

## Robotica

Robots en avatars die ons glazig grijnzend aankijken? Facebook wil daar een eind aan maken door machines te trainen op gezichtsuitdrukkingen.

Door **Laurens Verhagen**

# Een menselijk gezicht

**H**et was een opvallende noviteit, eerder deze maand, toen de jury van de Taylor Wessing Portretprijs het portret van ene Erica nomineerde. Erica is een jonge vrouw met een wat afwezige blik. Nee, Erica lijkt op een vrouw, maar ze is een robot, gecreëerd door Hiroshi Ishiguro, professor robotica aan de universiteit van Osaka. Maar inderdaad: Erica lijkt griezelig veel op een mens.

Vandaar dat haar portret mag meedoen aan de prijs, die eigenlijk voor menselijke portretten is bedoeld. Wie langer naar robots als Erica kijkt, krijgt een ongemakkelijk gevoel. Het klopt nèt niet. Levensecht en tegelijk levenloos. Dat ongemakkelijke gebied staat bekend als de *uncanny valley*. Het is het gebied waar afkeer het wint van affectie, al zal Erica's maker daar anders over denken.

**Facebook probeert** het ongemakkelijke gevoel weg te nemen door

robots natuurlijker te laten reageren. Onderzoekers van Facebooks AI-lab hebben een systeem gemaakt met door algoritmen gedreven reacties. Die algoritmen zijn getraind door de machine naar honderden Skype-video's te laten kijken. Door heel exact te bekijken hoe de ene gesprekspartner op de ander reageert met minime veranderingen in de gezichtsuitdrukking, kan de robot leren, zo was het idee. In het trainingsmodel werd het gezicht opgedeeld in 68 belangrijke stukjes. In een normale conversatie reageren mensen op elkaar via knikjes, kleine veranderingen bij de ogen of mondbewegingen. Al deze veranderingen geven de spreker het prettige gevoel – als het goed is – dat er naar hem wordt geluisterd. Concreet: als iemand een grap maakt, is het de bedoeling dat de ander lacht of op z'n minst zijn mond tot een glimlach vormt.

De trainingen hadden succes, claimen de onderzoekers. Een test-

panel bekeek de reacties van de robot op een mens en andersom. Het panel betitelde de animaties van de robot als ongeveer even realistisch als die van een mens.

Theo Gevers, hoogleraar computervisie aan de Universiteit van Amsterdam en gespecialiseerd in automatische gezichtsanalyse, reageert terughoudend op de claims. 'Het klinkt indrukwekkend: honderden Skype-video's. Maar voor het trainen van zogenoemde recurrente neurale netwerken voor het herkennen van dynamische expressies is die hoeveelheid data toch heel beperkt. Daarom hebben ze gekozen voor een ander soort neurale netwerken, die getraind zijn met kleine stukjes video's. Het nadeel daarvan is dat je harde overgangen ziet tussen de verschillende uitdrukkingen.'

**Deze weinig vloeiende** overgangen én beperkte data maken dat Gevers 'niet zo heel veel vertrouwen' in het onderzoek van Facebook heeft. Daarnaast benadrukt hij dat de tests plaatsvinden onder laboratoriumomstandigheden. 'Ze hebben niet voor niets Skype-sessies gebruikt, met de juiste belichting en een gezicht netjes frontaal in beeld. In het café gaat dat niet werken, dit is echt een eerste stapje.'

Zelf doet Gevers ook onderzoek naar gezichtsherkenning en expressies via zelflerende netwerken. De ontwikkelingen gaan hard: 'Een paar jaar geleden waren we al heel blij als een machine zes basisemoties kon onderscheiden op basis van een video die hem werd getoond.' Daarna baarde hij opzien door met zijn software de *Mona Lisa* te analyseren. De conclusie destijds: ze kijkt voor 83 procent vrolijk.

Het onderzoek naar het analyseren van emoties via video's loopt door, maar ondertussen werkt Gevers net als een aantal andere bedrijven aan software die het mogelijk maakt videobeelden met elkaar te mengen. Op die manier kun je bijvoorbeeld Trump iets laten zeggen wat hij nooit heeft gezegd, inclusief de juiste gezichtsuitdrukking. Daar kun je enge dingen mee doen, geeft Gevers toe, maar voor hem is de ultieme uitdaging om de software zo goed te maken dat een 'nep-Trump' er geheel natuurlijk uitziet op video. 'De leek ziet het verschil niet. Voor ons is het de heilige graal: als het lukt het gezicht zo nauwkeurig in kaart brengen dat je er zonder zichtbare overgangen iedere emotie op kunt projecteren.' Het is duidelijk: laat het lezen van gezichten maar aan de computer over. ●

**Je blijft harde overgangen zien tussen de uitdrukkingen**