

Groupaware Interfaces

Bram Meerten

Promotor: Dr. Davy Vanacken

Co-promotor: Prof. dr. Kris Luyten

2015-2016

Abstract

Deze thesis bespreekt hoe we, gegeven enkele personen, kunnen detecteren welke personen samen een groep vormen. Om groepen te kunnen detecteren hebben we een groepdetectie methode genaamd *the Gathering* ontwikkeld. Deze methode houdt rekening met de afstand tussen personen, naar waar personen aan het kijken zijn en met features (objecten zoals displays en tafels) in de omgeving. Om de posities van personen in een ruimte te bepalen maken we gebruik van een motion capture systeem genaamd Optitrack, gebruikers dragen markers op hun hoofd zodat we hun positie en oriëntatie kunnen bepalen.

Er is ook een gebruikerstest uitgevoerd om te testen hoe goed *the Gathering* werkt. In het eerste deel van de gebruikerstest hebben we zes personen vrij laten rondlopen in een ruimte terwijl ze drie vragen proberen te beantwoorden die in de ruimte verspreid liggen. In het tweede deel van de gebruikerstest hebben we vier personen situaties laten naspelen zodat we bepaalde scenario's die we wilden uitproberen konden testen. In meer dan 80% van de gevallen werden alle groepen correct gedetecteerd. Wanneer niet correct groepen gedetecteerd werden kwam dit enerzijds doordat de posities van enkele personen verkeerd bepaald werden omdat enkele markers niet goed zichtbaar waren voor de camera's. In de andere gevallen kreeg *the Gathering* wel de juiste informatie van het trackingssysteem, maar slaagde *the Gathering* er niet in de groepen correct te detecteren. Bijvoorbeeld wanneer twee personen elkaar passeerden werden ze even als een groep beschouwd.

In de toekomst kan er nog onderzocht worden of *the Gathering* nog rekening kan houden met andere parameters zoals spraak en lichaamstaal. En ook of *the Gathering* informatie over meerdere frames heen kan verzamelen en gebruiken (zoals beweging).

Woord vooraf

Het werken aan mijn thesis was een interessante en leuke ervaring. Ik had dit resultaat nooit alleen kunnen realiseren. Daarom zou ik graag mijn promotor dr. Davy Vanacken en co-promotor prof. dr. Kris Luyten bedanken voor hun inzicht en feedback. Ook zou ik graag Dr. Jo Vermeulen bedanken voor zijn advies.

Daarnaast ben ik iedereen erg dankbaar die tijd heeft vrij gemaakt om deel te nemen aan de gebruikerstest.

Verder ben ik ook nog alle professoren en assistenten dankbaar die me op de universiteit zoveel hebben bijgebracht. Tenslotte wil ik ook mijn medestudenten bedanken waarmee ik zo'n leuke tijd beleefd heb.

Inhoudsopgave

1	Introductie	8
2	Scenario's	10
3	Proxemic interaction	15
3.1	Proxemic zones	15
3.1.1	Raamwerk voor interactie fases	16
3.2	Fixed en semi-fixed features	18
3.3	Proxemic dimensions	18
3.4	De proximity toolkit	19
4	Bestaande tracking technieken	21
4.1	Diepte camera's	21
4.1.1	Diepte camera's gebruiken om personen te tracken	21
4.2	Optical motion capture	22
4.3	Tracking op basis van kleur	23
4.4	Vergelijking tracking methodes	24
5	Bestaande groepdetectie methodes	26
5.1	GroupTogether	26
5.2	SpiderEyes	28
5.3	Rule-based approach	29
5.3.1	De regels	29
5.3.2	Regels combineren tot een set	30
5.3.3	Scenario's	30
5.3.4	Demo: Whac-a-Mole	31
5.4	Vergelijking groepdetectie methodes	32
6	Toepassingen	34
6.1	SpiderEyes	34
6.2	Restaurant	36
6.3	Foto muur	37

7	The Gathering	38
7.1	Analyse scenario's	38
7.1.1	Conclusie analyse	40
7.2	Input verzamelen voor the Gathering	40
8	The Gathering: Algoritme	43
8.1	Interpersonal distance	43
8.2	Displays	44
8.3	Bepalen of personen met elkaar interageren	46
8.4	Accuraatheid	48
9	The Gathering: Architectuur	52
9.1	The Gathering gebruiken in een applicatie	53
9.1.1	Features toevoegen	55
9.2	Tool support	56
10	The Gathering: Evaluatie	60
10.1	Deelnemers	60
10.2	Set-up	60
10.3	Deel 1: Vrij rondlopen	61
10.3.1	Resultaten	62
10.3.2	Conclusie	67
10.4	Deel 2: Vooraf bepaalde situaties testen	67
10.4.1	Resultaten	70
10.4.2	Conclusie	74
10.5	Algemene conclusie	75
11	Conclusie en future work	76

Lijst van figuren

1.1	CityWall, een publieke display in Helsinki	8
2.1	Scenario waarin een groep een dagschema opvraagt.	11
2.2	Scenario waarin een gids foto's toont	12
2.3	Scenario waarin een nieuwe gebruiker een display benadert .	12
2.4	Scenario waarin een groep van twee personen split	14
3.1	De vier proxemic zones	16
3.2	Interactie raamwerk van Vogel en Balakrishnan	17
3.3	Feature moedigt interactie af	18
3.4	Proximity toolkit	20
4.1	Personen tracken met een overhead Kinect	22
5.1	Twee F-formations	27
5.2	SpiderEyes	29
5.3	Whac-a-Mole	31
6.1	Een gedachtenwolk die wordt toegevoegd aan het spiegelbeeld van een persoon	36
7.1	Optitrack camera's	41
8.1	Groepen vormen o.b.v. interpersonal distance	44
8.2	Displays als deel van een groep	45
8.3	Zwaartepunt van een groep	48
8.4	Focus gebieden van personen	49
8.5	Accuraatheid groepdetectie	50
8.6	Invloed van features op groepen	51
9.1	Overzicht architectuur the Gathering	52
9.2	Tool support	57
9.3	Scenario's afspelen via een CSV-file of realtime via NatNet .	57
10.1	Set-up eerste deel gebruikerstest	61
10.2	Oorzaken van het falen van tracking	63

10.3	Gebruiker loopt om andere gebruiker heen	64
10.4	Gebruiker loopt om andere gebruiker heen	65
10.5	Zeer klein focus gebied	65
10.6	Zeer groot focus gebied	66
10.7	Twee groepen benaderen elkaar	68
10.8	Een groep van vier personen praat met elkaar	69
10.9	Twee personen interageren met een display terwijl anderen interageren met een pakje kaarten	69
10.10	Twee groepen passeren elkaar	69
10.11	Een gebruiker gaat naar een andere groep	70
10.12	Display maakt deel uit van een groep	72
10.13	Twee groepen passeren elkaar	72
10.14	Verschil tussen twee kleine displays en één grote	73
10.15	Verschil tussen twee kleine displays en één grote	74

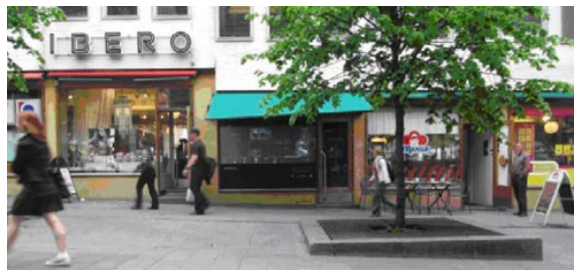
Lijst van tabellen

10.1 Resultaten gebruikerstest deel één	63
10.2 Resultaten gebruikerstest deel twee	70

Hoofdstuk 1

Introductie

De visie van ubiquitous computing wordt alsmaar meer een realiteit, we krijgen telkens maar meer en meer informatie ter onze beschikking, waar we ons ook bevinden. Een deel van die informatie wordt beschikbaar gemaakt door displays geplaatst in publiek, semipublieke en privé omgevingen zoals op straat, luchthavens, scholen, kantoren en huizen [1]. De grootte van deze displays zijn in twee richtingen aan het groeien. De kleinste displays worden alsmaar kleiner en de grootste displays worden alsmaar groter [2]. Façades van gebouwen kunnen zelfs dienst doen als display [3].



Figuur 1.1: CityWall, een interactieve publieke display in Helsinki. Gebruikers kunnen er foto's op bekijken, verplaatsen en roteren [4].

Vele grote displays zoals CityWall [4] hebben plaats genoeg om met meerdere personen te interageren. De grootte van displays is niet de enige reden dat er vaak verschillende personen tegelijk met de display interageren. In een publieke setting interageren slechts heel weinig mensen met een display wanneer ze alleen zijn [5]. Müller raadt daarom aan displays altijd zo te designen dat ze groepen ondersteunen.

Traditioneel wordt er meer gefocust op individuele interactie, maar voor grote displays in een publieke setting zoals media façades en CityWall gaan mensen meer samen interageren dan alleen. Moesten de displays weten

welke personen rondom hun bij elkaar horen, dan kunnen de displays hun content op basis hiervan aanpassen. Bijvoorbeeld in het volgende scenario waar in een museum een display van drie meter hoog en vijf meter breed staat. Gebruikers kunnen naar de display toelopen en dan verschijnt er voor de gebruiker een virtueel fotoalbum van enkele uitgestorven dieren. Als de display zich bewust is van welke personen samen staan, kan hij in plaats van voor elke persoon die voor de display staat een fotoalbum te tonen, een fotoalbum per groep tonen. Bijvoorbeeld wanneer er een persoon links voor de display staat en een groep van drie personen naar het midden van de display toe lopen, dan kan de display een fotoalbum tonen aan de persoon links en een fotoalbum tonen aan de groep van drie personen. Bovendien kan het fotoalbum dat gebruikt wordt door de groep van drie personen meer ruimte op de display gegeven worden dan het fotoalbum dat slechts door één persoon gebruikt wordt.

De onderzoeksvraag van deze thesis gaat als volgt: *Hoe kunnen we, gegeven verschillende personen, bepalen welke personen samen een groep vormen? En wat voor toepassingen kunnen gebruik maken van zo'n groepdetectie methode?* Met een groep personen bedoelen we enkele personen die bij elkaar horen. Bijvoorbeeld een koppel dat een etalage van een winkel binnen kijkt, enkele vrienden die samen over straat lopen, een gezin dat samen op een bankje in het park aan het eten is, enzovoorts. We bedoelen bijvoorbeeld niet een groep kinderen die op schoolreis zijn in een pretpark. Want zo'n groep is meestal verdeeld in verschillende kleinere groepen, bijvoorbeeld twee klasgenoten die op een bankje zitten, vier personen die samen aan het aanschuiven zijn aan een attractie, ... Met een groep bedoelen we in deze thesis ook niet erg grote groepen zoals een groep supporters in een stadion.

Onze methode focust op het detecteren van groepen met als doel deze informatie te gebruiken voor een toepassing op een display die ergens in een ruimte staat. Met een ruimte bedoelen we vooral een publieke ruimte zoals op straat, op een festival en in de inkomhal van het stadhuis. En met een display bedoelen we een grote display, geen kleine displays zoals die van een laptop, computer of smartphone, deze zijn niet zo geschikt om met meerdere personen mee te interageren.

Om een goede methode te kunnen ontwikkelen kijken we naar andere groepdetectie methodes en concepten zoals proxemic interaction. De groepdetectie methode die we ontwikkelt hebben heet *the Gathering* en de werking hiervan wordt uitgebreid besproken in deze tekst. We hebben ook een gebruikerstest uitgevoerd om deze methode te testen.

Hoofdstuk 2

Scenario's

Hieronder worden enkele scenario's gegeven waarin groepdetectie gebruikt kan worden. Dit geeft een goed beeld over hoe the Gathering gebruikt kan worden en over hoe de groepen in een ideaal geval gedetecteerd zouden moeten worden. In hoofdstuk 7.1 worden deze scenario's geanalyseerd om te kijken welke lessen eruit getrokken kunnen worden voor het groepdetectie algoritme.

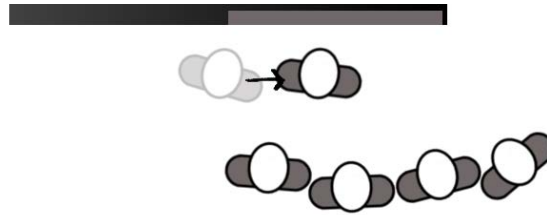
Scenario 1

In dit eerste scenario komt maar één groep voor. Sommige personen in de groep verplaatsen zich ten opzicht van elkaar maar blijven evenzeer één groep.

Op een evenement staat een display waar men onder andere het dagschema van het evenement kan opvragen. Een groep personen naderen de display. Ze willen het dagschema van het evenement bekijken. Eén persoon (Jan) loopt uit de groep naar voren om het dagschema op te vragen. Hoewel maar één persoon het schema opende, wil heel de groep het schema zien. Daarom wordt het dagschema in het groot voor heel de groep getoond.

Aangezien de groep aan de rechterkant van het scherm is gaan staan wordt het dagschema ook aan de rechterkant getoond. Jan was oorspronkelijk naar het midden van het scherm gelopen. Maar doordat de content rechts van hem verschijnt, moet hij zich een beetje naar rechts verplaatsen om verder te kunnen interageren. Op deze manier is er meer plaats vrij op de display voor andere personen of groepen die eventueel nog met de display zouden willen interageren (zie Figuur 2.1).

De groep personen bekijken het dagschema en bespreken het, ze geven aan welke activiteiten ze interessant vinden en Jan klikt op deze activiteiten in het dagschema om zo meer details erover weer te geven.



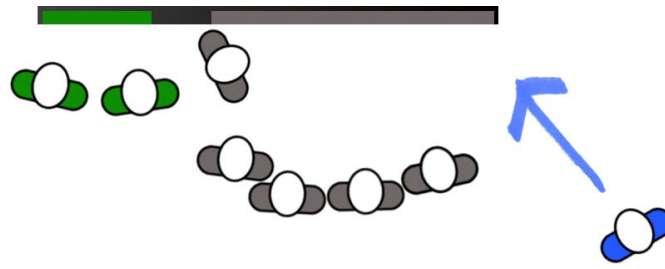
Figuur 2.1: Scenario 1. Jan interageert met een display om het dagschema op te vragen, de rest van de groep kijkt mee. De applicatie toont de content rechts van hem waardoor Jan een beetje moet opschuiven. Nu is er aan de linkerzijde van de display meer plaats voor andere gebruikers.

Scenario 2

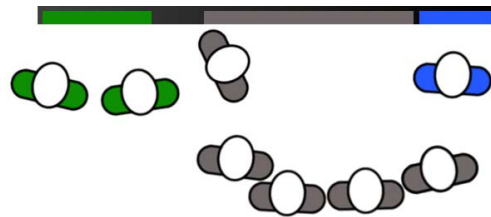
De volgende scenario's bevatten altijd meerdere groepen. Het volgende scenario toont hoe een display zijn content kan aanpassen op basis van het aantal groepen, de positie van de groepen en de grootte van de groepen rondom zich.

In een dinosauriër museum staat een grote display waarop foto's bekeken kunnen worden van hoe dinosauriërs er vroeger uitgezien zouden hebben. Een gids neemt een groep toeristen mee naar de display en opent er foto's om er vervolgens meer over te vertellen. De display ziet dat de gids (die interageert met de display) bij de toeristen hoort en toont de foto's groot genoeg, zodat heel de groep ze kan zien. Terwijl de gids uitleg aan het geven is aan de toeristen, komen er twee nieuwe personen de ruimte binnen en wandelen naar de display. Ze gaan links voor de display staan, om zo de gids niet in de weg te staan. De display herkent dat de twee personen samen horen en toont links een fotogalerij, net voor de twee gebruikers en veel kleiner dan voor de grote groep toeristen. Ze bladeren samen door de foto's.

Ondertussen is heel de display in beslag genomen door de twee fotogalerijen (één grote, en één kleinere). Maar er komt een nieuwe persoon naar de display toe gewandeld (zie Figuur 2.2), rechts van de gids. Er verschijnt een derde fotogalerij rechts, en de andere galerijen worden geherpositioneerd en herschaald om plaats maken voor de derde fotogalerij (zie Figuur 2.3). Maar deze laatste gebruiker raakt de foto's al snel beu en verlaat de display waarna de twee andere galerijen opnieuw geherpositioneerd en herschaald worden.



Figuur 2.2: Scenario 2. Een gids toont aan toeristen foto's van dinosaurïërs. Twee andere gebruikers bladeren links van de groep toeristen door enkele foto's.



Figuur 2.3: Scenario 2. Een nieuwe gebruiker wandelt naar de display toe. De content van de andere gebruikers wordt automatisch verplaatst en herschaald om plaats te maken voor de nieuwe gebruiker.

Scenario 3

Groepen kunnen niet altijd enkel op basis van afstand bepaald worden. In het volgende scenario komen twee groepen gelijktijdig aan en gaan ze dicht bij elkaar staan. Om groepen te detecteren moet er soms ook op meer subtiële zaken, zoals waar personen hun aandacht ligt, gelet worden.

Twee groepjes van elk twee personen benaderen ongeveer gelijktijdig een display. Ze gaan allemaal voor de display staan en vormen zo een rij van vier personen. Hoewel ze alle vier naast elkaar staan vormen ze evenzeer twee verschillende groepen. De display toont voor beide groepen content waar ze vervolgens mee interageren.

Scenario 4

Groepen blijven niet altijd hetzelfde. Personen kunnen een groep verlaten zoals in het volgende scenario.

Op een evenement staat een display waarop automatisch een filmpje van de vorige editie van het evenement getoond wordt wanneer er personen voorbij lopen. Maar er kan ook onder andere het dagschema van het evenement op-

gevraagd worden. Een groep personen naderen de display en blijven staan om het filmpje te bekijken. Eén persoon (Eric) loopt uit de groep naar de display toe om het dagschema op te vragen. Maar de focus van de andere groepsleden blijft gericht op het filmpje. Hierdoor blijft het filmpje voor de groep verder spelen en het dagschema wordt enkel vlak voor Eric getoond. Eric bladert door de activiteiten terwijl de rest van de groep naar het filmpje blijft kijken.

Na een tijdje heeft de groep genoeg van het filmpje en zijn ze ook benieuwd naar het dagschema. Eén persoon (Lize) loopt naar voren om het dagschema op te vragen. De rest van de groep focussen niet meer op het filmpje maar volgen hoe Lize het dagschema opent. Daarom wordt het dagschema voor de volledige groep getoond en verdwijnt het filmpje. De groep bespreekt welke activiteiten ze interessant vinden en Lize opent deze activiteiten om meer details weer te geven. Eric ziet dat de groep aan het overleggen is over naar welke activiteiten ze willen gaan en gaat bij hun staan om mee te overleggen. Het dagschema dat Eric aan het bekijken was verdwijnt en Eric wordt terug beschouwd als een deel van de groep.

Scenario 5

Net zoals in het vorige scenario kan het aantal personen in een groep veranderen. In dit scenario wordt een persoon, die oorspronkelijk geen deel van de groep uitmaakte, aan een groep toegevoegd.

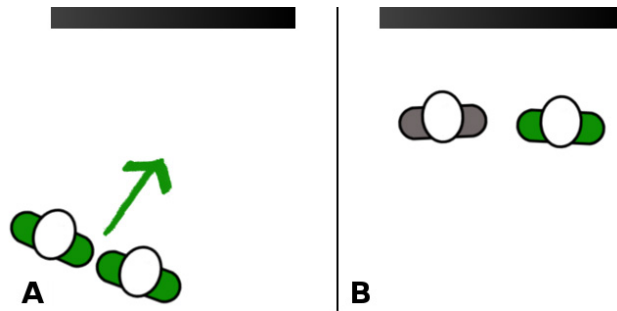
John en Sarah gaan samen naar een reisbureau. John gaat al naar binnen terwijl Sarah een parkeerplaats zoekt. John wandelt naar een display toe om wat informatie op te vragen, er zijn al andere personen aanwezig die ook met de display aan het interageren zijn. Terwijl John aan het interageren is met de display komt Sarah binnen en gaat bij John staan. Omdat John en Sarah samen een groep vormen, wordt de content zo gepositioneerd en geschaald dat ze voor beide duidelijk zichtbaar is.

Scenario 6

Het volgende scenario toont hoe gebruikers subtiel kunnen aangeven dat ze niet als één groep beschouwd willen worden, ook al worden ze oorspronkelijk als één groep gedetecteerd.

Piet en Joris komen elkaar tegen op de parking van een ziekenhuis, ze kennen elkaar van vroeger op school. Piet komt zijn vriendin bezoeken in het ziekenhuis, Joris komt zijn vader bezoeken. Ze lopen samen de inkomhal van het ziekenhuis binnen en lopen samen naar een display toe. Op deze display kunnen ze opzoeken welke patiënten in welke kamers liggen. De display toont één lijst van kamers voor hun beide, maar Piet wil het kamernummer

van zijn vriendin weten en Joris het kamernummer van zijn vader. Ze gaan daarom een beetje uit elkaar staan. De display toont nu voor beide een lijst waar ze elk mee kunnen interageren (zie Figuur 2.4).



Figuur 2.4: Scenario 6. Twee personen lopen samen naar een display maar gaan daarna een beetje uit elkaar te staan om elk apart met de display te interageren.

Hoofdstuk 3

Proxemic interaction

In dit hoofdstuk bespreken we proxemic interaction en hoe we proxemic interaction en concepten uit proxemic interaction kunnen gebruiken in groepdetectie.

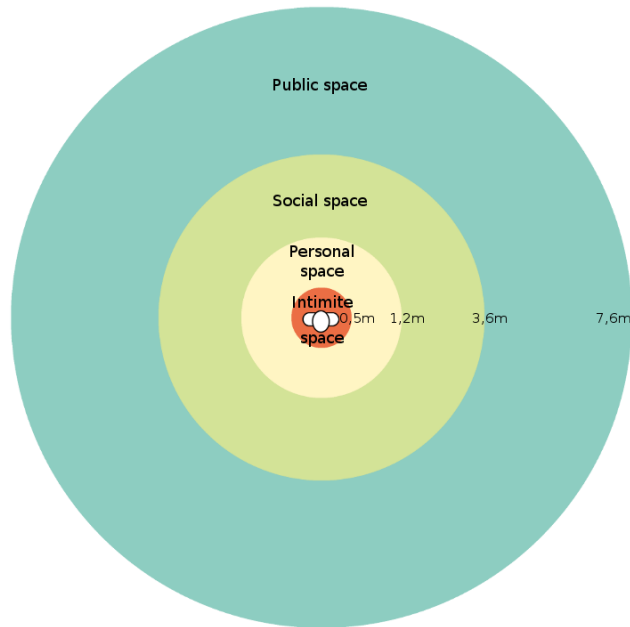
In proxemic interactions wordt de kennis van de relaties tussen personen, devices en objecten (proxemics) gebruikt om interactietechnieken te ontwerpen [6]. Deze relaties tussen personen zijn ook belangrijk bij het detecteren van groepen. Zo is bijvoorbeeld de afstand tussen verschillende personen een belangrijke indicatie dat ze samen een groep vormen. Ook andere concepten uit proxemic interaction zoals proxemic dimensions en fixed en semi-fixed features worden gebruikt in the Gathering. Ook voor applicaties die gebruik maken van the Gathering kan proxemic interactions gebruikt worden bij het ontwerpen van interacties. Bijvoorbeeld de afstand van een groep tot een display kan gebruikt worden. In hoofdstuk 6.1 beschrijven we een toepassing die hier gebruik van maakt, hoe dicht de groep bij een kaart staat, hoe meer detail er op die kaart getoond wordt.

3.1 Proxemic zones

Proxemics is een theorie ontworpen door antropoloog Edward T. Hall [7]. De theorie beschrijft hoe mensen afstand, houding en oriëntatie tussen personen (interpersonal spatial relationships) interpreteren en gebruiken. Hall definieerde vier proxemic zones die aangeven hoe mensen de afstand tussen elkaar interpreteren:

- Intimate (< 0,5 meter)
- Personal (0,5 tot 1,2 meter)
- Social (1,2 tot 3,6 meter)

- Public (3,6 tot 7,6 meter)¹



Figuur 3.1: Hall definieerde vier proxemic zones: Intimate, personal, social en public

Hoe kleiner de afstand tussen personen, hoe persoonlijker de interactie tussen de personen wordt. Mensen passen deze afstanden (onbewust) aan, afhankelijk van hun sociale activiteit. Deze zones kunnen ook uitgebreid worden naar interactie tussen bijvoorbeeld personen en devices (zie hoofdstuk 3.1.1).

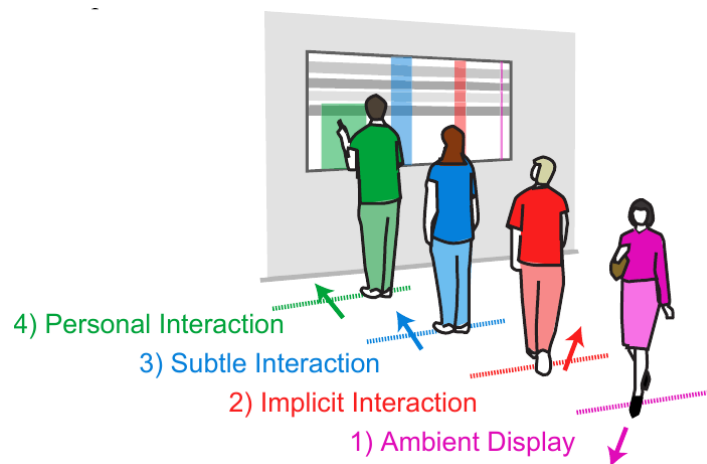
3.1.1 Raamwerk voor interactie fases

Vogel en Balakrishnan ontwikkelden een interactie raamwerk met vier continue interactie fases, die bepalen hoe verticale displays reageren op personen die het scherm benaderen (zie Figuur 3.2) [1]. Deze vier fases zijn gelijkwaardig aan de proxemic zones van Hall. Hall's proxemic zones gaan over de interacties tussen personen, terwijl deze fases over de interactie tussen persoon en device gaan.

- **Ambient display:** In deze fase is een display in zijn neutrale fase en toont een overzicht van beschikbare informatie, updates gebeuren maar langzaam.

¹Intimate: tot 1,5 feet; Personal: 1,5-4 feet; Social: 4-12 feet, Public: 12-25 feet

- **Implicit interaction:** Wanneer een persoon voorbijkomt gaat het systeem over naar deze fase. Het systeem moet uit de positie en de oriëntatie van de persoon, proberen af te lezen of deze open staat om informatie te ontvangen. Indien de persoon open staat om te communiceren moet het systeem een abstracte representatie van de persoon op het scherm weergeven. Dit helpt de gebruiker dichter naar het scherm te lokken en zo de volgende fase te activeren.
- **Subtle interaction:** Deze fase wordt geactiveerd wanneer een gebruiker het scherm benadert en impliciet aangeeft dat hij wil communiceren door bijvoorbeeld even te pauzeren. Nu kan meer gedetailleerde informatie getoond worden. Deze info kan eventueel aangevuld worden met persoonlijke informatie.
Vanaf deze fase kan ook expliciete interactie gebruikt worden, bijvoorbeeld met gestures.
- **Personal interaction:** Wanneer de gebruiker nog dichter naar het scherm stapt (nadat hij eventueel een informatie item geselecteerd heeft in de *subtle interaction* fase), wordt deze fase geactiveerd. In deze fase kan nog gedetailleerdere informatie getoond worden en nog meer persoonlijke informatie.



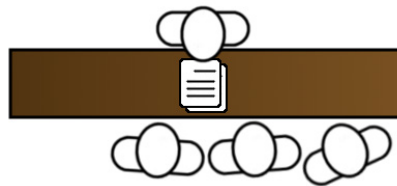
Figuur 3.2: Vogel en Balakrishnan ontwikkelden een interactie raamwerk met vier interactie fases die bepalen hoe verticale displays reageren op personen die het scherm benaderen [1].

Een belangrijk onderdeel van het raamwerk is de vlotte overgang tussen de verschillende fases. Eerst is de interactie impliciet (via de oriëntatie, positie en beweging van het lichaam), maar deze wordt geleidelijk aan explicieter (gestures en touch).

3.2 Fixed en semi-fixed features

Hall beschrijft ook hoe *features* in een omgeving mensen hun interactie beïnvloeden. Je hebt *fixed* features, dit zijn onder andere de features die grenzen markeren zoals muren, deuren en ramen [8]. Maar ook andere onverplaatsbare dingen zoals displays die op een vaste locatie staan. *Semi-fixed* features zijn objecten zoals stoelen, tafels en ander meubilair. Hoewel de meeste niet vaak verplaatst worden, is het wel mogelijk om ze te verplaatsen om zo aan een situatie aan te passen (bijvoorbeeld stoelen voor een tv plaatsen om een computerspelletje te spelen).

Informatie over de relaties tussen personen en features (afstand, oriëntatie, . . . , zie hoofdstuk 3.3) kunnen belangrijke indicaties zijn voor impliciete interacties. Bijvoorbeeld een tv die aanspringt wanneer iemand er naartoe loopt.



Figuur 3.3: De toonbank moedigt interactie tussen meer dan drie personen af omdat niet iedereen goed kan meekijken op de papieren.

Features kunnen ook interacties aanmoedigen of ontmoedigen. Marshall et al. onderzochten hoe fysieke structuren sociale relaties kunnen aanmoedigen of ontmoedigen [9]. Ze bestudeerde voor één dag mensen in een toeristisch informatie centrum. Hun analyse focuste op de locatie en compositie van F-formaties (voor meer informatie over F-formaties zie hoofdstuk 5.1), om zo bij te houden wat voor soort sociale interacties er zich voordeden en tussen wie. Ze kwamen tot de conclusie dat bijvoorbeeld de toonbank interactie tussen meer dan drie personen ontmoedigde omdat dan niet iedereen goed kon meekijken op papieren met toeristische informatie (zie Figuur 3.3). Een tafel moedigde groepsinteractie dan weer aan. Omdat features een invloed hebben op het vormen van groepen, houdt the Gathering hier ook rekening mee (zie hoofdstuk 8.2: displays).

3.3 Proxemic dimensions

Om proximity te gebruiken in een applicatie, moet je proximity kunnen meten. Ballendat et al. identificeren vier proxemic dimensies die volgens

hun essentieel zijn om proxemics tussen entiteiten (personen, devices en gewone objecten) te kunnen detecteren [6]:

- **Positie:** Deze kan zowel absoluut als relatief beschreven worden. De absolute positie wordt beschreven als de (drie dimensionale) positie ten opzichte van een vast punt. De relatieve positie wordt bepaald door middel van de spatiale relatie tussen twee entiteiten. Bijvoorbeeld tussen een object en een persoon of in een kamer.
- **Oriëntatie:** Ook de oriëntatie kan zowel absoluut (yaw, pitch, roll) als relatief (kijkt naar object) beschreven worden. Maar deze dimensie kan alleen gebruikt worden wanneer een entiteit een goed gedefinieerde voorkant heeft.
- **Beweging:** De verandering van de positie en oriëntatie over tijd. Zo kan er gemeten worden of twee personen naar elkaar toe wandelen, of juist van elkaar weg wandelen. Ook de snelheid van de beweging kan gemeten worden.
- **Identiteit:** Beschrijft een entiteit op een unieke manier, bijvoorbeeld een naam of ID. Dit kan ook iets minder precies zijn zoals een categorie (persoon, boek, ...).

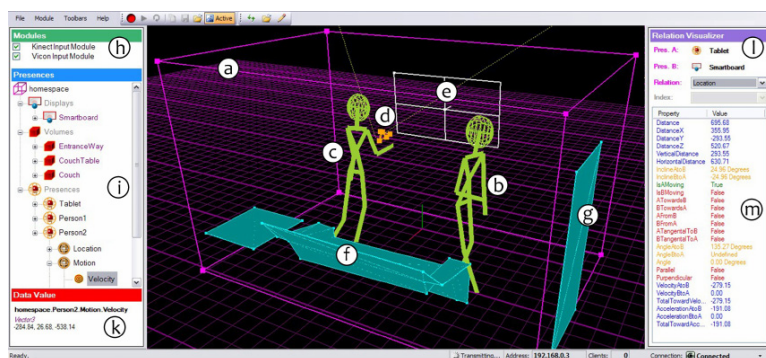
Greenberg et al. voegen hier nog een extra dimensie (afstand) aan toe [8]:

- **Afstand:** Typisch denken we aan afstand als een absolute meting (bijvoorbeeld afstand in meters). Maar de afstand kan ook discreet zijn, denk bijvoorbeeld aan Hall's proxemic zones (zie hoofdstuk 3.1).

In Hall zijn proxemic zones wordt slechts één proxemic dimension gebruikt namelijk de afstand tussen personen.

3.4 De proximity toolkit

De proximity toolkit is een toolkit ontworpen om meer te kunnen focussen op het ontwerpen van proxemic interacties en minder bezig te zijn met onderliggende zaken zoals het verwerken van sensor data [8][10]. De sensor hardware wordt op deze manier gescheiden van het proxemic data model gegenereerd uit deze sensors. Net zoals de proximity toolkit het ontwerpen van interacties probeert te vereenvoudigen, probeert the Gathering het ontwerpen van groepsinteracties te vereenvoudigen, door het detecteren van groepen en het verwerken van sensor data te scheiden van het ontwerpen van groepsinteracties.



Figuur 3.4: De proximity toolkit is een toolkit ontworpen om meer te kunnen focussen op het ontwerpen van proxemic interacties en minder bezig te zijn met onderliggende zaken zoals het verwerken van sensor data [10].

De proximity toolkit trackt objecten en de proximity relaties tussen deze objecten, hieruit wordt afstand, oriëntatie, identiteit en bewegingsinformatie gegenereerd. Deze informatie wordt via events gemakkelijk beschikbaar gemaakt. De toolkit bevat ook tools die de proxemic relaties in 3D visualiseert en waarin de proxemic relaties makkelijk bekeken, verkend en opgenomen kunnen worden (zie Figuur 3.4). Deze tool toont ook de fixed en semi-fixed features in de ruimte (zetel, muren, ...). Qua sensor hardware ondersteunt de toolkit voorlopig het Vicon motion capturing system en de Kinect sensor.

Hoofdstuk 4

Bestaande tracking technieken

Zoals besproken in het vorige hoofdstuk kunnen met proxemic dimensions de proxemics tussen verschillende entiteiten (personen, devices en objecten) gedetecteerd worden, maar om deze proxemic dimensions te meten moeten we eerste de positie en de oriëntatie van elke persoon in de ruimte kennen. Hieronder worden enkele technieken beschreven om personen te tracken en worden deze technieken met elkaar vergeleken.

4.1 Diepte camera's

Diepte camera's zijn camera's die niet alleen RGB informatie opnemen, maar ook voor elke pixel diepte informatie opnemen [11]. Een voorbeeld van zo'n diepte camera is de Microsoft Kinect, ontwikkeld voor de Xbox spel-console. Door de lage prijs en de goede beschikbaarheid van het product wordt het ook vaak gebruikt door onderzoekers in de informatica, robotica en elektronica[12][13]. Met de informatie die diepte camera's bieden kan niet alleen de positie getrackt worden, maar ook het skelet, de hoofdpositie en de gezichtsuitdrukkingen van personen.

Buiten de Kinect zijn er ook nog andere consumenten diepte camera's zoals de Asus Xtion PRO LIVE¹ en ZED².

4.1.1 Diepte camera's gebruiken om personen te tracken

De Microsoft Kinect kan onder andere gebruikt worden in de proximity toolkit (zie hoofdstuk 3.4). Diepte camera's worden ook gebruikt om de positie van personen te tracken door Beyer et al. in hun Puppeteer Display

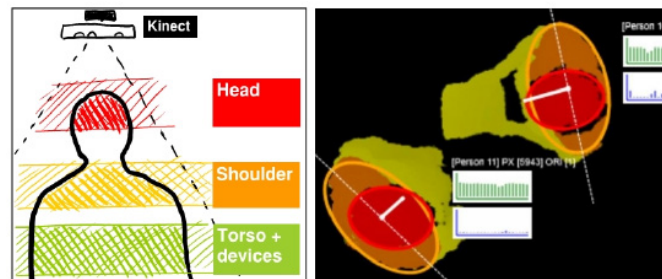
¹www.asus.com/be-nl/3D-Sensor/Xtion_PRO_LIVE

²www.stereolabs.com

[14] en door Winters in zijn rule-based groepdetectie library (zie hoofdstuk 5.3), beide gebruiken de Kinect.

Een probleem met de diepte camera's (wanneer deze horizontaal geplaatst worden) is dat personen voor elkaar kunnen gaan staan en de camera niet iedereen meer kan detecteren. Daarom kan de camera best zo hoog mogelijk opgehangen worden, in vogelperspectief. Maar ook wanneer de diepte camera hoger hangt in vogelperspectief kan occlusie nog regelmatig voorkomen [15].

Marquardt et al. gebruikte ook een diepte camera (de Kinect) om personen te tracken, maar plaatste deze naar beneden kijkend aan het plafond [16]. Op deze manier heb je minder last van occlusie. Marquardt et al. maakten geen gebruik van het skelet gegenereerd door de Kinect (omdat ze gebruik maakten van een overhead Kinect), maar detecteerden zelf personen uit de diepteafbeeldingen. Ze detecteren voor elke persoon het hoofd, de schouders en de torso (zie Figuur 4.1). Het nadeel van een overhead camera is dat het gebied waarin personen getrackt kunnen worden typisch niet zo groot is [17]. Marquardt et al. combineerden de diepteafbeeldingen van twee diepte camera's tot één diepteafbeelding om zo het gebied te vergroten.



Figuur 4.1: Personen tracken met een diepte camera. Bij elke persoon wordt het hoofd, de schouders en de torso gedetecteerd [16]. Rechts staat een diepteafbeelding van twee rechtstaande personen.

4.2 Optical motion capture

Optical motion capture systemen zoals de OptiTrack³ en Vicon⁴ gebruiken data verzameld door beeldsensoren om de 3D positie van personen of andere onderwerpen te berekenen via triangulatie. Bij het OptiTrack systeem en Vicon worden infrarood sensoren gebruikt, rondom deze infrarood sensoren zitten LEDjes die een infrarood signaal uitzenden. Het onderwerp dat getrackt wordt draagt markers die de infrarood signalen, uitgezonden door de LEDjes, reflecteert en dat signaal wordt dan opgevangen door de infrarood

³www.optitrack.com

⁴www.vicon.com

sensoren. Via triangulatie worden dan de posities van de markers berekend, bij een groep van drie of meer markers kan ook de oriëntatie berekend worden. In hoofdstuk 7.2 wordt de werking van het OptiTrack systeem nog meer in detail uitgelegd.

Markers zoals hierboven beschreven die enkel licht reflecteren zijn passieve markers. Er zijn ook markers die zelf licht uitzenden, dit zijn active markers. Een voordeel van active markers is dat elke marker uniek geïdentificeerd kan worden door de markers synchroon met het systeem één voor één licht te laten uitzenden. Passive markers kunnen niet uniek geïdentificeerd worden, maar een set van passieve markers kan wel een uniek patroon vormen (een rigid body) dat herkend kan worden. Het nadeel van active markers is dat de markers een energiebron nodig hebben.

In een optical motion capture systeem kunnen makkelijk extra camera's toegevoegd worden om zo occlusie te verminderen. Bij diepte camera's kun je niet zo gemakkelijk extra camera's toevoegen. De Kinect werkt bijvoorbeeld niet out of the box met meerdere camera's, maar er bestaan wel libraries voor [18]. Door de hoge USB bandbreedte kunnen er slechts twee of drie Kinects verbonden worden met één computer. Martinez-Zarzuola et al. lossen dit op door een client-server systeem te gebruiken en Kinects aan verschillende computers te koppelen [19].

Een nadeel van motion capture systemen zoals OptiTrack is dat de te tracken personen markers moeten dragen en dat het systeem veel duurder is dan diepte camera's.

4.3 Tracking op basis van kleur

McKenna et al. maakten gebruik van een kleurenmodel om personen op videobeelden te onderscheiden [20]. Personen worden gedetecteerd door met adaptive background subtraction de achtergrond te verwijderen. De overlappende voorgrond regio's vormen een persoon, mits ze aan een paar vereisten voldoen (ze moeten een minimum oppervlakte hebben en hun projectie op de x-as moet overlappen). De adaptive background subtraction houdt rekening met langzame verandering in de achtergrond en kan ook schaduwen verwijderen. Van elke persoon wordt vervolgens een kleurenmodel gebouwd, dit kleurenmodel wordt continu aangepast. Het kleurenmodel wordt gebruikt om de persoon te kunnen herkennen. Deze methode kan ook personen herkennen die in een groep staan en slechts gedeeltelijk zichtbaar zijn.

Het nadeel van deze methode is dat wanneer twee gebruikers gelijkaardige kleuren kleren dragen, het systeem niet meer weet wie welke persoon is nadat ze samen een groep hebben gevormd en weer uit elkaar gaan. Deze methode kan ook enkel personen in 2D tracken.

McKenna et al. gebruikten enkel beelden waarop personen van op zij gefilmd werden, het is niet zeker of deze methode ook met een top down camera zou werken. De minimum oppervlakte van regio's om als persoon beschouwd te worden zou dan een pak kleiner moeten zijn, hierdoor kunnen regio's die geen personen voorstellen makkelijker fout als een persoon beschouwd worden. Bovendien kan de tweede vereiste voor regio's om een groep te vormen, namelijk dat twee regio's hun projectie op de x-as moeten overlappen, niet meer gebruikt worden omdat personen top down zich in twee richtingen kunnen bewegen. Tenslotte is er doordat de camera top down geplaatst is veel minder kleureninformatie beschikbaar om een persoon te kunnen onderscheiden.

4.4 Vergelijking tracking methodes

Om groepen te kunnen detecteren moet de 2D positie (top down) van personen in de ruimte getrackt worden zodat we weten hoe personen ten opzichte van elkaar gepositioneerd staan. We moeten ook de oriëntatie van de personen weten zodat we weten naar waar personen zich richten. Dit alles moet realtime gebeuren, zodat realtime groepen gedetecteerd kunnen worden en applicaties meteen kunnen reageren op nieuwe groepen. Dit zijn de minimum vereisten om groepen te kunnen detecteren, zonder bijvoorbeeld de oriëntatie weten we niet of personen met de rug naar elkaar toe staan of niet.

Hieronder worden de verschillende voor- en nadelen van de besproken tracking methodes opgesomd en wordt er gekeken of ze aan de minimum vereisten voldoen. In de toekomst zou er nog meer informatie gebruikt kunnen worden zoals gestures. Twee personen die naar elkaar wuiven kan bijvoorbeeld een indicatie zijn dat ze elkaar kennen en een groep gaan vormen. Een tracking systeem dat nog andere informatie zoals gestures kan detecteren is dus een pluspunt.

Diepte camera's kunnen zowel de positie als oriëntatie van personen realtime tracken en voldoen dus aan de minimum vereisten. Diepte camera's kunnen ook het skelet van personen tracken, het skelet is een verzameling van gewrichten en elk gewricht wordt voorgesteld door een 3D-coördinaat. Met behulp van dit skelet kunnen dus makkelijk gestures zoals wuiven gedetecteerd worden. Nog een voordeel van diepte camera's is dat ze relatief goedkoop zijn, een Kinect kost ongeveer 150 euro. Een nadeel is occlusie, als twee personen achter elkaar staan kan de achterste persoon vaak niet gedetecteerd worden. Dit kan opgelost worden door meerdere camera's te gebruiken, maar extra camera's toevoegen is niet zo gemakkelijk als bij een motion capture systeem (zie hoofdstuk 4.2).

Met een Kinect die top-down geplaatst is kan ook de positie en oriëntatie van personen realtime gedetecteerd worden. Doordat de Kinect aan het plafond hangt heeft hij geen last van occlusie, maar gestures zijn veel moeilijker te detecteren doordat de Kinect top-down geen skelet kan detecteren. Het gebied waarin de Kinect personen kan tracken is meestal ook kleiner doordat de Kinect aan het plafond hangt [17].

Optical motion capture systemen kunnen de positie van personen tracken. Ze kunnen ook de oriëntatie van personen tracken, mits deze een rigid body (een uniek patroon van minstens drie markers) draagt. Ze kunnen ook gestures detecteren als markers op de juiste positie geplaatst worden. Om bijvoorbeeld een wuifgebaar te kunnen detecteren moeten er markers op de hand(en) van de gebruikers geplaatst worden. De markers zijn ook het nadeel van de optical motion capture systemen, elke persoon moet een uniek rigid body dragen (dus minstens drie markers). Active markers hebben nog eens als bijkomend nadeel dat de personen ook een batterij moeten dragen. Als er genoeg camera's geplaatst zijn heeft een optical motion capture systeem weinig last van occlusie. Een nadeel van een optical motion capture systeem is de prijs. Eén enkele infrarood camera voor een OptiTrack systeem kost tussen de 600 en 6000 dollar.

Het tracken op basis van kleuren, zoals gedaan door McKenna et al. is niet zo geschikt voor onze toepassing omdat er geen top down 2D positie bepaald kan worden van de gebruikers. Het voordeel van deze methode is dat hij zowel binnen als buiten werkt en dat er enkel een video camera nodig is.

Hoofdstuk 5

Bestaande groepdetectie methodes

Met de tracking informatie verzameld door tracking systemen kunnen groepen gedetecteerd worden. Hieronder worden drie bestaande groepdetectie methodes besproken: GroupTogether, SpiderEyes en een rule-based approach. GroupTogether gebruikt de afstand en de hoek tussen personen om groepen te detecteren. SpiderEyes detecteert groepen op basis van de afstand tussen personen en op basis van naar waar personen hun aandacht richten. Een rule-based approach detecteert groepen op basis van een set van regels waaraan elke groep moet voldoen. Op het einde van dit hoofdstuk worden ook de voor- en nadelen van elke methode besproken.

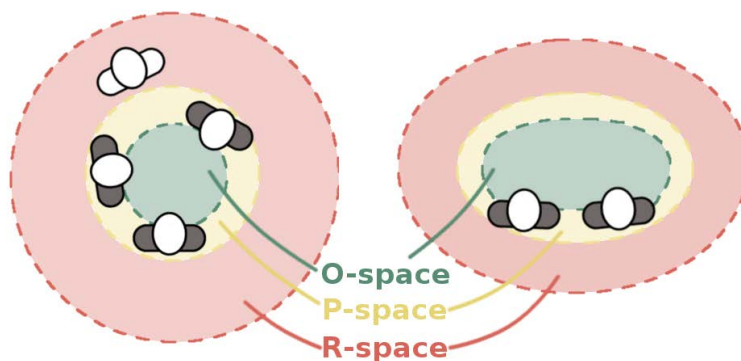
5.1 GroupTogether

GroupTogether is een systeem dat interactietechnieken ondersteunt om ter plaatse te kunnen samenwerken over devices heen [16]. Het systeem maakt gebruik van twee sociologische concepten. Ten eerste, F-formations. F-formations bekijken de afstand en relatieve lichaamsoriëntatie tussen verschillende gebruikers, deze geven aan wanneer en hoe mensen zich positioneren als een groep. Ten tweede, micro-mobility. Micro-mobility beschrijft hoe personen hun devices naar elkaar richten om informatie te kunnen delen tijdens het samenwerken.

Een F-formation bestaat uit twee of meer personen en is een theoretische lens waardoor groepsinteracties bestudeerd kunnen worden. F-formations houden niet enkel rekening met de afstand tussen personen maar houden ook rekening met hoe personen zich opstellen ten opzichte van elkaar tijdens interacties. Uit die opstelling ten opzichte van elkaar kunnen drie regio's afgeleid worden [21]: De O-space, de P-space en de R-space (zie Figuur 5.1). De O-space is de ruimte waar de focus van de groep ligt en de belangrijkste

ste activiteit van de groep plaatsvindt. De P-space is de ruimte rondom de O-space waar de groepsleden zich bevinden. De ruimte daarbuiten is de R-space, deze schermt de groep van de wereld buiten de groep af. Personen die zich in de R-space bevinden maken geen deel uit van de groep, maar de groep houdt deze personen wel in de gaten om te kijken of ze bij de groep willen komen. Fixed en semi-fixed features kunnen impliciet ook een deel uitmaken van een F-formation, bijvoorbeeld wanneer een groep personen voor een whiteboard staan, kan dat whiteboard als een deel van de groep beschouwd worden.

Het GroupTogether systeem detecteert F-formations wanneer twee personen a) niet achter elkaar staan, b) de hoek tussen hun oriëntatie vectoren niet groter is dan 180 graden (ze dus niet van elkaar weggijken) en c) de afstand tussen de twee personen klein genoeg is om makkelijk te kunnen communiceren en hun O-spaces overlappen. Het systeem gaat alle paren personen af en berekent de afstand en hoek tussen hun. Vervolgens kent het een F-formation type toe (zoals zij aan zij, L-vorm, oog in oog of geen) gebaseerd op tolerantie thresholds.



Figuur 5.1: Twee F-formations. Een F-formation bestaat uit twee of meer personen. Uit de opstelling van de personen ten opzichte van elkaar kunnen drie regio's afgeleid worden: De O-space, de P-space en de R-space.

Dat personen in een groep staan (een F-formation vormen) wil nog niet zeggen dat ze informatie met elkaar willen delen. Om te bepalen of personen informatie willen delen tussen devices wordt micro-mobility gebruikt. Micro-mobility van voorwerpen is het oriënteren en herpositioneren van objecten zodat ze zichtbaar, gedeeltelijk zichtbaar of verborgen zijn voor andere personen. Zo kan het dichterbij plaatsen van devices of het kantelen van devices naar elkaar toe een manier zijn om informatie met elkaar te delen. GroupTogether gebruikt micro-mobility om informatie te delen via devices.

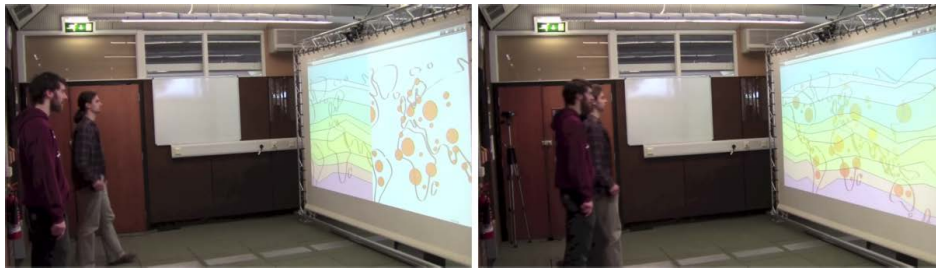
Ze ontwierpen verschillende interactietechnieken om informatie te delen. Door devices naast elkaar te houden, naar elkaar toe te richten of door devices te kantelen kan op verschillende manieren informatie gedeeld worden zoals informatie doorgeven, kopiëren, verspreiden over twee device, ... De positie van het device binnen de F-formation speelt ook een rol, zo kan een device enkel informatie doorgeven naar een ander device als het zich in de O-space bevindt. Informatie kan niet alleen met devices gedeeld worden, maar ook met fixed of semi-fixed features zoals een whiteboard op dezelfde manier als personen informatie delen tussen devices.

5.2 SpiderEyes

Dostal et al. introduceerden een concept genaamd collaborative proxemics [22]. Collaborative proxemics laat groepen personen toe om samen attention- en proximity-aware applicaties te gebruiken. Dit zijn applicaties die zich bewust zijn van de relaties tussen personen, objecten en devices (proximity-aware) en die zich bewust zijn van naar waar de gebruikers hun aandacht richten, bijvoorbeeld naar andere personen of naar displays (attention-aware).

Om designers te helpen bij ontwikkelen van attention- en proximity-aware interfaces voor grote displays, ontwikkelden ze een toolkit genaamd SpiderEyes. Deze toolkit kan tot en met vier gebruikers realtime tracken en detecteert ook de positie van de ogen van gebruikers. Voor de tracking wordt een dieptesensor gebruikt (de Kinect) in combinatie met computer vision technieken (toegepast op beelden gefilmd met een gewone RGB camera).

Het systeem detecteert ook groepen en doet dat op basis van de afstand tussen gebruikers. Applicatie ontwikkelaars kunnen een parameter instellen die de maximum afstand aangeeft tussen twee gebruikers om als een groep beschouwd te kunnen worden. Ze ontwikkelden ook een toepassing die gebruik maakt van deze groepdetectie. Wanneer twee gebruikers bij elkaar gaan staan en zo een groep vormen, worden de twee viewpoints van deze gebruikers gemerged.



Figuur 5.2: SpiderEyes is een toolkit voor het ontwikkelen van attention- en proximity-aware interfaces voor grote displays. Deze afbeelding toont hoe de viewpoints van twee personen gemerged worden wanneer ze bij elkaar gaan staan en een groep vormen [22].

Doordat de toolkit attention-aware is kan het een onderscheid maken tussen de personen die voorbijlopen en de personen die actief bezig zijn met het systeem (hun aandacht richten op het scherm). Op deze manier voorkomt men dat wanneer een persoon een groep voorbijloopt deze tijdelijk ook tot een deel van de groep beschouwd wordt.

5.3 Rule-based approach

Winters creëerde een rule-based groepdetectie library [15]. Deze library detecteert groepen op basis van of personen al dan niet aan set van regels voldoen.

5.3.1 De regels

De set van regels kan aangepast worden om zo optimale groepdetectie te kunnen bekomen voor specifieke situaties. De volgende regels kunnen allemaal gebruikt worden:

- **Interpersoonlijke afstand:** De maximale afstand tussen twee personen.
- **Maximaal verschil in afstand tot display**
- **Maximaal verschil in snelheid:** Met deze regel kan gecontroleerd worden of twee personen op ongeveer hetzelfde tempo aan het wandelen zijn.
- **Maximaal verschil in oorspronkelijke richting:** Deze regel vergelijkt de richting die twee personen uitgingen toen ze voor het eerst gedetecteerd werden door het systeem. De oorspronkelijke richting kan soms nuttiger zijn dan de huidige. Bijvoorbeeld wanneer twee personen samen een kamer komen binnengelopen en één van de twee

draait zich om om met de andere te praten. Hierdoor hebben ze een tegenovergestelde (huidige) richting, maar wel dezelfde oorspronkelijke richting.

- **Maximaal verschil in huidige richting**
- **Maximaal verschil in aankomst:** Met deze regel kan een onderscheid gemaakt worden tussen personen of groepen die op een ander tijdstip aangekomen zijn.
- **Tijd dat een overeenkomst gebeurt:** Deze regel controleert of de volledige set van regels voor een minimum tijdstip voldaan is. Deze regel dient om te voorkomen dat het systeem te snel een verkeerde conclusie trekt.

5.3.2 Regels combineren tot een set

Een set van regels kan gemaakt worden door verschillende regels te combineren en voor elk van die regels een threshold te kiezen. Het is mogelijk om de regels in runtime aan te passen. Regels kunnen gecombineerd worden met behulp van boolean operaties (AND, OR, NOT, ...). Bijvoorbeeld de interpersoonlijk afstand moet onder een bepaalde threshold liggen en (AND) het verschil in snelheid mag niet meer dan 0,15m/s zijn.

Door gewichten aan elke regel te geven kan men aanduiden hoe belangrijk elke individuele regel is. Voor elke aparte regel kan ook een percentage opgevraagd worden dat aangeeft hoe goed twee personen aan een regel voldoen. Voor dit percentage kan ook een threshold gezet worden.

5.3.3 Scenario's

Om de voor- en nadelen van elke regel te identificeren werden negen scenario's opgenomen. Doordat de scenario's opgenomen werden kon achteraf verschillende regels getest worden op exact dezelfde scenario's. Uit het testen van verschillende regels op deze scenario's blijkt dat, op één na, alle regels wel ergens nuttig kunnen zijn. Maar dat het noodzakelijk is om enkele regels te combineren om zo de error rate omlaag te trekken.

De *huidige richting* regel was de enige regel die nergens nuttig was doordat hij constant veranderde van waarde. De *interpersoonlijke afstand* regel is over het algemeen een zeer nuttige regel. De *tijd dat een overeenkomst gebeurt* regel is goed in het voorkomen dat verkeerde groepen gevormd worden wanneer personen elkaar voorbij lopen. De *tijd van aankomst* regel werkt goed wanneer in het scenario het meer waarschijnlijk is dat personen in een groep samen aankomen, maar dit is zeker niet altijd het geval. Een combinatie van regels die goed werkt is de *interpersoonlijke afstand* gecombineerd met de *tijd dat een overeenkomst gebeurt* regel.

Het nadeel van de rule-based approach is dat voor elk specifiek scenario wel een set van regels kan samengesteld worden die correct de groepen detecteert, het is dus makkelijk om nadat een scenario zich heeft afgespeeld een set regels voor dat scenario samen te stellen. Maar het is moeilijk om op voorhand te voorspellen wat er zich gaat voordoen, daarom zou het beter zijn een groepdetectie methode te hebben die voor vele scenario's werkt.

5.3.4 Demo: Whac-a-Mole



Figuur 5.3: Whac-a-Mole variant met groepdetectie, elke groep heeft een eigen kleur. Cirkel stelt een persoon voor, vierkant stelt het zwaartepunt van een groep voor. De spelers gebruiken het zwaartepunt als virtuele hamer om een mol terug de grond in te slaan [15].

Winters ontwikkelde een demo applicatie die gebruik maakt van de rule-based groepdetectie library. De demo is een spel gebaseerd op het arcade game Whac-a-Mole, waarbij de speler een mol die uit de grond springt terug de grond in moet slaan. Bij deze variant van het spel moeten twee of meer groepen voor het eerst de mol terug de grond in slaan. Het zwaartepunt van een groep wordt als virtuele hamer gebruikt (zie Figuur 5.3).

Verschillende sets van regels werden getest met behulp van het spel. Wanneer een regelset geen gebruik maakte van interpersoonlijke afstand vroeg een gebruiker: “Waarom vormen we geen Groep? We staan dicht bij elkaar.”. Dit geeft aan hoe belangrijk de interpersoonlijke afstand is, personen verwachten dat ze een groep vormen wanneer ze dicht bij elkaar staan. Wanneer de threshold van interpersoonlijke afstand op 100 cm stond voelden sommige personen zich soms ongemakkelijk zo dicht op elkaar te staan, bij 121 cm was er geen probleem. Verschil in snelheid bleek geen goede regel voor het spel omdat personen vaak stil staan of heel traag bewegen.

5.4 Vergelijking groepdetectie methodes

Hieronder worden de verschillende voor- en nadelen van de bovenstaande groepdetectie methodes besproken en wordt er bekeken welke technieken toegepast kunnen worden op the Gathering en wat the Gathering zou kunnen verbeteren.

Marshall et al. toonden aan hoe features sociale relaties kunnen aanmoedigen of ontmoedigen (zie hoofdstuk 3.2). Features zijn dus niet onbelangrijk bij het detecteren van groepen en the Gathering zou hier dus rekening mee moeten houden. GroupTogether houdt rekening met features in de omgeving, zo wordt wanneer een groep personen voor het whiteboard staan deze personen ook als een deel van de groep beschouwd. SpiderEyes is specifiek voor wall-sized displays bedoeld en houdt geen rekening met andere features. De rule-based approach houdt helemaal geen rekening met features.

Tijdens het testen van de demo applicatie van de rule-based approach gingen de groepsformaties vaak verloren door occlusie. Voor deze test werd maar één Kinect gebruikt. GroupTogether maakt gebruik van twee overhead Kinects om personen te tracken en heeft hierdoor minder last van occlusie. Hierdoor kunnen ze wel het skelet niet gebruiken om personen te tracken en kunnen gestures minder goed gedetecteerd worden. SpiderEyes heeft net zoals de rule-based approach last van occlusie doordat er slechts één Kinect en één RGB camera gebruikt wordt, die beide naast elkaar opgesteld staan.

Om occlusie te vermijden of te verminderen kan the Gathering best gebruik maken van een set-up met meerdere camera's zoals een motion capture systeem of door gewoon meerdere dieptesensoren te gebruiken.

Het nadeel van SpiderEyes is dat het specifiek voor wall-sized displays bedoeld is, het detecteert alleen groepen tussen mensen die voor de display staan. GroupTogether en de rule-based approach zijn niet gebonden aan één specifiek scenario. Hoewel de rule-based approach in veel verschillende scenario's kan werken, biedt het geen algemene oplossing aan die voor vele scenario's werkt. Het is enkel makkelijk om nadat een scenario zich heeft afgespeeld een set van regels op te stellen die voor dat scenario goed groepen zou detecteren.

Een voordeel van SpiderEyes is dat het attention-aware is, het houdt rekening met naar waar personen aan het kijken zijn. Op deze manier detecteert SpiderEyes dat een persoon die een groep passeert geen deel uitmaakt van de groep ook al komt de persoon dicht bij de groep in de buurt. SpiderEyes houdt wel alleen rekening met of de personen hun aandacht op de display richten of niet. Zo worden twee personen die bij elkaar staan en naar elkaar kijken niet als een groep beschouwd omdat ze niet daar de display aan het

kijken zijn. The Gathering zou dit kunnen verbeteren door te detecteren naar waar een persoon kijkt en niet alleen te detecteren of het naar een bepaalde display kijkt of niet.

Winters testte met de rule-based approach verschillende scenario's met verschillende sets van regels. De *interpersoonlijke afstand* gecombineerd met de *tijd dat een overeenkomst gebeurt* bleek vrij goed te werken. De interpersoonlijke afstand wordt ook door de andere twee groepdetectie methodes gebruikt en in Hall zijn proxemic zones (zie hoofdstuk 3.1). De interpersoonlijke afstand is dus een erg belangrijke indicatie dat personen samen een groep vormen en moet zeker gebruikt worden in the Gathering.

Hoofdstuk 6

Toepassingen

In dit hoofdstuk geven we een antwoord op het tweede deel van de onderzoeksvraag: *Wat voor toepassingen kunnen gebruik maken van zo'n groepdetectie methode?* We bespreken drie mogelijke toepassingen voor the Gathering.

6.1 SpiderEyes

De volgende toepassing is gebaseerd op een voorbeeld scenario van de SpiderEyes toolkit besproken in hoofdstuk 5.2.

In een kamer hangt een grote display van 2,5 meter hoog en 5 meter breed aan de muur. Er is niemand in de ruimte en de display staat uit. Een team van vijf analisten hebben informatie verzameld over de rijkdom en gezondheid van alle landen. De analisten hebben hun data ingegeven in een computer die in de kamer ernaast staat. Een analist genaamd Stijn gaat naar de kamer met de grote display. Zodra Stijn de kamer binnenwandelt springt de display aan en wordt er een wereldkaart weergegeven. Stijn gaat twee meter voor de display staan en er verschijnt automatisch informatie op de wereldkaart, een visualisatie van de data die ze zonet in de computer hebben ingegeven. Omdat Stijn twee meter van de display staat wordt een overzicht van de data gegeven, sommige data van enkele kleine aanliggende landen wordt samengenomen omdat deze te klein zijn om van op twee meter afstand duidelijk te kunnen zien. Zodra Stijn een algemene indruk heeft gekregen van de data stapt hij naar voren richting Europa er worden nu automatisch meer details getoond.

Stijn neemt enkele notities en tegelijkertijd komt een tweede persoon, Bob, binnengewandeld. Bob gaat rechts voor de display staan op zo'n twee meter afstand van de display en twee meter rechts van Stijn. The Gathering heeft gedetecteerd dat er een tweede persoon de ruimte is binnengewandeld en dat deze geen groep vormt met Stijn. De display toont nu twee verschil-

lende kaarten, een kaart voor Stijn en een kaart voor Bob. Stijn doet een stap naar links om Bob wat extra ruimte te geven. Bob zijn kaart is minder gedetailleerd dan de kaart van Stijn omdat Bob verder van de display verwijderd staat. De kaart past niet meer helemaal op de display omdat er nu twee kaarten afgebeeld staan. Bob zwaait met zijn hand naar links om de kaart naar rechts te scrollen. Stijn is nog altijd notities aan het nemen. Bob wandelt naar Stijn toe om te vragen wat zijn bevindingen zijn. The Gathering detecteert dat Bob en Stijn samen een groep vormen en de display toont weer één kaart in plaats van twee.

Terwijl Stijn en Bob aan het discussieren zijn komt een derde persoon, Sarah binnengewandeld. Bob kijkt even opzij om te kijken wie binnengekomen is. En ookal kijkt Bob even weg van de display en weg van Stijn, toch blijft the Gathering Bob en Stijn als één groep beschouwen. Sarah gaat voor de display staan, rechts van Stijn en Bob, er verschijnen weer twee kaarten op de display.

Een beetje later komen John en Sanne de kamer binnengewandeld, ze gaan achter de andere drie personen staan op vier meter van het scherm. Ze zijn met elkaar aan het praten en kijken niet naar de display. The Gathering detecteert nu drie groepen: Stijn & Bob, Sarah en John & Sanne. The Gathering detecteert dat Sanne en John geen aandacht hebben voor de display en er wordt geen derde kaart getoond.

Een beetje later roept Bob iedereen naar zich toe om zijn bevindingen te delen. Iedereen wandelt naar Stijn en Bob toe en er wordt weer één kaart getoond. De analisten discussiëren met elkaar en wijzen landen met interessante resultaten aan. Op een gegeven moment doet Sanne onbewust een paar stappen opzij. The Gathering detecteert nu twee groepen maar doordat Sanne nog vrij dicht bij de rest van de groep staat is het accuraatheid percentage vrij laag, 54%. Omdat het percentage zo laag is worden er niet meteen twee displays op de kaart getoond, maar wordt de kaart langzaam aan kleiner en verschijnt er een tweede kaart die oorspronkelijk onzichtbaar is en langzaam aan groter en minder transparant wordt. Sanne merkt meteen op dat ze niet meer als een deel van de groep beschouwd wordt en zet meteen een stap richting de groep. De tweede kaart verdwijnt.

Na al dat discussiëren krijgen de analisten dorst en verlaten ze de kamer om een koffie te drinken, de display springt vanzelf weer uit.

In tegenstelling tot de oorspronkelijke toepassing, worden voor personen die geen aandacht hebben voor de display, ook groepen gedetecteerd. The Gathering kan dit omdat the Gathering rekening houdt met de oriëntatie van gebruikers en niet enkel controleert of gebruikers naar een display kijken. Een andere meerwaarde die the Gathering in deze toepassing biedt is dat het een accuraatheid percentage geeft. Daardoor kan er in de toepassing aan gebruikers aangegeven worden dat het systeem niet zeker is of de gebruikers nu bij elkaar horen of niet.

6.2 Restaurant

In een restaurant in Hasselt hangt aan de voorgevel een reeks van grote displays van zo'n twee meter hoog. De displays zijn een soort van virtuele spiegels. Wanneer personen er voorbij wandelen wordt hun spiegelbeeld op de display getoond, aan het spiegelbeeld wordt een gedachtenwolkje (uit stripverhalen, zie Figuur 6.1) toegevoegd met daarin een maaltijd die het restaurant aanbiedt en de prijs van die maaltijd. Volgens Müller et al. is het spiegelbeeld van personen gebruiken een goede manier om de aandacht te trekken [5]. Wanneer een persoon even blijft staan en naar de display kijkt



Figuur 6.1: Een wolkje met daarin een maaltijd die verkocht wordt in het restaurant. Deze wolk wordt toegevoegd aan het spiegelbeeld van personen die voorbij wandelen om zo hun aandacht te lokken.

verandert het spiegelbeeld in een menu kaart met daarop de verschillende maaltijd en prijzen die er verkocht worden. De persoon kan op een maaltijd drukken om een afbeelding van die maaltijd op te vragen. Wanneer de persoon interesse verliest in de menu kaart en wegstijgt, veranderen enkele afbeeldingen op de menukaart en worden er andere maaltijden getoond. Dit is een techniek die ook door Wang et al. wordt gebruikt in hun Peddler framework [23]. Wanneer een gebruiker interesse verliest in de producten die worden aangeboden verschijnen er nieuwe aanbiedingen en proberen ze zo terug de interesse van de gebruiker te wekken.

Een gezin van vijf personen zijn net gaan bowlen en zijn nu onderweg naar hun auto. Onderweg passeren ze het restaurant, de jongste van het gezin merkt zijn spiegelbeeld op en ziet dat er een wolkje boven elk hun hoofd hangt met daarop vier verschillende maaltijden. Hij zegt tegen zijn ouders dat ze naar de display moeten kijken, het gezin blijft staan en er verschijnt een grote menu kaart op de display. Achter de menu kaart kunnen ze nog altijd hun spiegelbeelden zien. De ouders besluiten dat het wel een goed idee is om eerst een hapje te eten voor ze naar huis gaan. De pa gluurt door de deur naar binnen en ziet dat het vrij druk is. Ondertussen is er ook een koppel voor de display gestopt en het koppel krijgt ook een menu kaart te zien.

De vader zegt tegen zijn vrouw dat hij denkt dat ze geen plaats meer hebben voor vijf personen. Maar zijn vrouw wijst naar de menu kaart,

bovenaan de kaart staat geschreven dat er nog een tafel voor vijf personen vrij is. The Gathering detecteert uit hoeveel personen de groep bestaat en het systeem van het restaurant controleert automatisch of er nog een tafel voor zoveel personen vrij is en toont een boodschap boven de menukaart met of er nog een tafel vrij is of niet.

Het gezin wandelt het restaurant naar binnen en een ober krijgt een melding dat er een groep van vijf personen het restaurant is binnengewandeld. De ober gaat naar de ingang van het restaurant en biedt het gezin een tafel aan.

6.3 Foto muur

Op Pukkelpop, een muziekfestival, staan een reeks displays die samen één grote display van acht meter breed vormen. Boven de displays hangen een reeks videocamera's. Peter loopt naar de muur toe en ziet dat de muur volstaat met thumbnails van foto's van andere festivalgangers. Peter merkt dat hij de foto's op de muur kan verslepen en vergroten en speelt even met enkele foto's. Peter zet een stap opzij om wat andere foto's te bekijken en merkt dat een camera icoontje hem volgt. Hij drukt op het camera icoontje en er verschijnt een camera beeld op het scherm voor Peter, hij wordt met één van de camera's bovenop de display gefilmd. Een teller verschijnt en telt af van vijf naar nul. Peter begrijpt dat er een foto van hem getrokken gaat worden en poseert voor de camera. Op nul trekt de camera een foto en wordt deze foto getoond aan Peter, hij bekijkt de foto even en gooit hem dan opzij.

Een beetje verder staat een groepje van drie personen naar foto's te kijken. Voor de groep zweeft maar één camera icoontje, geen drie, dit is omdat the Gathering heeft gedetecteerd dat ze één groep vormen. De middelste persoon uit de groep, Anja, drukt op het camera icoontje en doordat het systeem weet dat Anja deel uitmaakt van een groep wordt het beeld van de camera zo gecropt dat ze alle drie op de foto staan en niet enkel Anja.

Sofie, de meest rechtse persoon uit de groep, wil ook nog een foto van zichzelf maken en doet een paar stappen opzij, er verschijnt nu een tweede camera icoontje, een knop voor Sofie en een knop voor de rest van de groep. Het systeem weet dat Sofie alleen is en het beeld wordt rondom haar gecropt.

Hoofdstuk 7

The Gathering

In de volgende hoofdstukken wordt the Gathering besproken. The Gathering is de groepdetectie methode die we ontwikkeld hebben en houdt rekening met de afstand tussen personen, de oriëntatie van personen en de features in de omgeving. We leggen uit hoe the Gathering precies werkt, bespreken enkele mogelijke toepassingen en voeren een gebruikerstest uit.

In dit hoofdstuk worden eerst de scenario's uit hoofdstuk 2 geanalyseerd om daaruit af te leiden wat the Gathering zoal zou moeten kunnen. Nadat we weten wat the Gathering zoal moet kunnen wordt er gekeken welke set-up er nodig is voor the Gathering om de nodige informatie te verzamelen. De set-up bestaat uit een motion capture systeem genaamd Optitrack, een Optitrack server en een pc waarop the Gathering wordt uitgevoerd.

7.1 Analyse scenario's

Door de scenario's uit hoofdstuk 2 te analyseren kunnen we zien wat the Gathering zoal zou moeten kunnen. We kunnen hieruit ook zien wat indicaties zijn dat personen een groep (of juist geen groep) vormen.

Scenario 1

In het eerste scenario stapte één persoon, Jan, uit een groep naar voren om op een display een dagschema op te vragen. De rest van de groep kijkt toe en discussieert over het dagschema.

Uit dit scenario blijkt dat een groep personen niet altijd mooi bij elkaar staan. Jan staat in dit scenario een beetje buiten de groep om met het scherm te kunnen interageren, maar maakt evenzeer deel uit van de groep. Er zijn verschillende indicaties dat Jan deel uit maakt van een groep die gedetecteerd zouden kunnen worden:

- Jan en de rest van de groep kijken naar een zelfde gebied op het scherm.
- Jan en de rest van de groep zijn samen aangekomen aan het scherm.

- De afstand tussen Jan en de rest van de groep is groter dan de afstand tussen de andere personen van de groep onderling, maar deze afstand is nog steeds klein genoeg om duidelijk met elkaar te kunnen communiceren.
- Waarschijnlijk keert Jan zich ook af en toe naar de rest van de groep om te overleggen of keert hij zich verticaal met het scherm zodat hij zowel het scherm als de rest van de groep kan zien.

Scenario 2

Net zoals in het vorige scenario staat er een persoon (de gids) een beetje buiten de groep, terwijl hij toch bij de groep hoort. Een beetje later komt er ook nog een tweede en een derde groep aan bij de display.

In dit scenario zijn er dezelfde indicaties dat de gids toch bij de groep hoort als in het vorige scenario. Er zijn ook verschillende indicaties die gedetecteerd kunnen worden dat de twee andere personen (groep 2) niet bij de eerste groep horen: Ze staan verder uit elkaar, groep 2 is later bij het scherm toegekomen en ze staan niet naar elkaar toe gericht.

Scenario 3

In het derde scenario naderen twee groepjes van elk twee personen ongeveer gelijktijdig een display. Ze gaan allemaal voor de display staan en vormen zo een rij van vier personen.

Hoewel in dit scenario vier personen naast elkaar staan vormen ze geen groep, dit is vrij moeilijk te detecteren omdat de signalen die aangeven dat ze niet bij elkaar horen vrij subtiel zijn. Indicaties hiervan zijn dat ze niet helemaal gelijktijdig aankomen en dat de focus van de twee groepen verschilt, de focus kan op verschillende gebieden van het scherm zijn. Bovendien staan ze waarschijnlijk niet perfect op een rij maar staan de personen binnen een groep iets dichter bij elkaar of iets meer naar elkaar toegericht.

Scenario 4

In dit scenario staat een groep voor een display, Eric verlaat de groep om alleen te interageren met de display. Een beetje later stapt Lize uit de groep naar voren om een dagschema op te vragen, Lize en de rest van de groep bekijken samen het schema.

In dit scenario moet the Gathering in staat zijn te detecteren dat wanneer Eric naar voren loopt hij de groep verlaat, maar wanneer Lize naar voren loopt ze de groep niet verlaat. Het enige verschil tussen deze twee gevallen is dat in het eerste geval Eric en de rest van de groep een andere focus hebben. En dat in het tweede geval Lize en de rest van de groep een gemeenschappelijke focus hebben.

Scenario's 5 en 6

In scenario 5 interageert John met een display, Sarah komt iets later toe en gaat bij John staan. In scenario 6 komen twee personen samen aan bij een display maar gaan een beetje uit elkaar staan om elk apart met de display te interageren.

Indien het algoritme geen rekening houdt met de geschiedenis tussen twee personen (of ze al dan wel niet samen zijn aangekomen), zijn deze scenario's vrij makkelijk te detecteren. Doordat de personen al dan niet bij elkaar gaan staan om naar het scherm te kijken wordt duidelijk of ze samen een groep vormen of niet.

7.1.1 Conclusie analyse

Hoewel het mogelijk is om voor elk scenario een algoritme te schrijven dat de groepen correct detecteert, is het erg moeilijk om een algoritme te schrijven dat in alle scenario's correct groepen detecteert. Sommige indicaties die erop wijzen dat personen een groep vormen werken in het ene scenario wel, maar in het andere dan weer niet.

Bijvoorbeeld in scenario 1 komen de personen gelijktijdig aangelopen en onder andere dit wijst erop dat ze één groep vormen. Maar in scenario 6 komen de personen ook samen aangelopen en vormen ze toch geen groep.

In scenario 1 en 2 (en bij Lize in scenario 4) wanneer een persoon uit een groep naar voren stapt naar het scherm is dit een teken dat de persoon het scherm gaat bedienen voor de hele groep. Terwijl bij Eric in scenario 4 dit niet het geval is en Eric de groep wel verlaat.

Verschillende applicaties in deze scenario's maken ook gebruik van de positie van personen of groepen. In scenario 1 wordt bijvoorbeeld de content rechts op het scherm weergegeven omdat de groep rechts staat. Het is dus ook belangrijk dat the Gathering behalve detecteren in welke groep een persoon hoort, ook doorgeeft waar deze groepen zich bevinden. Doorgeven waar een persoon of groep precies aan het kijken is op een scherm kan ook handig zijn om content op een gepaste locatie te plaatsen.

7.2 Input verzamelen voor the Gathering

Uit de analyse van de scenario's en uit de analyse van bestaande groepdetectie methodes (zie hoofdstuk 5.4) blijkt dat we de afstand tussen personen moet kunnen meten en dus hun positie moeten weten. Er moet ook gedetecteerd worden naar waar de personen zich richten, dit kan naar andere personen zijn maar ook naar een display toe. The Gathering zou ook moeten weten waar een persoon op een display aan het kijken is zodat the Gathering

kan detecteren of twee personen naar dezelfde regio op een display aan het kijken zijn.

We hebben ervoor gekozen een motion capture systeem (Optitrack) te gebruiken. Door markers op de hoofden van gebruikers te plaatsen kan het systeem de positie en oriëntatie (van het hoofd) van personen detecteren. We hebben voor Optitrack gekozen omdat we dit systeem beschikbaar hebben in ons lab en deze alle informatie kan verzamelen die we nodig hebben. Zoals in hoofdstuk 4.2 besproken is het een erg duur systeem, maar het kan makkelijk door een ander (goedkoper) tracking systeem vervangen worden. De enige vereiste is dat het de (2D) positie en oriëntatie (van het hoofd) van personen kan detecteren. Hiervoor kunnen ook diepte camera's gebruikt worden.



Figuur 7.1: Optitrack camera's

Het Optitrack systeem maakt gebruik van passieve markers, deze markers worden op het hoofd van de personen gedragen. Het systeem bestaat uit verschillende infrarood camera's (zie Figuur 7.1), rondom deze camera's zitten LEDjes die infrarood licht uitzenden, dit licht wordt gereflecteerd door de markers op het hoofd van de te tracken personen. Het gereflecteerd infrarood licht wordt dan weer opgevangen door de camera's. Vanuit de 2D afbeeldingen van de camera's worden de 2D posities berekend [24]. En doordat het systeem de positie en viewpoint van elke camera kent kan het vanuit de 2D posities de 3D posities van de markers berekenen via triangulatie [25]. Om de oriëntatie van een object te berekenen (in dit geval het hoofd van een persoon) worden er meerdere markers op een object geplaatst (minstens 3). Deze meerdere markers per object zijn ook nodig om de objecten (personen) van elkaar te kunnen onderscheiden. Zo'n set van markers wordt een rigid body genoemd.

Alle data van de camera's wordt doorgestuurd naar een Optitrack server. Op deze server staat software genaamd *Motive* waarmee het systeem gemakkelijk gekalibreerd en bestuurd kan worden (rigid bodies toevoegen, ...). *Motive* broadcast de tracking data over een netwerk. Een andere computer

is ook met dat netwerk verbonden. Op die computer staat the Gathering die de tracking data leest en gebruikt om groepen te detecteren.

Hoofdstuk 8

The Gathering: Algoritme

Hieronder wordt het algoritme van the Gathering toegelicht. Het algoritme gebruikt de afstand tussen personen (interpersoonlijke afstand), de features in de omgeving en de oriëntatie van personen (naar waar personen aan het kijken zijn) om te berekenen welke personen samen een groep vormen. In dit hoofdstuk wordt uitgelegd hoe dit precies berekend wordt. Omdat het algoritme niet altijd correct groepen kan detecteren wordt er een percentage berekend dat aangeeft hoe waarschijnlijk het is dat het algoritme de groepen correct gedetecteerd heeft.

De verschillende stappen van het algoritme worden verduidelijkt met enkele afbeeldingen. De afbeeldingen zijn (bewerkte) screenshots van de visualisatie en debug tool die nog besproken wordt in hoofdstuk 9.2. Elke persoon wordt voorgesteld door een cirkel met een rode streep in, de rode streep geeft aan in welke richting de persoon aan het kijken is. Personen die samen een groep vormen hebben allemaal dezelfde kleur. Features worden voorgesteld door een zwarte balk.

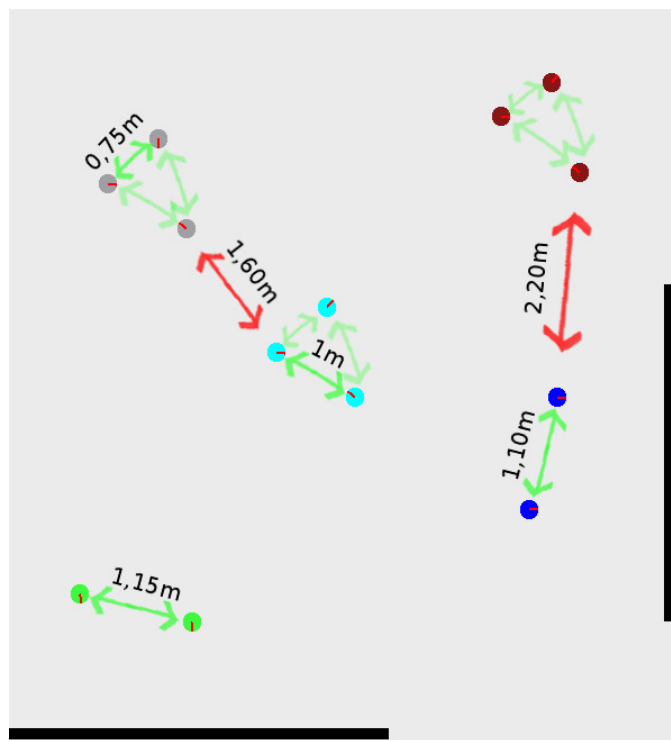
8.1 Interpersonal distance

Eerst worden er groepen gevormd op basis van de afstand tussen personen, personen die dicht genoeg bij elkaar staan worden samengenomen in een groep (zie Figuur 8.1). Deze interpersoonlijke afstand wordt gebruikt omdat een groep personen typisch bij elkaar staan. En uit een analyse van verschillende scenario's door Winters bleek ook dat de interpersoonlijke afstand een goede regel is bij het detecteren van groepen (zie hoofdstuk 5.3.3).

Om te bepalen welke personen dicht genoeg bij elkaar staan wordt gebruik gemaakt van Hall's social spaces[7]: *intimate space 0-0.45m*, *personal space 0.45-1.2m*, *social space 1.2-3.6m* en *public space 3.6-7.6m*. Twee personen staan dicht genoeg bij elkaar als ze in elkaars intimate of personal space staan. Als de afstand dus maximaal 1.2m bedraagt. In de

Whac-a-mole applicatie die gebruik maakt van een rule-based groepdetectie methode wordt deze afstand ook gebruikt en blijkt dit een comfortabele afstand te zijn, bij een kleinere afstand van 100cm voelen sommige mensen zich ongemakkelijk om zo dicht bij elkaar te staan (zie hoofdstuk 5.3.4).

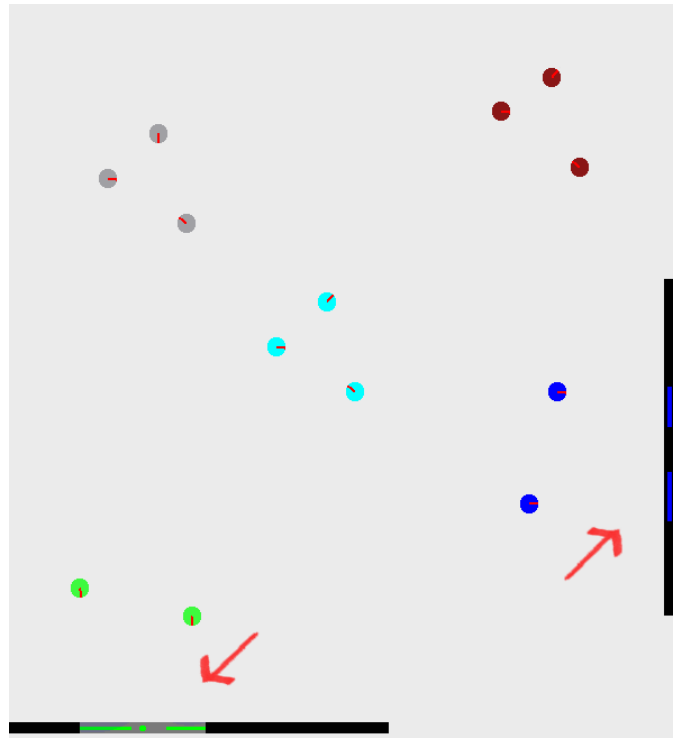
Het is mogelijk dat de afstand tussen twee personen binnen een groep groter is dan 1.2m. Dit kan bijvoorbeeld als persoon **A** 80cm van persoon **B** staat en persoon **B** 60cm van persoon **C** staat. Op deze manier is het mogelijk dat bijvoorbeeld drie personen die zij en zij staan een groep vormen. Bovendien zou anders de keuze gemaakt moeten worden of AB of BC een groep vormen.



Figuur 8.1: Groepen gevormd op basis van interpersonal distance. Twee personen staan dicht genoeg bij elkaar als ze in elkaars intimate of personal space staan (maximaal 1,2m van elkaar staan).

8.2 Displays

The Gathering beschouwt displays soms als een deel van een groep wanneer er interactie tussen een display en de andere groepsleden is. Marshall et al. toonden al aan dat een feature groepsinteractie kan beïnvloeden. Zo kan een tafel bijvoorbeeld groepsinteractie aanmoedigen terwijl een toonbank



Figuur 8.2: Displays worden als een deel van een groep beschouwt wanneer de meerderheid van een groep naar ongeveer dezelfde regio op dat scherm aan het kijken zijn. De grijze regio van de onderste display maakt deel uit van de groene groep. Maar de rechtse display maakt geen deel uit van de blauwe groep omdat de groepsleden op andere regio's van het scherm aan het focussen zijn, en deze regio's te ver uit elkaar liggen.

groepsinteractie tussen meer dan drie personen ontmoedigt (zie hoofdstuk 3.2). Wanneer de meerderheid van een groep naar ongeveer dezelfde regio op hetzelfde scherm aan het kijken zijn, worden deze regio's samengenomen tot één grote regio. Deze regio (deel van een scherm) wordt dan beschouwd als een deel van die groep (zie Figuur 8.2).

Grijze rechthoeken op features in Figuur 8.2 geven regio's op features aan waar groepen naar aan het kijken zijn. De groene stip op zo'n grijze regio geeft het middelpunt van die regio aan. De dunnere gekleurde balken op features (bijvoorbeeld de twee groene balken in Figuur 8.2) geven weer waar individuele personen aan het kijken zijn op features, de balk krijgt de kleur van de groep waarin die persoon zich bevindt.

Om te bepalen of een persoon naar een scherm aan het kijken is en zo ja naar welke regio op dat scherm hij of zij aan het kijken is, wordt eerst gecontroleerd of een persoon dichtbij genoeg staat. Deze minimum afstand

kan voor elk type display ingesteld worden (zie 9.1.1). Standaard is ingesteld dat een gebruiker dichtbij genoeg staat als de afstand tussen de persoon en een scherm maximum even groot is als de breedte van het scherm. Deze standaard afstand is gekozen door verschillende afstanden op verschillende vooropgenomen scenario's testen. Op deze manier hebben grotere displays een groter bereik.

Om te bepalen naar waar een persoon kijkt wordt naar de richting van het hoofd gekeken, er wordt enkel rekening gehouden met de yaw, niet met de pitch en de roll. De roll wordt niet gebruikt omdat het niet uitmaakt of een gebruiker zijn hoofd kantelt wanneer hij naar een scherm kijkt (of niet kijkt). De pitch wordt niet gebruikt omdat het algoritme enkel rekening houdt met de 2D positie (top down) van personen en om de pitch te kunnen gebruiken moet ook de z-positie (de vertical positie) van het hoofd geweten zijn. Een regio op een display wordt dus enkel in de breedte aangegeven en niet in de hoogte. Als kijkhoek wordt 20 graden genomen. Deze kijkhoek is gekozen door verschillende hoeken te testen op verschillende vooropgenomen scenario's. Deze hoek is vrij klein, we hebben ook grotere hoeken uitgetest zoals 125 graden. Seifried et al. gebruiken een hoek van 125 graden voor hun *Regional Undo/Redo Technique*[26]. Door de afstand van een persoon tot een display te meten en een field of view van 125 graden te nemen berekenen ze welke regio op dat scherm voor die persoon zichtbaar is. Onze hoek is een pak kleiner, dit is zo gekozen omdat we willen bepalen waar de focus van een gebruiker ligt, niet wat de gebruiker allemaal kan zien. Bovendien wordt de kijkhoek bij Seifried et al. gebruikt in een toepassing waarbij gebruikers dicht bij een display staan en de berekende regio's dus minder snel de hele display in beslag nemen, ondanks de grote hoek. In onze situatie kan een gebruiker ook vanop een grotere afstand naar een display aan het kijken zijn. En uit de gebruikerstest van Seifried et al. bleek ook dat de gebruikers van hun techniek de regio vaak te groot vonden en de field of view dus iets te groot was.

Als genoeg personen binnen een groep naar hetzelfde scherm aan het kijken zijn worden de regio's samengenomen mits ze dicht genoeg bij elkaar liggen. Deze samengenomen regio wordt dan toegevoegd aan de groep, met als positie het middelpunt van die groep. In Figuur 8.2 maakt de onderste display deel uit van de groene groep. De meest rechtse display maakt geen deel uit van de blauwe groep omdat de regio's op de display niet dicht genoeg bij elkaar liggen.

8.3 Bepalen of personen met elkaar interageren

Om te bepalen of personen binnen een groep met elkaar interageren wordt eerst gekeken hoever ze zich bevinden van het geometrisch zwaartepunt van de groep. Het geometrisch zwaartepunt van een groep is de gemiddelde

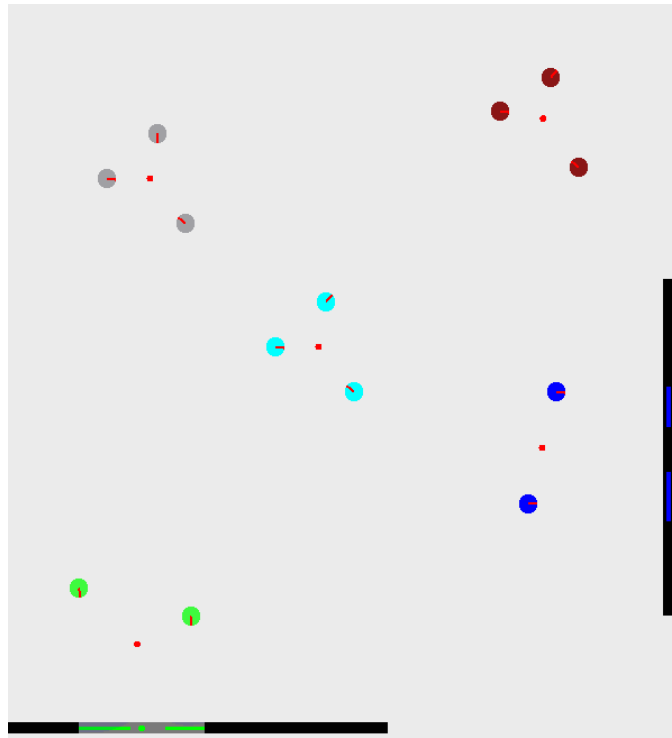
positie van alle personen en displays in die groep (rode punt in Figuur 8.3). Op basis van dit zwaartepunt wordt een focus gebied bepaald voor elke persoon in een groep. Het focus gebied van een persoon is een cirkel met als straal de afstand van de persoon tot het geometrisch zwaartepunt van de groep. Het middelpunt van het focus gebied ligt op dezelfde afstand van de persoon tot het zwaartepunt van de groep maar in de kijkrichting van die persoon (zie Figuur 8.4).

Het focus gebied is gebaseerd op de O-space van F-formations, dit is de ruimte waar de focus van de groep ligt en de belangrijkste activiteit van de groep plaatsvindt (zie hoofdstuk 5.1). Het verschil is dat de O-space het interactie gebied van de groep voorstelt, terwijl het focus gebied het mogelijke interactie gebied van één persoon voorstelt. De regio waar de focus gebieden van de verschillende personen overlappen kan wel gezien worden als de O-space.

Displays in een groep krijgen geen focus gebied. Maar ze beïnvloeden de focus gebieden wel doordat een display het geometrisch zwaartepunt beïnvloedt en dat zwaartepunt beïnvloedt dan weer de straal van de focus gebieden (zie Figuur 8.6).

Om deel te blijven uit maken van een groep, moeten de personen – die deel uitmaken van die groep – hun focus gebieden overlappen. Als de focus gebieden niet overlappen wordt de groep opgesplitst in meerdere groepen van personen wiens focus gebieden wel overlappen. Net zoals bij het vormen van groepen op basis van interpersonal distance kan het zijn dat twee personen in dezelfde groep zitten ook al overlapt hun focus gebied niet. Dit kan wanneer de focus gebieden van persoon **A** en **B** overlappen en ook de focus gebieden van persoon **B** en **C** overlappen, maar de focus gebieden van **A** en **C** niet overlappen. Op deze manier kunnen drie personen die zij aan zij staan ook een groep vormen.

Wanneer een persoon in een groep even opzij kijkt en zijn focus gebied daarom niet meer overlapt met de rest van de groep, wordt hij evenzeer niet uit de groep gegooid. Wanneer in minstens 30% van de laatste 480 frames (twee seconden) zijn focus gebied wel overlapt met de groep, blijft die persoon deel uitmaken van de groep. Dit is zo gedaan om te voorkomen dat wanneer een persoon even wegstijgt of afgeleid is door iets, hij niet meteen uit de groep wordt gegooid. Door te interpoleren over een langere tijd wordt een persoon nog minder snel uit een groep gegooid wanneer hij of zij wegstijgt, maar dat heeft als nadeel dat wanneer een persoon effectief de groep verlaat dit minder snel gedetecteerd wordt. De huidige instellingen zijn gekozen door verschillende waardes uit te proberen op vooropgenomen scenario's.

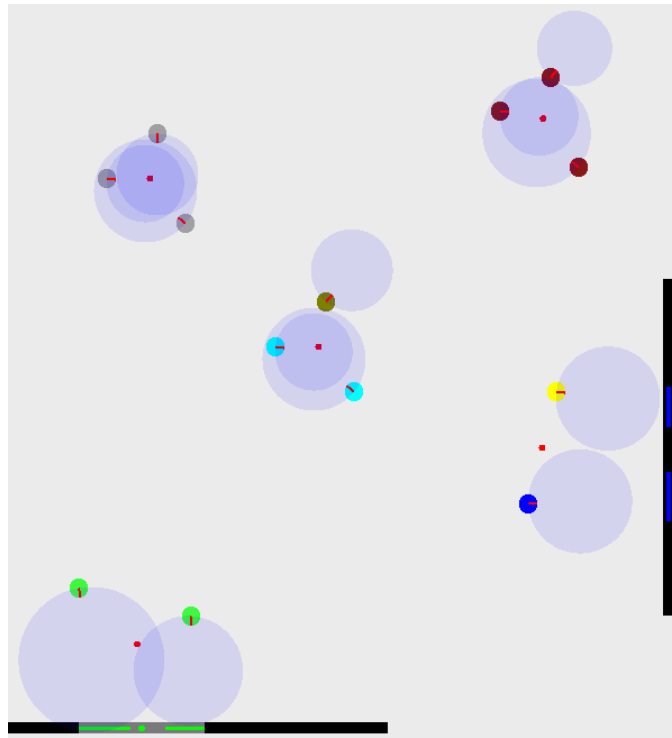


Figuur 8.3: Voor elke groep wordt het geometrisch zwaartepunt berekend (rode stip). De afstand van een persoon tot het zwaartepunt van zijn of haar groep wordt gebruikt om het zwaartepunt te berekenen (zie Figuur 8.4).

8.4 Accuraatheid

Het programma kan niet altijd correct groepen detecteren. Daarom wordt voor elke persoon in een groep een percentage berekend dat aangeeft hoe waarschijnlijk het is dat de persoon bij die groep hoort (Figuur 8.5). Dat percentage is het gemiddelde van twee andere percentages die eerst berekend worden.

Het eerste percentage is gebaseerd op basis van de afstand van de persoon tot het geometrisch zwaartepunt van de groep. Wanneer de afstand tot het zwaartepunt kleiner is dan 60cm (de helft van de minimum interpersonal distance), is dit percentage 100%. Wanneer de afstand gelijk is aan 1,20m (de minimum interpersonal distance), dan is het percentage 50%. Het eerste percentage heeft dus altijd een waarde tussen 50 en 100 procent. Het percentage ligt nooit onder 50% omdat een percentage onder de 50% zou aangeven dat de persoon eerder niet dan wel bij de groep hoort en het programma enkel groepen detecteert die waarschijnlijk wel groepen zijn.



Figuur 8.4: De blauwe cirkels geven de focusgebieden van personen aan. Om tot dezelfde groep te kunnen behoren moeten de focusgebieden overlappen.

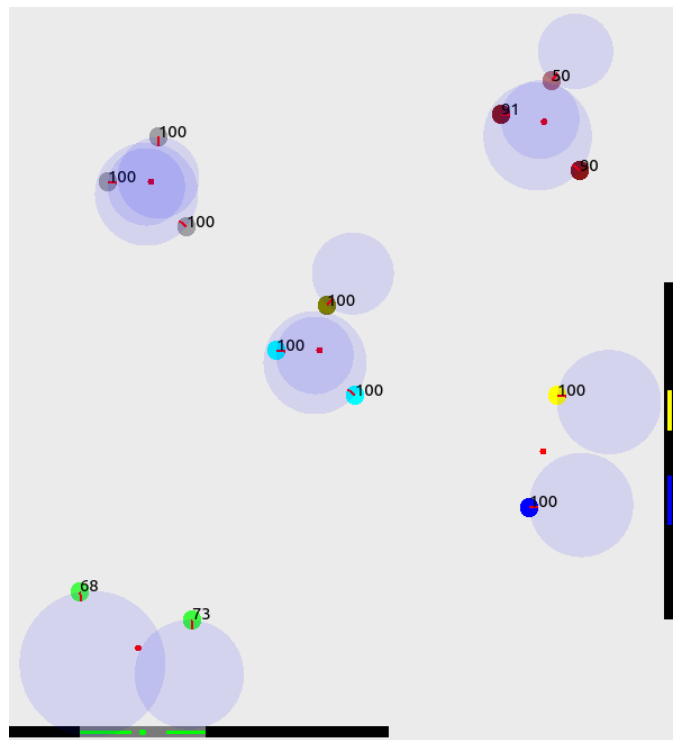
Het tweede percentage is gebaseerd op hoeveel het focusgebied van de persoon overlapt met de andere focusgebieden in die groep. Om te kijken hoeveel twee focus gebieden overlappen wordt niet naar het oppervlakte van de overlappende gebieden gekeken, maar naar de afstand tussen de twee middelpunten van de focusgebieden ($dist$). Twee focus gebieden overlappen pas wanneer $dist$ kleiner is dan de som van hun stralen (max). We berekenen het oppervlakte van de overlapping van twee focusgebieden niet omdat dit een ingewikkeldere berekening is dan de afstand tussen de twee middelpunten van de focusgebieden te berekenen terwijl de afstand berekenen voldoende is om te controleren of de focusgebieden overlappen. Wanneer de twee focusgebieden niet overlappen is het percentage 0%. Wanneer $dist$ kleiner of gelijk aan de helft van max is, is het percentage 100%. Het percentage ligt dus altijd tussen 0% ($dist \geq max$) en 100% ($dist \leq \frac{max}{2}$).

Voor elke persoon wordt op deze manier een percentage berekend met iedere andere persoon die ook in die groep zit. Hiervan wordt een gemiddelde genomen en dit gemiddelde is het tweede percentage voor die persoon.

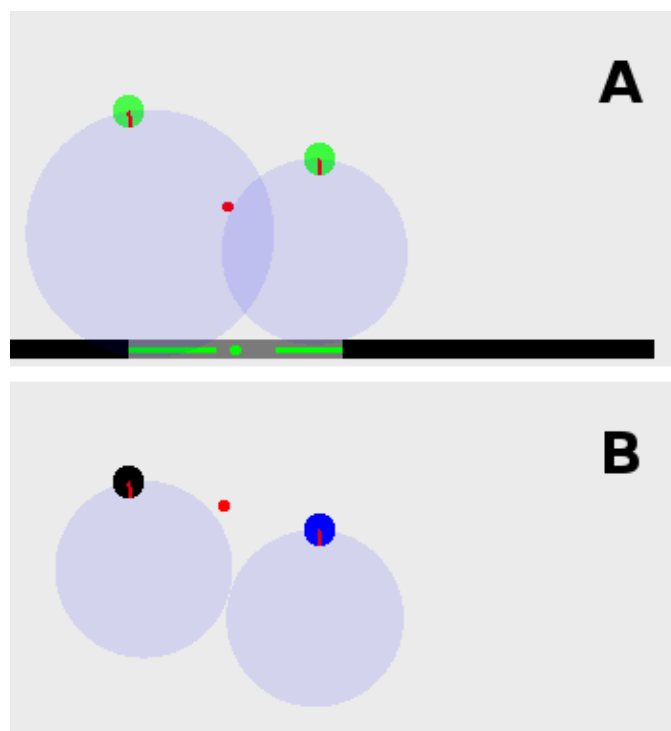
Het uiteindelijke percentage is dus het gemiddelde van deze twee percentages. Dit percentage kan gebruikt worden door ontwikkelaars, bijvoorbeeld

bij toepassingen waarvan men vrij zeker moet zijn of personen samen een groep vormen, dan kan de toepassing personen met een laag percentage vragen of ze wel degelijk een groep vormen. Zo'n interactie noemt een mixed-initiative interactie. Bij mixed-initiative interactie draagt elke agent van een systeem (mens en computer) bij aan een taak waar zij het beste in zijn [27].

Het accuraatheid percentage kan bijvoorbeeld in een toepassing gebruikt worden waarbij de personen hun silhouet gespiegeld worden op een grote display. Alle personen van dezelfde groep krijgen dezelfde kleur. Maar wanneer een persoon een erg laag accuraatheid percentage heeft wordt het silhouet van die persoon transparanter gemaakt. Wanneer de gebruiker ziet dat zijn silhouet doorzichtig is kan hij een stap richting de rest van de groep zetten om aan te geven dat hij deel uitmaakt van de groep.



Figuur 8.5: Voor elke persoon in een groep wordt een percentage berekend dat aangeeft hoe waarschijnlijk het is dat de persoon bij die groep hoort. De rode persoon rechtsboven heeft een erg laag percentage (50%), dit komt omdat hij wegstijgt van de groep en zijn focusgebied slechts amper met dat van de anderen overlapt.

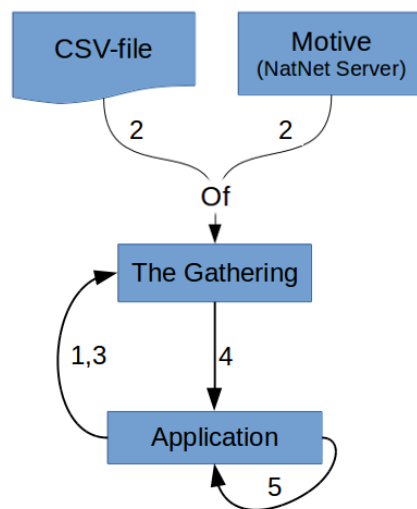


Figuur 8.6: Deze figuur geeft aan hoe een feature het vormen van groepen kan beïnvloeden. Zonder de display verandert het zwaartepunt van de groep en veranderen de focusgebieden waardoor ze niet meer overlappen en dus geen groep vormen.

Hoofdstuk 9

The Gathering: Architectuur

In dit hoofdstuk wordt een overzicht gegeven van de architectuur van the Gathering en wordt uitgelicht hoe ontwikkelaars the Gathering kunnen gebruiken in hun applicatie en welke informatie the Gathering aanbiedt aan ontwikkelaars (zoals uit welke personen een groep bestaat, het zwaartepunt van een groep, of een groep naar een feature aan het kijken is, ...). Verder wordt ook nog een tool toegelicht die kan gebruikt worden om the Gathering uit te proberen of om het algoritme beter te leren begrijpen.



Figuur 9.1: Overzicht architectuur the Gathering. Een applicatie die gebruik maakt van the Gathering werkt als volgt: 1) Een instantie van the Gathering aanmaken. 2) The Gathering haalt info over personen uit een CSV-file of leest realtime informatie gestreamd door Motive. 3) The Gathering wordt gestart. 4) Informatie over groepen wordt doorgestuurd via signal/slot mechanisme. 5) Applicatie verwerkt die informatie.

The Gathering maakt het mogelijk voor ontwikkelaars om gebruik te maken van informatie over groepen in de omgeving zonder dat ze zich zorgen moeten maken over het detecteren van personen en groepen. Deze informatie wordt via het signal en slot mechanisme van Qt¹ doorgestuurd. Zie Figuur 9.1 voor een overzicht van de architectuur. Het signal en slot mechanisme is een event-driven mechanisme. Een signal wordt verstuurd wanneer een bepaalde actie gebeurt en een slot is een functie die opgeroepen wordt wanneer een bepaalde signal verstuurd wordt. Op deze manier kunnen twee objecten die van elkaar losgekoppeld zijn toch nog berichten naar elkaar sturen. Het signal en slot mechanisme in the Gathering zou ook door een ander event systeem vervangen kunnen worden.

Momenteel kan the Gathering op twee manieren informatie ophalen over personen in de omgeving. De eerste manier is via het motion capture systeem Optitrack (zie hoofdstuk 7.2). De mensen in de ruimte dragen markers op hun hoofd, het systeem berekent vervolgens de positie en oriëntatie van deze markers. Deze informatie wordt dan realtime door Motive gestreamd en the Gathering leest deze informatie. De data wordt volgens het NatNet protocol gestreamd. De tweede manier om informatie op te halen is via een CSV-file. De personen worden ook met Optitrack getrackt maar dit wordt niet realtime doorgestuurd, alle informatie wordt naar CSV-file geschreven. The Gathering kan dan naderhand deze CSV-file inlezen. Zo'n CSV-file is handig bij het ontwikkelen van een applicatie, je kunt op voorhand enkele scenario's opnemen en dan hoeven er niet de hele tijd mensen rond te lopen om de applicatie te testen.

9.1 The Gathering gebruiken in een applicatie

In een applicatie die gebruik maakt van the Gathering gebeuren de volgende stappen (zie Figuur 9.1):

Stap 1: Een instantie aanmaken

De applicatie maakt een instantie van the Gathering aan, hierbij wordt al aangegeven wat als input methode gebruikt wordt.

Om gebruik te maken van een CSV-file, moet de locatie van die file meegegeven worden:

```
GroupDetection *detection;  
detection = new GroupDetection("frames.csv");
```

Om realtime data van Optitrack te gebruiken moet je je eigen IP-adres meegeven en het IP-adres van de NatNet server:

```
string localAddress = "192.168.1.108";
```

¹www.qt.io

```
string serverAddress = "192.168.1.101";
detection = new GroupDetection(
    localAddress, serverAddress);
```

Vervolgens kun je de signals ook al met een slot verbinden, zodat nadat the Gathering groepen heeft gedetecteerd hij hier informatie over kan doorsturen. Dit hoeft niet noodzakelijk in stap één te gebeuren, dit kan ook nog in stap drie gebeuren.

```
connect(detection,
    SIGNAL(frameProcessed(std::vector<Group*>)),
    groupsProcessor,
    SLOT(frameProcessed(std::vector<Group*>)));
```

Step 2: The Gathering leest info over personen

Vervolgens gaat de the Gathering starten met informatie te lezen over de personen in de omgeving via ofwel een CSV-file ofwel via de gestreamde Optitrack data.

Step 3: The Gathering starten

Daarna wordt in de applicatie the Gathering gestart, vanaf dan begint the Gathering met het detecteren van groepen en het doorsturen van informatie over deze groepen. Als realtime Optitrack data als input gekozen is, wordt informatie over de groepen ook realtime doorgestuurd. Als een CSV-file als input gekozen is, wordt na elke frame ongeveer 8,33 milliseconden gewacht, zodat er ongeveer 120 frames per seconde verwerkt worden.

```
g->start();
```

Step 4: Informatie over groepen wordt doorgestuurd

Als in een van de vorige stappen een signal met één of meer slots verbonden is, wordt informatie over de groepen doorgestuurd van the Gathering naar de applicatie via het signal en slot mechanisme.

Step 5: Informatie verwerken

De applicatie kan dan de informatie over de verschillende groepen verwerken.

```
void GroupsProcessor::frameProcessed(
    std::vector<Group*> groups) {
    // Verwerk informatie over groepen
    // ...
}
```

Deze doorgestuurde informatie bestaat uit een lijst van groepen en elke groep bevat de volgende informatie:

- Het zwaartepunt van de groep en uit hoeveel personen de groep bestaat.

```
Co Group::getCentroid();
size_t Group::size();
```

- Of de groep naar een feature aan het kijken is. En zoja, naar welke feature de groep aan het kijken is en naar welke regio op die feature.

```
bool Group::isLookingAtFeature();
Feature *Group::getFeature();
FeatureRegion *Group::getFeatureRegion();
```

- Uit welke personen de groep bestaat. Van elke persoon kan nog meer informatie op gevraagd worden zoals onder andere de positie en de waarschijnlijkheid dat de persoon tot de groep behoort.

```
GroupMember *Group::getMember(int index);
double GroupMember::getCertainty();
double GroupMember::getX();
double GroupMember::getY();
```

9.1.1 Features toevoegen

Als er zich features in de ruimte bevinden zoals displays kan dit op de volgende manier aan the Gathering doorgegeven worden:

```
void GroupDetection::addFeature(Feature *feature);
```

The Gathering ondersteund voorlopig nog maar één type feature een rechthoekige (in bovenaanzicht) display.

```
class RectangularDisplay : public Feature {
public:
    RectangularDisplay(
        Co p1, Co p2, double thickness);
    // ...
}
```

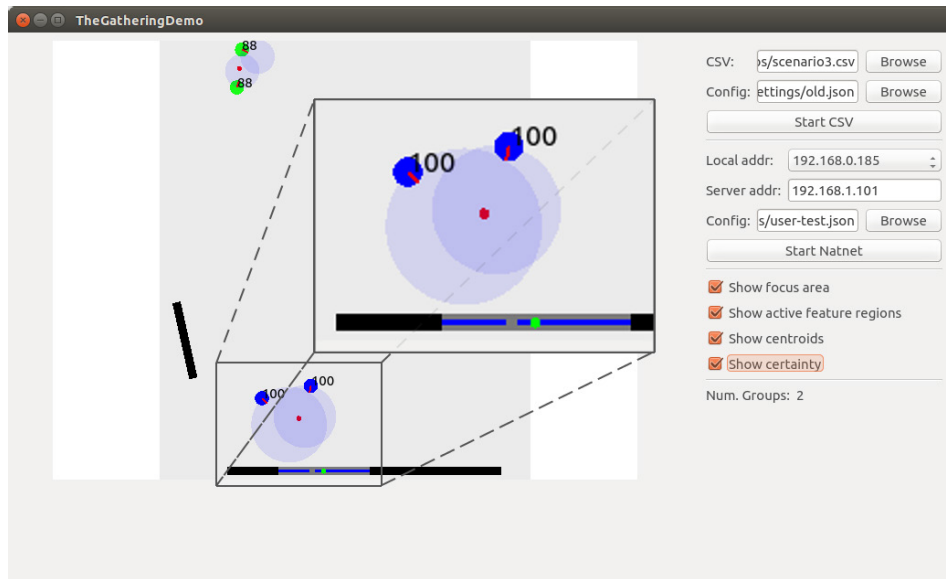
Applicatie ontwikkelaars kunnen wel hun eigen types features implementeren door een klasse aan te maken die overerft van de abstracte klasse *Feature*. In die klasse moeten wel nog enkele functies geïmplementeerd worden. Dit is onder andere functie die aangeeft hoe dichtbij een gebruiker moet staan om met het feature te kunnen interageren, en een functie die bepaalt of een persoon naar de regio aan het kijken is.

Het nadeel van deze manier van features toevoegen is dat de positie van elke feature manueel moet ingegeven worden. Een andere manier zou zijn om elke feature een marker te geven zodat de positie van elke feature automatisch gedetecteerd kan worden. Een rechthoekige feature kan op verschillende manieren met markers worden doorgegeven. Er kunnen twee markers gebruikt worden, zo weet het systeem zowel de grootte als de positie. Of er kan een marker in het midden gelegd worden en de hoogte en breedte moet daarna handmatig ingegeven worden. Bij de methode met twee markers moet wel nog handmatig aangegeven worden welke twee markers bij elkaar horen en bij de methode met één marker moet handmatig de dimensies ingegeven worden. Bij beide methodes moet ook nog manueel aangegeven worden welke markers features voorstellen en van welk type feature (rechthoekige feature in dit geval).

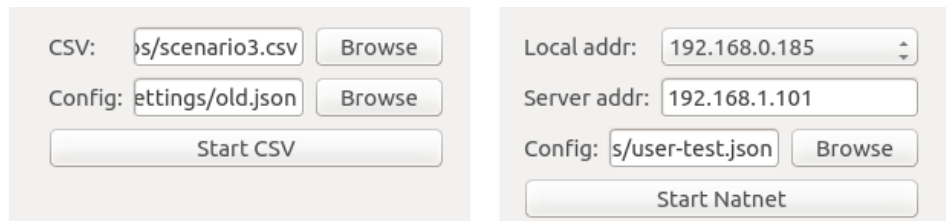
We hebben besloten om features niet met markers aan te duiden omdat er toch sowieso nog manueel informatie moet ingegeven worden en om te vermijden dat een feature plots niet meer gedetecteerd wordt nadat een gebruiker per ongeluk een marker bedekt en deze niet meer zichtbaar is voor de camera's.

9.2 Tool support

De library bevat ook een tool, deze kan gebruikt worden om the Gathering even uit te proberen en ook om het algoritme beter te leren begrijpen (zie Figuur 9.2). De tool heeft een GUI waarmee makkelijk een CSV-file ingelezen kan worden en waarmee makkelijk een verbinding met Optitrack gemaakt kan worden (zie Figuur 9.3). De applicatie visualiseert de groepen, personen en features in de ruimte, en kan ook virtuele elementen visualiseren die helpen bij het begrijpen van hoe groepen gevormd worden zoals de zwaartepunten van groepen, focusgebieden, ...



Figuur 9.2: Tool aanwezig in de library. De applicatie visualiseert de personen, groepen en features in de omgeving. De tool is ook handig om de werking van het groepdetectie algoritme beter te begrijpen omdat het ook onder andere focusgebieden en zwaartepunten van groepen kan visualiseren.



Figuur 9.3: Via de tool kun je gemakkelijk vooropgenomen scenario's, die geëxporteerd zijn naar een CSV-file inlezen (links). Je kunt er ook makkelijk realtime Optitrack data, die via het Natnet protocol gestreamd wordt mee inlezen (rechts). De tool leest ook een configuratie file (een JSON bestand). In dit bestand kunnen makkelijk features en nog andere configuratie elementen geconfigureerd worden.

Elke persoon wordt voorgesteld door een cirkel met een rode streep in, de rode streep geeft aan in welke richting de persoon aan het kijken is. Personen die samen een groep vormen hebben allemaal dezelfde kleur. Features worden voorgesteld door een zwarte balk. Grijs rechthoeken op features geven regio's op features aan waar groepen naar aan het kijken zijn. De groene stip op zo'n grijs regio geeft het middelpunt van die regio aan. De dunnere gekleurde balken op features (bijvoorbeeld de twee blauwe balken

op Figuur 9.2) geven weer waar individuele personen aan het kijken zijn op features, de balk krijgt de kleur van de groep waarin die persoon zich bevindt. De lichtblauwe transparante cirkels geven de focusgebieden van personen weer. De tekst naast elke persoon geeft de waarschijnlijkheid aan dat een persoon tot die groep behoort (van 0 tot 100, zie hoofdstuk 8.4).

Sommige van deze elementen kunnen afgezet worden via checkboxen (rechts op de applicatie, zie Figuur 9.2). Deze elementen zijn: De focusgebieden, de zwaartepunten, de dunne gekleurde balken en de percentages die de accuraatheid aangeven.

De applicatie leest ook een configuratie bestand in (zie Figuur 9.3), dit is een JSON bestand. Via dit bestand kunnen features ingelezen worden, er kan ook ingelezen worden hoe de assen van Optitrack geïnterpreteerd moeten worden, en er kan meegegeven worden welke regio de applicatie zeker moet visualiseren.

```
{
  "bounds": [
    {"x": -1.7, "y": 1.8},
    {"x": 1.2, "y": -2.2}
  ],
  "axis": {
    "x": "-x",
    "y": "-z"
  },
  "features": [
    {
      "type": "RectangularDisplay",
      "p1": {"x": -1.4, "y": -1.3},
      "p2": {"x": -1.45, "y": 0.25},
      "thickness": 0.1
    }
  ]
}
```

bounds verwacht twee (x, y) coördinaten, dit geeft aan welke regio de applicatie minstens moet visualiseren. Afhankelijk van de grootte van de window van de applicatie kan er ook een grotere regio weergegeven worden, maar de regio tussen de twee bounds coördinaten wordt altijd weergegeven.

axis verwacht twee strings, een string *x* en een string *y*. Deze strings moeten elk één van de volgende waardes hebben: x, -x, y, -y, z of -z. Dit is nodig omdat afhankelijk van hoe Optitrack gekalibreerd is, de assen anders kunnen liggen. Als *x* bijvoorbeeld de waarde “-x” bevat, dan wordt x-as van Optitrack ook in the Gathering als x-as beschouwd, maar wel in de

omgekeerde richting. De applicatie geeft de x-as (de the Gathering x-as, niet de Optitrack x-as) altijd horizontaal weer van links naar rechts, en de y-as (de the Gathering y-as) verticaal van boven naar onder.

features verwacht een array van features. Het voorbeeld hierboven bevat slechts één feature, deze is van het type *RectangularDisplay*, *p1* en *p2* stellen de twee uiterste posities van de display voor en *thickness* geeft aan hoe dik de display is. De tool ondersteund voorlopig enkel nog maar features van het type *RectangularDisplay* (net zoals the Gathering).

Hoofdstuk 10

The Gathering: Evaluatie

Om het systeem te evalueren hebben we zes personen uitgenodigd voor een gebruikerstest. De gebruikerstest bestond uit twee delen, in het eerste deel mochten de personen vrij rondlopen en in het tweede deel kregen ze instructies van ons die ze moesten uitvoeren. The Gathering detecteerde tijdens de gebruikerstest groepen en het systeem werd geëvalueerd door te vergelijken tussen hoe the Gathering de personen in groepen opsplijste en hoe wij de personen in groepen zouden opsplitsen.

10.1 Deelnemers

De testpersonen bestonden uit drie vrouwen en drie mannen, ze werken allemaal in ons lab en hebben een achtergrond in de informatica. Er werd hun op voorhand niet verteld dat het ging om het detecteren van groepen om te voorkomen dat ze zich te zelfbewust zouden worden van hoe ze zich ten opzichte van elkaar opstellen en dit hun gedrag zou beïnvloeden. Drie van de zes gebruikers werden speciaal uitgekozen voor de gebruikerstest omdat we wisten dat ze bevriend met elkaar waren en hopelijk samen een groep zouden vormen.

10.2 Set-up

De grootte van het gebied waarin de gebruikerstest uitgevoerd kon worden werd beperkt door het motion capture systeem. Met touw werd een gebied van ongeveer 3,5x5,5 meter afgebakend om aan te duiden waarbinnen het systeem personen kon tracken, de gebruikers mochten hier niet buiten gaan. De set-up van het motion capture systeem, Optitrack, is beschreven in hoofdstuk 7.2.

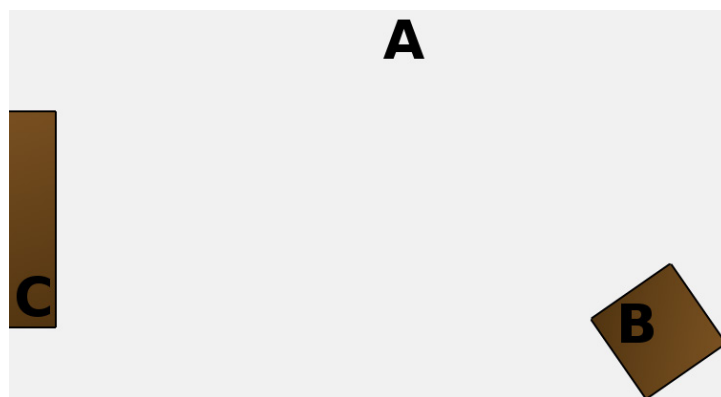
Elke gebruiker kreeg een pet met bovenop de pet een marker zodat iedereen getrackt kon worden door Optitrack. Tijdens de test werd verschillende data opgenomen om achteraf te analyseren. Twee video camera's werden

gebruikt om de gebruikers te filmen. De data verzameld door Optitrack werd opgenomen en geëxporteerd naar een CSV-file. En er werd een screenshot gemaakt van een applicatie die realtime de gedetecteerde groepen visualiseerde.

10.3 Deel 1: Vrij rondlopen

In het eerste deel van de gebruikerstest mochten de gebruikers vrij rondlopen. Om hun een doel te geven werd er in de ruimte drie vragen geplaatst, de antwoorden op deze vragen moesten ze op een antwoordblad noteren. Er werden ook enkele afgedrukte wikipedia pagina's verspreid over de ruimte, deze pagina's bevatten informatie nodig om de vragen te kunnen oplossen (zie Figuur 10.1).

De eerste vraag was "Rangschik deze films van oudste tot meest recente.", naast deze vraag hingen wikipedia pagina's over deze (en nog andere) films. Deze wikipedia pagina's bevatte ook informatie nodig om vraag drie op te lossen, dit was zo gedaan om ervoor te zorgen dat de gebruikers heen en weer moeten lopen. De tweede vraag was "Verbind de albumhoes met de overeenkomstige artiest.", naaste deze vraag lagen weer wikipedia pagina's die konden helpen bij het oplossen van die vraag. De derde vraag was "Welke artiesten spelen mee of worden vertolkt in deze films?", dit waren allemaal artiesten en films die ook in vraag één en twee voorkwamen.



Figuur 10.1: Set-up eerste deel gebruikerstest. Op plaats A hangt vraag één en enkele wikipedia pagina's met informatie over vraag één en drie. Op plaats B ligt vraag twee op een tafeltje samen met enkele wikipedia pagina's met informatie over deze vraag. En op plaats C hangt vraag drie op, geplakt tegen een display op een tafel.

Om interactie tussen gebruikers en het vormen van groepen aan te moedigen, mochten de gebruikers elkaar vragen stellen en helpen. Drie van de zes gebruikers werden zelfs speciaal uitgekozen voor de gebruikerstest omdat

we wisten dat ze bevriend waren en hopelijk dus samen een groep zouden vormen.

10.3.1 Resultaten

In het eerste deel werden zoals gehoopt verschillende groepen gevormd, de groepen veranderden regelmatig tijdens de test doordat mensen elkaar hielpen. Het grootste aantal groepen was drie (soms heel even vier of vijf, wanneer een persoon van de ene groep naar de andere liep) en het kleinste aantal groepen was één. Er deden zich veel verschillende situaties voor waarop we ons algoritme konden testen.

Om te kijken of the Gathering al dan niet correct een groep detecteerde, vergeleken we de oplossing van the Gathering met hoe wij de personen in groepen zouden verdelen. Bij het indelen van personen in groepen hielden we onder andere rekening met parameters waarmee the Gathering ook rekening houdt zoals interpersoonlijke afstand en naar waar personen aan het kijken zijn. Andere parameters die we gebruikten waren praten en lichaamstaal, bijvoorbeeld twee gebruikers die beide iets aanwezen op het antwoordenblad. Ook enkele acties die over een bepaalde tijdspanne gebeurden gaven aan dat gebruikers samen een groep vormden, bijvoorbeeld twee personen die samen naar een tafel toe liepen. The Gathering kan dit niet detecteren omdat the Gathering alles frame per frame bekijkt en geen rekening houdt met vorige frames (behalve voor de oriëntatie, daarvoor wordt wel gekeken naar de vorige 480 frames, zie hoofdstuk 8.3).

Het aantal seconden dat onze indeling van gebruikers in groepen overeenkwam met de indeling van the Gathering werd opgeteld. Dit resultaat werd dan gedeeld door het totaal aantal seconden en dat geeft een percentage dat aangeeft hoe accuraat the Gathering gewerkt heeft.

$$result = \frac{t_{match}}{t_{match} + t_{mismatch}}$$

Hieruit bleek dat 84% van de tijd the Gathering correct groepen detecteerde. 11% van de tijd werden groepen fout gedetecteerd, maar werden deze fouten veroorzaakt doordat Optitrack niet correct de posities van personen kon bepalen. 5% van de tijd werden groepen verkeerd ingeschat door the Gathering (zie tabel 10.1). De momenten waarop één of meerdere groepen fout gedetecteerd werden, maar één of meer groepen correct gedetecteerd werden, zijn evenzeer meegeteld bij incorrect gedetecteerd.

Correcte groepdetectie	84%
Incorrect, te wijten aan tracking	11%
Incorrect, te wijten aan the Gathering	5%

Tabel 10.1: Resultaten gebruikerstest deel één. Percentage van de tijd dat groepen correct gedetecteerd en incorrect gedetecteerd werden.

Problemen tracking systeem

Het gebeurde regelmatig dat het tracking systeem de posities van personen niet kon bepalen. Soms had dit gevolgen voor het correct detecteren van groepen (zie tweede percentage in tabel 10.1). Maar soms had het ontbreken van een persoon geen invloed op het correct detecteren van groepen.

De meest voorkomende situaties wanneer het tracking systeem personen niet kon detecteren waren wanneer: Personen voorovergebogen stonden om informatie te lezen, de markers zijn dan niet altijd zichtbaar voor de camera's. Of wanneer een persoon gehurkt zat en er personen rondom hem stonden, hierdoor werd de marker op die persoon zijn hoofd verborgen door de lichamen van de personen rondom hem. Of wanneer er veel personen dicht bij elkaar staan, dit komt waarschijnlijk doordat grotere personen tussen de camera's en markers van kleinere personen stonden (zie Figuur 10.2).



Figuur 10.2: Veel voorkomende oorzaken voor het niet kunnen tracken van personen waren: Hurken, vooroverbuigen, en te veel personen die dicht bij elkaar staan.

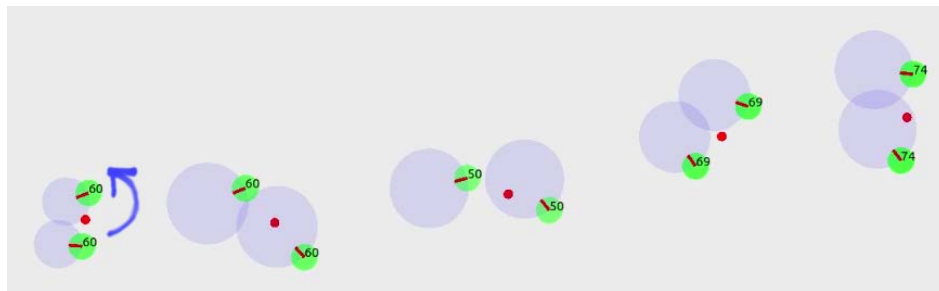
Analyse van enkele situaties

Er kwamen verschillende situaties voor waarbij personen binnen een groep even van elkaar wegkeken. The Gathering bleef in deze situaties meestal correct groepen detecteren. Vaak keken personen even om naar een andere

groep om te zien wat ze aan het doen waren, of werd er naar de andere groep gekeken om daarna om hulp te vragen.

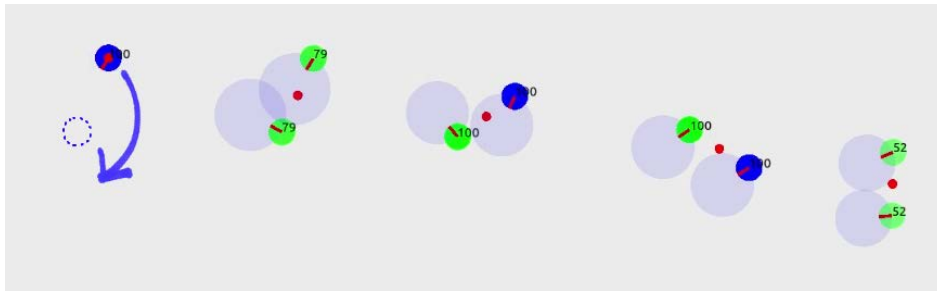
Het toelaten van focus gebieden die even niet overlappen was niet alleen nuttig bij het even wegstaren, maar had ook nut bij situaties die we zelf niet voorzien hadden. Zo wilde een persoon, Jan¹, die eerst links van Sofie stond, rechts van Sofie gaan staan om zo bepaalde informatie beter te kunnen lezen (zie Figuur 10.3). Doordat Jan achter Sofie door loopt, overlappen hun focusgebieden even niet, maar the Gathering blijft ze evenzeer als één groep zien.

Dezelfde situatie deed zich nog eens voor, deze keer wilde Jan van rechts naar links gaan. Maar the Gathering beslist deze keer (incorrect) dat ze niet tot dezelfde groep behoren. the Gathering maakte deze fout omdat net voor deze situatie zich afspeelde, Sofie even niet gedetecteerd werd door het tracking systeem en daardoor geen geschiedenis met Sofie kon opbouwen (zie Figuur 10.4). Dit soort problemen met tracking kunnen verholpen worden door the Gathering robuuster te maken. Zo kan wanneer een persoon plots niet meer gedetecteerd wordt de laatst gekende coördinaten gebruikt worden voor een tijdje. Een nieuw probleem dat dan kan optreden is wanneer een persoon het trackbaar gebied verlaat, in dit geval is het correct dat de persoon niet meer gedetecteerd wordt en gaat the Gathering toch even de laatst gekende positie van de persoon die er niet meer is bijhouden en gebruiken. Een andere oplossing zou kunnen zijn dat the Gathering de laatst gekende positie van een persoon even bijhoudt, maar als dit een positie dicht bij de rand van het trackingsgebied is, dat deze dan minder lang wordt bijgehouden.



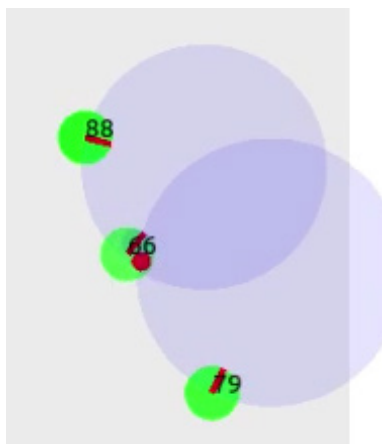
Figuur 10.3: Jan loopt achter Sofie om, om zo beter bepaalde informatie te kunnen lezen. Hun focus gebieden overlappen even niet, maar the Gathering beslist dat ze evenzeer tot dezelfde groep behoren.

¹De namen van de gebruikers zijn veranderd, dit zijn niet hun echte namen.



Figuur 10.4: Jan loopt achter Sofie om. Maar doordat net hiervoor Sofie niet gedetecteerd werd door het tracking systeem, is er geen geschiedenis tussen Jan en Sofie opgebouwd en worden Jan en Sofie niet als één groep gezien wanneer hun focusgebieden even niet overlappen.

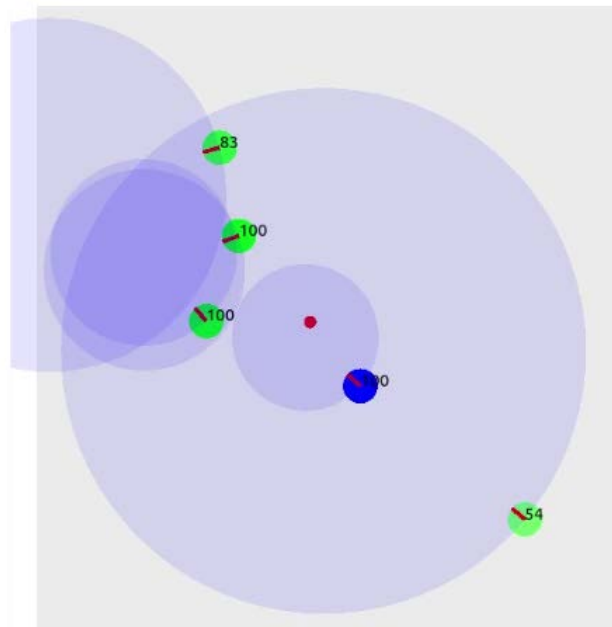
Soms staan enkele personen in een rij naast elkaar, ook in deze situatie detecteert the Gathering correct groepen. Maar bij de analyse van deze situatie is wel opgevallen dat gelijkaardige situaties wel problemen kunnen geven. Wanneer drie personen in een rij naast elkaar staan, bevindt de middelste persoon zich dicht bij het zwaartepunt van die groep (zie Figuur 10.5). Doordat deze persoon zo dicht bij het zwaartepunt staat is de straal van zijn focus gebied erg klein en heeft hij amper een focus gebied. Tijdens de gebruikerstest hebben er zich geen situaties voorgedaan waarbij dit een probleem was. Maar het zou wel problemen kunnen geven, wanneer de andere twee personen een kleiner focus gebied hebben en iets meer naar buiten kijken.



Figuur 10.5: De middelste persoon staat dicht bij het zwaartepunt en heeft daardoor bijna geen focus gebied, dit zou in andere situaties problemen kunnen geven.

De meest voorkomende situaties waren het verlaten en toetreden tot groepen, dit gebeurde vooral wanneer iemand klaar was met een vraag of wanneer personen naar elkaar toe gingen om elkaar te helpen. Soms was het één persoon die zich bij een groep voegde en soms was het een hele groep die zich bij een andere groep voegde.

Over het algemeen worden in deze situaties goed groepen gedetecteerd, maar soms gaat er tijdens het toetreden tot een groep iets fout, maar dit duurt nooit meer dan enkele seconden. Op Figuur 10.6 bijvoorbeeld lopen twee personen naar een groep toe. Eén van die twee personen, degene die het dichtst bij de groep is, wordt nog niet beschouwd als een deel van die groep. Terwijl de andere persoon, die verder verwijderd staat van de groep wel als een deel van de groep wordt beschouwd. Deze fout treed op door een bug in de implementatie van the Gathering, de gebruiker zou wel een deel moeten uitmaken van de groep omdat zijn focus gebied overlapt met een andere persoon in die groep. Maar zelfs met deze bug opgelost kunnen er zich problemen voordoen doordat het focus gebied van de onderste persoon zo groot is. Stel dat er nog een persoon is, Piet, die niet naar de groep toe wandelt maar met iets anders bezig is. Doordat het focus gebied van de onderste persoon zo groot is, overlapt dat focus gebied met dat van Piet en wordt Piet incorrect beschouwd als een deel van de groep.



Figuur 10.6: Door een bug wordt één persoon niet tot de groep gerekend. Maar zelfs zonder de bug kan deze situatie problemen geven door het grote focus gebied van de onderste persoon.

10.3.2 Conclusie

The Gathering detecteerde 84% van de tijd correct alle groepen, het grootste deel van de fouten werden veroorzaakt door slechte tracking van de gebruikers. Enkele zaken van het algoritme moeten herbekeken worden, moet er bijvoorbeeld een limiet staan op de grote van het focus gebied? Of moet de grootte van het focus gebied niet bepaald worden door het zwaartepunt van de groep? Dit laatste zou ook een oplossing kunnen zijn voor het probleem in Figuur 10.5, waar een gebruiker bijna geen focus gebied heeft doordat hij dicht bij het zwaartepunt staat.

10.4 Deel 2: Vooraf bepaalde situaties testen

In het tweede deel van de gebruikerstest moesten vier personen onze instructies opvolgen. De bedoeling van het tweede deel was het creëren van enkele situaties die we wilden uittesten. De reden waarom we hier maar vier personen in plaats van zes gebruikten was omdat vier personen genoeg was om de situaties te creëren die we wilden uittesten, en omdat vier personen instructies geven vlotter gaat dan zes personen instructies geven.

De ruimte bevatte een tafel met twee schermen op en een tafel met een pakje kaarten op. De volgende situaties werden gespeeld door de vier gebruikers.

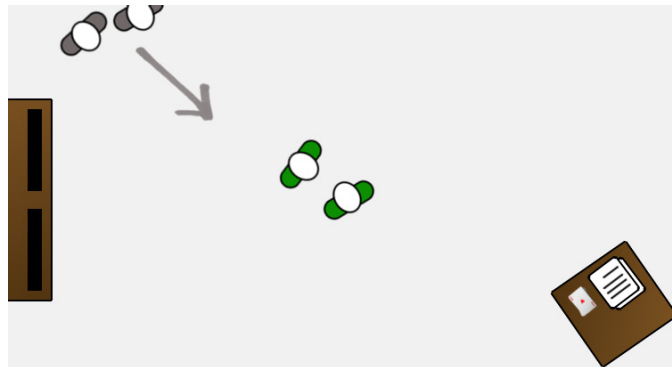
- Twee personen staan in het midden van de ruimte naar elkaar toe gericht en zijn aan het praten. Om de gebruikers een gespreksonderwerp te geven werd hun gezegd elkaars hand te schudden en elkaar voor te stellen. Vervolgens komen de twee andere personen de ruimte binnengelopen en wandelen naar de andere groep toe en mengen zich in het gesprek (zie Figuur 10.7 en Figuur 10.8).
- Eén persoon, Jan, verlaat de groep en wandelt naar het tafeltje met kaarten. De gebruiker krijgt de instructie om het blad dat naast het pakje kaarten ligt te lezen. Hierop staat dat de gebruiker het pakje kaarten in vier hoopjes moet verdelen, een harten stapel, een ruiten stapel, een klaveren stapel en een schoppen stapel. Dit is gedaan om de gebruiker iets te doen te geven.

Een meer realistische situatie is dat een gebruiker aan een drankautomaat of een drankstand iets koopt. Deze situatie aan het tafeltje bootst dat een beetje na.

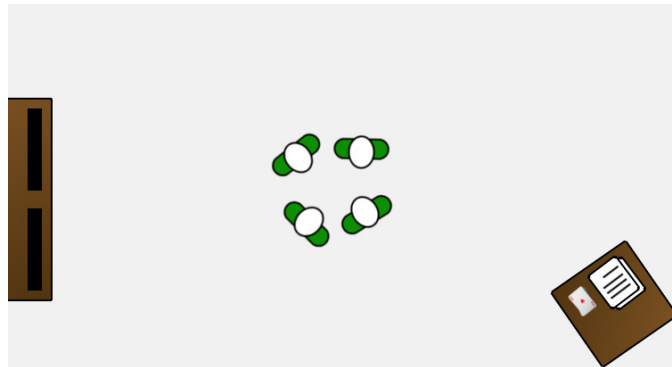
- Vervolgens gaat Sofie bij het tafeltje bij Jan staan. De twee andere personen, Eric en Lize, lopen naar de linker display op de andere tafel. Op deze display staat een slideshow van “Where is Waldo?” puzzels, Eric en Lize moeten Waldo zoeken. Om de 20 seconden komt er een nieuwe puzzel te staan. Dit is weer gedaan zodat de gebruikers iets

te doen hebben en ook effectief met het scherm interageren (zie Figuur 10.9).

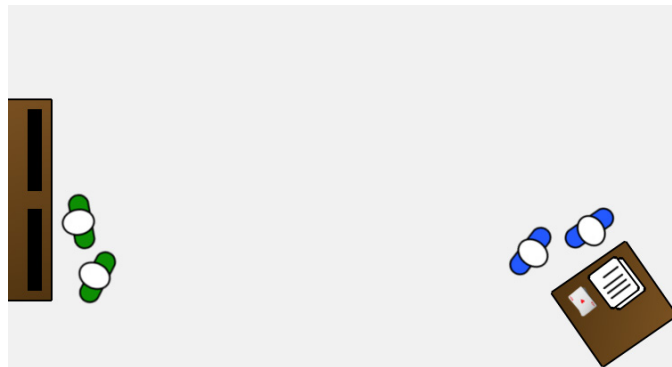
- Vervolgens wandelen Jan en Sofie naar de rechter display toe en wandelen Eric en Lize naar het tafeltje met kaarten toe. Ze kruisen elkaar in het midden van de kamer. Op de rechter display wordt ook een slideshow met “Where is Waldo?” puzzels afgespeeld (zie Figuur 10.10).
- Na een tijdje gaan Eric en Lize weer terug naar de linker display om “Where is Waldo?” puzzels op te lossen. Jan gaat vervolgens bij Eric en Lize staan om hun te helpen hun puzzel op te lossen (zie Figuur 10.11).
- Sofie, die nu alleen aan de rechter display staat, wandelt terug naar het tafeltje met kaarten. Even later gaat Sofie weer terug naar de rechter display en wandelt Eric naar het tafeltje met kaarten. Ze kruisen elkaar in het midden.
- Weer even later gaat Sofie terug bij Eric staan, aan het tafeltje met de kaarten. Vervolgens gaan de twee groepen van positie wisselen. Sofie en Eric wandelen naar de displays toe en Jan en Lize wandelen naar het tafeltje toe, ze kruisen elkaar in het midden.
- Ten slotte lopen de twee groepen naar het midden van de kamer, waar ze elkaar begroeten.



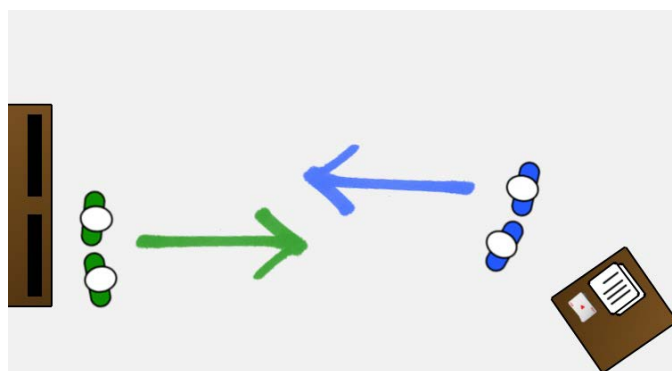
Figuur 10.7: Een groep van twee personen wandelt naar een andere groep van twee personen toe.



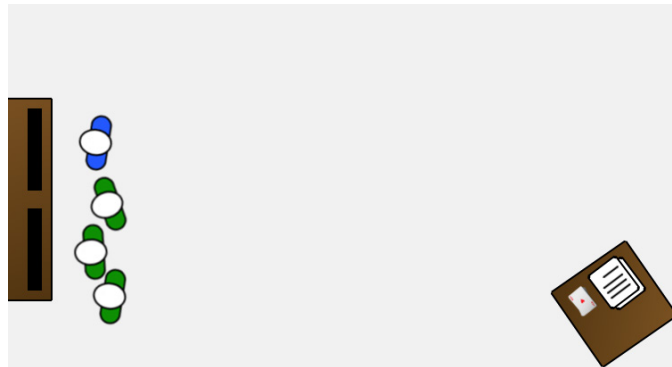
Figuur 10.8: Een groep van vier personen praat met elkaar.



Figuur 10.9: Twee personen interageren met een display. De andere twee personen interageren met een pakje kaarten op een tafeltje.



Figuur 10.10: Twee groepen passeren elkaar.



Figuur 10.11: Jan verplaatst zich van de rechter display naar de linker, hij maakt nu deel uit van de groene groep.

Verschillende interacties komen aan bod, twee personen die samen een groep vormen, twee groepen die samen één groep vormen, een persoon die een groep verlaat, een groep die interageert met een feature, twee groepen die elkaar passeren, ...

10.4.1 Resultaten

The Gathering detecteerde 82% van de tijd correct de groepen, 4% van de tijd werden er fout groepen gedetecteerd, maar deze fouten waren te wijten aan het tracking systeem, 14% van de tijd werden er fout groepen gedetecteerd door the Gathering (zie tabel 10.2). The Gathering detecteert nog steeds redelijk goed groepen (82% van de tijd), maar er worden meer fouten gemaakt die te wijten zijn aan the Gathering zelf (14% in plaats van 5%). Hier zijn twee redenen voor, the Gathering maakte in een bepaalde situatie een fout, en deze situatie duurde vrij lang, deze situatie wordt hieronder nog besproken. Een tweede reden is dat doordat er maar vier personen waren het tracken van gebruikers beter werkte. In deel 1 van de gebruikerstest waren 11% van de fouten te wijten aan het tracken van gebruikers. Maar dat wil niet zeggen dat moest de tracking wel goed hebben gewerkt, er geen fouten door the Gathering zouden zijn gemaakt.

Correcte groepdetectie	82%
Incorrect, te wijten aan tracking	4%
Incorrect, te wijten aan the Gathering	14%

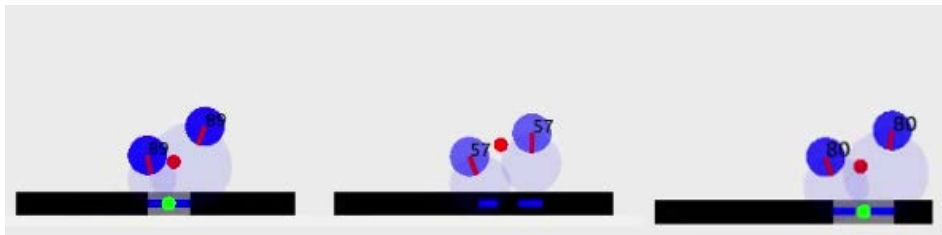
Tabel 10.2: Resultaten gebruikerstest deel twee. Percentage van de tijd dat groepen correct gedetecteerd en incorrect gedetecteerd werden.

Analyse situaties

In de eerste situatie lopen twee personen naar het midden en beginnen met elkaar te praten, een beetje later komen daar nog eens twee personen bij. Deze situatie wordt perfect gedetecteerd door the Gathering, er gaat niets fout, behalve bij het opkomen van de eerste twee personen verspringt de oriëntatie van de eerste gebruiker continu, maar dit is een tracking fout en geen fout van the Gathering.

In de tweede situatie krijgt Jan de opdracht om naar het tafeltje te gaan en de instructies op het blad uit te voeren. Maar één persoon begreep de opdracht verkeerd en ging samen met Jan naar het tafeltje. De persoon wordt teruggedroepen en keert terug naar de groep in het midden van de ruimte. De situatie was dus een beetje anders dan verwacht, maar de groepen werden evenzeer correct gedetecteerd.

Vervolgens gaat Sofie bij Jan aan het tafeltje staan en gaan de andere twee personen naar de linker display. De groepen worden weer correct gedetecteerd. Maar het was de bedoeling dat display ook een deel van de groep zou uitmaken, dit gebeurt maar heel even, de meerderheid van de tijd kijken ze elk op een andere regio op de display en wordt de display niet als een deel van de groep beschouwd (zie Figuur 10.12). Dit komt omdat de gebruikers erg dicht bij de display staan om Waldo te zoeken, hierdoor is de regio op het scherm die voor elke persoon berekend wordt erg klein en overlappen ze minder snel. Een oplossing hiervoor zou een grotere kijkhoek zijn, die staat nu op 20 graden (zie hoofdstuk 8.2). Maar een grotere hoek geeft dan weer problemen wanneer de gebruikers verder van een display staan, dan nemen hun regio's op displays te veel plaats in en kan een display niet met twee personen gedeeld worden. Een tweede reden waarom de regio's niet overlappen is omdat de display een beetje verplaatst is nadat hun positie aan the Gathering is doorgegeven, de display staat eigenlijk iets meer naar achter en iets meer naar links. Nadat de positie van de features gecorrigeerd zijn, krijgen we betere resultaten, maar nog steeds wordt de meerderheid van de tijd de display niet als een deel van de groep beschouwd.



Figuur 10.12: De display zou een deel moeten uitmaken van de groep, soms is dit het geval (links), maar meestal niet (midden). Eén van de redenen is dat de display een beetje verschoven is en eigenlijk maar naar achter moet staan. Wanneer de display correct wordt doorgegeven aan the Gathering maakt de display al vaker deel uit van de groep (rechts), maar nog altijd vaker niet dan wel.

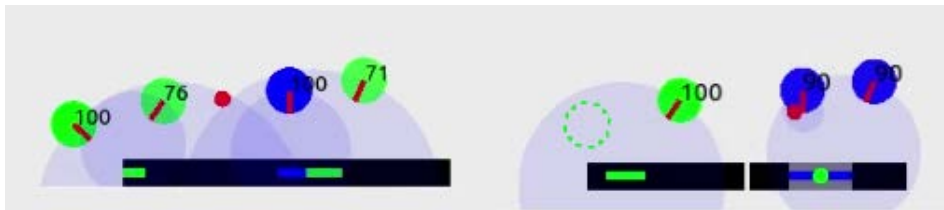
In de volgende situatie gaan Eric en Lize naar het tafeltje en Jan en Sofie naar de rechter display, ze kruisen elkaar in het midden van de ruimte. Wanneer de twee groepen elkaar passeren worden ze als één groep beschouwd door the Gathering (zie Figuur 10.13). Dit was wat we verwacht hadden dat the Gathering zou doen, maar het zou eigenlijk moeten gedetecteerd worden als twee verschillende groepen. Hiervoor zou het algoritme rekening moeten houden met de richting die de personen uitgaan, maar dat zou dan weer in andere situaties problemen geven bijvoorbeeld wanneer twee personen naar elkaar toe wandelen.



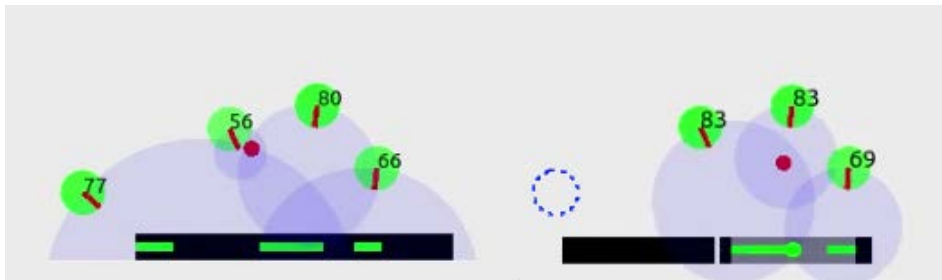
Figuur 10.13: Twee groepen passeren elkaar, the Gathering zou nu twee groepen moeten detecteren, niet één.

Na een tijdje gaan Eric en Lize weer terug naar de linker display. Maar the Gathering detecteert nu niet twee groepen zoals hij zou moeten doen. Jan, Sofie en Lize worden door the Gathering als één groep beschouwd en Eric wordt ook als één groep beschouwd ook al staat Eric tussen Lize en Jan in (zie Figuur 10.14). Dit heeft te maken met dezelfde bug als in hoofdstuk 10.3.1. Zonder deze fout zou the Gathering de vier personen als één groep beschouwd hebben, wat ook niet correct is. Dit komt omdat de personen erg dicht bij elkaar staan en ongeveer in dezelfde richting kijken. Bovendien zijn de twee displays die naast elkaar staan aan the Gathering doorgegeven als één grote display. Moest dit als twee display ingegeven zijn dan worden de personen als twee groepen beschouwd omdat ze elk op een andere display aan het kijken zijn. Doordat we de Optitrack data naderhand naar een CSV-file geëxporteerd hebben kunnen we dit testen. Maar Optitrack heeft de data niet perfect geëxporteerd en de positie van Sofie ontbreekt in deze situatie. Wanneer we de situatie met twee displays testen worden de groepen correct gedetecteerd, maar Sofie maakt nu wel geen deel uit van de situatie (zie Figuur 10.14).

Na een tijdje gaat Jan bij Eric en Lize staan om hun te helpen. Er zou nu een groep van drie personen en een groep van één persoon moeten zijn. Maar the Gathering detecteert één grote groep van vier personen (zie Figuur 10.15), dit komt zoals hierboven al vermeld doordat de twee displays als één grote display ingegeven zijn. Wanneer we dit veranderen naar twee displays lijkt het wel te werken, maar dat kunnen we niet met zekerheid zeggen omdat Sofie nog steeds niet goed getrackt wordt wanneer we de CSV-file gebruiken.



Figuur 10.14: Verschil tussen wanneer de displays als één display worden samengenomen en wanneer niet. De stippelijijn cirkel stelt een persoon voor die niet getrackt wordt maar daar zou moeten staan.



Figuur 10.15: Verschil tussen wanneer de displays als één display worden samengenomen en wanneer niet. De stippelijijn cirkel stelt een persoon voor die niet getrackt wordt maar daar zou moeten staan.

Vervolgens gaat Sofie naar het tafeltje met kaarten, even later gaat Eric naar het tafeltje met kaarten terwijl Sofie naar de rechter display wandelt. Deze situatie wordt goed gedetecteerd, behalve wanneer Sofie en Eric elkaar passeren, dan worden ze even als één groep beschouwd. Een mogelijke oplossing zou zijn om ook rekening te houden met in welke richting personen aan het bewegen zijn. The Gathering zou dan zien dat Sofie en Eric beide een andere richting opgaan.

Daarna gaat Sofie terug bij Eric staan. Een beetje later gaan Sofie en Eric terug naar displays en gaan Jan en Lize naar het tafeltje, ze kruisen elkaar in het midden. Net zoals in de andere situaties waar personen elkaar kruisen wordt er even één grote groep gedetecteerd in plaats van twee groepen.

Ten slotte wandelen beide groepen naar het midden van de ruimte en ontmoeten ze elkaar, dit wordt correct gedetecteerd door the Gathering.

10.4.2 Conclusie

The Gathering detecteerde 82% van de tijd correct alle groepen, dit is iets minder dan in deel 1. Het grootste deel van de fouten werden veroorzaakt doordat twee displays die naast elkaar stonden aan the Gathering doorgegeven werden als één grote display. The Gathering detecteert ook slecht twee groepen wanneer deze elkaar passeren.

Wanneer personen erg dicht bij een display staan worden voor deze personen een slechts erg kleine regio op dat scherm aangeduid, dit geeft soms problemen. De kijkhoek kan eventueel naar een grotere hoek dan twintig graden veranderd worden, maar dit kan problemen geven wanneer personen ver van een scherm staan, een tussenoplossing is eventueel de kijkhoek vergroten wanneer een persoon dicht bij een display staat.

10.5 Algemene conclusie

The Gathering detecteerde in het eerste deel van de gebruikerstest 84% van de tijd correct de groepen en in het tweede deel 82% van de tijd. Wanneer het detecteren van groepen fout ging lag dat vaak aan de tracking. In het eerste deel werd 11% van de tijd slecht groepen gedetecteerd door trackingsfouten en in het tweede deel was dat 4%. The Gathering zou robuuster moeten zijn en zou beter met deze fouten moeten omgaan door bijvoorbeeld de laatst gekende positie van een persoon te gebruiken wanneer deze plots niet meer getrackt wordt.

In het eerste deel werd 5% van de tijd verkeerd groepen gedetecteerd en kwam dit niet door de tracking. In het tweede deel was dat 14%. Eén van de fouten was dat the Gathering twee personen die elkaar passeerden even als één groep beschouwde in plaats van twee verschillende groepen. Een oplossing hiervoor is om niet enkel frame per frame groepen te detecteren maar ook rekening te houden met de voorbij frames. Op deze manier kan beweging gedetecteerd worden en kan the Gathering detecteren dat de twee personen beide in een tegenovergestelde richting aan het wandelen zijn en dus waarschijnlijk niet bij elkaar horen. Een ander probleem dat soms voorkwam was de grootte van de focusgebieden. Bij de huidige methode wordt de grote van de focus gebieden bepaald door het zwaartepunt van de groep. Dit zorgt er soms voor dat focus gebieden heel erg groot kunnen zijn en dus met heel veel andere focus gebieden overlappen. Het omgekeerde komt soms ook voor, namelijk dat de afstand tussen een persoon en het middelpunt zo klein is dat er bijna geen focus gebied is. Een mogelijke oplossing is de grootte van het focus gebied blijven baseren op de afstand tot het zwaartepunt, maar een maximum en minimum grootte instellen.

Tijdens de gebruikerstest deelden we als observanten de gebruikers niet enkel in groepen op basis van de interpersoonlijke afstand en op basis van naar waar ze keken. We keken ook naar gebaren, spraak en naar wat de personen aan het doen waren over een periode (zoals ergens naartoe wandelen). Dit zijn nog drie mogelijke parameters die later aan the Gathering toegevoegd kunnen worden.

Het nadeel van een lab studie is dat het de deelnemers uit hun natuurlijke omgeving haalt. Om the Gathering nog beter te kunnen evalueren moeten we in de toekomst nog een in the wild studie uitvoeren. Zo kunnen we beter de impact van the Gathering begrijpen in ons dagelijks leven [28]. Om een in the wild studie uit te voeren moet er eerst nog een toepassing voor the Gathering ontwikkeld worden en moeten we een ander trackingssysteem gebruiken dat personen kan detecteren zonder dat ze daarvoor markers moeten dragen.

Hoofdstuk 11

Conclusie en future work

Het doel van de thesis is te bepalen hoe we best kunnen detecteren welke personen samen groepen vormen. Tijdens de literatuurstudie ontdekten we dat features (objecten in de omgeving zoals displays en tafels) invloed hebben op de interactie tussen personen. Ze mogen dus niet genegeerd worden tijdens het detecteren van groepen. Behalve features gebruiken we ook nog de afstand tussen personen en de oriëntatie van personen bij het bepalen van groepen. Dit zijn allemaal concepten uit proxemic interaction. Fixed en semi-fixed features zijn concepten rechtstreeks afkomstig uit de proxemic interaction, en de oriëntatie en positie (en dus ook afstand tussen personen) van personen zijn beide proxemic dimensions, deze worden onder andere gebruikt bij het meten van proxemic interaction.

We hebben ook gekeken naar andere groepdetectie methodes. Eén van deze methodes is een rule-based approach, het grote verschil met onze methode is dat ontwikkelaars zelf kunnen bepalen hoe groepen worden gedetecteerd door zelf regels samen te stellen. Een nadeel hierbij is dat dit telkens slechts werkt voor een zeer specifieke situatie. Om ervoor te zorgen dat the Gathering in de meeste situaties goed werkt hebben we het aantal regels beperkt (features, afstand en oriëntatie) en geven we de ontwikkelaars geen mogelijkheid om de regels aan te passen.

Uit de gebruikerstest blijkt dat the Gathering in meer dan 80% van de gevallen alle groepen correct detecteert. Wanneer sommige groepen niet correct gedetecteerd werden kwam dit enerzijds doordat het tracking systeem soms de posities van enkele gebruikers verkeerd bepaalde omdat enkele markers niet goed zichtbaar waren voor de camera's. In de andere gevallen kreeg the Gathering wel de juiste input van het trackingssysteem, maar slaagde the Gathering er niet om om de groepen correct te detecteren.

In de toekomst kan er nog onderzocht worden of the Gathering nog rekening kan houden met andere parameters zoals spraak en lichaamstaal. En ook of the Gathering informatie over meerdere frames heen kan verzamelen en gebruiken (zoals beweging). Cutler et al. kunnen met behulp van een

camera en een microfoon detecteren wanneer iemand aan het praten is en wie aan het praten is [29]. Dit zou mogelijk toegepast kunnen worden op the Gathering. Behalve bepaalde gestures zoals wuiven te detecteren zou the Gathering ook lichaamstaal kunnen proberen te lezen. Tan et al. ontwikkelden een systeem genaamd mASqUE dat uit de lichaamstaal van een persoon de affective state (emotie, humeur, ...) en openheid probeert af te lezen, hiervoor wordt enkel een dieptecamera gebruikt [30].

Een tweede deel van de onderzoeksvraag was te onderzoeken in wat voor toepassingen een groepdetectiemethode gebruikt kan worden. We hebben drie voorbeeldtoepassingen gegeven, maar in de toekomst zou er ook nog een toepassing geïmplementeerd moeten worden zodat we the Gathering nog beter kunnen testen.

In de toekomst moet er ook nog een in the wild studie uitgevoerd worden. Het nadeel van een lab studie is dat het de deelnemers uit hun natuurlijke omgeving haalt en met een in the wild studie kunnen we beter de impact van the Gathering begrijpen in ons dagelijks leven [28]. Om een in the wild studie uit te voeren moet er eerst nog een toepassing voor the Gathering ontwikkeld worden en moeten we een ander trackingsysteem gebruiken dat personen kan detecteren zonder dat ze daarvoor markers moeten dragen.

We besluiten dat we een groepdetectie methode ontwikkeld hebben die tijdens een gebruikerstest in meer dan 80% van de tijd correct groepen detecteert, en rekening houdt met afstand, oriëntatie en features in de omgeving. The Gathering kan een aanzet zijn om meer toepassingen te ontwikkelen, voornamelijk voor grote displays, die rekening houden met personen die samen interageren en die niet enkel focussen op individuele interactie.

Bibliografie

- [1] Daniel Vogel and Ravin Balakrishnan. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. In *UIST '04 Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 137–146. ACM, 2004.
- [2] Daniel Michelis and Florian Resatsch. Unlocking the interactive capabilities of large outdoor displays. *Information Display*, 23(3):24, 2007.
- [3] Martin Brynskov, Peter Dalsgaard, Tobias Ebsen, Jonas Fritsch, Kim Halskov, and Rune Nielsen. *Human-Computer Interaction – INTERACT 2009: 12th IFIP TC 13 International Conference, Uppsala, Sweden, August 24-28, 2009, Proceedings, Part I*, chapter Staging Urban Interactions with Media Façades, pages 154–167. Springer Berlin Heidelberg, 2009.
- [4] Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. It’s mine, don’t touch!: Interactions at a large multi-touch display in a city centre. In *CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1285–1294. ACM, 2008.
- [5] Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. Looking glass: A field study on noticing interactivity of a shop window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 297–306. ACM, 2012.
- [6] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. Proxemic interaction: Designing for a proximity and orientation-aware environment. In *TS '10 ACM International Conference on Interactive Tabletops and Surfaces*, pages 121–130. ACM, 2010.
- [7] E.T. Hall. *The Hidden Dimension*. Doubleday, 1966.

- [8] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. Proxemic interactions: The new ubicomp? In *Interactions*, pages 42–50. ACM, 2011.
- [9] Paul Marshall, Yvonne Rogers, and Nadia Pantidi. Using f-formations to analyse spatial patterns of interaction in physical environments. In *CSCW '11 Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 445–454. ACM, 2011.
- [10] Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. The proximity toolkit: Prototyping proxemic interactions in ubiquitous computing ecologies. In *UIST '11 Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 315–326. ACM, 2011.
- [11] Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. Rgb-d mapping: Using depth cameras for dense 3d modeling of indoor environments. In *The 12th International Symposium on Experimental Robotics*. ISER, 2010.
- [12] Wenjung Zeng. Microsoft kinect sensor and its effect. In *IEEE Multi-Media (Volume 19, Issue 2)*, pages 4–10. IEEE, 2012.
- [13] M.R. Andersen, T. Jensen, P. Lisouski, A.K. Mortensen, M.K. Hansen, T. Gregersen, and P. Ahrendt. Kinect depth sensor evaluation for computer vision applications, 2012.
- [14] Gilbert Beyer, Vincent Binder, Nina Jäger, and Andreas Butz. The puppeteer display: Attracting and actively shaping the audience with an interactive public banner display. In *DIS '14 Proceedings of the 2014 conference on Designing interactive systems*, pages 935–944. ACM, 2014.
- [15] Yannick Winters. A rule-based approach to detect groups using a depth sensing camera, 2014.
- [16] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. Cross-device interaction via micro-mobility and f-formations. In *UIST '12 Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 13–22. ACM, 2012.
- [17] Michal Kepski and Bogdan Kwolek. Detecting human falls with 3-axis accelerometer and depth sensor. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 770–773. IEEE, 2014.

- [18] Thomas Diewald. `dlibs_freenect`, a library for the programming environment processing. www.thomasdiewald.at/processing/libraries/dLibs_freenect.
- [19] Mario Martínez-Zarzuela, Miguel Pedraza-Hueso, Francisco Javier Díaz Pernas, David González Ortega, and Miriam Antón-Rodríguez. Indoor 3d video monitoring using multiple kinect depth-cameras. *CoRR*, abs/1403.2895, 2014.
- [20] Stephen J. McKenna, Sumer Jabri, Zoran Duric, Azriel Rosenfeld, and Harry Wechsler. Tracking groups of people. In *Computer Vision and Image Understanding (Volume 80, issue 1)*, pages 42–56. Elsevier, 2000.
- [21] Tian Gan, Yongkang Wong, Daqing Zhang, and Mohan S. Kankanhalli. Temporal encoded f-formation system for social interaction detection. In *Proceedings of the 21st ACM International Conference on Multimedia*, pages 937–946. ACM, 2013.
- [22] Jakub Dostal, Uta Hinrichs, Per Ola Kristensson, and Aaron Quigley. Spidereyes: Designing attention- and proximity-aware collaborative interfaces for wall-sized displays. In *IUI '14 Proceedings of the 19th international conference on Intelligent User Interfaces*, pages 143–152. ACM, 2014.
- [23] Miaosen Wang, Sebastian Boring, and Saul Greenberg. Proxemic peddler: A public advertising display that captures and preserves the attention of a passerby. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, pages 3:1–3:6. ACM, 2012.
- [24] Optitrack tutorial. <http://brouet.imag.fr/fberard/Docs/Optitrack>.
- [25] Optitrack faqs. <http://www.optitrack.com/support/faq/general.html>.
- [26] Thomas Seifried, Christian Rendl, Michael Haller, and Stacey D. Scott. Regional undo/redo techniques for large interactive surfaces. In *CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2855–2864. ACM, 2012.
- [27] Marti A. Hearst. Mixed-initiative interaction. In *IEEE Intelligent Systems and their Applications (Volume 14, Issue 5)*, pages 14–23. IEEE, 1999.
- [28] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. Research in the wild: Understanding ‘in the wild’

- approaches to design and development. In *DIS '12 Proceedings of the Designing Interactive Systems Conference*, pages 795–796. ACM, 2012.
- [29] Ross Cutler and Larry Davis. Look who’s talking: Speaker detection using video and audio correlation. In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on (Volume 3)*, pages 1589–1592. IEEE, 2000.
- [30] Chiew Seng Sean Tan, Johannes Schöning, Kris Luyten, and Karin Coninx. Informing intelligent user interfaces by inferring affective states from body postures in ubiquitous computing environments. In *IUI '13 Proceedings of the 2013 international conference on Intelligent user interfaces*, pages 235–246. ACM, 2013.

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:

Groupaware interfaces

Richting: **master in de informatica-Human-Computer Interaction**

Jaar: **2016**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Meerten, Bram

Datum: **15/06/2016**