

NEMO Science Live 2011

19 februari-13 maart / 22 april-08 mei 2011



Onderzoekers:

Theo Gevers, Albert Ali Salah, Hamdi Dibeklioglu en Sezer Karaoglu.

I. UvA-NEMO Database

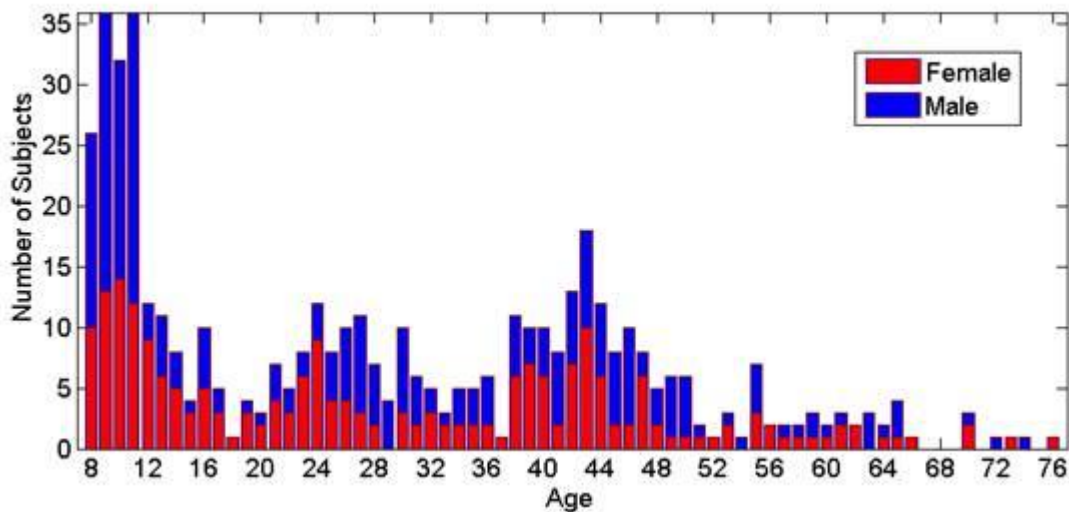
Een grote collectie van video's is opgenomen van gezichten met verschillende gelaatsuitdrukkingen. De video's zijn opgenomen met een kleurencamera en de Microsoft Kinect-sensor (dieptebeelden). Beide opnames zijn tegelijkertijd gemaakt op ongeveer 1.5 meter afstand van de deelnemers (zie Figuur 1). De videobeelden hebben een (hoge) resolutie van 1920x1080 pixels.



Figuur 1. Opname set-up.

De opnames van de deelnemers bestaan uit (1) gezichtsuitdrukkingen van de zeven basisemoties (neutraal, geluk, verdriet, woede, walging, angst en verrassing), (2) spontane lachsessies (tijdens het kijken naar grappige video's), (3) verschillende gezichts- en lipbewegingen tijdens het lezen van een lijst van woorden, en (4) verschillende hoofdbewegingen. Verder hebben de deelnemers video opnames geanalyseerd van andere deelnemers (zoals hoe blij, boos, trots, enz. andere deelnemers eruitzien). Deze menselijke annotaties worden gebruikt om te onderzoeken hoe goed mensen in staat zijn om het gedrag van anderen te beoordelen. Verder wordt het gebruikt om te toetsen hoe nauwkeurig de door ons ontwikkelde computersoftware is in vergelijking met de menselijke waarneming.

Er waren 481 deelnemers (waarvan 221 vrouwen en 260 mannen). De leeftijd van de deelnemers varieert van 8 tot 76 jaar. Leeftijd en geslacht van de deelnemers zijn weergegeven in Figuur 2.



Figuur 2. Distributie van leeftijden en geslacht van de deelnemers.

II. Onderzoek gedaan op de UvA-NEMO Database

II.A Het schatten van leeftijd

Het schatten van leeftijd blijkt vaak een moeilijke taak voor ons te zijn. Soms schat je iemand ouder dan hij/zij is. Of jonger. Hoe komt dat nou? Welke gezichtskenmerken bepalen hoe oud iemand eruitziet?

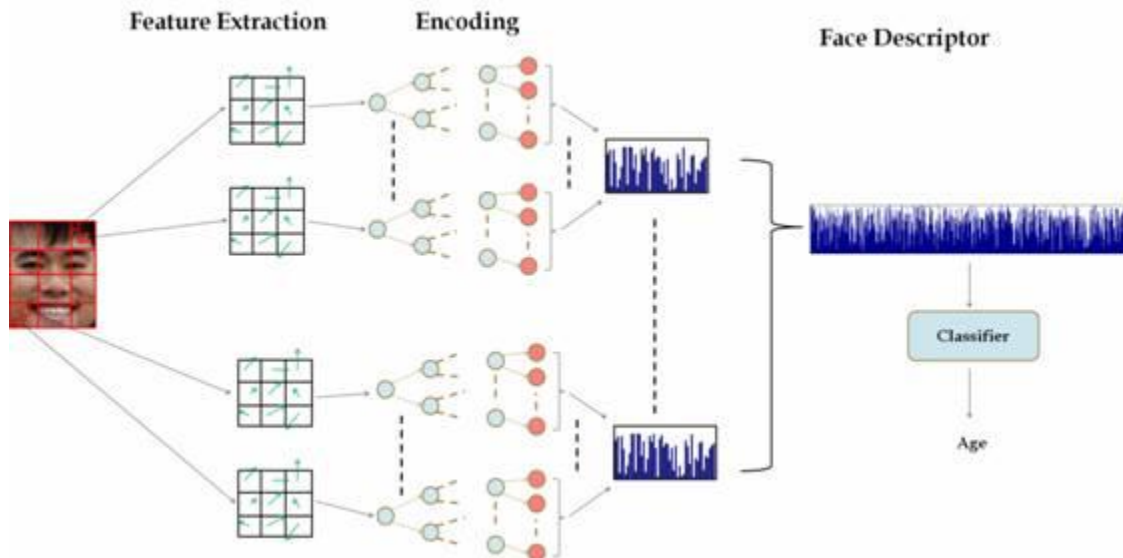
In de vakliteratuur is er onderzoek gedaan hoe goed mensen zijn in het schatten van de leeftijd van anderen. Hierbij worden op een scherm plaatjes van gezichten getoond waarvan de deelnemers (observanten) de leeftijd moeten bepalen. Onder normale beeldopname-omstandigheden, schat de mens de leeftijd van anderen met een nauwkeurigheid van plus of min zeven jaar. Dat betekent dat we er gemiddeld zeven jaar naast zitten! Het schatten van leeftijd wordt natuurlijk moeilijker als de beelden (van gezichten) zijn opgenomen onder slechte belichting, compressie of het dragen van baard of bril. Ook gezichtsexpressies hebben invloed op het schatten van leeftijd.

In dit onderzoek hebben we het volgende onderzocht:

1. Is het mogelijk om computersoftware te ontwikkelen die de leeftijd van mensen automatisch kan schatten?
2. Wat is de invloed van een expressie (zie je er met een glimlach er jonger uit)?
3. Wat is de invloed van de kleur van de lichtbron (zie je er onder een blauwere lamp ouder uit)?
4. Welke gezichtskenmerken (zoals voorhoofd (rimpels), rond de ogen (kraaienpoten) of wangen) bepalen je leeftijd het meest?

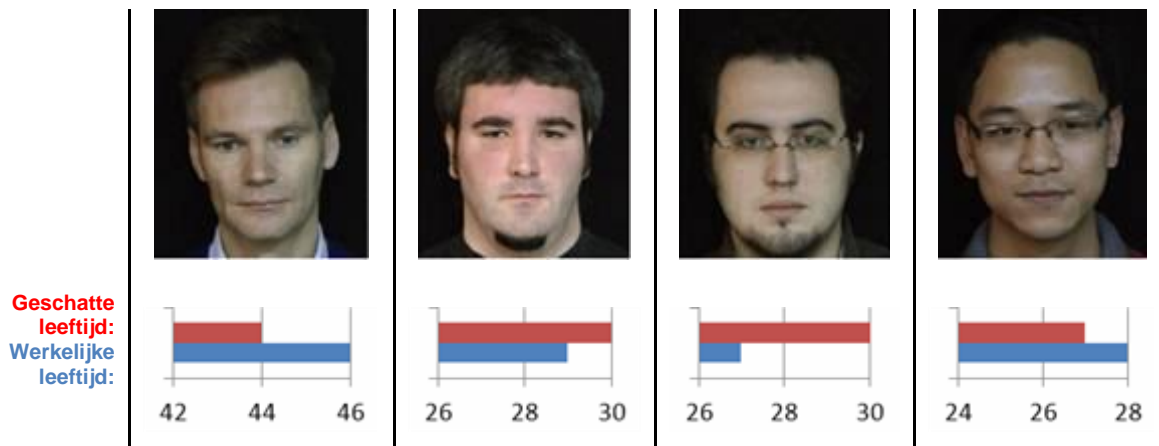
Voor het automatisch bepalen van de leeftijd van mensen hebben we eerst gekeken welke gebieden in het gezicht bepalend zijn voor het berekenen van leeftijd, zoals voorhoofd, gebieden rondom de ogen en wangen. Er is software ontwikkeld die de verschillende kenmerken berekent, zoals hoeveelheid rimpels (textuur), patronen (kraaienpoten), huid (ouderdom vlekjes) etc. Statistische patroonherkenning is uitgevoerd om te bepalen in hoeverre elk kenmerk bijdraagt

aan de leeftijdsschatting. Het algoritme werkt op basis van een gecodeerd gezichtsbeeld dat opgedeeld is in een raster van vierkante gebieden. Een codehistogram wordt berekend voor elk gebied. Hierdoor worden lokale beschrijvingen gegenereerd voor elk gebied. Deze worden dan samengevoegd tot een globale beschrijving van het hele gezicht. Een grafische weergave van het algoritme voor het bepalen van leeftijd zie je in Figuur 3. De software werkt in real-time.



Figuur 3. Grafische weergave van het leeftijdschattingsalgoritme.

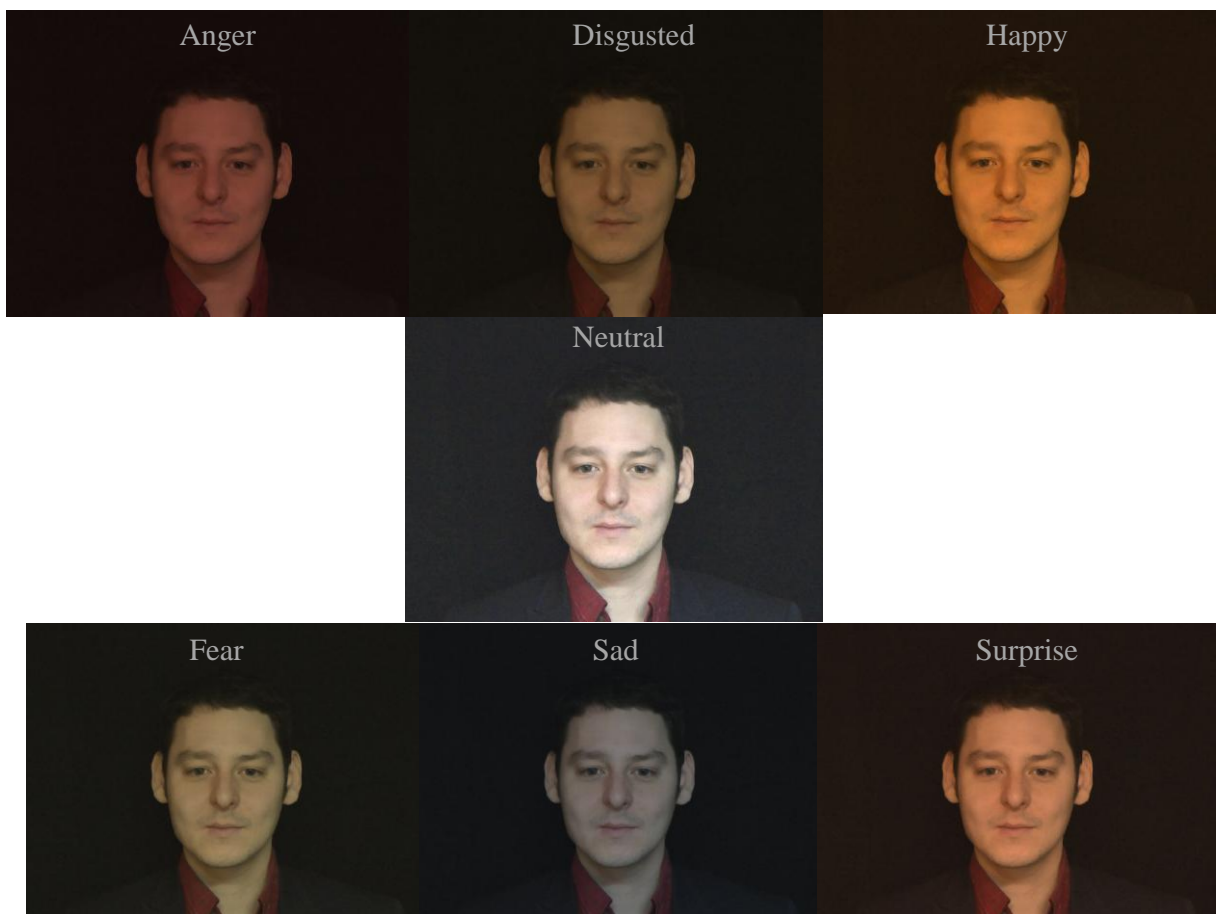
Experimenten op basis van de UvA-NEMO database met bovenstaande algoritme tonen aan dat de software in staat is om de leeftijd van proefpersonen te bepalen met een gemiddelde fout van 7,07 jaar. Figuur 4 toont de echte en de door de computer geschatte leeftijd van verschillende deelnemers. Leg je hand over de echte en (computer-)geschatte leeftijd en probeer zelf de leeftijd te schatten. Hoever zit je ervan af? Ben je beter dan de computer?



Figuur 4. De werkelijke en de (computer-)geschatte leeftijd van verschillende deelnemers.

Voor het bepalen van de invloed van expressies hebben we gebruik gemaakt van gezichten met verschillende emotionele gelaatsuitdrukkingen, zoals blijdschap, triestheid, verbazing etc. De verschillende expressies hebben invloed op de inschatting van leeftijd. Daarom hebben we software ontwikkeld die rekening houdt met de waargenomen gezichtsexpressie. De software stuurt in feite de leeftijdschatting bij, afhankelijk van de gelaatsuitdrukking. Als je triest kijkt, zal de software andere gebieden in je gezicht gebruiken dan als je verbaasd kijkt.

Voor het bepalen van leeftijd onder verschillende lichtbronnen is er gebruik gemaakt van beelden (gezichten) onder verschillend gekleurde lichtbronnen. De kleuren die de lichtbronnen kunnen aannemen komen overeen met de kleuren die men associeert met een emotie. Een kleur die overeenkomt met een positieve emotie (blijdschap) kan bijvoorbeeld een positievere impressie geven (zie figuur 5). Leg in onderstaande afbeeldingen je hand op de tekst en bepaal zelf de emotie die de lichtbron oproept.



Figuur 5. Het centrale beeld is opgenomen onder een witte lichtbron (D65). Vervolgens zijn de beelden gereconstrueerd onder verschillend gekleurde lichtbronnen die overeenkomen met de basisemoties.

In eerste instantie hebben we onderzoek gedaan naar de invloed van een positieve (happy) en een negatieve (sad) lichtbron. Deze twee lichtbronnen komen overeen met de kleuren oranjeachtig (happy-positief, $L^* = 67.5$, $a^* = 23.5$, $b^* = 50$) en grijsblauw (sad-negatief, $L^* = 56$, $a^* = -1.5$, $b^* = -4$). Verder is er gekeken of men er met een glimlach jonger uitziet dan met een neutraal gezicht.

De conclusie is dat (gemiddeld) de geschatte leeftijd boven de echte leeftijd uitkomt. Onder de leeftijd van 40 jaar, hebben lachende gezichten een hogere geschatte leeftijd dan neutrale gezichten, en vice versa voor leeftijden boven 40 jaar. Dit effect wordt verstrekt door de positief gekleurde illuminant. De negatief gekleurde lichtbron heeft nauwelijks effect.

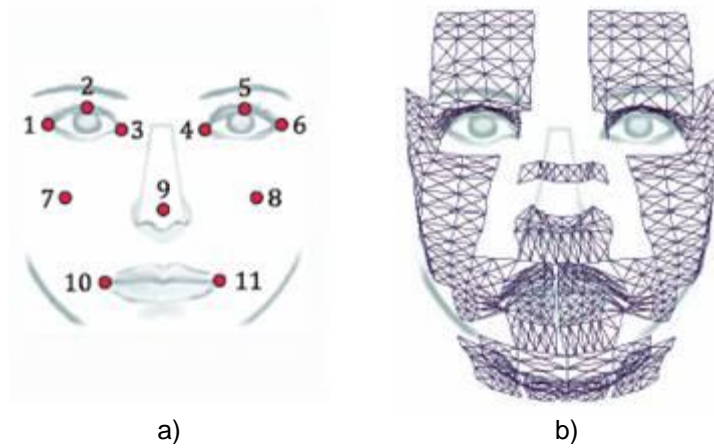
De voornaamste conclusies van dit onderzoek zijn:

1. Computersoftware is ontwikkeld voor het automatisch berekenen van leeftijd van mensen met een gemiddelde fout van 6 jaar. In het algemeen, hebben mensen een lagere nauwkeurigheid (rond de 7 jaar)!
2. Expressies van mensen hebben invloed op het schatten van leeftijd.
3. De kleur van de lichtbron heeft invloed op leeftijdschatting.

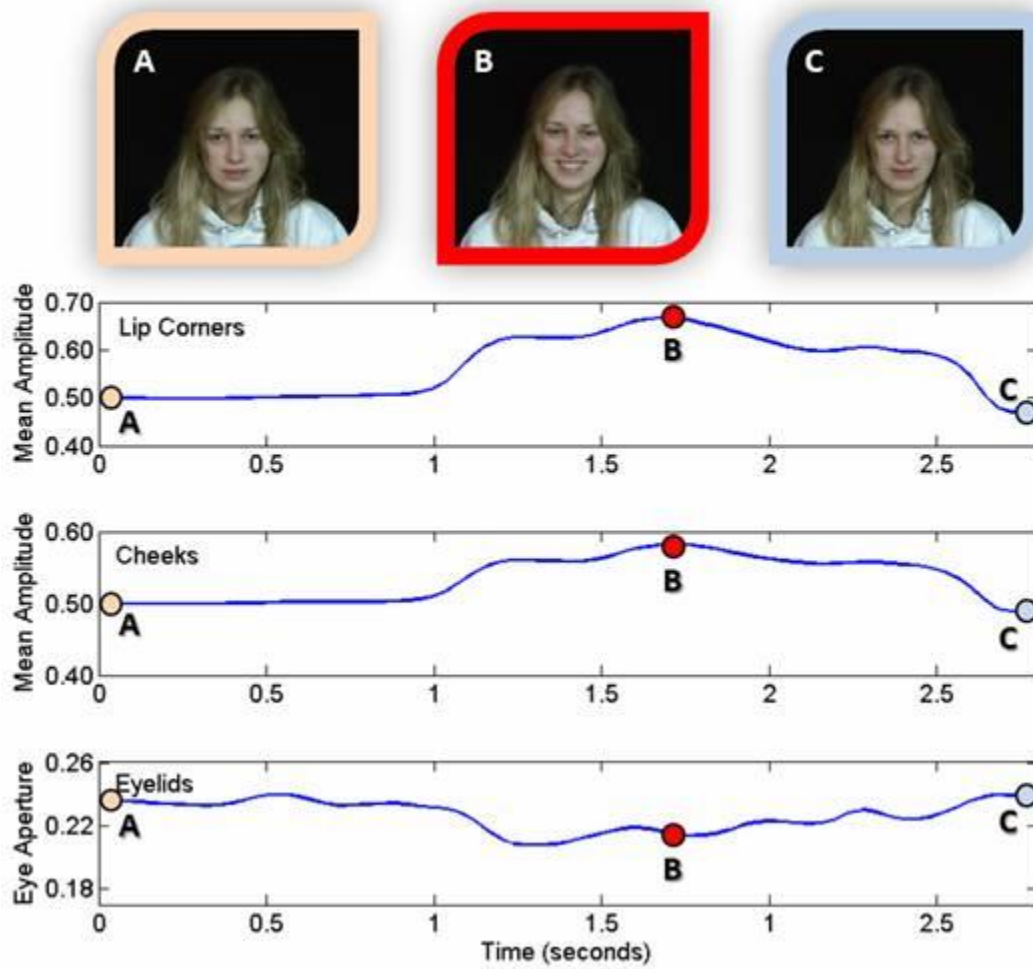
II.B Onderscheid tussen een echte en een nepglimlach

De (glim)lach is een onmisbaar element van onze non-verbale sociale interactie. Zo af en toe lachen we ons rot. Of toch niet? Het is voor mensen vaak lastig te bepalen of iemand echt (glim)lacht of nep. Aangezien er meer dan twaalf verschillende vormen zijn van een (glim)lach (cynisch, trots, verlegen etc.) hebben we in dit onderzoek de positieve "glimlachsvorm" als uitgangspunt genomen: de "vreugde"-lach. Lachen om iets wat leuk is!

We hebben onderzocht of er gelaatskenmerken bestaan voor het herkennen van een echte en een nepglimlach. Hieronder vallen: de ooghoeken, de oogleden, de wangen, de neus en mondhoeken, zie ook Figuur 6.a. Aangezien we video's tot onze beschikken hebben (beweging), is er ook naar de dynamiek van deze gezichtskenmerken gekeken, zoals snelheid, acceleratie, symmetrie en bewegingspatroon. We hebben software ontwikkeld die automatisch gezichtskenmerken herkent in video's en deze volgt met behulp van een gezichtmodel (zie Figuur 6.b). Hierdoor kunnen we automatisch, accuraat en objectief de bewegingen van de kenmerken berekenen van de gezichten van de testpersonen. Bijvoorbeeld in Figuur 7 worden de bewegingen (intensiteit) van de mondhoeken, wangen en oogleden over tijd (1-3 seconden) van een echte glimlach weergegeven. Bij punt A wordt de glimlach ingezet en bij punt B is de maximale expressie. C is het punt waar de lach eindigt. Door middel van statistische patroonherkenning worden de belangrijke verschillen tussen een echte en nepglimlach geleerd.



Figuur 6. (a) Gebruikte gelaatskenmerken (b) het 3D computer model



Figuur 7. Het verloop van de mondhoeken, wangen en oogleden tijdens een spontane glimlach.

Om de software te testen is er gebruik gemaakt van de UvA-NEMO Database met daarin in totaal 400 deelnemers met 597 echte en 643 nep-lachuitingen (1240 video's in totaal) met leeftijden tussen 8 - 76.

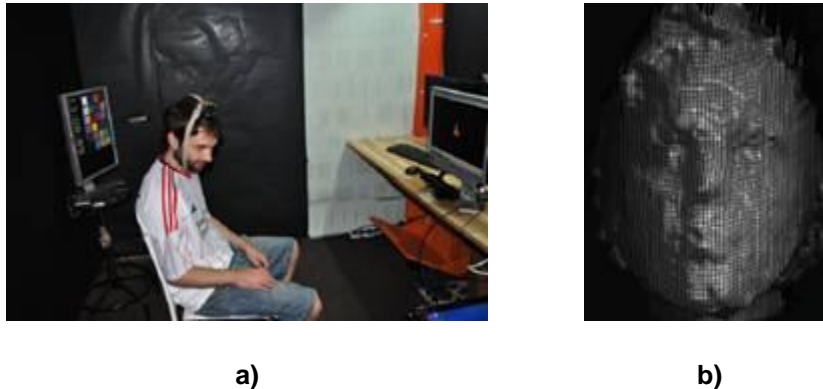
Op basis van de dynamiek van de gezichtskenmerken die berekend wordt door de computer en statistische patroonherkenning zijn er de volgende conclusies:

1. Het meest onderscheidende kenmerk voor het onderscheiden van een echte en neplach is de dynamiek van de oogleden. Als men nep-lacht, zullen de oogleden (in het algemeen) niet aangestuurd worden. Als men echt lacht wel. Dit komt door de samenhang van spieren in het gezicht.
2. Er is een verschil in intensiteit en expressiekracht voor verschillende leeftijden.
3. De computersoftware classificeert 86,21% van de glimlachen correct. In het algemeen hebben mensen een lagere nauwkeurigheid: 70%!

II.C Het automatisch schatten van hoofdbewegingen

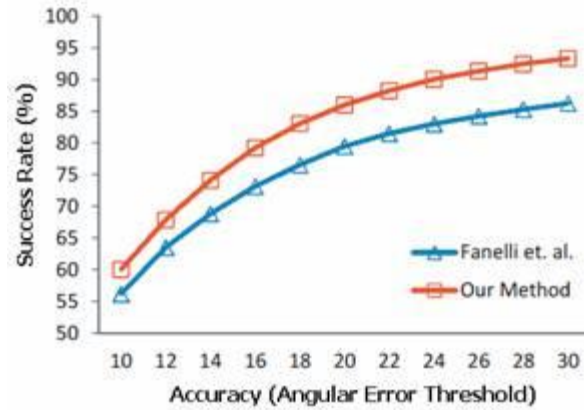
Door middel van hoofdbewegingen kunnen mensen verschillende bedoelingen uiten, zoals verwarring, instemming, of verschil van mening. In feite kan de mens vrij eenvoudig de intentie van anderen aflezen aan hoofdbewegingen. In dit experiment gaan we na hoe de computer de bedoeling van de mens kan bepalen door te leren van voorbeelden. Als de computer in staat is om dit automatisch te doen dan zijn er een aantal applicaties voorhanden, zoals mens-machine interactie. Ook is er onderzocht of we automatisch kunnen bepalen waar de mens naar kijkt (blikrichting) en het herkennen waar iemand mee bezig is (activiteitsherkenning). De moeilijkheid van het automatisch herkennen schuilt in de veranderingen van belichting, identiteit (gezichtshaar, haar, bril, enz.), en gezichtsuitdrukkingen. Op basis van gedegen onderzoek, hebben we intelligente algoritmes ontwikkeld die verschillende hoofdbewegingen automatisch en real-time kunnen bepalen.

Om te kunnen meten hoe goed de algoritmes werken, is gebruik gemaakt van de 3D-versnellingsmeter en 3D-gyroscoop sensoren op de Iphone 4 (zie Figuur 8) .



Figuur 8. (a) Hoofdpose-opname, (b) het dieptebeeld van de Kinect.

De dataset bevat 48 proefpersonen (30 mannen en 18 vrouwen). De nauwkeurigheid van onze methode is weergegeven in Figuur 9.



Figuur 8. Nauwkeurigheid van onze methode.

De conclusie van dit onderzoek is:

1. Computersoftware is ontwikkeld voor het automatisch bepalen van hoofdbewegingen.
2. De software is nauwkeurig met een succesratio van 90%. Dit is een verbetering van 8.64% ten opzichte van de best bestaande methode.