

# De humanoid: van statisch naar dynamisch

*Ooit al een humanoid van dichtbij gezien? Had je er geen afkeer van? Sterk!*

Robots worden vandaag de dag meer en meer gebruikt om te communiceren met de mens. Om de kwaliteit van de communicatie te verhogen zijn expressieve features voorhanden die goed moeten worden aangestuurd.

## Probleemschetsing

Het is niet eenvoudig om een realistische robot te ontwerpen, vooral in het geval van een mensreplica. Kijk maar naar de wassen beelden van Madame Tussauds (fig. 1). Dit zijn zeer realistische poppen die statisch worden opgesteld. Jaren onderzoek hebben ertoe geleid om op de mens lijkende robots, zogenaamde humanoids, te ontwerpen die gelijken op de mens, die emoties en/of kunnen nabootsen.



Figuur 1: Een wassen beeld van de wereldberoemde George Clooney

Gewoon een humanoid bouwen, is niet genoeg. Een humanoid kan enkel overtuigend overkomen als deze zich ook realistisch gedraagt. Vandaag de dag is dit ook de grootste uitdaging, de opdracht van de robotbouwers is pas voltooid indien ze er in slagen een robot te bouwen en aan te sturen die geen afkeergevoel opwekt bij de mens. Dit fenomeen staat bekend

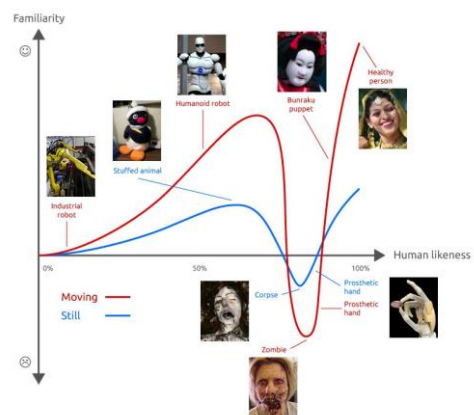
als het Uncanny Valley effect en wordt het meest ervaren bij robotreplica's van dieren en mensen. Kijk maar naar een slecht opgezette kat en de dinosaurus op figuur 2.



Figuur 2: Een slecht opgezette kat en een dinosaurus. De kat wekt afkeer op door zijn afwijkende snoet, terwijl de dinosaurus geen afkeer wekt.

De kat wekt enige afkeer op terwijl dit niet of minder het geval is bij de bewegende dinosaurus. De opgezette kat kunnen we vergelijken met een echte levende kat, terwijl dit niet het geval is bij de dinosaurus. Dit effect is nog erger bij humanoids, gezien afwijkingen in hun bewegingen en gedrag meteen worden opgemerkt. Deze afwijkingen spelen een zeer belangrijke rol, zoals aangegeven op de grafiek in figuur 3.

Op deze grafiek kan dit effect duidelijk worden gezien, zo wekt een zombie nog meer afkeer op dan een opgebaard lijk. Hetzelfde fenomeen is ook zichtbaar wanneer mensen flauwvallen, bij zichtbare breuken of zich ongemakkelijk voelen bij onnatuurlijke houdingen en bewegingen.



Figuur 3: Op deze grafiek wordt het zogenaamde Uncanny Valley effect afgebeeld. Dit fenomeen wordt versterkt bij beweging en gedrag.

## Klassieke aanpak

Om dit afweermechanisme te ontwijken trachten de ontwikkelaars hun humanoids zo goed mogelijk aan te sturen en dit is net waar de voorgestelde thesis om draait.

Vandaag de dag gebeurt zo een calibratie vaak manueel, zowel voor spraak als emotie. Elke actuator wordt dan iteratief ingesteld om een realistisch en niet afkeerwekkend resultaat te bekomen. Het effect van een niet-ideale calibratie kan gezien worden in figuur 4.



Figuur 4: Een niet-ideale calibratie van een humanoid die afkeer opwekt. De afgebeelde emotie is angst.

Een manuele calibratie duurt lang en is zeer kostelijk. Bij elke aanpassing moet deze dan ook opnieuw gebeuren. Daarom wordt er gereikt naar andere oplossingen. Deze alternatieven maken gebruik van complexe computermodellen. Hiermee trachten ze voor elke emotie of mondvorm de configuratie voor elke actuator te bepalen. Deze methodes hangen echter af van de kwaliteit van het gebruikte computermodel. Deze modellen geven echter maar een benaderend resultaat waardoor een verdere manuele calibratie noodzakelijk is.

Om dit probleem op te lossen, wordt er in deze thesis een algoritme voorgesteld dat gebruik maakt van de ware humanoid. Dit in plaats van een computermodel.

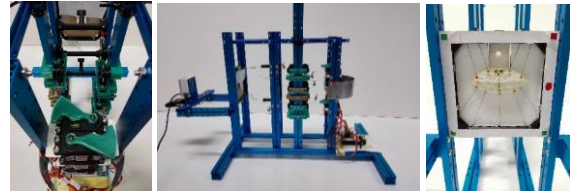


Figuur 5: Een aantal referentiemondvormen die worden gebruikt in de nieuwe voorgestelde calibratiemethode

## Een nieuwe aanpak

De techniek die wordt voorgesteld werd getest op spraakcalibratie. Een aantal mondvormen worden weergegeven in figuur 5.

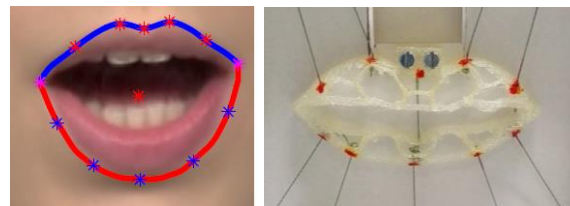
Deze mondvormen tracht men na te bootsen met de humanoid, die gecalibreerd dient te worden. Om de werking aan te tonen werd een opstelling opgebouwd (zie figuur 6).



Figuur 6: De opstelling dat werd gebruikt om de nieuwe calibratiemethode te testen

Deze opstelling omvat een modulair actuatie mechanisme dat een 3D-geprinte mond kan aansturen door touwtjes aan te trekken via een set servomotoren. Op de mond worden een aantal punten aangeduid. Hiermee wordt de complete vorm op een compacte manier weergegeven.

Dezelfde punten worden aangeduid op de referentie figuren uit figuur 5, op een goed gedefinieerde, automatische manier (zie fig. 6).

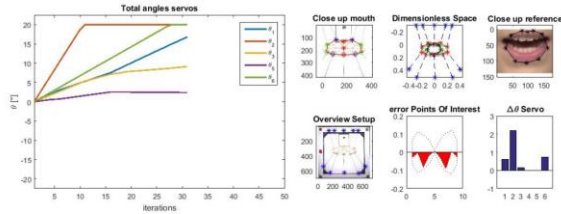


Figuur 7: De referentiepunten op een referentiemond (links) en diezelfde punten die worden afgebeeld op de echte mond (rechts).

Nu al deze punten gekend zijn, op de humanoid en de referentiemonden, kan het algoritme in werking treden.

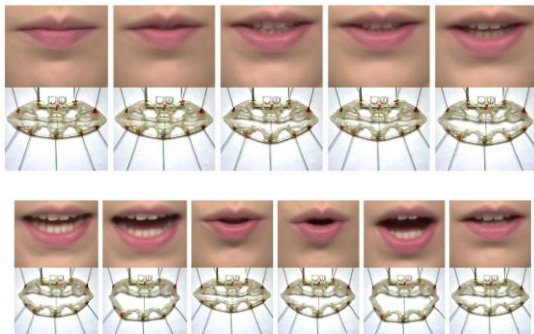
Door wiskundige transformaties is het mogelijk om beide monden te vergelijken en de verschillen aan te duiden (door een set van

pijlen). Die pijlen geven dan aan of de touwen die de mond aansturen meer moeten trekken, lossen of goed ingesteld zijn. Deze aanpak wordt iteratief uitgevoerd, tot de methode convergeert. Dit wil zeggen dat beide monden zo goed mogelijk op elkaar gelijken. Het iteratief programma dat werd gebruikt wordt afgebeeld in figuur 8.

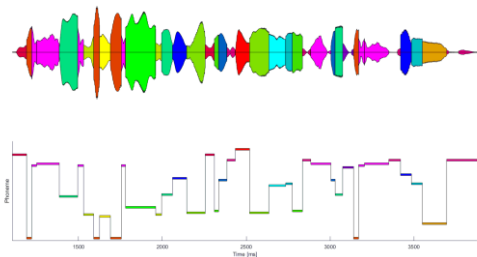


Figuur 8: Een afbeelding van de calibratiesoftware dat werd ontworpen

De calibratie duurde een uur (voor een totaal van 52 afbeeldingen), wat aanzienlijk korter is dan een manuele calibratie. Een paar van de resultaten van deze calibraties staan afgebeeld in figuur 9.



Figuur 9: Een paar calibratieresultaten dat werden bereikt met de zelfgeschreven calibratiemethode

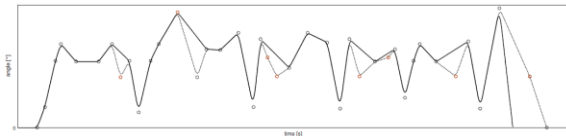


Figuur 10: Een zin wordt in verschillende onderdelen opgesplitst, waarbij elk onderdeel bij een welbepaalde mondvorm hoort.

## Resultaten

Om de calibratieresultaten en de kracht van de methode mooi in kaart te brengen, is een text-to-robot pakket ontwikkeld dat ons in staat stelt om ingesproken zinnen vanop de computer synchroon uit te spreken. Hiervoor werden de ingesproken zinnen in stukken geknipt zoals te zien is in figuur 10.

Deze stukken krijgen een bepaalde mondvorm toegekend die op het juiste moment worden uitgevoerd. Om een vloeiende spraak te garanderen wordt er geïnterpoleerd tussen de configuraties van elke mondvorm die nodig zijn om een complete zin uit te spreken, zoals in figuur 11. Hier wordt een bepaalde motorconfiguratie afgebeeld dat in de tijd verandert om de zin uit figuur 10 uit te spreken.



Figuur 11: Een voorbeeld van hoe het text-to-robot motorconfiguraties interpoleert om vloeiende spraak te garanderen.

## Conclusie

De voorgestelde techniek werkt en heeft veel potentieel om in toekomstige humanoid calibratie gebruikt te worden. De aanpak werd enkel getest voor spraakcalibratie maar kan zeker worden toegepast voor emotiecalibratie.

Mogelijke verdere informatie en filmpjes van het resultaat kunnen verkregen worden bij de auteur op het volgende e-mail adres:

[gavdevel@vub.ac.be](mailto:gavdevel@vub.ac.be)