

Raoul Privenau

Mustererkennung mit Neokognitron und Anwendungen

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Dieses Werk sowie alle darin enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsschutz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen, Auswertungen durch Datenbanken und für die Einspeicherung und Verarbeitung in elektronische Systeme. Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen Wiedergabe (einschließlich Mikrokopie) sowie der Auswertung durch Datenbanken oder ähnliche Einrichtungen, vorbehalten.

Copyright © 2011 Diplomica Verlag GmbH
ISBN: 9783842820364

Raoul Privenau

Mustererkennung mit Neokognitron und Anwendungen

Raoul Privenau

Mustererkennung mit Neokognitron und Anwendungen

Raoul Privenau

Mustererkennung mit Neokognitron und Anwendungen

ISBN: 978-3-8428-2036-4

Herstellung: Diplomica® Verlag GmbH, Hamburg, 2011

Zugl. Martin-Luther-Universität Halle-Wittenberg, Halle, Deutschland, Diplomarbeit, 2011

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die Informationen in diesem Werk wurden mit Sorgfalt erarbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden und der Verlag, die Autoren oder Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für evtl. verbliebene fehlerhafte Angaben und deren Folgen.

© Diplomica Verlag GmbH

<http://www.diplomica.de>, Hamburg 2011

Inhaltsverzeichnis

Abbildungsverzeichnis	V
Abkürzungsverzeichnis	VII
Tabellenverzeichnis	IX
1 Einleitung	1
1.1 Motivation	1
1.2 Ziele der Arbeit	2
1.3 Aufbau der Arbeit	2
2 Mustererkennung	5
2.1 Mustererkennungsprozess	6
2.2 Wesentliche Mustererkennungsaufgaben	8
2.3 Methoden der Mustererkennung	10
3 Künstliche Neuronale Netze	13
3.1 Standardneuronenmodell	13
3.2 Grundlegende Netzarchitekturen	15
3.3 Lernverfahren	16
3.4 Einsatz im Bereich der Mustererkennung	17
4 Hopfield-Netz	19
4.1 Netzstruktur	19
4.2 Mustererkennung	20
4.2.1 Rekursive Berechnung	21
4.2.2 Energiefunktion	21
4.3 Musterspeicherung	22
4.4 Musterlöschung	24
4.5 Verschiedene Netzkonfigurationen	24
4.6 Bewertung	26
5 Neokognitron	29
5.1 Grundlegende Netzstruktur	29
5.1.1 Eingabeschicht	30
5.1.2 Stufensystem	30

5.1.3	Verbindungsstruktur	31
5.2	Mustererkennung	32
5.3	S-Zellen	34
5.3.1	Verbindungsstruktur	34
5.3.2	Mathematische Beschreibung	37
5.3.3	Merkmalsextraktion auf Basis der Ähnlichkeit	39
5.3.4	Steuerung der Selektivität	40
5.4	C-Zellen	42
5.4.1	Verbindungsstruktur	42
5.4.2	Mathematische Beschreibung	43
5.4.3	Positionsinvarianz und Unschärfe	44
5.5	Lernverfahren	46
5.5.1	Wahl der Netzgröße	46
5.5.2	Nicht-überwachtes Lernen	47
5.5.3	Überwachtes Lernen	51
5.5.4	Vergleich der Lernverfahren	52
5.6	Bewertung	53
5.6.1	Integration in den Mustererkennungsprozess	53
5.6.2	Geeignete Muster	54
5.6.3	Erkennungsleistung und Störungsinvarianz	54
5.6.4	Fazit	55
6	Erweiterungen des Neokognitrons	57
6.1	Erweitertes Neokognitron	57
6.1.1	Kontrastgewinnung	57
6.1.2	Hemmendes Umfeld von C-Zellen	58
6.1.3	Hybrider Lernprozess	59
6.1.4	Weiterentwicklung	60
6.1.5	Computersimulation	63
6.2	Selektive Aufmerksamkeitssteuerung und Autoassoziation	64
6.2.1	Routenwahl	66
6.2.2	Autoassoziation	66
6.2.3	Verstärkung der Aufmerksamkeit	67
6.2.4	Segmentierung	67
6.2.5	Wechsel der Aufmerksamkeit	67
6.2.6	Antwortkontrolle	68
6.3	Rotationsinvarianz	69

6.3.1	Version 1 – Rotationsinvariantes Neokognitron	69
6.3.2	Version 2 – Hybrides Neokognitron entsprechend der mentalen Rotation	70
6.4	Doppelte C-Zellen-Schicht	72
7	Anwendungen	75
7.1	Zeichenerkennung	76
7.1.1	Erkennung numerischer und alphanumerischer Zeichen	76
7.1.2	Weitere Anwendungen im Bereich der Zeichenerkennung . . .	78
7.2	Erkennung handgeschriebener Musiknoten	79
7.3	Gesichtserkennung	80
7.4	Automatische Zielerkennung	85
7.5	Analyse von Echokardiogrammen	87
7.6	Erkennung von Draht-Modellen	88
7.7	Suche struktureller Einheiten auf Mikrochips	89
7.8	Weitere Anwendungen	90
8	E-Learning Modul für binäre Hopfield-Netze	93
8.1	Wahl der Software-Plattform und Entwicklungsumgebung	94
8.2	Softwarearchitektur und allgemeine Modulverwendung	95
8.3	Kritische Aspekte der Entwicklung	97
8.3.1	Hopfield-Logik	97
8.3.2	Benutzeroberfläche und Ablaufsteuerung	99
9	Resümee und Ausblick	101
	Anhang	104
A	Hopfield Netz	105
A.1	Gewichtsmatrix	105
A.2	Beispiel für eine Energiefunktion	105
A.3	Beispiel zum Lernen der Schwellen	106
B	Software zur Simulation eines Neokognitrons	107
B.1	Beholder 2.0b	107
B.2	NeoCognitron 1.0.0 Beta	112
C	Muster bei der Erkennung alphanumerischer Zeichen	117

D CAPTCHA-Beispiele	123
E E-Learning Modul	125
E.1 Abbildung des zweidimensionalen Eingabebereiches auf ein eindimensionales Feld	125
E.2 Abbildung der zweidimensionalen Gewichtsmatrix auf ein eindimensionales Feld	125
E.3 Einsatz einer Datenbank zur Speicherung der Gewichtsmatrix	127
E.4 Ermittlung der maximalen Mustergröße in Abhängigkeit von der Prozessorarchitektur	128
Literaturverzeichnis	131

Abbildungsverzeichnis

2.1	Mustererkennungsprozess	6
3.1	Vergleich biologisches und künstliches neuronales Netz	13
5.1	Grundlegende Netzstruktur eines dreistufigen Neokognitrons	30
5.2	Unterscheidung direktes und indirektes rezeptives Feld	31
5.3	Beispiel für die Mustererkennung mit dem Neokognitron	33
5.4	Verbindungsstruktur von S-Zellen der ersten S-Schicht	34
5.5	Verbindungsstruktur eines Stufenübergangs	36
5.6	Ein- und Ausgabecharakteristika einer S-Zelle	37
5.7	Bewertung der Ähnlichkeit durch S-Zellen im mehrdimensionalen Merk- malsraum	40
5.8	Verbindungsstruktur von C-Zellen	42
5.9	Ein- und Ausgabecharakteristika einer C-Zelle	43
5.10	Positionsinvarianz und Unschärfe-Operation	44
5.11	Störungstoleranz einer C-Zelle in Abhängigkeit von der Größe ihres rezeptiven Feldes	45
5.12	Hypercolumns als Basis des Lernverfahrens	49
6.1	Kontrastgewinnungsschicht	58
6.2	Funktionen des hemmenden Umfeldes von C-Zellen	59
6.3	Enthemmung von C-Zellen	62
6.4	Grundlegende Netzstruktur mit integrierter selektiver Aufmerksamkeits- steuerung	65
6.5	Struktur einer Stufe mit selektiver Aufmerksamkeitssteuerung	66
6.6	Rotationsinvariantes Neokognitron (Version 1)	69
6.7	Grundlegender Ablauf der Mustererkennung mit dem rotationsinvarianten Neokognitron (Version 2) auf Basis der mentalen Rotation	71
6.8	Gleichmäßige und ungleichmäßige Unschärfe-Operation im Vergleich . .	73
7.1	Netzstruktur zur Gesichtserkennung und Segmentierung	84
7.2	Drahtmodelle für das Training	88
8.1	Softwarearchitektur des E-Learning Moduls	95
8.2	Programmbildschirm des E-Learning Moduls	97
A.1	Beispiel für eine Energiefunktion	105

B.1	Beholder – Konfiguration der Netzstruktur	107
B.2	Beholder – Konfiguration der Parameter	108
B.3	Beholder – Mustermenge	108
B.4	Beholder – Mustereditor	108
B.5	Beholder – Aktivierungen der Rezeptorzellen bei angelegtem Muster . .	109
B.6	Beholder – Programmfenster mit angelegtem gelerntem Muster	110
B.7	Beholder – Programmfenster mit angelegtem nicht gelerntem Muster . .	111
B.8	NeoCognitron – Programmfenster mit trainiertem Neokognitron	112
B.9	NeoCognitron – Konfiguration der Netzstruktur	113
B.10	NeoCognitron – Mustermenge	113
B.11	NeoCognitron – Mustereditor	114
B.12	NeoCognitron – Aktivierung von S-Zellen auf der zweiten S-Schicht bei angelegtem Muster	114
B.13	NeoCognitron – Programmabsturz	115
C.1	Trainingsmuster der ersten S-Schicht	117
C.2	Trainingsmuster der zweiten S-Schicht	118
C.3	Trainingsmuster der dritten S-Schicht	119
C.4	Trainingsmuster der vierten S-Schicht	120
C.5	Korrekt erkannte Muster bei der Mustererkennung	121
D.1	Registrierungsformular eines Google Mail Kontos (Ausschnitt)	123
D.2	Registrierungsformular zur Webanwendung „Delicious“ (Ausschnitt) . . .	123
D.3	CAPTCHA-Erzeugung durch den freien CAPTCHA Webdienst „reCAPTCHA“	124
D.4	CAPTCHA zur Objektidentifikation und manuellen Abgrenzung	124
E.1	Abbildung der zweidimensionalen Mustereingabe auf ein eindimensionales Feld	125
E.2	Abbildung der zweidimensionalen Gewichtsmatrix auf ein eindimensio- nales Feld	126
E.3	Java-Code des Zugriffs auf das eindimensionale Feld der Gewichtsmatrix	126

Abkürzungsverzeichnis

ALU	arithmetic logic unit
API	application programming interface
ATR	automatic target recognition
AWT	Abstract Window Toolkit
BNN	biologisches neuronales Netz
CAPTCHA	completely automated public turing test to tell computers and humans apart
CMYK	Cyan, Magenta, Yellow und Key
CNN	convolutional neural network
CPU	central processing unit
CUDA	compute unified device architecture
DB	Datenbank
DBS	Datenbanksystem
EM	E-Learning Modul
GPU	graphics processing unit
GUI	graphical user interface
HC	Hypercolumn
HN	Hopfield-Netz
HRC	high resolution channel
HTML	hypertext markup language
ICBM	intercontinental ballistic missile
JAR	Java Archive
JDK	Java Development Kit
JFC	Java Foundation Classes
JRE	Java Runtime Environment

JVM	Java Virtual Machine
KNN	künstliches neuronales Netz
LRC	low resolution channel
LSM	logarithmic spiral mapping
MRNN	multi-resolution neural network
NZ	Nervenzelle
OffZZ	Off-Zentrum-Zelle
OnZZ	On-Zentrum-Zelle
PA	Prozessorarchitektur
RAM	random access memory
RGB	Rot, Grün und Blau
SDI	Strategic Defense Initiative
UFSCar	Universidade Federal de São Carlos

Tabellenverzeichnis

4.1 Verschiedene Konfigurationen eines binären Hopfield-Netzes	25
5.1 Vergleich der Lernverfahren des Neokognitrons	53
7.1 Vergleich der Netzstrukturen zur Erkennung handgeschriebener numeri- scher und alphanumerischer Zeichen	76
7.2 Netzstruktur zur Erkennung handgeschriebener Zahlen mit dem erweiterten Neokognitron	77
7.3 Vergleich der Erkennungsgeschwindigkeiten bei der Gesichtserkennung bzgl. verschiedener Systemumgebungen	83
7.4 Netzkonfiguration zur Erkennung der Drahtmodelle	88

1 Einleitung

1.1 Motivation

Das menschliche Gehirn empfängt eine Fülle unterschiedlicher Reize über verschiedene Sinnesorgane. Ein bedeutendes Gebiet in diesem Zusammenhang ist die visuelle Wahrnehmung¹, die als „... Aufnahme und die zentrale Verarbeitung von visuellen Reizen ...“² definiert ist. Der Mensch ist in der Lage aus einer großen Menge visueller Reize bestimmte Signale innerhalb kürzester Zeit herauszufiltern und richtig zu interpretieren. Probleme ergeben sich allerdings bei der Portierung dieser Fähigkeit der natürlichen Mustererkennung auf Computersysteme durch die erheblichen Unterschiede hinsichtlich der Leistungsfähigkeit und der Architektur. Nach dem derzeitigen technischen Stand können Computer Daten um ein vielfaches schneller verarbeiten, als unser Gehirn, sind aber dennoch mit Aufgaben überfordert, die durch unser Gehirn in kürzester Zeit erfolgreich durchgeführt werden.³ Charakteristisch für das menschliche Gehirn, als komplexes biologisches Netzwerk, ist seine hochgradige parallele Signalverarbeitung.⁴ Dank seiner ca. 10^{12} Nervenzellen⁵, die über ca. 10^{15} Verbindungen⁶ miteinander verknüpft sind, verarbeitet es Millionen von Reizen innerhalb weniger Millisekunden.⁷ Die serielle Datenverarbeitung auf einer Von-Neumann-Architektur ist damit nicht vergleichbar.⁸ Es existieren allerdings stark vereinfachte Modelle zur Nachahmung dieses komplexen biologischen natürlichen Nervensystems, die als künstliche neuronale Netze (KNN) bezeichnet werden.⁹ Auch im Bereich der Mustererkennung hat sich der Einsatz von KNN bewährt, wobei besonders das Neokognitron eine gute Nachahmung der natürlichen Mustererkennung im visuellen Bereich verspricht.¹⁰

1 Prozess der visuellen Wahrnehmung siehe [66], S. 162 ff. und [153], S. 345 ff.

2 [25], S. 173.

3 Vgl. [10], S. 1; [133], S. 13; [154], S. 13 und [57], S. 318.

4 Vgl. [154], S. 13.

5 Vgl. [97] und [172] S. 1.

6 Vgl. [97].

7 Vgl. [133], S. 13.

8 Vgl. [133], S. 13.

9 Vgl. [133], S. 13 und [12], S. 15.

10 Vgl. [12], S. 11 und [133], S. 14.