

University of Groningen

Contour detection and shape recognition in image analysis

Grigorescu, Cosmin

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:

2004

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

Citation for published version (APA):

Grigorescu, C. (2004). *Contour detection and shape recognition in image analysis*. 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

De onderwerpen behandeld in deze dissertatie kunnen het best worden weergegeven door wat Leonardo Da Vinci gekarakteriseerd heeft als een inleidende vorm van artistieke uitdrukking: "Het eerste beeld bestond uit een omtreklijn die de schaduw van een man omsingelde zoals de zon het op een wand heeft afgebeeld". Hoewel over de artistieke eigenschappen van deze kunstvorm zullen veel critici in debat treden, wordt een discussie daarover aan hen overgelaten. In plaats daarvan zullen wij proberen om vraagstukken over deze vorm van uitbeelding vanuit een ander perspectief aan te pakken. Wij zullen namelijk proberen te beantwoorden of het mogelijk is om een dergelijke representatie, bestaand uit contouren van objecten, met behulp van computer-beeldanalyse technieken op te bouwen, en hoe deze contourgebaseerde representatie gebruikt kan worden om verschillende objecten in digitale beelden te vinden en te herkennen.

Op het eerste gezicht, kan een dergelijke contourgebaseerde representatie niet veel zeggen over het object dat afgebeeld is. Maar toch bevat deze representatie het wezen van onze visuele waarneming, namelijk de *contouren*, als de voornaamste representatie, en de *vorm van objecten*, gegeven door de ruimtelijke posities van contouren, als een bijkomende, afleidende representatie. Wij zullen vooral onze aandacht vestigen op twee vaak voorkomende onderwerpen in digitale beeldanalyse, waaruit de twee delen van dit proefschrift ontstaan: contourdetectie en vormgebaseerde objectherkenning.

(a) Contour detectie

Hoewel bijzonder veel vooruitgang is gemaakt tijdens de laatste drie decennia om contouren van objecten in digitale beelden te vinden is de belangstelling van informatica onderzoekers in dit onderwerp nogal groot. Een bewijs daarvan is het enorme aantal publicaties dat erover gepubliceerd is.² Ondanks deze ontwikkelingen ruimte voor verbetering van prestaties van deze beeldanalyse operatoren bestaat.

²Een zoekopdracht met het woord "edge detection" (randdetectie) gegeven aan de IEEE Xplore on-line database (<http://ieeexplore.ieee.org>) zal ongeveer 270 tijdschrift artikelen vinden die in de laatste 5 jaar alleen door de IEEE associatie gepubliceerd zijn. Deze artikelen gaan niet alleen over nieuwe methoden van randdetectie maar ook over de toepassingen daarvan.

Huidige systemen zijn gericht op verschillende applicaties (contourdetectie van organen in medische beelden, analyse van satellietbeelden, etc.) of ze functioneren in een kunstmatige, nauwkeurig gecontroleerde omgeving (automatische foutinspectie). Ze hebben beperkte bekwaamheden voor detectie van randen van objecten in beelden die natuurlandschappen of omgevingen bevatten, zoals wij in het dagelijks leven zien (*b.v.*, als er contouren van een dier op een achtergrond van gras gedetecteerd moeten worden). Eén van de redenen waarom volle prestaties van deze operatoren niet volledig bereikt zijn komt door onze beperkte kennis over hoe het menselijk visueel systeem natuurbeelden analyseert. Inderdaad hebben de meeste studies nauwelijks relaties gemaakt met neurofysiologische mechanismen van het visuele systeem en tevens zijn studies uit de neurofysiologie meestal gericht op kwantitatieve data-analyse, met ontoreikende nadruk op de principes die achter de werkwijze van hersengebieden zitten.

Het is ons doel om deze achterstand in te lopen door de introductie van een aantal contourdetectie-operatoren die het gedrag simuleren van bepaalde types neuronen die in het visuele systeem voorkomen. Neurofysiologische metingen, gesteund door studies uit de psychofysica, tonen aan dat het visuele systeem een onderscheid maakt tussen randen die ontstaan uit textuurgebieden, bijvoorbeeld gras, en contouren van objecten. Van een dergelijke onderscheiding maken de meeste randdetectie modellen weinig, of geen gebruik om contouren in natuurbeelden te vinden.

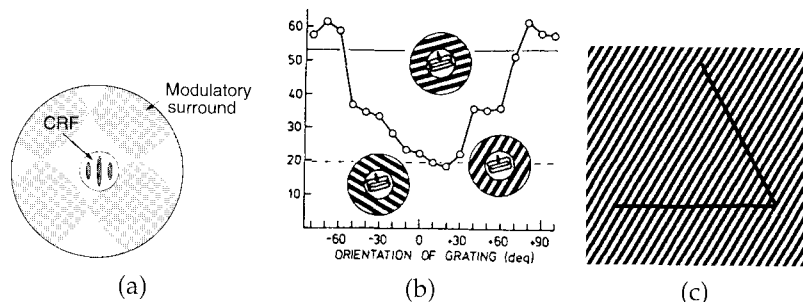


Figure B.1: (a) De niet-klassieke receptieve veld (non-CRF) inhibitie ontstaat in de omgeving van het klassieke receptieve veld (CRF). (b) De non-CRF inhibitie vermindert de respons van een neuron (gemeten in spikes per seconde) wanneer andere randen met dezelfde oriëntatie buiten het gebied van de optimale stimulus, of zijn klassieke receptieve veld, voorkomen (Courtesy of Blakemore and Tobin (1972)). (c) Onze perceptie stemt overeen met het non-CRF gedrag: we zien slechts nog twee van de randen van de driehoek en in sterk verminderde mate de rand die door de georiënteerde textuur omgeven is.

Na een kort overzicht van randdetectie-operatoren, gemaakt in hoofdstuk 2, zullen wij in hoofdstukken 3 en 4 een aantal biologisch-gemotiveerde operatoren voor contourdetectie introduceren. Deze operatoren zijn geïnspireerd op het responsgedrag van bepaalde typen neuronen dat bekend staat als "nonclassical receptive field (non-CRF) inhibition" (niet-klassieke receptieve veld inhibitie). Dit gedrag karakteriseert de respons van meer dan 80% van de neuro-

nen van de primaire visuele cortex van apen en het is aangetoond dat de visuele perceptie van mensen daar ook door wordt beïnvloedt. In essentie, manifesteert dit effect zich in een vermindering van de respons van een neuron op een rand die in zijn receptieve veld voorkomt wanneer andere randen buiten dit receptieve veld aanwezig zijn, Fig. B.1.

Twee verschillende types van dit inhibitiegedrag zullen worden gemodelleerd: *anisotrope* en *isotrope* non-CRF inhibitie. De anisotrope inhibitie ontstaat wanneer andere randen met dezelfde oriëntatie in het non-CRF gebied van een neuron aanwezig zijn. Daarentegen toont een neuron met isotrope inhibitie een vermindering van de respons op alle randen in zijn non-CRF gebied, zonder een onderscheid in de oriëntatie van deze randen te maken.

Meer precies stellen we in hoofdstuk 3 twee contourdetectie-operatoren voor, gebaseerd op het non-CRF effect. De modellen van deze operatoren hebben een aantal bewerkingstappen. In de eerste stap, wordt de neurale respons binnen het klassieke receptieve veld gemodelleerd door een nonlineaire combinatie van Gabor functies, een operator bekend in de beeldanalyse als de "Gabor energy" operator. Aan deze operator voegen we een onderdrukkingstap toe die zich baseert op het bovengenoemde non-CRF inhibitie effect. We beschouwen deze stap aan de ene kant afhankelijk en aan de andere kant onafhankelijk van de oriëntaties van andere randen in de non-CRF omgeving. We spreken dan over anisotrope en, respectievelijk, isotrope non-CRF contourdetectie. In de laatste bewerkingstap worden binaire contour beelden verkregen met een postprocessing procedure (suppressie van nonmaxima gevolgd door thresholding met hysteresis).

We gebruiken natuurbeelden met handmatig geannoteerde "ground truth" contouren van objecten om de prestaties van deze contourdetectie-operatoren vast te stellen. De resultaten tonen aan dat de nieuwe operatoren beter presteren in vergelijking met bekende randdetectie-operatoren vanwege de onderdrukking van valse randen afkomstig uit textuurgebieden en betere detectie van object contouren. Ons model heeft een predictieve waarde in zoverre dat het verklaart waarom ons visuele systeem gevoelig is voor contouren tussen twee textuurgebieden met verschillende oriëntaties.

In hoofdstuk 4, wordt het effect van anisotrope en isotrope non-CRF inhibitie toegevoegd als een aanvullende bewerkingstap aan één van de bekendste methoden van randdetectie in beeldverwerking, de gradiëntgebaseerde methode. Bovendien stellen we een nieuwe methode van postprocessing voor die randen van textuur nog verder onderdrukt en tot betere contourdetectie in aanwezigheid van textuur leidt. Met dezelfde beoordelingsmethode uit hoofdstuk 3 vergelijken we nogmaals de prestaties van deze nieuwe gradiëntgebaseerde contourdetectie-operatoren met de prestaties van twee bekende randdetectie-operatoren, Canny en SUSAN. In natuurbeelden, de contourdetectie-operatoren veel beter presteren dan deze twee randdetectie-operatoren, Fig. B.2.

(b) Vormgebaseerde objectherkenning

Verskillende technologieën, bijvoorbeeld automatische documentanalyse, inhoud-gebaseerd vinden van een beeld in een beeld-database ("image database retrieval"), herkenning van objecten in satellietbeelden, etc. eisen snelle en betrouwbare methoden voor detectie en herken-

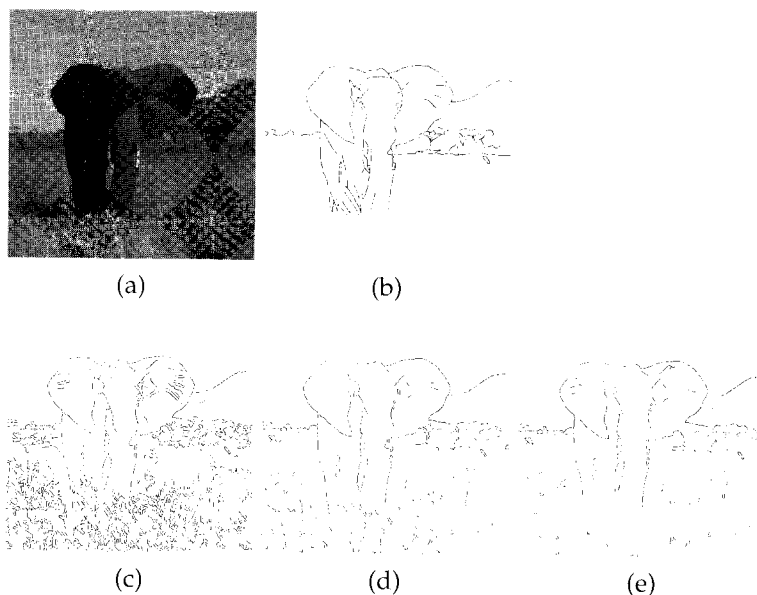


Figure B.2: (a) Natuurbeeld. (b) Handmatig geannoteerde "ground truth" contouren. (c) Beste respons van de Canny operator. (d)-(e) Beste respons van de anisotrope en, respectievelijk, isotrope contourdetectie-operatoren. De prestaties worden geëvalueerd door vergelijking van de respons met de handmatig geannoteerde contour map; de beste prestatie wordt bereikt wanneer het aantal goed-gedetecteerde contouren maximaal is en het aantal gedetecteerde randen van textuur minimaal is.

ning van verschillende typen van objecten in beelden. In het detectie- en herkenningsproces is vorm een van de distinctieve eigenschappen van objecten die ons visuele systeem gebruikt om objecten met elkaar te vergelijken. Het tweede deel van deze dissertatie gaat over vormrepresentaties en methoden van objectherkenning.

In hoofdstuk 5 geven we een korte samenvatting van vormgebaseerde representaties gebruikt in huidige computersystemen. Numerieke of niet-numerieke representaties die, aan de ene kant, gebruikmaken van getallen of, aan de andere kant, van symbolen om de vorm van objecten te beschrijven zijn bekend in de literatuur onder de naam "shape descriptors" – methoden voor vorm beschrijven, of *vormdescriptoren*. Met behulp van vormdescriptoren kunnen twee aspecten die tezamen vormen wat we noemen de *herkenning* van een object worden aangepakt: object *detectie* en object *categorisatie*.

In object detectie vindt een beoordeling plaats over de aanwezigheid van een bepaald object in het beeld, terwijl in objectcategorisatie, gegeven een aantal objectklassen, een gedetecteerd object wordt toegekend aan een bepaalde klasse van objecten die in overeenstemming zijn qua vorm-gelijkenis met andere objecten van de betreffende klasse. Deze twee processen impliceren

dat een vergelijking moet worden gemaakt tussen de vormdescriptoren van een reeds bekend object en het object dat gedetecteerd of geclassificeerd moet worden.

Zelfsprekend zijn de bovengenoemde detectie- en classificatieprocessen afhankelijk van zowel robuuste en betrouwbare berekening van vorm descriptoren als van de prestaties van behorende vergelijksmethoden tussen vormdescriptoren. Een goede vormdescriptor moet onafhankelijk, of weinig afhankelijk zijn van transformaties (rotatie, translatie, herschaling, spiegeling, deformaties) die met een object kunnen worden uitgevoerd. Een goede ongelijkenismaat tussen vormdescriptoren moet overeenkomen met menselijke perceptie van vormverschil tussen objecten.

In hoofdstuk 6 stellen wij een nieuwe "shape descriptor" voor, samen met een methode voor objectherkenning. Deze vorm descriptor behoort tot de zogenaamde klasse van geometrische vormdescriptoren, die de ruimtelijke verhouding tussen bepaalde beeldpunten karakteriseert. Deze beeldpunten, die perceptueel belangrijk of in andere opzichten interessant zijn, worden vaak *kenmerk punten* genoemd. Ze kunnen worden bepaald en geïsoleerd door toepassing van beeldbewerkingsoperatoren, zoals bij voorbeeld contourdetectie-operatoren, textuurdetectie-operatoren, etc.

Meer precies te zijn, berekenen we vanuit elk kenmerkpunt een specifieke datastructuur die de ruimtelijke verhouding tot andere dergelijke kenmerkpunten karakteriseert. De betreffende datastructuur is hiërarchisch georganiseerd zodanig dat aan elk kenmerkpunt een verzameling van afstanden tot een gegeven aantal buurkenmerkpunten wordt toegewezen. Er kunnen verschillende types van kenmerkpunten in een beeld gelokaliseerd worden (contouren, textuur, hoeken, etc.), waarbij elk type gevonden en geïsoleerd wordt door een bepaalde beeldbewerkingsoperator. Daarom groeperen we de afstanden in deelverzamelingen overeenkomstig met het type van de beeldbewerkingstechniek waarmee een kenmerkpunt wordt bepaald. Verder beschrijven we de vorm van een object door de verzameling van alle afstandsverzamelingen die geassocieerd zijn met de kenmerkpunten afkomstig van het betreffende object.

Verder introduceren we ongelijkenismaten tussen kenmerkpunt-descriptoren en verzamelingen van punt-descriptoren die gebruikt kunnen worden om objecten te lokaliseren en te classificeren. Experimenten met herkenning van gedrukte en handgeschreven letters, evenals de detectie van verkeersborden zijn uitgevoerd om de betrouwbaarheid van de detectiemethode te beoordelen. De prestaties van de categorisatiemethode zijn beoordeeld in experimenten met handgeschreven letters, objectherkenning (COIL-20 database) en herkenning van objecten gebaseerd op hun silhouetten (MPEG-7 silhouette database retrieval).

Alhoewel op dit moment geen vorm representatie bekend is die een dergelijke werkwijze in biologische systemen zal tonen, bestaat er een zekere cognitieve motivatie achter onze modellen. De kenmerkpunt-gebaseerde vormdescriptor toont gelijkenis met één van de geaccepteerde cognitieve modellen van perceptie, het zogenaamde "feature map" model. Volgens dit model wordt de vorm van een object in onze perceptie waargenomen als een collectie van kenmerken samen met hun ruimtelijke posities. Deze representatie heeft ook een zekere fysiologische geloofwaardigheid omdat de hersengebieden waarin de eerste stappen van visuele informatie worden verwerkt netvlies-georganiseerde neuronen bevatten die zich bezighouden met hetzelfde

type informatieverwerking. Een ander principe van organisatie waar we in onze detectie- en categorisatie-algoritmen impliciet rekening mee houden is dat informatiebewerking in de visuele hersengebieden hiërarchisch georganiseerd is, en dat lokale bewerking eerder dan de globale bewerking plaatsvindt.