

University of Groningen

## Strategies for motion detection and texture recognition in artificial visual systems

van Deemter, Jacobus Henk

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

1994

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

van Deemter, J. H. (1994). *Strategies for motion detection and texture recognition in artificial visual systems*. s.n.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Samenvatting

Dit proefschrift beschrijft onderzoek aan kunstmatige visuele systemen. Op het eerste gezicht lijkt het construeren van een dergelijk systeem misschien een triviaal probleem. Als wij naar een film kijken, zien we in een oogopslag wat zich daar afspeelt. Het is meteen duidelijk welke fragmenten een samenhangend geheel vormen en welke objecten waarheen bewegen. Aan deze snelle interpretatie van beelden gaat echter heel wat 'reken'werk vooraf. Licht van onze omgeving wordt in ons oog door een lens op het netvlies, dat bezaaid is met fotoreceptoren, geprojecteerd. Deze fotoreceptoren zetten de lichtinformatie om in elektrische informatie. In onze ogen en hersenen voeren grote netwerken van neuronen vervolgens complexe operaties uit, die uiteindelijk leiden tot een bruikbare interpretatie van een beeld. Het mag hieruit duidelijk zijn dat het bouwen van een kunstmatig visueel systeem verder gaat dan slechts het registreren van beelden met camera en videorecorder. Kunstmatige visuele systemen worden bijvoorbeeld gebruikt in robots en navigatiesystemen maar ook toegepast in de medische diagnostiek en tal van industriële processen.

Bij de meeste kunstmatige visuele systemen worden eerst met een camera beelden opgenomen. Deze beelden zijn opgedeeld in pixels, die de intensiteiten (en evt. kleuren) van kleine stukjes beeld representeren met een getal. Met deze pixels worden dan berekeningen uitgevoerd volgens bepaalde beeldverwerkingstechnieken om aspecten (bijv. randen) in beelden te benadrukken. Uiteindelijk moet dit leiden tot een zekere interpretatie van de beelden. Bij het doen van onderzoek naar dergelijke technieken wordt er vaak met een schuin oog naar de natuur gekeken. Immers, in miljoenen jaren van evolutie hebben visuele systemen van mensen en veel dieren zich ontwikkeld tot systemen die verreweg superieur zijn aan elk tot dusver ontwikkeld kunstmatig visueel systeem. Daarom ligt het voor de hand om te onderzoeken hoe de visuele systemen van mens en dier zijn georganiseerd, daarvan te leren en de principes te gebruiken in kunstmatige visuele systemen. Dit is een methode die ik ook in dit proefschrift nastreef. Verder maak ik gebruik van kunstmatige neurale netwerken. Dit zijn computermodellen die zijn opgebouwd uit elementen, kunstmatige neuronen, die via verbindingen volgens een bepaalde architectuur informatie aan elkaar doorgeven. Een kunstmatig neuron is een sterk vereenvoudigd model van een echt neuron: het ontvangt stroompjes van andere (kunstmatige) neuronen via verbindingen, telt die op, transformeert de uitkomst volgens een overdrachtsfunctie en stuurt het resultaat via verbindingen naar andere kunstmatige neuronen. Kunstmatige neurale netwerken hebben met natuurlijke neurale netwerken aantrekkelijke rekenkundige eigenschappen gemeen, die met conventionele technieken niet kunnen worden bereikt. Een illustratie

hiervan is het kunnen leren van voorbeelden door volgens bepaalde leerregels de sterkten van verbindingen tussen de kunstmatige neuronen aan te passen.

In dit proefschrift worden twee aspecten van kunstmatige visuele systemen behandeld, nl. bewegingsdetectie en textuurclassificatie. In beide aspecten wordt gebruik gemaakt van kunstmatige neurale netwerken.

## **Bewegingsdetectie**

Bij bewegingsdetectie in kunstmatige visuele systemen is het de bedoeling om uit momentopnamen van een bewegende scène te bepalen hoe voorwerpen hebben bewogen. Uit psychofysisch onderzoek bestaan er aanwijzingen dat er in het menselijk visueel systeem bewegingsdetectie plaats vindt volgens het zgn. "token matching" principe: Op verschillende tijdstippen worden uit beelden bepaalde elementen geëxtraheerd en vervolgens worden de elementen uit opeenvolgende beelden op een of andere manier bij elkaar gezocht. Dit is een techniek die ook in het eerste deel van dit proefschrift, "Motion Detection", wordt gebruikt. In dit deel worden slechts twee momentopnamen genomen en de elementen uit het beeld worden gekarakteriseerd door slechts drie kenmerken, nl. plaats, oriëntatie en grootte. Verder worden in het onderzoek alleen de 2-dimensionale bewegingstypes translatie, rotatie, schaling en combinaties van deze 3 bewegingstypes beschouwd. Om uit te rekenen hoe er bewogen is tussen twee momentopnamen, wordt een statistische correlatie techniek (SCT) ontwikkeld. Deze techniek maakt gebruik van de groepeigenschappen van de drie kenmerken van de elementen in de momentopnamen. Er kan worden afgeleid dat uitgaande van twee momentopnamen, voor en na de beweging, een willekeurige enkelvoudige of gecombineerde beweging van translaties, rotaties en schalingen kan worden herleid tot een beweging met een enkel centrum (m.u.v. alleen translaties, omdat daarbij geen sprake is van een centrum). Met de SCT kunnen uit de momentopnamen alle benodigde bewegingsparameters worden bepaald, zodat de elementen uit de eerste momentopname correct kunnen worden afgebeeld op de elementen uit de tweede momentopname.

Vervolgens werd de SCT geïmplementeerd in twee kunstmatige neurale netwerken. In deze netwerken bestond de mogelijkheid tot interactie tussen de kenmerken van de elementen, wat in de SCT veel minder goed mogelijk was. Verder konden in de netwerken relatief gemakkelijk randvoorwaarden worden ingebouwd, zoals de één-op-één randvoorwaarde, die ervoor zorgt dat elk element uit de eerste momentopname precies bij één element uit de tweede momentopname mag horen en vice versa. Vergelijking van de prestaties van de SCT en de neurale netwerken bij enigszins verstoorde elementen (bijv. door ruis) liet duidelijk zien dat de netwerken beduidend beter presteerden.

## **Textuur**

Voor het  
van de ach  
ligt aan h  
dit proef  
textuurcla  
kunstmati  
is een cor  
pixel, maa  
oplossing  
signaal, i  
golflengte  
welke ver  
verkregen  
verhoudin  
Gabor fun  
de positie  
benaderin  
visuele sy  
positie een  
kunnen w  
twee textu  
functie, d  
geval is v

Voor v  
netwerken  
geconclud  
ondersche  
texturen l  
Gabor res  
duidelijk  
gemaakt,  
van genoe  
gebruikt l  
uitstek ge  
van belan

## Textuurclassificatie

Voor het interpreteren van beelden is het van groot belang dat voorwerpen van elkaar en van de achtergrond kunnen worden onderscheiden. Omdat dit een taak is die ten grondslag ligt aan hogere processen in (kunstmatige) visuele systemen, wordt in het tweede deel van dit proefschrift, "Pre-processing", ingegaan op randdetectie en wordt met name textuurclassificatie onderzocht. Bij textuurclassificatie is het de bedoeling dat een kunstmatig visueel systeem onderscheid maakt tussen verschillende texturen in beelden. Dit is een complex probleem, omdat een textuur niet kan worden gekarakteriseerd door één pixel, maar door de samenhang tussen een tevoren onbekend groot aantal pixels. Voor een oplossing voor dit probleem maakte ik gebruik van de Fourier analyse, die stelt dat elk signaal, in dit geval een beeld, kan worden opgebouwd uit golven van verschillende golflengten, richtingen en uitwijkingen. Door op elke positie van een textuur te bepalen in welke verhoudingen verschillende golven aanwezig zijn, kan belangrijke informatie worden verkregen over de karakteristieken van een textuur. Voor het bepalen van deze verhoudingen op verschillende posities, werden de texturen gefilterd met een verzameling Gabor functies. Deze Gabor functies zijn bij uitstek geschikt voor deze taak, omdat ze zowel de positie als de kenmerken van een golf goed representeren en daarnaast een goede benadering leveren voor de responsies van de zgn. "simple cells", die in ruime mate in de visuele systemen van primaten voorkomen. Het filteren levert op deze manier op elke positie een ruw spectrum op van de golven die in de texturen aanwezig zijn. Texturen kunnen worden onderscheiden op grond van verschillen tussen hun Gabor responsies: Als twee texturen nauwelijks verschillen op grond van de responsies van een bepaalde Gabor functie, dan zal die functie nauwelijks bijdragen aan de classificatie, terwijl dit wel het geval is voor Gabor functies met sterk verschillende responsies.

Voor verschillende groepjes van 8 texturen werd een groot aantal kunstmatige neurale netwerken getraind om de aanwezige texturen te onderscheiden. Uit de resultaten kon geconcludeerd worden dat de netwerken de verschillende texturen goed konden leren onderscheiden. Daarnaast bleek dat de netwerken niet alleen details van aangeboden texturen leerden, maar ook algemenere eigenschappen van texturen (generalisatie). Door de Gabor responsies in verschillende vormen aan de neurale netwerken te presenteren, werd duidelijk dat op grond van uitwijkingen van golven een goede classificatie kan worden gemaakt, terwijl fases van golven hierop een versturende invloed hebben. Uit de resultaten van genoemde experimenten kan worden geconcludeerd dat de ontwikkelde methode goed gebruikt kan worden voor het onderscheiden van geleerde texturen en dat ze daarnaast bij uitstek geschikt is om te onderzoeken welke combinaties en kenmerken van Gabor functies van belang zijn bij het classificeren van texturen.