U.S.I.R

Utility Stylus with Identity Recognition

Prosjektrapport INF2260
Høsten 2013

Av
Bjørn Anders Reutz, Jørgen Valen, Peter Havgar, Thomas Schwitalla

25.11.2013

INNHOLDSFORTEGNELSE

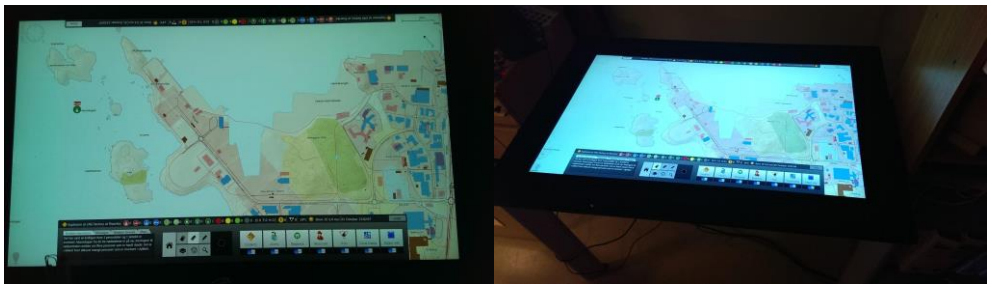
1. Innledning……. 1
    1.1. Vårt mål og vår visjon……. 2
2. Prosjektplanen…… 2
    2.1. Vår prosess……... 3
3. Konseptuelt design……. 3
    3.1. Brainstorming og tankekart……...…… 3
    3.2. Parallell prototyping: Skisser og storyboard……. 3
    3.3. Forarbeid før idéutvikling……. 3
    3.4. Idéutviklingen…… 4
    3.5. Undersøkelse med sluttbrukere……. 5
    3.6. Forarbeid før intervju…… 6
    3.7. Intervju…….. 6
    3.8. Analyse av intervjudata……. 7
    3.9. Konklusjon av undersøkelsen…… 7
    3.10. Reliabilitet og validitet…… 8
    3.11. Oppsummering av den konseptuelle designfasen…… 8
4. Fysisk design……... 9
    4.1. Første iterasjon med high-fidelity prototyper…… 9
    4.2. Andre iterasjon med high-fidelity prototyper…… 9
    4.3. Tredje iterasjon med high-fidelity prototyper…… 10
    4.4. Programvarekomponenten……. 10
    4.5. Fjerde iterasjon med high-fidelity prototyper…… 11
5. Evaluering med brukere…….. 12
    5.1. Del 1, Brukbarhetstest og ekspertvurdering med SINTEF…… 12
        5.1.1. Resultater, del 1…… 13
        5.1.2. Tilbakemeldinger fra deltagerne…….. 13
        5.1.3 Konklusjon av evaluering del 1……. 14
    5.2. Del 2, Hovedstudien: Eksperiment…… 14
        5.2.1. Det eksperimentelle designet…… 15
        5.2.2. Variablene i eksperimentet…… 15
        5.2.3. Utvalg av deltagere…….. 16
        5.2.4. Retningslinjer for testingen…….. 16
        5.2.5. Resultater og konklusjoner fra evaluering del 2…….. 16
        5.2.6. Analyse av experimentdata…….. 16
        5.2.7. Statistiske tester…… 17
        5.2.8. Konklusjon av ANOVA-test……. 17
        5.2.9. Tolkning av observasjoner og tilbakemeldinger…….. 17
        5.2.10. Reliabilitet og validitet…….. 18
6. Konklusjon……. 18
7. Kilder…….. 20
    7.1. Trykte……. 20
    7.2. Internett…….. 20

1. Innledning

Vår oppdragsgiver var SINTEF. Vår oppgave het (løst oversatt fra engelsk) Én kontroller vs. flere kontrollere i en flerbrukersetting. Den dreide seg om å utvikle en prototype som kan brukes til å teste forskjeller i brukbarhet mellom én og flere samtidige brukere av et Microsoft PixelSense-bord[1], som "MASTER"-systemet kjører på. MASTER er en del av et større prosjekt kalt BRIDGE, som SINTEF er prosjektleder for.

BRIDGE (*Bridging Resources and Agencies in large-scale Emergency management*) er et prosjekt med mål å støtte koordinasjonen mellom nødetater (brann, politi, helse) i situasjoner som for eksempel terrorhandlinger, naturkatastrofer eller andre hendelser av kritisk natur. Det er et system som fra et overordnet plan kan bidra til å etablere kontroll og oversikt. Systemet skal kunne brukes til å tilrettelegge kommunikasjonen mellom de sentrale aktørene i en gitt krisesituasjon. BRIDGE er et EU-prosjekt i samarbeid med andre forskningsinstitutter Europa over.

MASTER-systemet et interaktivt kart som tilbyr brukere et felles situasjonsbilde, "common operational picture", på tvers av geografiske avstander og apparater. Systemet lar brukerne bevege, tegne, og markere på et bilde av et kriseområde, og slik gi direkte ordre og beskjeder til andre ledere eller mannskap i felten. Mannskapet kan bli utstyrt med håndholdte enheter som kjører en egen versjon av systemet. Alt som gjøres på én instans av systemet, vil være synlig på alle andre til enhver tid. MASTER tilbyr også ulik informasjon fra kriseområdet, som for eksempel status av pasienter, informasjon om hendelsen og annet fra eksterne tjenester slik som sosiale medier og værmelding.



MASTER på MS PixelSense

Microsoft PixelSense er en 40" multi-touch-skjerm med en oppløsning på 1920x1080 punkter. Skjermen er orientert horisontalt og har fire ben, som et bord. Noe som gjør PixelSense spesielt er at det bygger på optisk, infrarød touch-teknologi, og ikke kapasitativ teknologi som finnes i de fleste touch-skjermer vi ser idag. PixelSense støtter opptil 52 samtidige touch-inputs. Den fungerer ved at når noe blir plassert på skjermen, stråler en infrarød bakgrunnsbelysning (Microsoft, 2012) gjennom skjermen og opp i objektet, som igjen reflekterer lyset ned i integrerte sensorer. Lyset blir så tolket som et signal og prosessert. Den tolker gjenstanden som én av tre ulike typer objekter: Fingre, "Blobs" (vilkårlige former som skiller seg fra fingre på ulike måter, eks størrelse), og "Tags", bestemte mønstre som kan identifiseres av bordet og brukes til for eksempel spillbrikker.

[1] Vi kommer tilbake til spesifikasjonene av MS PixelSense.

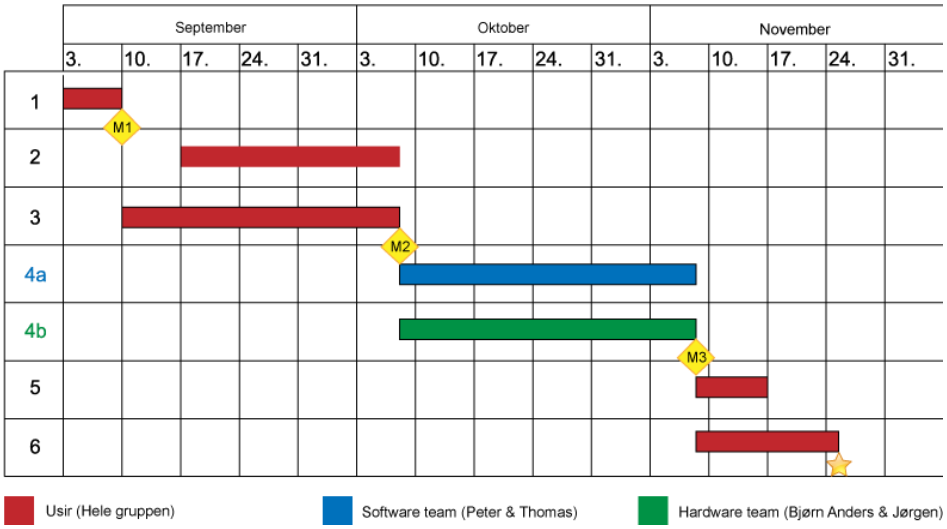
1.1. Vårt mål og vår visjon

Vårt fokus har vært på å utforske brukbarheten til PixelSense-bordet og MASTER med særlig oppmerksomhet rettet mot flerbrukersituasjoner. Vi har utforsket løsninger på å gjøre bruken mer intuitiv for noen som ikke har god kjennskap til MASTER, og heller ikke har brukt et PixelSense-bord før. Vi hadde som visjon at vårt design kan gjøre at hvem som helst skal kunne bruke MASTER effektivt med minimalt av opplæring. Vår ambisjon var at de siste prototypene vi utviklet skulle være evolusjonære, slik at SINTEF skulle ha muligheten til å utvikle dem videre dersom de ønsket det.

Våre interessenter var primært representanter fra utviklerne av MASTER ved SINTEF IKT og systemets sluttbrukere ved brann --og politivesen. Disse sluttbrukerne er utelukkende fagledere og innsatsledere, da det er de som har ansvar for ledelse av operasjoner (Stølen & Nilsson. 2011). Et eksempel på sekundære interessenter er mannskap i felten som mottar ordre og beskjeder fra lederne gjennom MASTER.

2. Prosjektplanen

Planen ble lagt innenfor rammene satt i «Visual Map of the Course» på kurshjemmesiden. Vi tok utgangspunkt i denne fordi den ga et worst-case tidsestimat for de viktigste milepælene. Av erfaring vet vi at presis estimering av tidsforbruk er vanskelig på tidlige stadier, så vi ønsket derfor å heller overestimere enn å underestimere. Vi oppdaterte planen inkrementelt etterhvert som vi fikk et klarere bilde av prosessen. Den endelige planen er her gjengitt i et flytdiagram:



Nøkkel:

1. Designbrief og prosjektplan. Resulterer i milepæl 1, designbrief og plan.

2. Etablering av krav, studier med interessenter.

3. Konseptuelt design: Lo-fi prototyper. Resulterer i milepæl 2, kravspesifikasjon. Går parallelt med kravetablering og bruksundersøkelser. Sammenfaller med midtveispresentasjonen.

4. Fysisk design: Bygging og testing av hi-fi prototyper. Resulterer i milepæl 3, prototyper som er egnet for brukertesting. Dvs. bygging av pennene og programmering

5. Evaluering, brukbarhetstesting med brukere.

6. Skriving av prosjektrapport og planlegging av sluttpresentasjon. Resulterer i prosjektrapporten og avslutningen av kurset.

2.1. Vår prosess

Vi valgte å følge den tradisjonelle prosessmodellen for designaktiviteten: undersøkelse av bruk, design av alternativer, prototyping og evaluering (Preece, Rogers and Sharp 2011). Dette gjorde vi intuitivt fordi vi stadig søkte godkjenning fra interessentene, og da særlig de potensielle sluttbrukerne av MASTER-systemet, slik at våre designvalg ville understøtte deres behov og mål.

Under den konseptuelle designfasen lagde vi lo-fi-prototyper og presentasjon av disse for interessentene fra SINTEF i to iterasjoner, før vi i en tredje iterasjon involverte fagledere fra Romerike Brannkorps.

Under den fysiske designfasen ble prototypene evaluert og testet av oss selv underveis, helt til vi oppnådde et produkt som ytet tilstrekkelig godt for brukerbasert testing. Deretter involverte vi SINTEF og studenter i to ulike brukbarhetstester under kontrollerte forhold. Dessverre, på grunn av tidsbegrensninger, fikk vi aldri anledning til å teste de siste prototypene med folkene fra Romerike Brann og Redningsvesen.

3. Konseptuelt design

Vi designet våre konsepter gjennom gjentatte perioder av idéutvikling og resulterende prototyper i lav oppløsning (lo-fi). Våre metoder var som følger:

3.1. Brainstorming og tankekart

Brainstorming var vår hovedmetode for idéutvikling. Vår fremgangsmåte var løst basert på Osborns metode[2]. På rundgang skrev vi ned det første vi tenkte om et gitt begrep i et tankekart, i korte, tidsbegrensede økter. Dette ble alltid gjort i plenum. Idéene ble diskutert først i etterkant, aldri under selve økten; vi ønsket ikke kritisere noen underveis, da dette kunne hemme flyten.

3.2. Parallell prototyping: Skisser og storyboard

Med parallell prototyping mener vi at designkonsepter blir utviklet i parallell før de utsettes for kritikk. Vi benyttet dette innad i gruppen ved å tegne hver våre skisser av ett og samme konsept, for å så diskutere dem i plenum etterpå. Slik fikk vi se ulike forestillinger av de samme idéene, noe som førte til både nye og bedre idéer. Deretter lagde vi storyboard, korte tegneserier som illustrerer et tenkt brukstilfelle med vårt hypotetiske produkt. Disse, sammen med skisser og tankekart, ble brukt til å formidle våre konsepter til interessentene.

3.3. Forarbeid før idéutvikling

Før idéutviklingen startet fikk vi anledning til å se på MASTER-systemet. Vi prøvde ut grensesnittet og identifiserte noen potensielle problemer ved flerbrukersituasjoner (gitt et vilkårlig antall brukere):

- Det kan fort bli trangt om plassen rundt bordet.
- Bordets rektangulære form gjør at brukere som står på kortsidene og ved hjørnene får vanskeligheter med å nå hele skjermen.
- Ved bruk av MASTERs innebygde tegnefunksjon og ikonfunksjoner, er det umulig å se hvilken bruker som gjorde hva, noe som kan skape forvirring underveis.

[2] http://en.wikipedia.org/wiki/Brainstorming#cite_note-2

    - Dersom ledere fra to eller flere ulike etater hypotetisk sett skal jobbe samtidig på ett og samme bord, ville det blitt umulig for mannskapet i felten å se hvilke ledere som gir hvilke beskjeder (og vice versa).
- Man er begrenset til et lite område på kartet dersom man ikke ønsker å zoome ut, hvilket ville innebære at man mister mye detalj. Dette blir et problem dersom brukernes ansvarsområder befinner seg et stykke fra hverandre.

3.4. Idéutviklingen

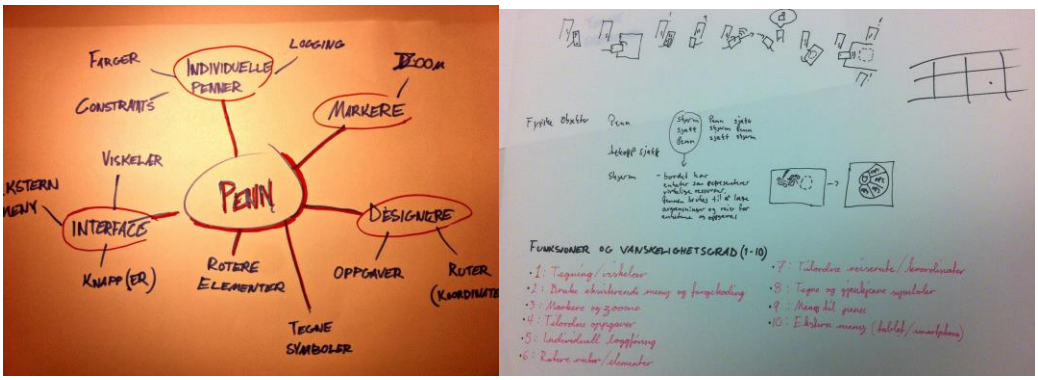
Idéutviklingen ble drevet fram av behovet for en løsning på en "kontroller", slik oppgaven fra SINTEF krevde. Med *kontroller* tolket vi en fysisk gjenstand som kunne brukes til å interagere med bordet. Vi forsøkte å tenke utenfor rammene som begrenser MASTER, å tenke "utenfor bordet". Vi drøftet også hva som egentlig menes med "god brukbarhet", og satte dette i sammenheng med MASTER. Her er de mest sentrale konseptene vi utarbeidet i den første iterasjonen, og som ble presentert:

- Brikker som benytter tag-funksjonaliteten til PixelSense. Denne idéen kom fra tanken om at MASTER-systemet minnet oss om et kart som generaler og hærførere samler seg rundt for å planlegge sine slag.
- "Split screen". En idé inspirert av flerbrukersettinger i videospill. Dette dreide seg om at arbeidsflaten kunne deles inn i sektorer som tilhører hver bruker, slik at arbeidene deres ikke overlapper og kommer i konflikt. Dette var ikke et forslag til en fysisk kontroller, men en funksjon som kunne implementeres i programvare.
- Benytte bevegelsesfølsomme kontrollere slik som Leap Motion og Kinect. Dette kunne forbigått begrensningene med plass rundt bordet, men bare hvis det var mulig å henge skjermen opp på veggen.
- Øvrige forslag til kontrollere, uten noen bestemt funksjonalitet:
    - Hansker
    - Armbånd og ringer
    - Penner

Da vi presenterte konseptene for vår oppdragsgiver ved SINTEF IKT første gang, ble vi mer eller mindre skutt ned. Idéene ble vurdert som enten upraktiske på grunn av begrensninger ved PixelSense, eller som lite relevante fordi de ville gjort hele argumentet for en delt arbeidsflate ugyldig. Vi gikk derfor tilbake til tegnebrettet.

Vi måtte se på enklere løsninger som ikke gjorde bordet overflødig. Først drøftet vi idéene våre på nytt og diskuterte fordeler og ulemper ved dem. En av de enkleste og mest innlysende forslagene stod fram som det beste, nemlig *pennen.* Konseptet vi fokuserte på heretter var altså en penn som kunne fungere på PixelSense, med tilhørende funksjonalitet ivaretatt fra våre tidligere idéer. Dette var våre argumenter for hvorfor en penn kunne være relevant:

- Pennen ville trolig være mer presis under bruk av bordet, enn det å bruke det med fingrene.
- En penn er noe de fleste kan relatere seg til, og kan derfor gjøre det enklere å bli kjent med et nytt system. Vi antok at våre sluttbrukere hadde mye erfaring med vanlige papirkart, og tenkte derfor at vi kunne få MASTER til å føles mer naturlig ved hjelp av en penn.
- Det kunne løse noen av problemene vi forestilte oss med brukervennlighet, særlig det at pennene ga oss en mulighet å identifisere brukerene unikt.



Vi skrev en prioritert liste med funksjoner som vi ønsket å forsøke å prototype. Tanken var å fokusere på de viktigste, men levere flest mulig. Listen ble presentert for vår oppdragsgiver hos SINTEF sammen med resten av skissene som utgjorde konseptet.

Før denne presentasjonen gjorde vi også et utforskende eksperiment hvor vi tok i bruk penner. Dette var et veldig enkelt oppsett, hvor vi fikk testet grov funksjonalitet og samarbeid mellom brukere. Vi designet eksperimentet med tanke på løsningsorienterte oppgaver til brukerene, og fikk dem da altså til å løse labyrinter i fellesskap. Vi kan ikke si å ha hentet noen essensiell data fra dette eksperimentet, men det belyste enkelte ting, som for eksempel begrenset arbeidsområde, som vi så kunne arbeide videre med i utformelsen av prosjektarbeidet.

Vi fikk god respons på dette konseptet. Grunnen til at pennen var interessant for MASTER-utviklerne, var nettopp det at brukerne trolig ville oppleve den som en mer presis og intuitiv interaksjonsmåte. Vår kontakt hadde faktisk på et tidligere tidspunkt utforsket en løsning på penner selv, men la det på hyllen blant annet av følgende grunner:

- Han hadde ikke funnet noe egnet materiale til en penn.
- Det ville blitt et problem for bordet å filtrere ut input fra pennene og alt annet (fingre og løse ermer med mer).
- Han mente at noen brukere ville ha en tendens til å lene seg på bordet med én hand mens de bruker pennen med den andre.

Vår interessent utfordret oss til å løse de ovennevnte problemene, og dermed gikk vi allerede her løs på det fysiske designet (vi kommer tilbake til dette under kapitlet om fysisk design).

3.5. Undersøkelse med sluttbrukere

Selv om SINTEF syntes konseptet vårt var lovende, hadde vi ingen illusjoner om at våre sluttbrukere mente det samme. Dessuten manglet vi domenekunnskap; ingen av oss visste noe om hverdagen til en innsatsleder fra politi eller brannvesenet. Vi ønsket derfor å gjøre en formativ evaluering av konseptet med sluttbrukerne før vi tok flere designavgjørelser. SINTEF rekrutterte ledere fra Romerike Brann og Redningsvesen til vår disposisjon. Noe ventetid gikk før vi fikk muligheten til å treffe de, naturlig nok på grunn av deres travle hverdag.

3.6. Forarbeid før intervju

I mellomtiden ble vi forsynt med dokumenter fra tidligere undersøkelser som SINTEF hadde gjort med nødetatene, i tillegg til dokumentasjon som angår utviklingen av MASTER. Vi ble oppfordret til å bruke disse dokumentene til å få innsikt i domenet, samt etablere krav på bakgrunn av behov som allerede var blitt avdekket.

Særlig verdifull var studien *Generic functionality in user interfaces for emergency response* (Nilsson, Stølen, 2011). Den identifiserte overlappende behov for funksjonalitet i brukergrensesnitt hos aktører i nødrespons, det vil si blålysetatene, samt uavhengige organisasjoner som Røde Kors. Studien avdekket 11 kategorier av funksjonalitet som aktørene deler behov for i ulik grad.

I tillegg dro vi nytte av rapporten "*Situation awareness in crisis management" (Stiso et al, 2013)* som beskriver utviklingen av MASTER, særlig fordi den viste oss skopet for systemets fremtidige utvikling. Dette var interessant fordi vi kunne forsøke å bygge inn støtte for noen av de planlagte funksjonene i våre prototyper.

3.7. Intervju

Vi designet et intervju for brannlederne. Det overordnede målet for undersøkelsen var å sette oss inn i en innsatsleders kontekst i håp om å se hvordan vårt konsept kunne støtte deres arbeid, samt å utsette designet for kritikk. Intervjuet var semistrukturert, med en blanding av åpne og lukkede spørsmål og en utforskende hensikt. Vi ble forberedt på at det kunne møte opp et vilkårlig antall ledere, så vi skrev to utgaver av intervjuet, ett som var ment for en enkeltperson, og ett som var mer i retning av et fokusgruppeintervju.

Spørsmålene dekket følgende emner:

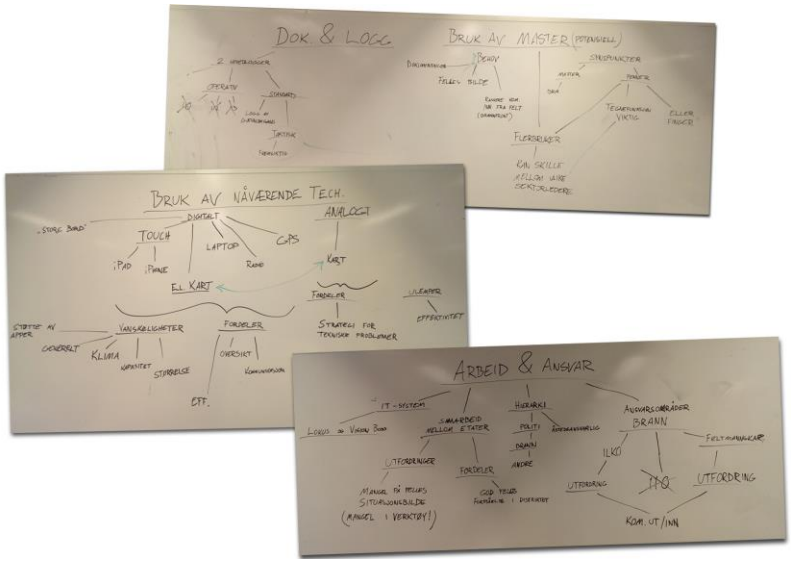
- Innsatsledernes arbeid og hverdag.
  - Generelt.
  - Spesifikke oppgaver, rutiner og prosesser.
  - Om samarbeid med andre etater under kriseoperasjoner.
- Deres individuelle ferdighetsnivå med teknologi, spesielt touch-apparater.
  - Foretrekninger og behov, både på jobb og fritid
- Vi presenterte vårt konsept med skisser og andre prototyper mot slutten av intervjuet for å invitere til diskusjon og meningsutveksling.

Kun én innsatsleder var tilgjengelig for intervju da vi ankom Romerike Brannstasjon. Vi stilte med to mann fra gruppen; én ledet intervjuet og den andre sto for praktiske saker. Jan-Håvard Skjetne fra SINTEF varmet opp med en kort presentasjon av BRIDGE og MASTER. Intervjuobjektet samtykket til lydopptak (et samtykkeskjema utformet av SINTEF og tilpasset vår undersøkelse ble presentert). Vi transkriberte opptaket så ordrett som mulig.

3.8. Analyse av intervjudata

Vår hovedmetode for analyse av kvalitative data var *content analysis* (innholdsanalyse). Det er en teknikk som systematisk identifiserer spesifikke karakteristika til tekstlig innhold ved å kategorisere det etter bestemte regler (Lazar, Feng, Hochheiser 2010). Etter intervjuet hadde vi begrensede forhåndskunnskaper om emnet og domenet og valgte derfor å ikke benytte et teoretisk rammeverk (taxonomy), men heller benytte åpen koding.

Vi utarbeidet en hierarkisk kodestruktur (nomenclature) med klare regler for hva som kjennetegnet hver kode:



3.9. Konklusjon av undersøkelsen

Av intervjuet med innsatslederen fra Romerike Brann og Redningsvesen lærte vi følgende:

Innsatsledere benytter touch-apparater i tjeneste, spesifikt iPad og iPhone. De opplever til tider vanskeligheter ved disse tilknyttet klimaforhold samt kompitabilitet med nødvendig programvare (operative systemer). Dessuten er de små formatene til tablets og smartphones uhåndterlige i felt:

> "*Utfordringa vår er at noen ganger jobber vi i 20 kuldegrader og vann og fuktighet. Da er det utfordrende. Da må jeg ofte gå inn i bilen for å få det til å funke."*

Vi trekker fra dette at touch-teknologi ikke er fremmed for våre sluttbrukere, og at bruk av MASTER muligens ikke vil innebære en stor læringskurve.

Videre benytter innsatslederne elektroniske kart i tjenesten, som fungerer bra så lenge kapasitetene på trådløst nettverk er tilstrekkelig og at kartene er oppdatert. Svært ofte må de imidlertid ty til tradisjonelle kart som en løsning på problemene. Våre penner kan replikere brukeropplevelsen til de tradisjonelle kartene. Da innsatslederen ble bedt om å kommentere vårt konsept, uttrykte han sterk interesse for en tegnefunksjon (som allerede er innebygd i MASTER, men dette visste han trolig ikke), men uttrykte ikke

et eksplisitt behov for en penn. Han mente likevel at det kunne føles mer naturlig, særlig hvis det lignet på et *SmartBoard:*

"Det kan føles mer naturlig, hvis det ligner på et smartboard."

Faglederen foreslo faktisk selv et bruksområde av MASTER, nemlig å loggføre taktiske avgjørelser slik som sektorinndeling:

"*Tegne opp brannfronter [og] bruker enhetlig ledelsessystem. Det gir mulighet for å dele i sektorer og å tildele disse til sektorledere i felt, da vet vi i KO hvem som er leder, dette gir bedre oversiktlighet også når brannen sprer seg".*

Behovet for dokumentasjon og loggføring av øvelser og operasjoner ble sterkt vektlagt av innsatslederen. Kodene for dette hadde mange forekomster, da faglederen gjentok seg selv for å understreke poenget:

"*Så har vi røyk og kjemikal, dykkerøvelser, så det er alltid planer og alt skal dokumenteres [...]. Så vi er helt strukturert og alt skal dokumenteres [...]. Så vi legger opp året og øver og er forberedt på de situasjoner som vi kan møte. Og alt skal dokumenteres."*

Dette er i kontrast til SINTEFs egen undersøkelse, som ikke fant dette behovet hos brann, bare hos politi og i mindre grad ambulanse. Det skal sies at SINTEF *ikke* intervjuet brannvesen, de gjorde kun observasjoner (Nilsson, Stølen 2011).

3.10. Reliabilitet og validitet

Teksten ble kodet separat av to kodere. Vi sjekket vår *inter-coder reliability* (Lazar, Feng, Hochheiser 2010) ved å konstruere en *agreement matrix* og beregne Cohen's Kappa ut ifra denne. Beregningen ga en K på 0,76 (76%), som antyder at koderne hadde en høyst tilfredsstillende enighet.

Da vi kun har snakket med én fagleder, er vi forsiktige med å generalisere for brannvesen i andre distrikter. Vårt subjekt har over 30 års erfaring som fagleder. Det er derfor rimelig å anta at informasjonen han har bidratt med er høyst gyldig.

3.11. Oppsummering av den konseptuelle designfasen

Vi besluttet på bakgrunn av intervjuet samt studiene av eksisterende litteratur at vi ville ta konseptet med pennene videre. Våre argumenter:

- Pennene kunne behjelpe vanskelighetene med kulde og størrelse; en tilstrekkelig stor penn kan lett holdes med hansker, og vil samtidig være mer presist enn å bruke fingrene. Dessuten vil pennene sannsynligvis oppleves mer naturlige å bruke på MASTERs interaktive kart.
- Vi innså at vår idé med unik identifikasjon av brukere gjennom pennene kunne ha et større potensiale enn vi opprinnelig trodde. Dersom man skulle ha en flerbrukersituasjon med ledere av ulike etater på MASTER, bør det finnes en slags loggføring på individuell basis for å kunne skille mellom hvem som gjorde hva til enhver tid. Våre penner kunne tilby slik funksjonalitet, muligens i tandem med loggføringsfunksjoner som allerede er planlagt for utvikling (Stiso et al 2013).

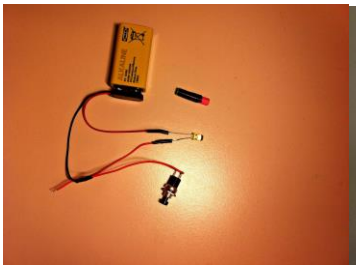
4. Fysisk design

Under går vi igjennom iterasjonene av fysisk design for å vise hvordan penneprototypen har utviklet seg i løpet av prosjektet. Vi regner programvaredelen av prototypene som fysisk design, da denne sørger for pennenes funksjonalitet. Hardware og software ble utviklet i parallell, slik det framgår i prosjektplanen.

4.1. Første iterasjon med high-fidelity prototyper

Målet for den første iterasjonen var å finne ut hvilke materialer som kunne brukes til å lage en penn som gjenkjennes av PixelSense-bordet. Dette skjedde litt på sparket, da vår oppdragsgiver utfordret oss til å finne en fungerende penn.

Med den kunnskap at PixelSense-bordet reagerer på IR-lys, bygde vi først en krets bestående av en IR-LED (infrarød lysdiode) med et 9-volts batteri og en trykkbryter. Vi testet så denne på bordet med MASTER-systemets innebygde tegnefunksjon. Dette fungerte beundringsverdig, så vi valgte å bygge videre på denne løsningen.



Første versjon

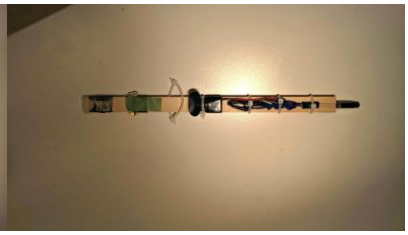
Andre versjon

4.2. Andre iterasjon med high-fidelity prototyper

Denne iterasjonen ble påbegynt så snart vi var ferdige med bruksundersøkelsen. Her valgte vi å teste forskjellige typer av/på-mekanismer, forskjellig spenning og forskjellige LEDs. Vi laget derfor 3 forskjellige prototyper, alle med omtrent samme krets, (et LED, et 9V batteri og en 470 ohms motstand), med ulikheter på hver av disse.

Prototype med av/på-switch

Prototype med potentiometer



Prototype med utskiftbare LEDs

Den venstre prototypen hadde en av/på-bryter som vi testet for å se om den var mer bekvem å bruke enn knappene fra første iterasjon, som man måtte holde inne for å slå på lyset. Neste prototype (midten) hadde et potentiometer slik at vi kunne teste forskjellig spenning/lysstyrke og se hvordan bordet reagerte. Høyre prototype utskiftbare dioder, slik at vi kunne teste hvilke farger som ga best respons.

Testing av andre iterasjon pennene viste at IR-LED fungerte best, fordi de har en høyere bølgelengde i det elektromagnetiske spekteret (950nm), kontra grønn, gul eller rød (570nm, 590nm og 620nm). Sensorene i

PixelSense registrerer hovedsakelig IR-spekteret (Microsoft, 2012) og nærliggende frekvenser, så dette var egentlig ingen stor åpenbaring. Potentiometeret gjorde at vi kunne regulere størrelsen på prikken; den var tydeligvis relativ til styrken på lyset fra dioden.

Dessuten viste andre iterasjon at IR-diodene trengte lite spenning for at bordet skulle kunne se lyset, og vi så dermed at vi i stedet kunne bruke et 1,5V batteri med en mindre motstand. Vi konkluderte også med at av/på-bryteren var tungvint å bruke og at vi måtte prøve andre løsninger.

4.3. Tredje iterasjon med high-fidelity prototyper

Tredje iterasjon av pennene startet med å teste en av/på-mekanisme som var innebygd i tuppen, slik at lyset slår seg på når man trykker den mot en flate akkurat som en vanlig penn. Dette prøvde vi å løse med en fjær bak dioden, som gjorde at når tuppen ble presset mot bordet ble den skjøvet bakover i penhylsteret og kom i kontakt med resten av kretsen. Hylsteret er laget av en whiteboard-tusj.

Penn med pressmekanisme

Pressmekanismen fungerte ikke fordi den påførte bordet for mye trykk under bruk. Vi la denne løsningen på is og gikk tilbake til trykknappen fra første iterasjon. Hylsteret viste seg å være for lite, da vi ikke fikk plass til hele kretsen. Vi måtte legge motstanden og noe ledning på utsiden, og den var dessuten vanskelig å holde.

4.4. Programvarekomponenten

Kodingen foregikk i C# og XAML (Xtensible Application Markup Language), som er brukt for resten av MASTER. Systemet benytter .NET-rammeverket, hovedsakelig klasser fra bibliotekene Windows Presentation Foundation (WPF) og Surface 2.0 Software Developer Kit (SDK). Vi hadde ingen forhåndskunnskaper om .NET og C# og brukte derfor noe tid på å lære oss dette fortløpende. Heldigvis fikk vi svært god starthjelp fra vår mann på SINTEF.

Vi skrev et tegneprogram som kan identifisere hver penn som en unik bruker. Selve tegnefunksjonaliteten var allerede ferdiglaget av klasser fra WPF-biblioteket, men identifisering av unike brukere fantes ikke i Surface SDKet. Dette måtte vi selv programmere en løsning på. Vi startet med å undersøke klassebibliotekene for å finne en måte å lese av "rå" data fra våre lysdioder, det vil eksempelvis si verdier for lysintensitet eller bølgelengder. Vi håpet at slike verdier kunne leses av, og dermed benyttes til å gi hver penn unike karakteristikker, men konkluderte med at dette ikke var mulig

Vi visste på forhånd at PixelSense kunne skille mellom fingre og andre gjenstander basert på blant annet størrelsen (i bildepunkter). Videre undersøkelse av dette viste at størrelsen til en berøring kunne leses av fra et *strokeCollectedEvent-objekt*[3]. Takket være dette, kunne vi skille mellom ulike penner ved å se størrelsen på lysprikkene deres og sette disse som en betingelse for hvilken farge som angis strøkene. Vi så deretter at det var ingen grense for ting som kunne defineres på individuell basis; fargen er bare ett

[3] Ett "stroke" ("strøk" på godt norsk) defineres som én sammenhengende input-output-hendelse, en strokeCollectedEvent. Det vil si fra en bruker tar på bordet med en finger eller gjenstand til han løfter den av. I vårt tilfelle betyr dette for eksempel at noen tegner en sammenhengende strek.

eksempel. Vi valgte å fokusere på å finjustere identifikasjonen heller enn å lage ekstra funksjonalitet. Det var helt avgjørende at denne fungerte 100% av tiden.

Vi implementerte også en løsning for å filtrere ut input fra andre objekter enn pennene, slik som at noen legger en arm over mens de tegner og dermed lager store flekker. Slike uønskede flekker og lignende vil vi heretter referere til som "støy". Denne funksjonen sjekker også størrelse og er derfor begrenset i hva den klarer å filtrere ut.

Som alle prototyper, har vår programvare visse begrensninger og kompromisser; den er som antydet vertikal. De viktigste kompromissene er som følger:

- Fargen på et strøk settes først *etter* at det er blitt ferdig tegnet, og er sort i mellomtiden.
- Tykkelsen på strøkene vil variere fra penn til penn. For at de skal oppleves som like, blir størrelsen for hvert strøk endret til å være like samtidig som at fargene settes.
- Vi rakk ikke implementere en innebygd viskelærfunksjon i pennene, så programvaren bruker en vanlig, global "eraseByPoint"-løsning, slik som MASTERs egen tegnefunksjon gjør.

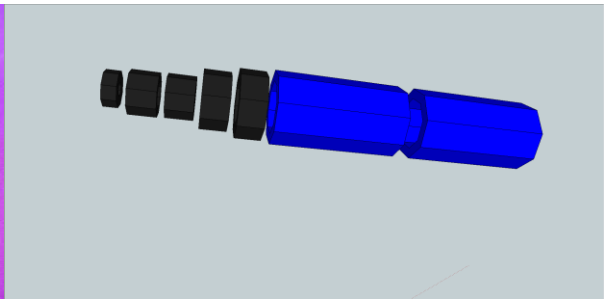
4.5. Fjerde iterasjon med high-fidelity prototyper

Den fjerde og endelige iterasjonen av pennene skulle bli egnet for brukertesting. På dette stadiet hadde vi bestemt valgene av innvendig elektronikk. Denne iterasjonen fokuserte på det overfladiske heller enn innmaten. Vi bestemt å 3D-printe pennhylsteret fordi vi da kunne få den akkurat slik vi ville raskt og billig.

Den første modellen vi prøvde å 3D-printe lignet på en vanlig penn. Det viste seg etter å ha printet den ene halvdelen at 3D-printeren ikke klarer å printe særlig runde kanter. Vi måtte derfor lage en modell med færre kanter. Vi tegnet en åttekantet modell i seks deler. Denne modulmodellen gjorde det mye lettere å installere kretsen etter print, og gjorde det også lettere å gjøre endringer i ettertid ved at vi kunne legge til flere moduler.

Den ferdige modellen ble den desidert sørste og mest robuste modellen. Vi måtte legge til en modul etter litt testing for å øke avstanden mellom knappen og bordet, fordi lyset fra bordet reflekterte i neglene til de som holdt pennen, noe som forårsaket støy. Ved å legge til en modul og øke avstand mellom hånd og bord, minsket vi sannsynligheten for støy på bordet. Vi oppdaget samtidig enda et problem; materialet fra pennene reflekterte også lys. Dette løste vi ved å male tuppen av pennene svart.

De endelige prototypene



3D-modell av pennene

Vi brukte en 10 ohms motstand i den ene pennen som skulle gi de største strøkene, en 47 ohms motstand i den andre som ga en litt mindre strøk, og i den siste valgte vi en 380 ohms motstand. Vi behøvde et sprik i motstandene fordi programvaren leter etter størrelsesverdier mellom bestemte intervaller fordi en helt spesifikk størrelse er umulig å produsere konsistent.

5. Evaluering med brukere

I denne delen vil vi beskrive evalueringen av hi-fi-prototypene.

5.1. Del 1, Brukbarhetstest og ekspertvurdering med SINTEF

Målet for denne delen av evalueringen var primært å teste hvor godt den seneste iterasjonen av prototypene fungerte, og samtidig få tilbakemelding på disse av SINTEF.

Første del av evalueringsfasen ble utført i form av en kombinert ekspertvurdering og brukbarhetstest hos SINTEF. Hensikten var å få luket ut ytterligere funksjonelle problemer i forkant av hovedstudien.

Fire personer fra SINTEF, hvorav tre var medlemmer av BRIDGE MASTER-teamet, deltok. Vi vurderte opprinnelig å strukturere ekspertvurderingen som en "cognitive walkthrough", det vil si at ekspertene simulerer et brukstilfelle slik de forestiller seg at de reelle sluttbrukerne ville opplevd det (Lazar, Feng, Hochheiser, 2010). Imidlertid gjorde tegneprogrammet vårt det vanskelig å simulere et realistisk scenario. Hadde funksjonaliteten vår vært en del av MASTER, kunne vi gjennomført det. I tillegg vurderte vi en heuristisk vurdering, men passende heuristikker fantes ikke for våre design. Isteden ble testen gjort mindre formelt, der ekspertene gav oss mer eller mindre subjektive tilbakemeldinger på brukeropplevelsen.

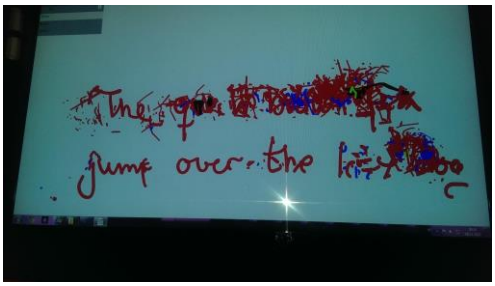
Testen vi designet besto av en kort skriveoppgave som gikk ut på å skrive setningen "The quick brown fox jumped over the lazy dog" på PixelSense-bordet med våre prototyper. Denne setningen er normalt brukt for å teste tastaturer, da den består av hele det engelske alfabetet. Teksten ble valgt på grunn av denne sammensetningen, slik at en mengde ulike former kunne testes, og fordi lengden var passende i forhold til størrelsen på skriveoverflaten. Deltagerne ble instruert til å skrive i sitt eget tempo og å skrive på den måten som faller seg naturlig for dem. Slik håpet vi på å se fremtredende problemer som vi ikke klarte å provosere fram under egen testing. Vi hadde altså en formodning om at ulike skrivestiler kunne ha betydning.

I forkant av testen ga vi hver deltager en kort opplæring der de ble forklart visse begrensninger ved vår prototype, slik som at de må holde pennen forholdsvis rett under skriving. I tillegg sjekket vi at klærne deres ikke reflekterte for mye lys (én hvit bluse ga vanskeligheter).

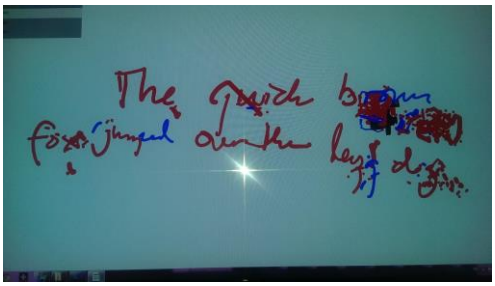
Vi målte ytelsen til pennene og tilhørende programvare ved å se på følgende parametere:

- Antall strøk som fikk feil farge (alt skulle bli rødt).
- Mengder av støy, det vil si streker og andre former som ikke ble produsert av pennen.
- Hvilke bokstaver og ord der feil oppsto oftest.
- Tiden deltageren brukte på å skrive setningen. Dette er relevant fordi skrivehastigheten ser ut til å påvirke PixelSense' gjenkjenning av strøk.

Fotografier ble tatt av “tegningene”, og undersøkt i etterkant. Ikke alle resultater kunne kvantifiseres nøyaktig, særlig med tanke på støy som kunne ta svært ulike former. Tolkningene av bildene vil derfor være uformelle. Her er to eksempler:



Mye støy



Mindre støy, fler feil

Tilbakemeldingene fra deltagerne ble notert for hånd. De ble bedt om å kommentere følgende:

- Generell opplevelse av pennens utforming, størrelse, tyngde osv.
- Hvordan de opplevde å holde inne knappen mens de skrev.
- Hvorvidt det oppleves mer presist og intuitivt enn å skrive med fingertuppene (Vi gjør oppmerksom på at disse brukerne har mye erfaring med PixelSense på forhånd; de vet allerede hvordan det er å bruke det med fingrene.).

5.1.1. Resultater, del 1

Deltager	**Tid**	**Sekunder**	**Støy**
1	01:18	78	>50
2	00:32	32	24
3	00:52	52	>50
4	00:39	39	20

5.1.2. Tilbakemeldinger fra deltagerne

Vi gjorde en uformell analyse av intervjuene, med koder kategorisert etter temaene som ble diskutert. Tilbakemeldingene var konkrete og åpnet lite for tolkning, så vi så ikke behov for en streng innholdsanalyse. Koder og antall forekomster er presentert i følgende tabell:

UTFORMING	**KNAPPEN**	**OPPLEVELSE**
Riktig tyngdepunkt (2). Bra lengde (4). Mangekantet form var behagelig å holde (2). Solid byggkvalitet (3)	Vanesak, men helt brukbart med litt øvelse (4). Trolig slitsomt å holde inne i lengden (2). Trolig bedre med bryter innebygd i tuppen (1)	Mer presist og intuitivt enn å bruke fingrene (4). Bør forbedre støyfiltreringen til programvaren (1). Fargen må være riktig 100% av tiden (1).

Våre kommentarer til deltagernes ytelse

Deltager 1	Deltager 2	Deltager 3	Deltager 4
Lite støy, men mange feil fordelt over hele teksten. Skrev i blokkbokstaver, og måtte derfor løfte pennen mellom hver bokstav, som muligens var årsaken. Tok lengst tid av alle; brukte sannsynligvis ikke sin naturlige håndskrift.	Jevnt over noe støy, få feil. Skrev formskrift. Den høyeste av alle deltagerne, og holdt pennen mer eller mindre loddrett hele tiden. Venstrehendt.	Fargen var jevnt over korrekt, men enorme mengder støy på første linje, økende jo lenger til høyre i tegningen. Skrev løkkeskrift. Personen holder pennen med ringfinger og lillefinger hengende nær bordet, som åpenbart er kilden til støy. Hurtig skrift er sannsynligvis også en faktor.	Den hurtigste, men den mest feilfrie av alle deltagerne. Skrev løkkeskrift. Feil farge mest konsentrert i ordene "jumped", "brown" og "lazy", hovedsakelig på slutten av ordene. Stor støy rundt "brown", trolig på grunn av refleksjon fra håndleddet da personen strakk armen og dermed kom nærmere bordet.

5.1.3. Konklusjon av evaluering del 1

Tilbakemeldingene var svært positive og oppmuntrende. I tillegg fikk vi konkrete forslag på hva (og hvordan) funksjonaliteten kunne forbedres. Til tross for at samtlige mente at trykknappen fungerte bra med litt øvelse, ble det klart at vi likevel burde forsøke å bygge inn en trykkfølsom aktivering i pennens spiss, da dette vil sansynligvis føles mer naturlig.

Undersøkelse av fotografiene fra testene viste oss dessuten hovedårsaken til at pennene iblant gir feil farge, nemlig at brukerne slipper knappen for sent slik at den siste størrelsesverdien som leses av havner i feil interval. En trykkfølsom løsning er derfor vårt hovedpunkt for fremtidig forbedring av pennen.

5.2. Del 2, Hovedstudien: Eksperiment

Til hovedstudien valgte vi å utføre en brukbarhetstest i form av et eksperiment. Målet med dette var å svare på oppgaven for SINTEF, nemlig å undersøke forskjeller i brukbarhet mellom fler-- og enkeltbrukersituasjoner på Microsoft PixelSense. Mer konkret formulert, målet var å undersøke sammenhengen mellom antallet samtidige brukere og forandring i brukbarhet. Etter forespørsel fra vår oppdragsgiver, gjorde vi samtidig en kvalitativ undersøkelse (gjennom observasjon) av hvordan samarbeidet til deltagerne utartet seg under flerbrukertesten.

Vi ønsket ikke å se på sammenhenger mellom testernes demografiske kriterier eller andre variabler og ytelse; målet var ikke å generalisere for en populasjon. I tillegg understreker vi at testen var av PixelSense-bordet og våre egne prototyper, ikke MASTER-systemet og dets sluttbrukere. Opprinnelig var tanken å gjøre en tredje studie på lokasjon Romerike Brannstasjon, men som tidligere nevnt rakk vi ikke dette.

Tanken bak oppgaven i eksperimentet var å skape et relativt ekstremt brukstilfelle, der testerne måtte jobbe på hele arbeidsflaten med opptil to medarbeidere. På grunn av plassbegrensninger i "testlaben" vår, var det vanskelig å teste flere enn tre samtidig. Vi valgte å ikke påføre testerne unødvendig stress som et følge av et trangt og innestengt kontor. For å simulere ytterligere begrenset plass rundt bordet, var det ikke tillatt for brukerne å bevege seg fritt. Oppgaven gikk ut på å tegne et forhåndsbestemt sett av geometriske figurer inn i et rutenett på 6x3 celler. Det ble utført tre tester for hver av brukerne:

1. Deltagerne jobber enkeltvis.

2. De jobber samtidig med én av hver av de andre testerne.
3. Samtlige testere jobbet samtidig.

Test 1 og 2 ble kjørt 3 ganger per deltager (-1), test 3 ble kjørt 3 ganger, for totalt 20 løp. Hver tester fikk sin unike farge.

Stedet for testen ble SINTEFs kontorer, "objekt-rommet" hvor PixelSense-bordet står. Vi var bundet til dette kontoret gjennom hele prosjektet. Testingen ble moderert av oss selv.

5.2.1. Det eksperimentelle designet

Det må igjen understrekes at dette er først og fremst en brukbarhetstest som skal avdekke brukbarhetsproblemer. Vi benytter derfor kun et lite antall testere, og stoler på at testmetodene effektivt avdekker eventuelle problemer. Formen "eksperiment" er valgt for å strukturere og veilede testingen på en kontrollert måte som kan isolere sammenhenger mellom forskjeller i ytelse til prototypene og ulike variabler. Vårt forskningsspørsmål og våre hypoteser er som følger:

Hvordan påvirkes brukbarheten til våre prototyper når desto flere brukere skal jobbe på bordet samtidig?

H_0: Det er ingen forskjell i brukbarhet mellom 1 og flere brukere.
H_A: Det er en forskjell i brukbarhet mellom 1 og flere brukere.

For klarhets skyld, H_0 sier at det ikke er noen signifikant forskjell på om én eller flere bruker prototypen samtidig, mens H_A sier det motsatte, at det er en forskjell. Vi ser altså etter forskjeller i begge retninger, men forventer at brukbarheten reduseres.

5.2.2. Variablene i eksperimentet

Eksperimentet undersøker én uavhengig variabel, nemlig antallet samtidige brukere. De ulike tilstandene (conditions) er naturligvis antallet brukere, som går fra 1 til 3.

De avhengige variablene vil være tidsbruk og antall feil per testløp. Med feil mener vi her at en bruker tegner feil symbol i feil celle. Vi noterer fortsatt forekomster av feil farger i tillegg til støy. Begge deler skal loggføres individuelt for hver bruker.

De kontrollerte variablene er som følger:

- Settene av symboler som brukerne skal tegne (se eksempel lenger ned).
- Hver deltager vil ha faste plasseringer i matrisen for sine symboler for å hindre at avstander mellom ulike plasseringer tar forskjellig tid å tegne. Rekkefølgen av symbolene skal imidlertid variere fra test til test, for å begrense læringskurve som en faktor.
- Testernes plasser rundt bordet skal være faste av samme grunn som at symbolenes plasseringer skal være like.

5.2.3. Utvalg av deltagere

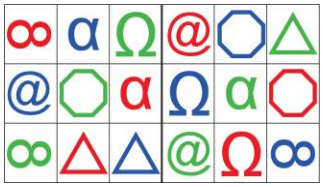
Som nevnt innledningsvis, er det ingen grunn til å velge testere som er representative for MASTER-systemets sluttbrukere. Disse er som kjent svært vanskelig å rekruttere, så det er mer logisk å benytte andre. Testerne kan i prinsippet være hvem som helst. Vi gjør et tilfeldig utvalg av studenter på universitet[4].

5.2.4. Retningslinjer for testingen

For å utføre eksperimentet trengs det to moderatorer. Moderatorene skal vise fram bilder av symbolenes plassering for subjektene på hver sin side av bordet. Filmopptak av testen skal gjøres dersom deltagerne samtykker. Hvis ikke, tas fotografier av hvert testløp. Videre skal følgende punkter overholdes:

- Testerne skal få en kort opplæringsperiode med vår prototype i forkant av eksperimentet. De skal instrueres i bruken av pennene, samt få lov til å tegne og skrive litt på frihånd. Understrek for testerne at dette er ikke en test av dem, men en test av PixelSense.
- Testerne skal forklares at de vil få fremvist en plakat med sine symboler under testen, og at de skal tegne disse inn i matrisen så fort de klarer.
- Dette er matrisen for plasseringen av symbolene for hver bruker, samt deres farger og eksempel på plakat:

b1	b2	b3	b1	b2	b3
b2	b3	b1	b2	b3	b1
b3	b1	b2	b3	b1	b2



- Rekkefølgen på symbolene skal variere for hver bruker for hvert testløp (men plasseringene skal være de samme).
- Testerne skal *ikke* fortelles på forhånd at plasseringene er like.
- Tid skal tas individuelt. Derfor må testerne forstå at de ikke bør vente på hverandre. Oppfordre gjerne til samarbeid. De skal likevel bare tegne sine egne symboler.
- Testerne skal ha faste plasser rundt bordet gjennom hele eksperimentet. De skal stå på langsidene; 2 på den ene siden, 1 på den andre. De får ikke lov til å bevege seg rundt bordet.

5.2.5. Resultater og konklusjoner fra evaluering del 2

Gjennomføringen av eksperimentet gikk forholdsvis problemfritt, kun med små forsinkelser og ingen svikt i utstyret. Det største problemet ble at én av de tre deltagerne fra pulje 2 uteble, og vi rakk ikke finne en erstatning. I analysen av statistikkene måtte vi derfor skape "dummy-data" for test 3 i denne puljen. Vi tok mangelen på resultatene i betraktning under tolkningen av dataene.

5.2.6. Analyse av eksperimentdata

For å kunne bekrefte det vi observerte under eksperimentet og vitenskapelig avkrefte eller bekrefte nullhypotesen vår, gjennomførte vi en statistisk analyse av dataene. Først gjorde vi beskrivende analyse av dataene for å se etter klare forskjeller. Dette gjorde vi ved å finne gjennomsnitt, avvik, varians og standardavvik for hver bruker, og ved å finne gjennomsnittet for alle brukerne for hver tilstand.

[4] Vi valgte å gå bort ifra IFI-studenter pga bias knyttet til kjennskap til vårt prosjekt.

Gjennomsnittlig tid for de ulike testene

	Test med 1 bruker	Test med 2 brukere	Test med 3 brukere
Sekunder i gjennomsnitt	43	29,19	23,94

Det mest interessante med den beskrivende statistikken var den store variasjonen mellom gjennomsnittene til brukerne (fra 14,67 sek for den raskeste til 56,25 sek for den tregeste), og at det var tendens til at gjennomsnittet minket for hver test brukerne var igjennom. Dette tyder på at det var en læringskurve, at brukerne ble mer kjent med prototypen, bordet og symbolene for hver test.

5.2.7. Statistiske tester

For å teste hypotesen vår videre, valgte vi å gjøre noen statistiske tester for å se om vi fant noen signifikante forskjeller mellom tilstandene. Da vi hadde et within-group ekseriment med én uavhengig variabel og tre tilstander valgte vi å gjøre en Repeated-measure ANOVA-test med et signifikansnivå på 5 prosent. Vi valgte 5 prosent signifikansnivå fordi det da er veldig liten sannsynlighet for å gjøre type-1 feil, men likevel mulig å finne statistisk signifikans (Ottar Hellevik, 2011).

Repeated measure ANOVA-testresultater

Source of variation	df	F	P-value	F crit
Colums	2	10,25	0,0018	3,74

5.2.8. Konklusjon av ANOVA-test

Det er en signifikant forskjell i brukbarhet mellom 1 eller flere brukere, $F(2) = 10,24$, $p = 0,0018$.

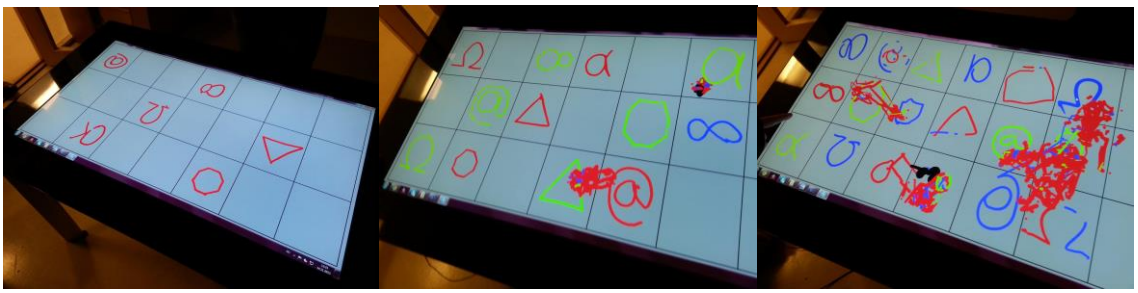
Resultatet tilsier at det er en statistisk signifikant forskjell mellom tilstandene, og vi kan dermed forkaste null hypotesen og godta den alternative hypotesen. Siden denne testen bare tester om det er statistisk forskjell og ikke tester hvilke tilstander det er forskjell mellom, ble vi nødt å kjøre en *t*-test mellom to og to av tilstandene, for å finne ut hvor forskjellene lå.

I t-testene måtte vi teste med et signifikansnivå på 0,017 i stedet for 0,05, fordi jo flere signifikanstester man gjennomfører jo større sannsynlighet er det å finne et signifikant resultat uten at det faktisk er det. Derfor deler vi 0,05 på tre og får et signifikansnivå på 0,017 (Ottar Hellevik, 2011).

T-testene viste en signifikant forskjell mellom 1 og 2 brukere og mellom 1 og 3 brukere, men ingen signifikant forskjell mellom 2 og 3. Dette antyder at læringskurven jevner seg ut fra andre til tredje test.

5.2.9. Tolkning av observasjoner og tilbakemeldinger

Disse tolkningene er basert på subjektive tilbakemeldinger fra brukerne og observasjonene fra videoopptak og fotografier vi tok under eksperimentet.



Test med 1 bruker Test med 2 brukere Test med 3 brukere

I kontrast til vår konklusjon i den kvantitative analysen, kan vi påstå at brukbarheten påvirkes negativt av flere samtidige brukere. Vi har observert en klar sammenheng mellom antallet brukere og mengder støy og feil. Imidlertid skal det sies at vi kunne se konkurranse mellom testerne på tross av at vi understreket at testen ikke var ment som en konkurranse. Dette kom fra brukerenes eget initiativ, og kan ha forskøvet resultatene i retning av hurtighet på bekostning av feil. Vi måtte trolig ha holdt tidtagningen skjult for testerne for å eliminere konkurranseinstinktet.

Den generelle konsensus blant deltagerne var at det var lettere å jobbe parvis eller enkeltvis, heller enn alle tre samtidig. Imidlertid fikk vi klar beskjed at de to som sto ved siden av hverandre fikk vanskeligheter i alle settingene. Observasjonene fra videoopptaket underbygger dette; man ser at de må krysse armer og lene seg over hverandre når de skal tegne symboler på motsatte sider av bordet. Dette forårsaket også de verste tilfellene av støy (se bildet av 3-testen over). De vanskelige forholdene framprovoserte altså til ufrivillige feil, hvilket er akkurat det vi ønsket.

"Det hadde gått bedre om det ble avtalt hvor man skulle jobbe."

I rak kontrast til tilbakemeldingene fra brukertesten med SINTEF, var ikke disse brukerne begeistret for trykknappen på pennen. Dette gir nok en grunn til å finne en løsning på en innebygd bryter i tuppen til en senere iterasjon:

"Det er vanskelig å vite når man skal trykke inn knappen."

5.2.10. Reliabilitet og validitet

Våre testresultater skal ikke generaliseres til bruk av PixelSense og MASTER i en reell arbeidskontekst der brukerne er ledere fra brann eller politi. Vi har kun demonstrert hvordan brukeropplevelsen og brukbarheten til vår penneløsning kan bli påvirket under generelle flerbrukersituasjoner med trange forhold og tidspress.

6. Konklusjon av Prosjektet

Denne rapporten har beskrevet prosjektet gjort av U.S.I.R høsten 2013. Oppgaven vi har svart på har vært *"Én kontroller vs. flere kontrollere i en flerbrukersetting"*, som baserer seg på Microsoft PixelSense og er en forlengelse av BRIDGE MASTER system-prosjektet SINTEF har pågående.

Vi har designet og bygd et redskap med tilhørende programvare som kan benyttes som en alternativ interaksjonsform på Microsoft PixelSense. Pennen tilbyr unik identifikasjon og automatisk loggføring av

strøk, noe som potensielt kan underbygge fremtidig utvikling og bruk av MASTER i reelle brukskontekst med nødetater.

Vi har testet prototypene med brukere, hvor vi prøvde å simulere vanskelige arbeidsforhold. Vårt eksperiment viser at effektiviteten til arbeidet under flerbrukersituasjoner ikke påvirkes negativt, men imidlertid at brukeropplevelsen av PixelSense lider under slike forhold. Pennen gjør situasjonen bedre, ved at man kan se tydelig hvem som gjorde hva etter et brukstilfelle. Arbeidsflaten blir også lettere tilgjengelig ved bruk av pennene, takket være økt rekkevidde og presisjon.

Vi har observert at to brukere som står ovenfor hverandre er en optimal fordeling av brukere. Dersom SINTEF ønsker å implementere vår løsning i MASTER med tanke på hypotetiske flerbrukersituasjoner med flere ledere fra flere ulike etater (eller ulike sektorledere innenfor samme etat), anbefaler vi på bakgrunn av de samme observasjonene, samt tilbakemeldinger fra testsubjekter, at brukerne bør fordeles jevnt rundt bordet. Vår støyfiltrering håndterer ikke strøk som er innenfor størrelsesverdiene som gjenkjennes som penner av vår programvare. Derfor må en viss periode med opplæring i bruk av pennene, anbefalt ca 20-30 min, tilbys eventuelle sluttbrukere slik at støy ikke forstyrrer arbeidet.

For videre forbedring av pennene, bør en løsning på en trykkfølsom aktivering bygges inn i tuppen, istedenfor den eksterne trykknappen som benyttes i den seneste versjonen. I tillegg bør en viskelærfunksjon bygges inn; det står bare på å finne en måte å slette strokes på uten å bruke global *eraseByPoint* i programvaren.

Vi har dessverre ikke rukket å utføre summative evalueringer med reelle sluttbrukere. Vi håper likevel på at dette kan arrangeres på et senere tidspunkt, og tilbyr vår tjeneste i et eventuelt videre samarbeid.

7. Kilder

7.1. Trykte

1. I. Sommerville (2011) Software Engineering 9th ed.
2. Lazar, Feng and Hochheiser (2010) Research Methods in Human-Computer Interaction
3. J. Preece, Y. Rogers, and H. Sharp (2011), Interaction Design beyond Human Computer Interaction, John Wiley & Sons.
4. Ottar Hellevik (2011), Forskningsmetode i sosiologi og statsvitenskap, Universitetsforlaget
5. Nilsson, Stølen. 2011. Generic functionality in user interfaces for emergency response
6. Stiso et al. 2013. Situation awareness in crisis management.

7.2. Internett

Iterative design (2013) Wikipedia:
http://en.wikipedia.org/wiki/Iterative_design (lest 06.11.2013)

Bridge master system (2013), Bridge project:
http://www.bridgeproject.eu/content/bridge_master_system_flyer.pdf (lest 06.11.2013)

Parallel prototyping (2012), Michael Straker:
http://www.cardinalpath.com/web-design-the-case-for-parallel-prototyping/ (lest 06.11.2013)

Parallel prototyping (2010), Dow et al:
http://hci.stanford.edu/publications/2010/parallel-prototyping/ParallelPrototyping2010.pdf (lest 06.11.2013)

PixelSense Technology (2012) Microsoft:
http://www.microsoft.com/en-us/pixelsense/pixelsense.aspx (lest 21.11.2013)