombelyse rommet med IR-lys også fungere. IR-komponenten fra det varierende lyset vil da trolig viskes ut og drukne i lyset fra IR-flombelysningen.

For å kunne teste systemet i en reell sammenheng ble det programmert patcher i Pure Data for tilbakemelding til publikum i form av lyd og grafikk. Systemet fungerte tilfredsstillende før brukertesten, men under brukertesten på Mesnali oppsto det forandringer i belysningen etter at eksperimentet var startet. Dette førte til at systemet måtte re-kalibreres under selve eksperimentet. Systemet fungerte derfor ikke så godt som det skulle og mye tyder derfor på at resultatene fra denne testen er dårligere enn hva de kunne ha vært.

Deltakerne fylte ut et spørreskjema etter eksperimentet og svarene viste at bare noen få av deltakerne opplevde interaktivitet. De deltakerne som svarte at de opplevde interaktivitet svarte også mer positivt på alle de andre spørsmålene. På tross av dette lyktes heller ikke disse deltakerne i å tidfeste det tidsintervallet hvor de faktisk hadde påvirkningsmulighet. Trolig hadde deltakernes innstilling like mye å si som den faktiske interaktiviteten som systemet ga. Muligens ble systemet for transparent slik at mange av deltakerne ikke var bevisste på at de kunne påvirke lyd og bilde.

På grunn av komplikasjonene under brukertesten kan ikke resultatene tillegges for stor verdi, men testingen var likevel verdifull fordi den ga erfaring rundt problemstillingene. Dette gjør det mulig å gjennomføre nye tester som mest sannsynlig vil være mer vellykket. I tillegg ga DJ-en som spilte under testen positive tilbakemeldinger. Han fortalte at systemet gav han en ekstra bekreftelse i form av lydsamplene når aktiviteten til publikum økte. Dette førte til at han ønsket å "gi" ennå mer. Han mente også at systemet var et positivt bidrag til hele stemningen og kunne tenke seg å benytte systemet ved flere anledninger hvis det var enkelt å bruke.

6 Videre arbeid

En naturlig videreutvikling av prosjektet vil være å gjennomføre flere eksperimenter med filming i IR-området. Det ble nevnt at det kan være en løsning å flombelyse hele lokalet med IR-lys for å drukne lysvariasjonen fra lysscannere. Bruk av termisk kamera som fanger opp kroppsvarmen til publikum, kan også være en mulighet for å unngå forstyrrelse fra lysscanneren. Begge disse løsningene er interessante å teste i praksis.

En annen utvidelse er å detektere bevegelsen i forskjellige deler av lokalet og la dette påvirke tilbakemeldingen til publikum. Publikum kan deles inn i flere soner som kontrollerer forskjellige elementer (lyd, bilde, animasjon, lys) eller en sammenkobling av disse som tilhører hver sone. Publikum kan også være med å styre lysscannere og de mønstrene som disse lager. Andre kontrollmuligheter kan være styring av videoavspilling, mulige filtre på denne og delvis kontroll av musikernes instrument slik som cut-off frekvensen på trommelopper (cut-off frekvens = terskelverdi på et lavpass-filter som bestemmer hvor mye av frekvensområdet som passerer). Det er også interessant å teste bruk av flere kamera for å kunne bruke systemet på større lokaler. Dette kan eventuelt kombineres ved å dele opp publikum i flere soner.

Automatisk skalering av aktivitetsverdien kan også være ønskelig. Denne kan skaleres relativt i forhold til aktivitetsnivået i lokalet. En slik løsning vil fjerne behovet for manuell justering og fører til at systemet selv kalibrerer seg i forhold til hvor aktive deltakerne er.

Siden innstillingene på systemet under brukertesten ikke var optimale bør det gjennomføres flere brukertester under mer kontrollerte lysforhold. Systemet bør også testes med publikum som har større spredning i alder og ikke representerer en så homogen gruppe som konfirmanter er. I tillegg bør det gjennomføres eksperimenter der publikum ikke vet at de har påvirkningsmulighet for å se om de oppdager interaktiviteten selv og "spiller" på denne. Dette eksperimentet kan for eksempel sammenlignes med et eksperiment der publikum er bevisst på interaktiviteten.

Musikken som ble brukt under brukertesten var house-gospel, men tilbakemeldingen på spørreskjemaet viste at publikum ønsket mer variert musikk. Det er derfor nyttig å teste systemet sammen med mer kjent musikk. Et samarbeid med flere musikere og i bandsammenheng vil også være ønskelig.

Bibliografi

- [1] Haugsbakk, G. 2000. *Interaktivitet, teknologi og læring*. Forskning- og kompetansenettverk for IT i utdanning, Universitetet i Oslo.
- [2] Interaktivitet - ett begrep med flere betydninger. www.infodesign.no/artikler/interaktivitet280401.htm. Besøkt juni 2005.
- [3] Feldmeier, M. C. Large group interaction using disposable wireless motion sensors. Master's thesis, Massachusetts Institute of technology, 2003.
- [4] Jensenius, A. R. 2004. The interactive mass. An environment for multimedia improvisation. In *Poster presentation, Leeds International Music Technology Education Conference. Leeds*.
- [5] Nettsted for Interaktiv messe. www.messen.no. Besøkt mai 2005.
- [6] CINEMATRIX interactive entertainment systems. www.cinematrix.com. Besøkt februar 2005.
- [7] Maynes-Aminzade, D., Pausch, R., & Seitz, S. 2002. Techniques for interactive audience participation. In *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces (ICMI02)*.
- [8] Nettsted for Ars Electronica. www.aec.at/en/. Besøkt juni 2005.
- [9] Picard, R. W. & Scheirer, J. 2001. The galvactivator: A glove that senses and communicates skin conductivity. In *Proceedings from the 9th International Conference on Human-Computer Interaction, New Orleans*. vismod.media.mit.edu/tech-reports/TR-542.pdf.
- [10] Ulyate, R. & Bianciardi, D. 2002. The interactive dance club: avoiding chaos in a multi-participant environment. *Computer music journal*, 26:3, 40–49.
- [11] RS Online nettbutikk. www.rsonline.no. Besøkt mai 2005.
- [12] Nettsted for Electrotap. www.electrotap.com/sensors/T303G.shtml. Besøkt mai 2005.
- [13] Richardson, B., Leydon, K., Fernström, M., & Paradiso, J. A. 2004. Z-tiles: Building blocks for modular, pressure-sensing floorspaces. In *Late Breaking Results Paper, Vienna, Austria*. www.idc.ul.ie/ztiles/.
- [14] Løvhaug, J. W. 2004. Konsert for joystick og fullt orkester. *Apollon - Forskningsmagasin fra Universitetet i Oslo*, 4, 22–23.
- [15] Komplett nettbutikk. www.komplett.no. Besøkt mai 2005.
- [16] Sandberg IT, Sandberg NightCam. www.sandberg.it. Besøkt april 2005.

- [17] Gonzales & Woods. 2002. *Digital Image Processing, Second Edition*. Prentice Hall.
- [18] Irlight.com, An IR light primer. www.irlight.com/whatisir.html. Besøkt april 2005.
- [19] Wikipedia. Netleksikon. www.netleksikon.dk/g/gl/glodelampe.html. Besøkt april 2005.
- [20] Instrumentcompaniet AS. www.instrumentcompaniet.no. Besøkt april 2005.
- [21] Yoneyama, A., Nakajima, Y., Yanagihara, H., & Sugano, M. 1999. Moving object detection and identification from mpeg coded data. volume 2, 934–938. *Image Processing, 1999. ICIP 99*.
- [22] ZEN, H., WA, T. H., & WA, S. O. 1999. Moving object detection from mpeg coded picture. volume 4, 25–29. *Image Processing, 1999. ICIP 99*.
- [23] Harris, C. & Stephens, M. 1988. A combined corner and edge detector. *Plessey Research Roke Manor*, 147–151.
- [24] Lu, T. & Suganthan, P. 2004. An accumulation for video shot boundary detection. *Multimedia Tools and Applications, Kluwer Academic Publishers*, 22, 89–106.
- [25] Lipton, A. J., Fujiyoshi, H., & Patil, R. S. 1998. Moving target classification and tracking from real-time video. In *IEEE Image Understanding Workshop*, 129–136.
- [26] Pure Data community site. www.puredata.info. Besøkt februar 2005.
- [27] Cycling'74 Max/MSP. www.cycling74.com. Besøkt februar 2005.
- [28] Puckette, M. 2003. Theory and techniques of electronic music, draft. www.crea.ucsd.edu/msp/techniques.htm. Besøkt desember 2004.
- [29] Lab. InfoMus, U. o. G. Eyesweb. www.eyesweb.org. Besøkt mars 2005.
- [30] Wright, M., Freed, A., & Momeni, A. 2003. OpenSound control: State of the art. Conference on new interface for musical expression (NIME-03), Montreal, Canada. citeseer.ist.psu.edu/wright03opensound.html.
- [31] Wright, M. & Freed, A. 1997. OpenSound control: A new protocol for communicating with sound synthesizers. International Computer Music Conference. cnmat.cnmat.berkeley.edu/ICMC97/papers-html/OpenSoundControl.html.
- [32] Wright, M. OpenSound control specification, version 1.0. Technical report, Berkeley University of California, Center for New Music and Audio Technologies, 2002.
- [33] The Norwegian color research laboratory. colorlab.hig.no. Besøkt mai 2005.
- [34] The Filter Connection. www.2filter.com/prices/products/bw093.html. B+W Circular and Linear Polarizers, Besøkt april 2005.

A Spørreskjema

? SPØRRESKJEMA ?

svar ved å sette ring rundt ditt alternativ

1 alder _____
 2 kjønn jente
 gutt

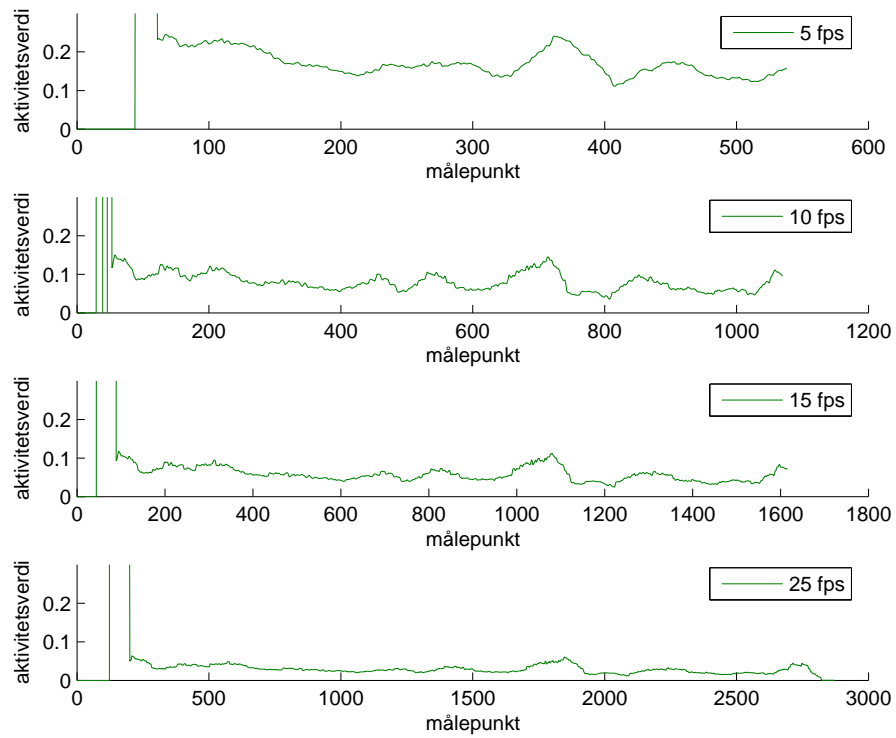
3 jeg liker å danse ja litt nei
 4 jeg hører ofte på elektronisk musikk ja litt nei
 (house, techno, elektronika, dance)

vennligst svar på punktene under:	uenig		nøytral		enig
5 jeg likte denne interaktive house-gospelen	1	2	3	4	5
6 jeg danset hele tiden i stedet for å bare henge rundt	1	2	3	4	5
7 jeg danset heller alene enn med andre	1	2	3	4	5
8 det som skjedde på skjermen og i lyden fikk meg til å danse mer	1	2	3	4	5
9 jeg følte jeg kunne påvirke musikken	1	2	3	4	5
10 jeg følte jeg kunne påvirke skjermen	1	2	3	4	5
11 jeg ønsker å kunne påvirke musikken enda mer	1	2	3	4	5
12 jeg ønsker å kunne påvirke skjermen enda mer	1	2	3	4	5
13 variasjonen i musikken var stor nok til å holde på interessen min	1	2	3	4	5
14 jeg ville likt opplevelsen bedre med en annen musikkstil	1	2	3	4	5
hvilken musikkstil: _____					
15 når følte du at du hadde størst påvirkning på skjermen og musikken ?	starten		midten		slutten

Takk :o)

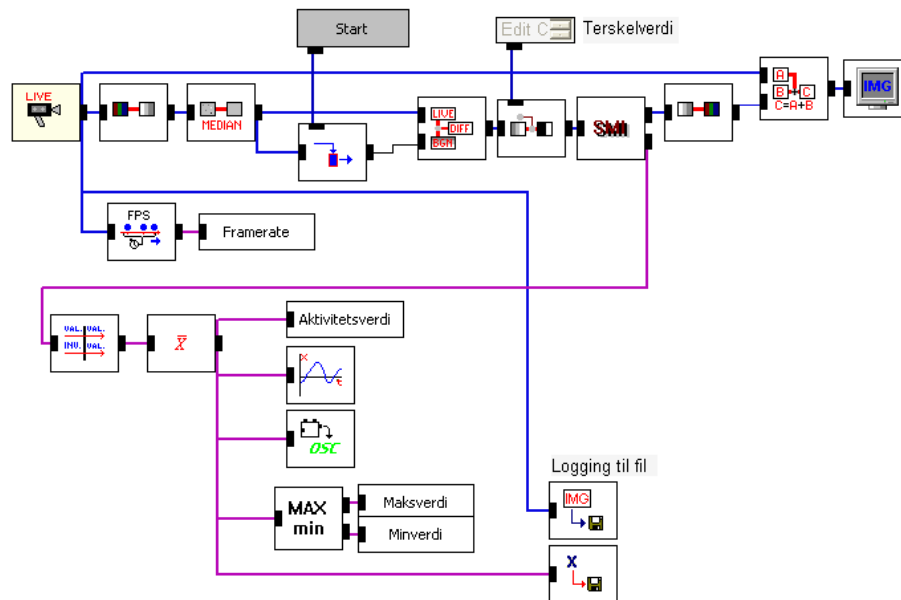
Figur 37: Spørreskjemaet som ble brukt under brukertesting

B Rådata fra rammefrekvenseksperimentet

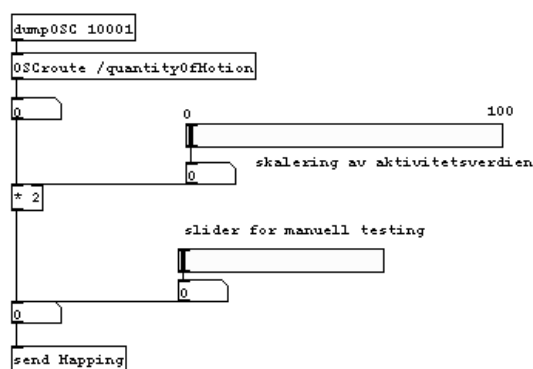


Figur 38: Aktivitetskurvene ved forskjellig rammefrekvens

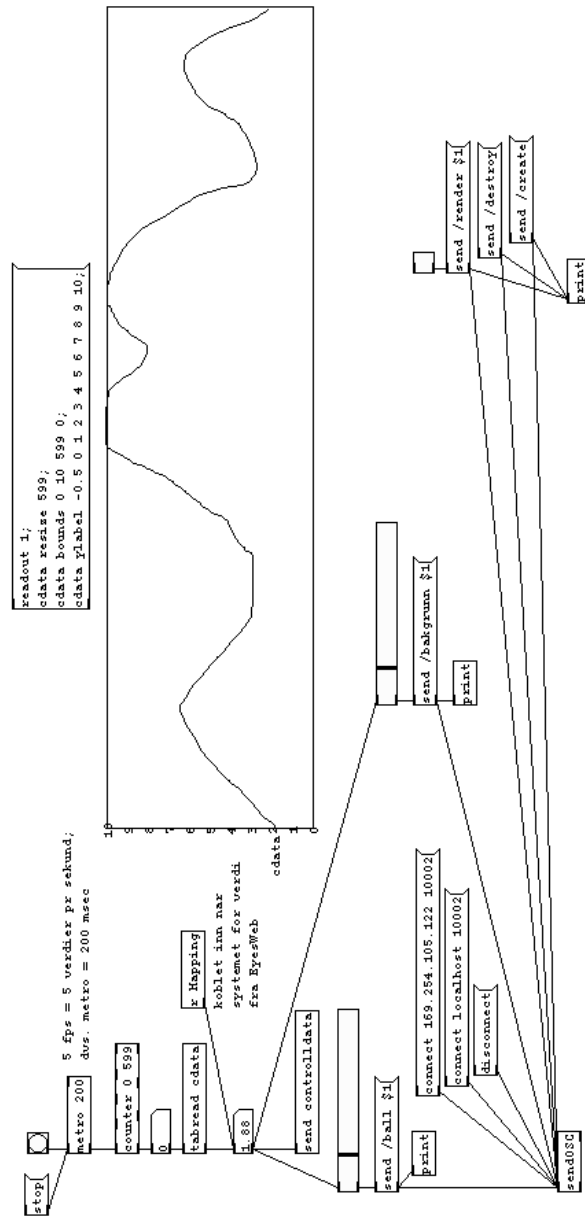
C Programkode fra brukertesten



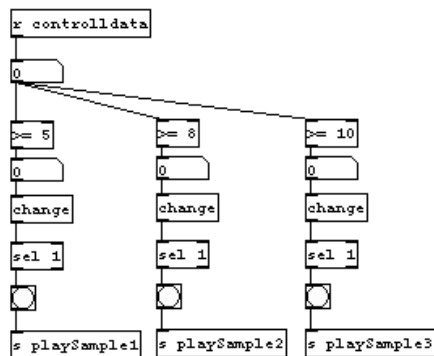
Figur 39: Analysepatch i EyesWeb som beregner aktivitetsverdi og sender den til PD over OSC-protokollen. Filnavn: **LiveMesnali12mars.eyw** Plassering: **Maskin 1**



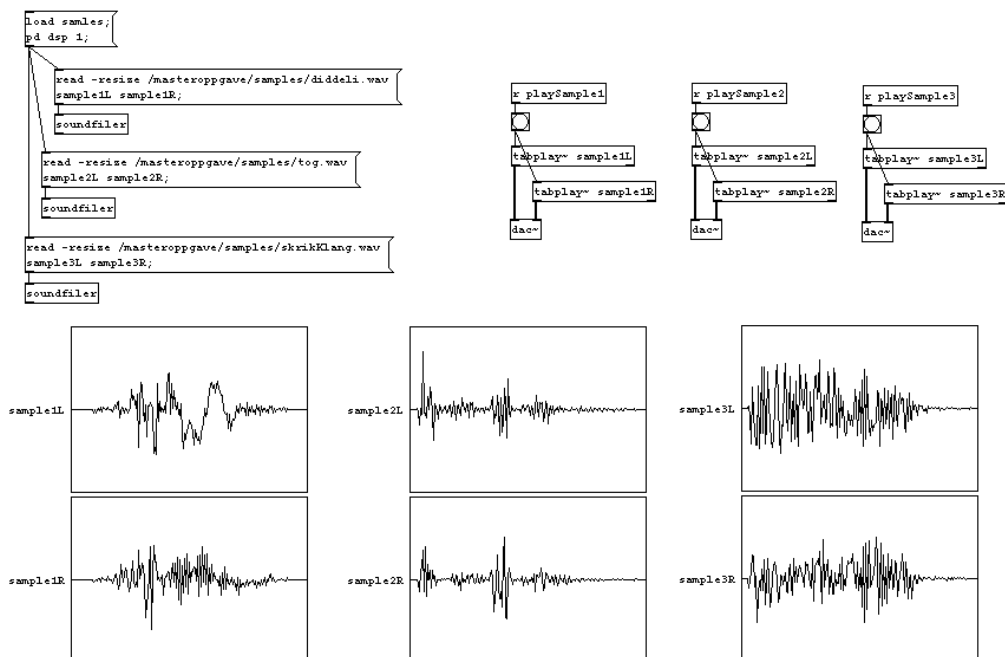
Figur 40: PD patch for mottak av aktivitetsverdi fra EyesWeb over OSC-protokollen internt på maskin 1. Filnavn: **pdLiveOscReceive.pd** Plassering: **Maskin 1**



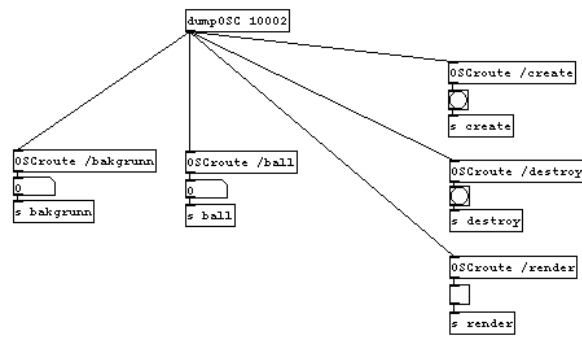
Figur 41: PD patch med virtuelt publikum og styring av rendring på maskin 2 over OSC. Filnavn: **LiveVirtueltPublikum.pd** Plassering: **Maskin 1**



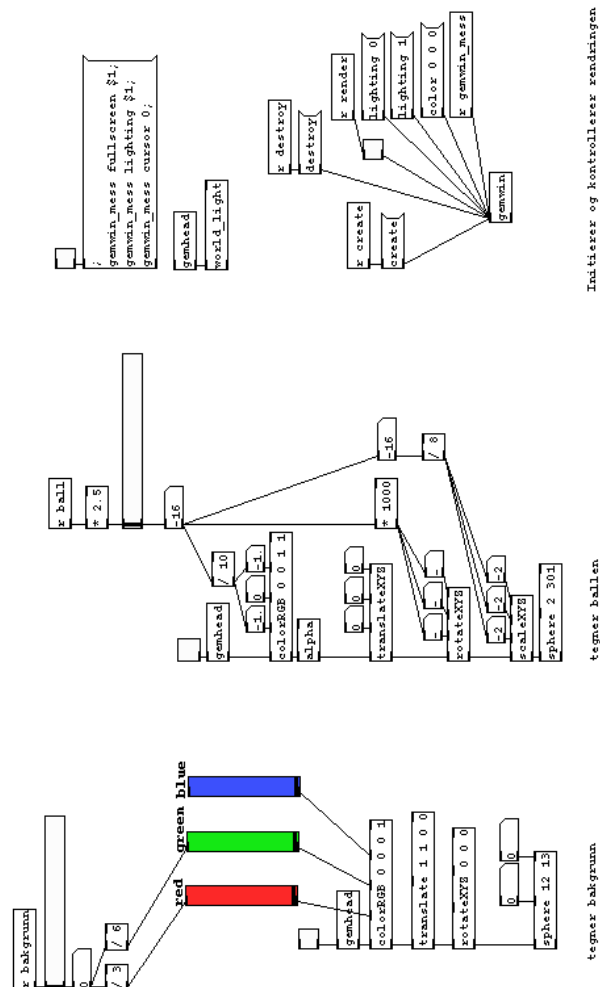
Figur 42: PD patch for styring av sampleavspilleren. Filnavn: **LiveStyringAvSamplePlayer.pd**
Plassering: **Maskin 1**



Figur 43: PD patch med sampleavspiller. Filnavn: **LiveLiveSamplePlayer.pd** Plassering: **Maskin 1**

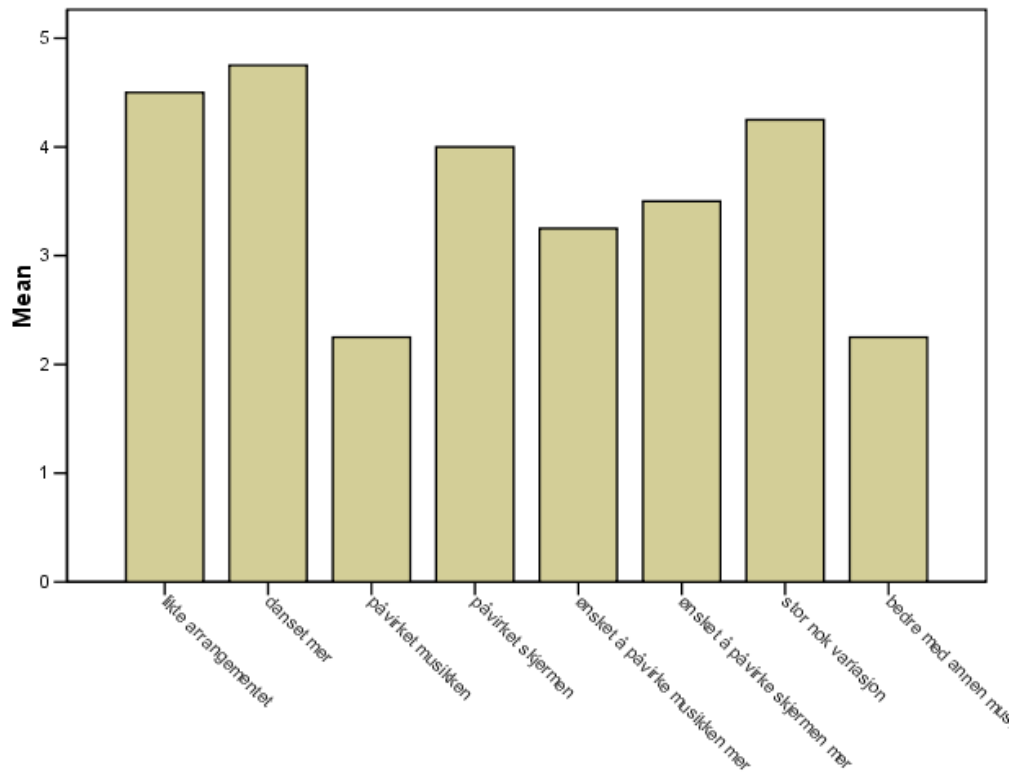


Figur 44: PD patch som tar imot aktivitetsverdien fra maskin 1 over OSC. Filnavn: **pdLiveOscReceiveBallBakgrunn.pd** Plassering: **Maskin 2**

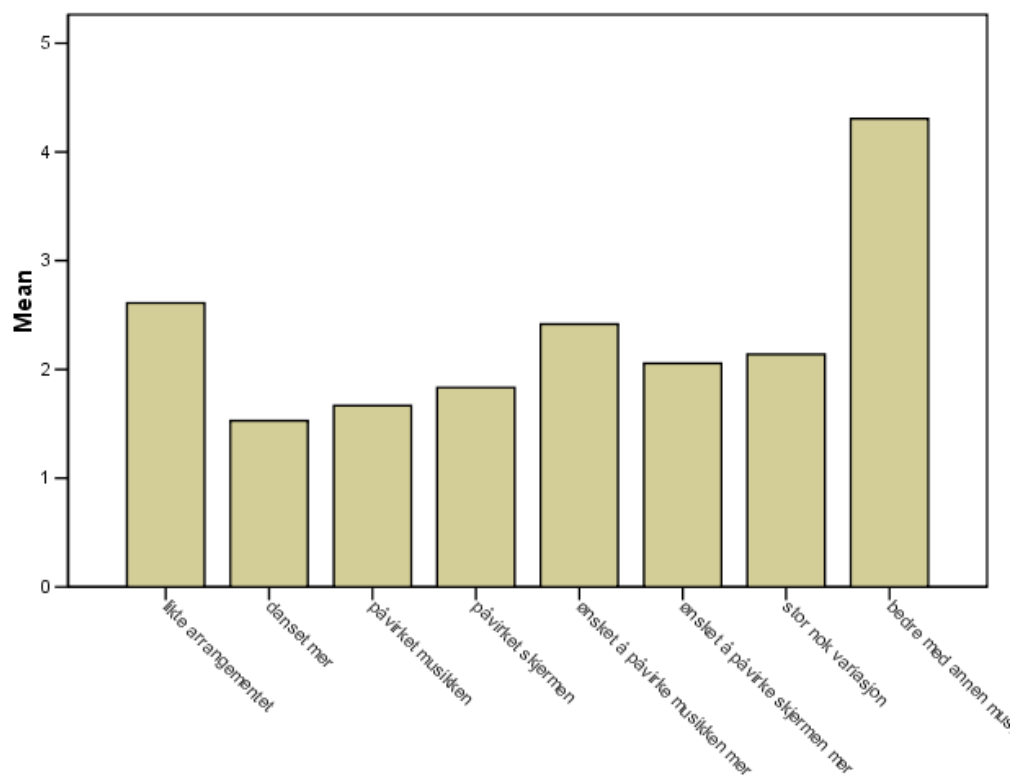


Figur 45: PD patch som renderer animasjon. Filnavn: **LiveGem.pd** Plassering: **Maskin 2**

D Diagram fra brukertesten



Figur 46: Diagrammet viser svarene til de 5 deltakerne som svarte positivt på spørsmålet om det som skjedde på skjermen og i lyden fikk dem til å danse mer.



Figur 47: Diagrammet viser svarene til de 43 deltakerne som svarte negativt på spørsmålet om det som skjedde på skjermen og i lyden fikk dem til å danse mer.