

Kurzfassung der Diplomarbeit

Thema:

Erkennen und Lokalisieren von
Objekten in komplexer Umgebung
anhand von Videobildern

Schlagwortkatalog:

Digitale Bildverarbeitung, Objekterkennung, Modellbasiert,
Hausdorff-Distanz, Korrelationsmethode, Positionsbestimmung

1 Ziel

Im Projekt DAVID wird die Objekterkennung verwendet, um ein Zielobjekt im Raum zu lokalisieren, so daß es vom Roboter aufgenommen und an seinen Bestimmungsort gebracht werden kann. Dabei soll das System in der Lage sein, ohne großen Aufwand, also möglichst nur durch einen veränderten Aufruf, nach verschiedenartigen Objekten zu suchen.

2 Methodik zur Objekterkennung

Zur Objekterkennung wurde das auf der Hausdorff-Distanz basierende Verfahren verwendet. Grundidee dabei ist, eine Distanz zwischen zwei Punktmengen zu definieren. Übertragen auf das Problem der Objekterkennung sind diese Punktmengen zum einen das kantenextrahierte Abbild der zu untersuchenden Szene und zum anderen die Kantenpunkte eines gegebenen Referenzmodells. Ziel bei der Suche ist, diese sogenannte Hausdorff-Distanz zwischen Bildpunkten und Modellpunkten zu minimieren. Kann eine Transformation der Modellpunkte so gefunden werden, daß der Abstand zu den Punkten des Bildes innerhalb tolerierbarer Grenzen liegt, besteht ausreichende Ähnlichkeit zwischen Modell und Bild und das Objekt gilt als gefunden.

Die Hausdorff-Distanz in ihrer ursprünglichen Definition ist als Maßstab für die Übereinstimmung zwischen Modell und Bild zu restriktiv. Daher wird eine Erweiterung der Hausdorff-Distanz verwendet, die als Korrelationskriterium für eine Objekterkennung genügend flexibel ist.

3 Entwickelte Objekterkennung

3.1 Modelldatenbank

Eine einzelne Ansicht eines Referenzmodells für die Objektsuche ist durch einen Array dargestellt, in dem alle Kantenpunkte dieser Ansicht abgelegt sind. Für jedes zu suchende Objekt wird eine Modelldatenbank angelegt, in der alle stabilen Ansichten eines Referenzmodells gespeichert sind, die für eine Suche relevant sind. Durch die Angabe, welche Modelldatenbank bei der Objektsuche verwendet werden soll, läßt sich dann das

gesamte System leicht auf unterschiedliche Modelle anwenden. In einer Datenbank können auch mehrere Modell abgelegt sein, nach denen dann gleichzeitig gesucht wird.

3.2 Bildaufbereitung

Wie schon in Abschnitt 2 beschrieben, ist das Ausgangsbild für die Objekterkennungsroutine die Kantenextrahierte der Originalaufnahme. Zur Reduzierung des Kantenbildes auf wesentliche Teile wird zusätzlich eine Farbklassifizierung durchgeführt. Aus der Verknüpfung der Kantenextraktion mit der Farbklassifizierung resultiert ein reduziertes Kantenbild, das nur noch die Kanten enthält, deren Bezug zum gesuchten Objekt wahrscheinlich ist. Die Schritte im einzelnen sind wie folgt.

3.2.1 Bildglättung

Durch die Bildglättung führt eine spätere Kantenextraktion zur Reduzierung des Anteils falscher Kanten im Bild. Verwendet wird dafür ein 3×3 -Medianfilter.

3.2.2 Farbklassifizierung

Das geglättete Originalbild wird einer Farbklassifizierung unterzogen. Diese fest dimensionierte unüberwachte Klassifizierung bildet Segmente, deren Färbung mit der des gesuchten Objekts ausreichend gut übereinstimmt. In Bereichen, die der klassifizierten Farbe nicht zugeordnet werden können, wird die Existenz des gesuchten Objekts ausgeschlossen.

3.2.3 Kantenextraktion

Die zur Objektsuche benötigte Kanteninformation wird ebenfalls aus dem geglätteten Originalbild gewonnen. Zur Extraktion wird ein 3×3 -Sobeloperator mit besonderer Gewichtung der Mittellinie verwendet. Die Ableitung des Bildes erfolgt jeweils in horizontaler und vertikaler Richtung. Aus den Ableitungen in x - und y -Richtung wird der Gradientenbetrag ermittelt und als Intensität für den berechneten Bildpunkt abgespeichert.

3.2.4 Kombination von Farbklassifizierung und Kantenextraktion

Um aus dem Kantenbild die Bereiche zu extrahieren, in denen die Existenz des gesuchten Objekts wahrscheinlich ist, wird das Ergebnis der Farbklassifizierung mit dem der Kantenextraktion verknüpft. Alle Kantenpunkte, die innerhalb einer festgelegten Entfernung d um ein Farbsegment liegen, werden in das Ausgangsbild übernommen. Durch die Erweiterung der Farbsegmente werden die Auswirkungen von Fehlern erster Art bei der Farbklassifizierung minimiert, in der Regel sogar eliminiert. Ergebnis ist ein reduziertes Kantenbild, das nur noch die Kanten enthält, deren Zugehörigkeit zum gesuchten Objekt wahrscheinlich ist.

3.2.5 Distanztransformation

Für eine effiziente Berechnung der Hausdorff-Distanz während der Objektsuche wird aus dem reduzierten Kantenbild im voraus die Distanztransformierte gebildet und diese anstatt

des reduzierten Kantenbildes als Ausgangsbild verwendet. Die anschließende, immer wiederkehrende Bestimmung von Punktabständen wird dadurch erheblich beschleunigt.

3.3 Objektsuche

Gesucht wird die Transformation der Modellpunkte, die diese mit minimaler Hausdorff-Distanz zu den Bildpunkten anordnet. Gültige Transformationen sind die Translation in x - und y -Richtung und die Skalierung der Modellpunkte von ihrer ursprünglichen Größe bis hin zu einer Verkleinerung auf 30% getrennt für x und y . Daraus ergibt sich ein vierdimensionaler Transformationsraum für die Objektsuche.

Im gesamten Suchraum werden die Transformationen ermittelt, deren *partial forward distance* von Modell- zu Bildpunkten innerhalb der tolerierten Grenzen liegen. Für diese Transformationen wird die *partial box-reverse distance* berechnet. Ist auch sie ausreichend klein, so gilt das Objekt an der durch die Transformation festgelegte Stelle als gefunden.

Eine erschöpfende Suche im kompletten Transformationsraum ist aus Effizienzgründen nicht praktikabel. Für die Suche existieren jedoch optimierte Verfahren, durch die eine Suche im Raum teilweise erheblich beschleunigt wird. Für die Bestimmung der Objekthypothesen, also den Transformationen, die die Kriterien für die *partial forward distance* erfüllen, wird die *cell decomposition method* verwendet. Die Berechnung der *partial box-reverse distance* wird durch *early rejection* beschleunigt.

Ist eine Objekthypothese durch die Verifikation bestätigt worden, so werden die Anforderungen an die Güte der Korrelation verschärft. Daraus ergibt sich, daß als Ergebnis nur die Transformation ausgegeben wird, die zur besten Übereinstimmung zwischen Modell und Bild geführt hat.

3.4 Performanz

Die Bearbeitungsdauer für den gesamten Objekterkennungsprozeß, also die Bildaufbereitung und die Suche nach dem Objekt selbst, liegt im Minutenbereich. Je nach Komplexität des Modells und der Menge von Kanten im reduzierten Kantenbild also zwischen 30 Sekunden und 3 Minuten.

4 Objektlokalisierung

Dem Referenzobjekt wird ein Referenzpunkt zugeordnet, der in einer bekannten Ebene im Raum liegen muß. Beispiele für eine solche Ebene sind der Fußboden und eine Tischoberfläche in bekannter Höhe. Einzige Einschränkung für die Referenzebene ist, ihr Normalenvektor darf auf keinem Sichtstrahl der Kamera senkrecht stehen.

Aus dem Ergebnis des Objekterkennungsprozesses kann der Sichtstrahl aus der Kameraperspektive abgeleitet werden, auf dem dieser Referenzpunkt liegt. Wird die Orientierung der Kamera mit einbezogen, so ergibt sich eine Gerade im Raum. Aus dieser Geraden und der Kenntnis über die bekannte Ebene, in der sich der Referenzpunkt ebenfalls befinden muß, kann die Position dieses Referenzpunktes im Raum ermittelt werden.

Die Information über die Lage des gefundenen Objekts im Raum liefert die Objekterkennung.