PS: Einführung in die Bioinformatik (SS99)

Fast Object Recognition

Obermair Christoph

cober@cosy.sbg.ac.at

Schmidbauer Josef

jschmid@cosy.sbg.ac.at

Stieglbauer Gerald

gstiegl@cosy.sbg.ac.at

Version 1999-04-13

Gruppe 3

Zeit: 17.45 - 18.15

Lehrveranstaltungsleiter: Schwaiger Roland

Inhaltsverzeichnis

Zusammenfassung

Ziel des Projekts ist eine ausführliche Auseinandersetzung mit dem Thema „Objekterkennung“. Dabei werden eingangs grundlegende Methoden der Objekterkennung betrachtet und verglichen. In folge soll eine intensivere Auseinandersetzung mit der Objekterkennung mit Hilfe von künstlichen neuronalen Netzen erfolgen. Daraus soll ein möglichst schnelles Verfahren zur Objekterkennung gesucht werden, welches langfristig auf dem am Institut für Computerwissenschaften der Universität Salzburg entwickelten Roboter E.M.M.A. Einsatz finden soll.

Kapitel 1

Einleitung

1.1 Was ist Objekterkennung?

Informationsverarbeitende Systeme (IVS) können ihre Umwelt nur über bestimmte Schnittstellen wahrnehmen. Diese Schnittstellen empfangen Informationen von außen (Input) und werden als Sensoren bezeichnet. Die Sensordaten können auf verschiedene Weise ausgewertet werden. Ein wichtiges Gebiet ist dabei die „Object Recognition“.
Ein Objekt ist in diesem Kontext durch seine vordefinierten Eigenschaften eindeutig bestimmt und grenzt sich dadurch von seiner Umwelt ab. Man verwendet bestimmte Eigenschaften des Objektes um es von anderen unterscheiden zu können.
Dieses Dokument beschränkt sich darauf, nur jene Eigenschaften zu verwenden, die aus Bildern bestimmt werden können, wobei ein Bild in diesem Zusammenhang eine zweidimensionale optische Abbildung der Umwelt ist.

Im weiteren verstehen wir im speziellen unter IVS Roboter. Damit diese intelligent und effizient mit ihrer Umwelt interagieren können, muß der Roboter Objekte in dieser Umwelt erkennen und lokalisieren können, was wie bereits erwähnt über Sensoren erfolgt. Diese Aufgabe läßt sich in drei Gebiete unterteilen [1]:

- Identitifikation der Gestalt des gesuchten Objektes in einer natürlichen Umgebung.
- Untersuchung des gefundenen Objektes auf seine definierenden Merkmale.
- Feststellung der Position des Objektes bezüglich einer räumlichen Abbildung der Umwelt.

Weitere wichtige Aspekte der Objekterkennung [1]:

- Welche Sensoren sind relevant, um ein bestimmtes Objekt zu erkennen?
- Wie soll man die definierenden Merkmale eines individuellen Objekt bestimmen?
- Wie können große Mengen von Klassifizierungsmerkmalen effizient gespeichert werden?
- Welche Adressierungsmethoden werden verwendet, um bestimmte Objekte aus einer solchen Datenbank ableiten zu können?

- Welche Methoden werden verwendet, um eine Beziehung zwischen Sensordaten und Objekteigenschaften herzustellen?
- Wie kann man aus Sensordaten die Position und Lage eines Objektes herleiten?
- Wie werden Eigenschaften neuer Objekte erkannt und in der Datenbank gespeichert?
- Welche Rolle spielt die Gewichtung bei der Auswahl derjenigen Sensordaten, die für das Objekt wesentlich sind?
- Welche Rolle spielt die Bewertung einer speziellen Situation bzw. die Wahrscheinlichkeit von Ereignissen bei der Erkennung?
- Welchen Beitrag leistet die Funktion oder der Zweck eines Objektes zu dessen Identifikation?
- Welche Rolle spielen die geometrischen Eigenschaften und Merkmale bei dessen Lokalisierung?

1.2 Erkennung von Objekten in Bildern

Die Erkennung von Objekten läßt sich wie in 1.1 beschrieben in zwei Gebiete einteilen:

- Abgrenzen der Objekte von ihrer Umwelt
- Klassifizieren von Objekten

1.2.1 Abgrenzen der Objekte von ihrer Umwelt

Eine wichtige und grundlegende Komponente der Objekterkennung ist das Abgrenzen der Objekte, die sogenannte Segmentierung. Hier werden einzelne Pixel in einer Bildmatrix als Teile von verschiedenen Objekten bzw. Bereichen einer Szene identifiziert.
Objekte können durch verschiedene physikalische Gesetzmäßigkeiten gekennzeichnt sein und somit aus einem Bild herausglöst werden[2]:

- **Formkonturen**
  Kontinuierliche Kanten markieren meist die Grenze eines Objekts, das andere Gegenstände oder einen Hintergrund verdeckt. Solche Konturen hängen in der Regel mit abrupten diskontinuierlichen Tiefenänderungen zusammen. Dazu gehören aus Bewegungen resultierende, parallele Verschiebungen. Kontinuierliche Kanten, abrupte Tiefenänderungen und Bewegungsparallaxen grenzen einen Gegenstand von seiner Umgebung ab.
- **Textur**
  Die Textur einer Oberfläche ist durch Bereiche ähnlicher Helligkeitsdichten gekennzeichnet. Es kommen keine Tiefenänderungen vor.
- **Helligkeit**
  Besondere Helligkeitseffekte entstehen besonders beim Zusammentreffen mehrerer Oberflächen in einer Ecke. Die Helligkeitsunterschiede auf den verschiedenen Oberflächen sind graduierlich und nicht abrupt.

- **Schatten**
  Schatten sind auch kontinuierliche Kanten, gehen jedoch nicht mit abrupten Tiefenänderungen einher.

- **Bewegung**
  Sich parallel bewegende Teile gehören meist zum gleichen Objekt.

Textur geht über Helligkeitseffekte und Schatteneffekte hinweg, nicht jedoch über Formkonturen.
Es ist schwierig die Segmentierung als einen generellen Lösungsansatz zu implementieren, da die Auswahl der angeführten Gesetzmäßigkeiten typischerweise sehr problemabhängig sind [3].

1.2.2 Vorverarbeitung der Objekte

Vor der Klassifizierung wird eine Vorverarbeitung durchgeführt. Dabei werden wichtige Eigenschaften von Objekten extrahiert und in einem n-dimensionalen Raum, den sogenannten Feature Raum (oder Eigenschafts-Raum), eingetragen [3], [4]. Methoden zum Klassifizieren von Objekten setzen auf diesen Feature Raum auf und versuchen anhand der darin enthaltenen Daten Objekte einzuteilen.
Jede verwendete Eigenschaft eines Objektes wird auf eine Achse im Feature Raum abgebildet. Je größer die Anzahl der verwendeten Eigenschaften (Features) ist, desto besser können die Objekte im allgemeinen differenziert werden, aber die Performance (Geschwindigkeit) sinkt mit der Anzahl der verwendeten Features [5]. Durch eine vorangehende Reduzierung des Feature Raums (Veringerung der Dimension des Raumes) auf die jeweils wesentlichsten Eigenschaften kann man die Komplexität verringern und die Geschwindigkeit erhöhen. W. Siedlecki und J. Sklansky [3] beschreiben genetische Ansätze um die Dimension des Feature Space zu reduzieren.
Wesentliche Eigenschaften sind vom Anwendungsfall abhängig. So ist z.B. die Verwendung der Objekthelligkeit als Feature für die Objektklassifizierung in einer Umgebung mit konstanter Helligkeit sinnvoller als in Umgebungen mit veränderlichen Beleuchtungsverhältnissen.
Eine gängige Methode zur Extraktion von Features ist die Fourier Transformation [6]. Dabei werden aus dem Bild des Objektes Frequenz und Phaseninformationen gewonnen. Je nach Anwendungsfall kann man dadurch erreichen das z.B. beim Verwenden der Frequenzinformation und Weglassen der Phaseninformation ein Objekt auch bei einer verschobenen Abbildung die selbe Position im Feature Space hat [4]. Weitere Feature-Extraktionsverfahren sind z.B. Moment Feature Space [4], Chord Distribution Feature Space, die Hough Transformation [3], Wigner Distribution Feature Space [3], Orthogonal Transformations [4] und Gabor Funktionen [3].
Bei verschiedenen Feature-Extraktionsverfahren lassen sich neuronale Netze einsetzten. So beschreibt z.B. Kulkarni [4] Modelle künstlicher neuronaler Netze (Artificial Neural Networks, ANN) für die Feature-Extraktion via Fourier Transformation und via Walsh-Hadamard Transformation.
Im Feature Space wird nun jedem Objekt aufgrund seiner Eigenschaften ein eindeutiger Punkt zugeordnet. Je geringer die Distanz zwischen zwei benachbarten Punkten im Feature Space ist desto ähnlicher sind die Eigenschaften der Objekte die sie repräsentieren. Das bedeutet aber im allgemeinen nicht, daß benachbarte Objekte unbedingt der selben

Klasse angehören müssen [5]. Die Feature Space Abbildungen einer Klasse von Objekten befinden sich in einem Teilraum des Features Space. Im Idealfall überschneiden sich die Teilräume verschiedener Objektklassen nicht und jedes Objekt lässt sich direkt zu dem entsprechenden Teilraum und damit einer eindeutigen Objektklasse zuordnen [5].
Klassifizierungstechniken unterteilen anschließend den Feature Space in diese Teilräume und schließen durch den, dem Punkt zugeteilten Teilraum auf die Klasse des Objektes (z.B. [4],[7]).

1.2.3 Klassifizieren von Objekten

Die konventiellen Klassifizierungstechniken für Objekte lassen sich in Anlehnung an [4] grob in zwei Kategorien eingeteilt werden:

1. **Statistische Methoden**
   Verwendet statistische Eigenschaften im Objekt-Eigenschafts-Raum (Feature Space) um Objekte zu klassifizieren.

2. **Syntaktische Methoden**
   Verwendet eventuell wichtige, strukturelle Beziehungen zwischen Eigenschaften um Objekte zu klassifizieren.

Weiters kann man die Klassifizierungstechniken auch in

1. **Überwachte (supervised) Methoden**
   Hier steht für jede Klasse eine Anzahl von Übungsbeispielen zur Verfügung, welche verwendet werden um das System zu trainieren

2. **Nicht-überwachte (unsupervised) Methoden**
   Bei nicht überwachten Verfahren stehen keine Beispiele zur Verfügung. Das System versucht die Daten selbst zu klassifizieren.

einteilen. Beispiele für die Verwendung von überwachten und nicht-überwachten Methoden in den konventionellen Klassifizierungstechniken (nicht ANN-basierende Verfahren) finden sich in [7, 58 ff., 109 ff.]

Anhang A

Protokolle

A.1 Besprechungsprotokolle

A.1.1 09.03.1999

Teilnehmende Personen: Schwaiger Roland
Obermair Christoph
Schmidbauer Josef
Stieglbauer Gerald

Protokollführer: Stieglbauer Gerald
Ort: Büro von Schwaiger Roland
Zeit: 17.55 - 18.10

Einführende Besprechung über das Projekt und erste Aufgabenstellung

Spezifikation des abzugebenden Projektberichtes:

- Zu verwendendes Satzsystem: LaTeX
- Eigene Titelseite
- Abstract (Übersicht)
- Einleitung (Um was geht es?)
- Literaturübersicht
- Literaturverzeichnis

Abgabe der Projektberichte jeweils bis Montag 13.00 Uhr

Erste Fragestellung:

- Was ist Object Recognition?

A.1.2 16.03.1999

Teilnehmende Personen: Schwaiger Roland
Obermair Christoph
Schmidbauer Josef
Stieglbauer Gerald

Protokollführer: Schmidbauer Josef
Ort: Büro von Schwaiger Roland
Zeit: 17.45 - 18.05

Vorzunehmende Verbesserungen am Paper:

- Angaben auf der Titelseite um Lehrveranstaltung, Version, Zeit, Gruppe, E-Mail-Adressen und Lehrveranstaltungsleiter ergänzen
- Abstract hinzufügen
- Einführung von Abkürzungen für häufig verwendete Begriffe
- Angaben im Protokoll um Zeit, Ort und teilnehmende Personen ergänzen
- Abgabe via E-Mail mit Subject „Bioinformatik, Gruppe 3“

Aufgaben bis zum 23.03.1999:

- Suche nach Beschreibung für „Object Recognition“ in anderen Büchern
- Suche nach weiteren Verfahren zur Objekterkennung
- Ausführlichere Beschreibung der statistischen Objekterkennung
- Warum ist die statistische Objekterkennung so wichtig?
- Stufen der Vorverarbeitung (Pre-Processing) eines Bildes,
  Analyse in einem anderen Abbildungsraum

A.1.3 23.03.1999

Teilnehmende Personen: Schwaiger Roland
Obermair Christoph
Schmidbauer Josef
Stieglbauer Gerald

Protokollführer: Stieglbauer Gerald
Ort: Büro von Schwaiger Roland
Zeit: 17.45 - 18.10

Vorzunehmende Verbesserungen am Paper:

- E-Mail-Adressen auf Titelseite unter die Namen schreiben

- Datum als Versionsnummer

Aufgaben bis zum 13.04.1999:

- Abgrenzung der Objekte von der Umwelt, Image Segmentation
- Was sind überwachte Methoden bei dem statistischen Verfahren? (Beispiel)
- Ansätze für die Transformation in den Feature-Raum
- Was sind wichtige Eigenschaften bei der Ermittlung der Features von Objekten?
- Was bedeuten Nachbarschaftsbeziehungen im Feature-Raum?

Literaturverzeichnis

[1] W. E. L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints.* The MIT Press, 1990.

[2] G. Kleiter, *Skriptum zur VL Allgemeine Psychologie I.* Institut fuer Psychologie, Universitaet Salzburg, 1998.

[3] C. H. Chen, L. F. Pau, and P. S. P. Wang, *Handbook of Pattern Recognition and Computer Vision.* World Scientific Publishing, 1993.

[4] A. D. Kulkarni, *Artificial Neural Networks for Image Understanding.* Van Nostrand Reinhold, 1994.

[5] C. M. Bishop, *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, fourth ed., 1997.

[6] E. Herter and W. Loercher, *Nachrichtentechnik.* Carl Hanser Verlag, seventh ed., 1994.

[7] R. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches.* John Wiley and Sons Inc., 1992.