

Kinderleicht

Grundlagen der Bildverarbeitung: Formerkennung

Das menschliche visuelle Wahrnehmungsvermögen ist ausgezeichnet in der Lage, Formen von Objekten zu erkennen. Schon kleine Kinder unterscheiden mühelos Kreise, Quadrate oder Dreiecke bei Holzklötzen oder Süßigkeiten. Auch bei Anwendungen der Bildverarbeitung liegt es nahe, quantitative Merkmale zu verwenden, die die Form von Objekten in der Szene charakterisieren. Beispielsweise können Verkehrszeichen schon weitgehend aufgrund ihrer Form klassifiziert werden, und bei manchen Tabletten ist die Form ein wesentliches Prüfkriterium. Dieser Beitrag beschreibt einige einfache, etablierte Methoden zur Erkennung von Formen.

Eine Methode zu Formerkennung soll ein quantitatives Merkmal liefern, das die Form eines Objekts eindeutig charakterisiert, unabhängig von seiner Drehlage, seiner Position und seinen Abmessungen. Wir beschränken uns hier auf den zweidimensionalen Fall: die Objekte liegen in einer Ebene und haben nur geringe Ausdehnung senkrecht zur Ebene. Die Abbildung soll außerdem formtreu sein, d.h. perspektivische Verzerrungen, optische Verzerrungen, Inhomogenitäten der Beleuchtung, Randabfall der optischen Abbildung und Verzerrung durch den Bildeinzug sollen vernachlässigbar sein oder sind bereits herausgerechnet. Wir gehen außerdem zunächst davon aus, dass wir uns nur für die Formen von zusammenhängenden Objekten ohne Löcher interessieren, die sich im Bild deutlich vom Untergrund abheben. Das entstehende Bild könnte dann so aussehen wie in Abbildung 1. Ein menschlicher Betrachter kann die dort abgebildeten Objekte sofort anhand ihrer Form beschreiben: Rechteck, Quadrat, Kreis, Dreieck, Mond, Herz, Ellipse. Eine quantitative Kenngröße zur Charakterisierung dieser Formen ist die Rundheit, auch als Kompaktheit bekannt. Ihre Konstruktion beruht auf der Überlegung, dass der Kreis diejenige geometrische Figur ist, die bei gegebener Fläche den kleinsten Umfang hat. Im Vergleich dazu ist ein Quadrat weniger kompakt, denn sein Umfang ist bei gleicher Fläche größer als bei einem Kreis. Das Verhältnis von Umfang U und Fläche A ergibt beim Kreis mit Radius r den Wert $2\pi r/\pi r^2$, also $2/r$; beim Quadrat mit Kantenlänge a folgt $4a/a^2$, also $4/a$. Diese Größe ist zwar rotationsinvariant,

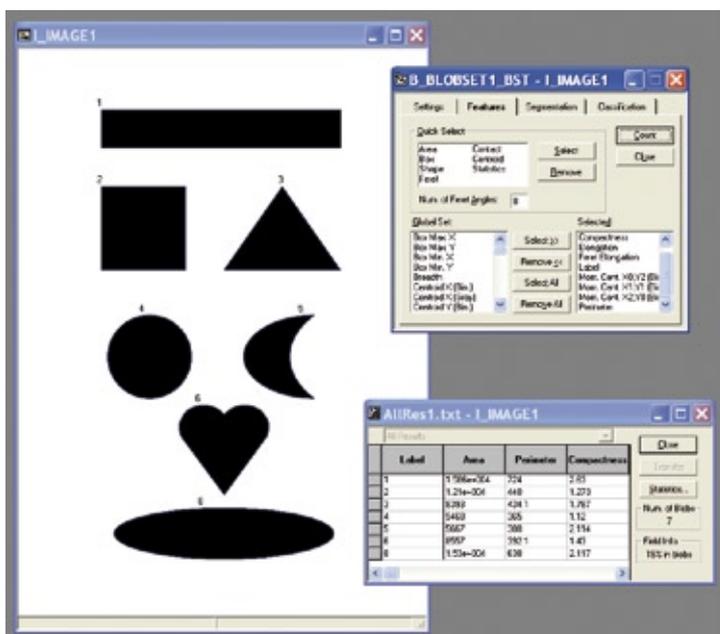


Abb. 1: Einige einfache Formen im Binärbild und Ergebnisse der Blobanalyse

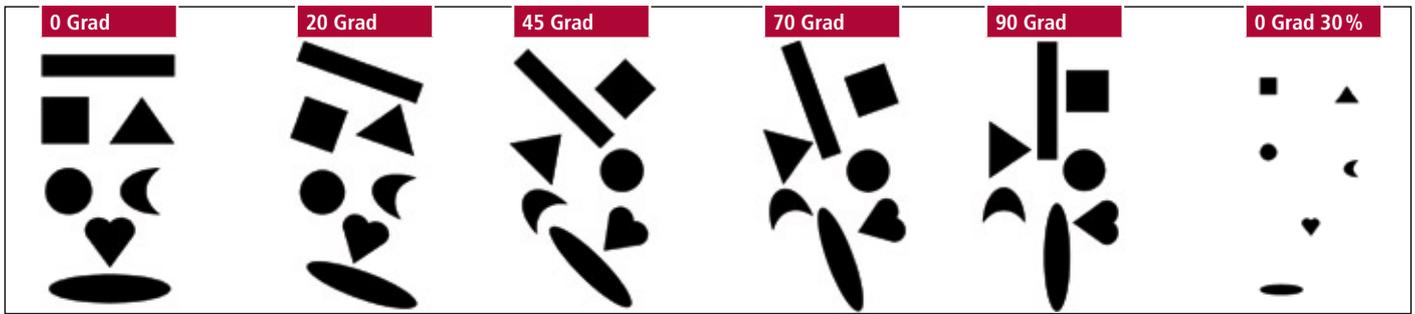


Abb. 2: Das Binärbild aus Abbildung 1 in verschiedenen Drehlagen und auf 30% verkleinert

aber die Linearabmessung geht ein. Geschickter ist ein Merkmal der Form U^2/A , so dass r bzw. a herausfallen. Für den Kreis ergibt sich der Wert 4π , für das Quadrat der Wert 16, unabhängig von der Größe, der Orientierung und der Position im Bild.

Kompaktheit

Die meisten Bibliotheken für Bildverarbeitungsfunktionen liefern dieses oder ein ähnliches Formmerkmal bei der Blobanalyse. Oft wird das Formmaß noch so normiert, dass es für den Kreis den Wert 1 annimmt. Das Programm, mit dem Abbildung 1 erzeugt wurde, verwendet als „compactness“ den Ausdruck $U^2/4\pi A$. Formen, die vom Kreis abweichen, haben dann eine „compactness“ größer als 1. Auch der Kehrwert ist gebräuchlich. Ein Blick auf die Ergebnisliste der Blobanalyse in Abbildung 1 zeigt, dass sich Rechteck, Quadrat, Dreieck, Kreis und Herz offenbar recht gut in der Kompaktheit voneinander unterscheiden. Der Wert für den Kreis weicht jedoch deutlich vom theoretischen Wert 1 ab. Dafür sind zwei Effekte verantwortlich. Einerseits gibt es verschiedene Methoden, den Umfang eines Objekts zu berechnen. Manche Verfahren benutzen einfach die Zahl der Pixel auf der Kontur, manche gewichten Diagonalschritte mit dem Faktor $\sqrt{2}$, andere bestimmen den Umfang aus dem Konturcode als Gesamtlänge der Verbindungslinie zwischen den Mittelpunkten der Konturpixel. Ähnliche Unterschiede gibt es bei der Berechnung der Fläche. An einem Extremfall werden die Unterschiede besonders deutlich: für ein Objekt, das nur aus einem einzigen Pixel besteht, kann eine Methode die Fläche 1 und den Umfang 4 ergeben, eine andere die Fläche 0 und den Umfang 0. Andererseits besteht jedes Objekt in der digitalen Ebene aus Pixeln, es ist diskretisiert. Kreise werden in der digitalen Ebene stets durch Vielecke angenähert. Daher kommen für die Merkmale, die den Umfang und die Fläche eines Objek-

Tabelle 1: Kompaktheit und normiertes Trägheitsmoment für die Objekte aus Abbildung 2

0 Grad	U^2/A	Iz/A^2	$100Iz/A^2$
Rechteck	33,05	0,5217	52,17
Quadrat	16,00	0,1667	16,67
Dreieck	22,45	0,1956	19,56
Kreis	14,07	0,1592	15,92
Mond	26,57	0,2168	21,68
Herz	17,97	0,1747	17,47
Ellipse	26,60	0,3523	35,23
20 Grad			
Rechteck	38,40	0,5199	51,99
Quadrat	18,54	0,1665	16,65
Dreieck	23,69	0,1955	19,55
Kreis	14,18	0,1592	15,92
Mond	26,46	0,2168	21,68
Herz	18,16	0,1747	17,47
Ellipse	27,58	0,3536	35,36
45 Grad			
Rechteck	33,13	0,5229	52,29
Quadrat	15,96	0,1666	16,66
Dreieck	22,32	0,1956	19,56
Kreis	14,07	0,1592	15,92
Mond	26,46	0,2165	21,65
Herz	17,91	0,1746	17,46
Ellipse	26,75	0,3550	35,50
70 Grad			
Rechteck	38,46	0,5195	51,95
Quadrat	18,64	0,1667	16,67
Dreieck	23,90	0,1956	19,56
Kreis	14,06	0,1592	15,92
Mond	26,57	0,2163	21,63
Herz	18,18	0,1747	17,47
Ellipse	27,72	0,3538	35,38
90 Grad			
Rechteck	33,05	0,5217	52,17
Quadrat	16,00	0,1667	16,67
Dreieck	22,27	0,1955	19,55
Kreis	14,07	0,1592	15,92
Mond	26,55	0,2163	21,63
Herz	18,02	0,1748	17,48
Ellipse	26,71	0,3529	35,29
0 Grad 30%			
Rechteck	32,18	0,5035	50,35
Quadrat	16,00	0,1666	16,66
Dreieck	21,93	0,1954	19,54
Kreis	14,41	0,1591	15,91
Mond	26,15	0,2178	21,78
Herz	18,32	0,1741	17,41
Ellipse	26,29	0,3433	34,33

tes beschreiben, Abweichungen von den idealen Verhältnissen vor. Beide Effekte beeinflussen die Berechnung der Kompaktheit. Die Diskretisierung wirkt sich umso deutlicher aus, je weniger Pixel für das Objekt oder die relevanten Strukturen zur Verfügung stehen. Weiter fällt in der Ergebnisliste der Abbildung 1 auf, dass sich die Kompaktheit für das sichelförmige Objekt und die Ellipse nur geringfügig unterscheiden. Für eine Anwendung in der industriellen Bildverarbeitung oder im Außenraum muss jeweils geprüft werden, ob die Stabilität der verwendeten Merkmale im Kontext der jeweiligen Erkennungsaufgabe ausreicht. Da der Umfang empfindlich auf das Rauschen des Bildsignals reagiert, liegt die Vermutung nahe, dass die Kompaktheit für eine sichere Trennung dieser beiden Formen in einer Industrienanwendung nicht stabil genug ist.

Robustheit

Ein Blick auf die Bildsequenz in Abbildung 2 und die Merkmalswerte in der zugehörigen Tabelle 1 gibt einen ersten Eindruck von der Stabilität der Größen. Die Formen aus Abbildung 1 wurden jeweils um 20°, 45°, 70° und 90° gegenüber der Ausgangslage verdreht. Ein weiteres Teilbild enthält dieselben Formen, jedoch auf 30% der ursprünglichen Größe verkleinert. In Abbildung 3 sind schließlich zwei der Formen, Herz und Ellipse, in verschiedenen, willkürlich gewählten Drehlagen und Größen abgebildet. Die zugehörigen Merkmale sind Tabelle 2 aufgelistet. Als Maß für die Kompaktheit wurde hier die Größe U^2/A berechnet. Für U und A wurden die Werte aus der Blobanalyse des Programms verwendet, mit dem Abbildung 1 erzeugt wurde. Hier wird deutlich, dass es Formen gibt, bei denen die Drehlage erheblichen Einfluss auf den Wert für die Kompaktheit hat. Beispielsweise variiert das Kompaktheitsmaß für das Quadrat zwischen den Werten 16,0 und 18,6, also um 15%. Dieselbe Spannweite der Werte ergibt sich für das Recht-

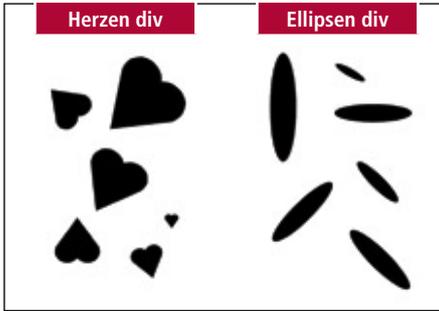


Abb. 3: Herzen und Ellipsen in willkürlichen Größen und Drehlagen

eck. Auch die Vermutung, dass sich „Mond“ und „Ellipse“ nicht trennen lassen, wird bestätigt. Für die Ellipse liegen die Werte zwischen 26,60 und 27,72, für das sichelförmige Objekt zwischen 26,46 und 26,57. Für die Objekte, die auf 30% verkleinert sind, ergeben sich zudem die Werte 26,29 für die Ellipse und 26,15 für die Sichel. Bei unabhängiger Variation von Drehlage und Größe werden sich die Wertebereiche folglich überdecken. Abbildung 3 und Tabelle 2 zeigen, dass die Situation bei den Herzen und Ellipsen in dieser Hinsicht deutlich besser ist. Die Spannweite des Kompaktheitsmaßes beträgt hier etwa 2,5%, bei Berücksichtigung der Werte aus Abbildung 2 insgesamt etwa 5%, ohne dass es zu einer Überdeckung der Werte kommt.

Trägheitsmoment

Ellipse und Sichel kann man mit Hilfe eines anderen, ebenfalls häufig verwendeten Formmerkmals unterscheiden: dem Trägheitsmoment. Momente sind statistische Kenngrößen. Allgemein ist das Moment m_{pq} einer Punktgruppe mit den Koordinaten (x,y) von der Form $\sum x^p y^q$; p und q sind ganzzahlig. Beispielsweise ist das Moment m_{00} gleich der Zahl der gesetzten Pixel und wird meist mit der Fläche des Objekts assoziiert. Die Momente m_{02} und m_{20} haben die Form $\sum y^2$ und $\sum x^2$; sie entsprechen den Trägheitsmomenten in Bezug auf die x-Achse bzw. die y-Achse. Für die Formerkennung werden diese Momente zunächst auf den Schwerpunkt (x_s, y_s) des Objekts bezogen („zentriert“), haben also die Form $\sum (x-x_s)^2$ bzw. $\sum (y-y_s)^2$. Das sind die Trägheitsmomente I_y bzw. I_x bezogen auf eine Achse parallel zur y- bzw. x-Achse durch den Schwerpunkt des Objekts. Ein Kreis hat für beide Richtungen dasselbe Trägheitsmoment, eine gestreckte Ellipse, deren große Halbachse parallel zur x-Achse liegt wie in Abbildung 1, hat bzgl. der x-Achse ein kleines, bzgl. der y-Achse ein großes Trägheitsmoment. Die Summe beider Momente ist rotationsinvariant und ent-

Tabelle 2: Kompaktheit und normiertes Trägheitsmoment für die Objekte aus Abbildung 3

Herzen div	Fläche A	Umfang U	U^2/A	I_z/A^2	$100I_z/A^2$
1	24970	668,7	17,91	0,1748	17,48
2	7403	362,9	17,79	0,1745	17,45
3	15690	532,3	18,06	0,1747	17,47
4	801	118,7	17,59	0,1740	17,40
5	8549	392,9	18,06	0,1747	17,47
6	4405	281,8	18,03	0,1745	17,45
Ellipsen div					
1	15300	636,9	26,51	0,3512	35,12
2	1417	196,0	27,11	0,3456	34,56
3	7586	449,4	26,62	0,3510	35,10
4	4044	328,4	26,67	0,3540	35,40
5	8667	481,6	26,76	0,3541	35,41
6	8673	482,1	26,80	0,3536	35,36

spricht dem Trägheitsmoment I_z bezüglich einer Achse, die senkrecht auf der Bildebene steht. Eine Normierung der Momente, in diesem Fall auf das Quadrat der Fläche, sorgt dafür, dass die Abmessungen eines Objekts nicht in das Merkmal eingehen. Man spricht dann von normierten, zentrierten Momenten. In Tabelle 1 und 2 ist das Merkmal I_z/A^2 ebenfalls aufgeführt. Die Daten zeigen, dass Sichel und Ellipse anhand dieses Merkmals sicher unterschieden werden können, auch bei Variation der Drehlage und der Größe. Auch für die anderen Formen ist dieses Merkmal im Vergleich zur Kompaktheit sehr stabil. Zwar sind die numerischen Unterschiede zwischen den Merkmalswerten für einige Formen relativ gering, das Merkmal ist jedoch so stabil gegen Rotation und Skalierung, dass es für eine Formunterscheidung auch in diesen Fällen in Betracht gezogen werden kann. Aus der Kombination von verschiedenen zentrierten, normierten Momenten kann man weitere Formmerkmale konstruieren [1]. Außerdem kann man durch Gewichtung mit dem Grauwert eines Objektpixels das Konzept der Momente auch auf Graustufenbilder übertragen. Hinzu kommt ein wesentlicher Vorteil der Momente gegenüber Parametern, die auf der Blobanalyse basieren: Momente kann man für beliebige Pixelgruppen berechnen, während die Blobanalyse immer vollständige, zusammenhängende Objekte voraussetzt.

Fazit

Formmerkmale müssen translations-, rotations- und skalierungsinvariant sein. Eine verlässliche Formerkennung ist bereits bei zweidimensionalen Objekten und optimal aufgenommenen Bildern

nicht trivial. Die Diskretisierung führt bei einfachen Formmerkmalen wie der Kompaktheit zu wesentlichen Abweichungen von den idealen Verhältnissen. Sie können geeignet sein, wenn bei der Prüfaufgabe nur wenige, definierte Formen vorkommen und die Stabilität des Merkmals sorgfältig getestet wird. Für Formparameter, die aus Programmbibliotheken entnommen werden, lohnt sich ein kritischer Blick auf die Details der zugrunde liegenden Methode. Aus normierten, zentrierten Momenten kann man Merkmale konstruieren, die zur Formerkennung oft gut geeignet sind und zudem für beliebige, nicht zusammenhängende Punktgruppen berechnet werden können.

Literatur

[1] W. Burger, M. J. Burge, Digitale Bildverarbeitung, Springer-Verlag 2005, S.229

► **Autor**
 Prof. Dr. Christoph Heckenkamp



► **Kontakt**
 Hochschule Darmstadt –
 University of Applied Sciences
 Studiengang Optotechnik und Bildverarbeitung
 heckenkamp@h-da.de
 www.fbmn.h-da.de