

Farbbildverarbeitung für fußballspielende Roboter

Thomas Erdmann und Volker Rehrmann
Labor Bilderkennen
Universität Koblenz-Landau
Rheinau 1, 56075 Koblenz
{terdmann, volker}@uni-koblenz.de

Kurzfassung

Seit einigen Jahren wird versucht, Robotern das Fußballspielen beizubringen. Eine Grundvoraussetzung dafür ist die sichere Erkennung und Lokalisierung des Balles und der Tore mittels Farbbildverarbeitung. Da viele bestehende Systeme diese Aufgaben nur ungenügend lösen, stellen wir in dieser Arbeit einen echtzeitfähigen und sehr robusten auf einer Farbsegmentierung basierenden Lösungsansatz für dieses Problem vor.

1 Einleitung

Die Idee, mit Robotern Fußball zu spielen und mit verschiedenen Teams gegeneinander anzutreten, entstand vor einigen Jahren in der KI-Gemeinde. Hintergrund für diese Idee ist es, die Methoden und Werkzeuge der verschiedenen Arbeitsgruppen in einem Wettbewerb aneinander zu messen und so die leistungsfähigsten Verfahren zu ermitteln und dadurch dem Gebiet der Robotik insgesamt einen Schub nach vorne zu geben. Ein weiterer, angenehmer Nebeneffekt liegt in der sehr großen Attraktivität dieses spielerischen Umgangs mit Wissenschaft, sowohl für die Öffentlichkeit als auch insbesondere für die an dem Projekt beteiligten Studenten. Seit drei Jahren werden regelmäßig Welt- und Europameisterschaften ausgetragen, RoboCup genannt, mittlerweile unter sehr großem Medieninteresse.

Die Spielqualität der Roboter ist auch heute noch auf einem erstaunlich niedrigen Niveau. Der Grund dafür ist vor allem in der nur schlecht funktionierenden Bildverarbeitung der Roboter zu sehen. Ohne eine sichere Lokalisation des Balles und der Tore kann sich keine vernünftige Spielweise entwickeln. Bisher haben fast ausschließlich KI-Gruppen an den Roboterteams gearbeitet und kaum Bildverarbeitungsspezialisten. Die Regeln des RoboCup sind dabei sehr entgegenkommend aufgestellt. Die wichtigen visuellen Objekte besitzen klare deutlich voneinander unterscheidende Farben. Der Ball ist leuchtend orange, das Spielfeld grün und die Tore blau bzw. gelb. Trotz dieser deutlichen Farbgebung versagen viele Roboter bereits bei der Aufgabe, den Ball sicher zu erkennen. Dies war für uns Anlaß, ein schnelles Farbbildverarbeitungssystem für den RoboCup zu entwickeln. Das System basiert auf den Erfahrungen und bereits bestehenden Programmpaketen aus diversen Farbbildprojekten.

Wir analysieren in dieser Arbeit die Schwächen der bisherigen Systeme und stellen unseren Lösungsansatz vor. Die Anforderungen an die Farbbildverarbeitung scheinen hier auf den ersten Blick gering zu sein, was sich allerdings bei näherem Blick nicht ganz bestätigt. Als echte *Real-World* Anwendung muss das System zwei wesentliche Anforderungen erfüllen: Effizienz und Robustheit. Da Fußball ein schnelles Spiel ist, müssen ca. 10 Bilder pro Sekunde verarbeitet werden. Um diese Frequenz zu erreichen, wird in den bisherigen Systemen fast ausschließlich mit einfachen Farbklassifikationsalgorithmen gearbeitet, die aber nur unter konstanten Lichtverhältnissen robust arbeiten. Da die Lichtverhältnisse

beim RoboCup aber, wie im folgenden Kapitel beschrieben, nicht konstant sind, arbeitet dieses Verfahren nur sehr unzuverlässig.

Unser Lösungsansatz basiert deshalb nicht auf Farbklassifikation, sondern auf Farbsegmentierung mit dem Color-Structure-Code (CSC), der sich bereits in vielen Projekten bewährt hat. Durch die gezielte Analyse der vom CSC gelieferten Farbregionen wird eine sehr robuste Erkennung auch unter wechselnden Beleuchtungsverhältnissen erreicht.

2 Analyse bestehender Ansätze

Durch die Regeln des RoboCups sind einige Randbedingungen für den Einsatz einer Farbbildverarbeitung für die fußballspielenden Roboter vorgegeben. Beim RoboCup werden verschiedene Ligen unterschieden, wobei in der sogenannten Middle-Size-Liga, die in dieser Arbeit betrachtet wird, die größten Herausforderungen an den Einsatz einer Bildverarbeitung gestellt werden [1]. In dieser Liga befinden sich immer fünf Roboter jeder Mannschaft auf einem ca. 5 m mal 9 m großen Spielfeld. Der zu erkennende Ball mit einem Durchmesser von 20 cm ist orange und die Tore sind blau bzw. gelb. Es werden vollkommen autonome mobile Roboter mit einem maximalen Durchmesser von 50 cm eingesetzt, in welche die gesamte Sensorik und die für die Steuerung und Informationsverarbeitung benötigten Rechner integriert werden müssen. Aus diesem Grund steht bei vielen Systemen nur eine begrenzte Rechenleistung zur Verfügung, weshalb in bestehenden Systemen häufig einfache Algorithmen eingesetzt werden, um die für das Roboterfußball erforderliche Echtzeitfähigkeit zu erreichen.

2.1 Farbklassifikation zur Objekterkennung

Das allen Ansätzen zugrunde liegende Verfahren ist die Farbklassifikation, die auch bei dem von vielen Teams eingesetzten kommerziellen Bildverarbeitungssystem Cognachrome zum Einsatz kommt [1]. Dabei wird jedes Pixel des digitalisierten Bildes anhand einer Tabelle (Lookup table, LUT) einer vorher definierten Farbklasse zugeordnet. Dazu werden zunächst die in der jeweiligen Anwendung relevanten Farben trainiert und in einer Tabelle abgelegt. In einem Durchlauf über das Bild kann dann jedes Pixel sehr effizient klassifiziert werden. Nun werden zusammenhängende Pixel einer Farbklasse zu Regionen verbunden, und es können verschiedene geometrische Merkmale, wie Fläche und Position, berechnet werden.

Die digitalisierten Bilder liegen üblicherweise als RGB-Werte mit jeweils 8 Bit vor. Daraus ergibt sich ein Farbraum mit über 16 Millionen Werten. Um nicht eine so große Tabelle anlegen zu müssen, wird häufig jeder Kanal in einer separaten Tabelle gespeichert, d. h. es lassen sich nur Quader im RGB-Raum beschreiben. Eine genauere Beschreibung erhält man durch die Verwendung der 2-dimensionalen Ebenen RG, RB und GB. Hiermit lassen sich auch komplexere Volumen im 3-dimensionalen Raum beschreiben. Eine Abwandlung dieses Verfahrens mit nur zwei Ebenen wird beim Cognachrome-System angewendet [8]. Ein Nachteil der Beschreibung der interessierenden Bereiche im RGB-Raum ist die fehlende Stabilität des RGB-Raumes gegen Helligkeitsänderungen.

Abhilfe schafft hier die Wahl eines anderen Farbraumes. Bei dem HSV- oder HIS-Farbraum ist der Farbkanal H (hue) invariant gegenüber Helligkeitsänderungen, so dass sich der interessierende Bereich im Farbraum leichter beschreiben lässt. Der Weltmeister von 1999 im RoboCup, das CS-Sharif Team, verwendet eine Farbklassifikation im HIS-Farbraum [2]. Zur Beschleunigung des Algorithmus wird eine Unterabtastung verwendet, die erst dann, wenn ein Pixel die gesuchte Farbe aufweist, auf das hochauflösende Bild zurückgreift.

2.2 Schwächen der bisherigen Systeme

Den beiden großen Vorteilen der im letzten Absatz beschriebenen Verfahren, der hohen Effizienz mit den daraus resultierenden geringen Anforderungen an die Hardware und der leichten Implementierbarkeit, stehen einige gravierende Nachteile gegenüber. Die größte Schwäche der bisher eingesetzten Systeme ist die mangelnde Robustheit gegenüber Beleuchtungsänderungen. Obwohl die Farben der zu erkennenden Objekte klar voneinander abgegrenzt und die Objekte einfarbig sind, erfolgt keine sichere Erkennung von Ball und Toren, da es zahlreiche störende Einflüsse auf die Lichtverhältnisse gibt. Dazu zählen Schattenwurf, Helligkeitsschwankungen innerhalb des Spielfeldes, Lichtreflexe und durch die Geometrie des Balles auftretende starke Variationen in der Helligkeit auch innerhalb eines Objektes. Zusätzlich kann es durch Sonnenlichteinfall zu Farbtemperaturschwankungen kommen.

Ein Problem bei dem Ansatz über die Farbklassifikation, das sich auch aus der mangelnden Robustheit gegenüber Beleuchtungsänderungen ergibt, liegt in dem hohen Aufwand zur Kalibrierung der Farbbildverarbeitung. An jedem neuen Einsatzort und nach jeder Veränderung der Lichtverhältnisse müssen zunächst aufwendig die verschiedenen Farbklassen eintrainiert werden. Jedes zu erkennende Objekt muss bei allen im Spielfeld auftretenden Helligkeiten eintrainiert werden. Selbst bei Wahl des HSV-Farbraumes ist dies ein zeitaufwendiger Prozess [7], der vor jedem Einsatz erfolgen muss, da sich die Lichtverhältnisse im Laufe des Tages ändern.

Ein weiterer Schwachpunkt liegt in der Wahl der Kamera. Bei allen Teams kommen Farbkameras mit S-VHS- oder sogar nur mit Composite-Ausgang zum Einsatz. Bei der Übertragung der Farbbilder werden dabei die Farbinformationen im Gegensatz zum RGB-Signal näherungsweise halbiert [3]. Zunächst erfolgt eine Umwandlung der RGB-Werte in das Luminanz- (Y) und die beiden Farbdifferenzsignale (U, V). Diese Umwandlung ist verlustfrei und reversibel. Bei der Übertragung über das S-VHS-Signal wird die Luminanz als eigenes Signal übertragen, während die Farbdifferenzen U und V zu einem gemeinsamen Farbsignal moduliert werden, wozu die Signale zuerst mit einem Tiefpass gefiltert werden müssen. Die dabei verlorene Information ist nicht mehr zu rekonstruieren.

Weiterhin fällt auf, dass bei vielen Systemen aus Effizienzgründen auf den Einsatz einer Formanalyse verzichtet wird. Sobald nun mehrere Objekte in eine Farbklasse fallen, was leicht passiert, da die Farbklassen sehr weiträumig gewählt werden müssen, kann keine eindeutige Entscheidung mehr getroffen werden.

2.3 Anforderungen an eine Farbbildverarbeitung für fußballspielende Roboter

Aus der Analyse der bestehenden Systeme können zwei Hauptanforderungen an die Farbbildverarbeitung hergeleitet werden. Das System muss echtzeitfähig und robust bei allen Beleuchtungsbedingungen sein. Sowohl starke Helligkeitsschwankungen als auch leichte Schwankungen in der Farbtemperatur sollten ausgeglichen werden können. Diese Anforderungen lassen sich von einer Farbklassifikation nicht erfüllen.

Zusätzlich wachsen die Ansprüche an die Leistungsfähigkeit der Farbbildverarbeitung bei den künftigen Wettbewerben dahingehend, dass sowohl die eigenen als auch die gegnerischen Roboter sicher erkannt werden sollen. Die Roboter werden dazu mit Farbmarken versehen. Die vorgesehenen Farben sind Hellblau und Magenta. Durch die starken Variationen der Helligkeit während eines Spieles, müssten die Farbklassen zur Farbklassifikation für die einzelnen Objekte sehr weiträumig gewählt werden. Dadurch würden die hellblauen Markierungen in die gleiche Klasse wie das blaue Tor fallen und ein Roboter vor dem blauen Tor wäre auch mit zusätzlicher Formanalyse nicht mehr zu detektieren. Abhilfe kann hier eine Farbsegmentierung schaffen, mit der auch solche Objekte noch zu erkennen sind.

3 Objekterkennung basierend auf Farbsegmentierung

Aus den im letzten Kapitel angesprochenen Gründen haben wir uns bei unserem Lösungsansatz für eine Farbsegmentierung entschieden. Bei der Farbsegmentierung werden zunächst, unabhängig von den gesuchten Farben, gleichfarbige Bereiche gesucht und zu Regionen zusammengefasst. Die gefundenen Regionen bekommen dann einen durchschnittlichen Farbwert zugeordnet und können klassifiziert werden. Der Vorteil der Farbsegmentierung ist darin zu sehen, dass eine gute Segmentierung bei Helligkeits- und selbst bei Farbtemperaturänderungen die gleiche Regioneneinteilung ergibt, nur die durchschnittlichen Farben verändern sich.

3.1 Hardware Architektur des Systems

Die zugrunde liegende Plattform bildet ein Pioneer 2-DX Roboter der Firma ActivMedia, der bei den letzten RoboCup-Meisterschaften von vielen Teams eingesetzt wurde. Der Pioneer 2 ist ein autonomer Roboter der Größe 44 cm x 38 cm x 22 cm. Zur Motoransteuerung und zur Abfrage der eingebauten Sensoren dient ein auf einem Siemens C 166 Prozessor basierendes Mikrocontrollerboard, das über eine RS-232-Schnittstelle angesprochen werden kann.

Diese Grundversion des Roboters wurde von uns um geeignete Komponenten zur Bildverarbeitung und Steuerung erweitert, wie in Abbildung 1 zu sehen ist. Zum Einsatz kommt eine kompakte 1-CCD-Kamera mit RGB-Ausgang, um keine Verluste bei der Übertragung der Bilder durch den Videostandard zu haben, und ein 3,6 mm Objektiv, um einen großen Ausschnitt des Spielfeldes überwachen zu können. Unser System besteht aus einem kompakten PC mit einem Pentium III

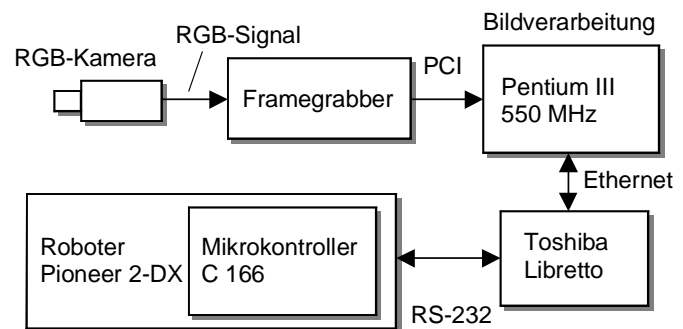


Abbildung 1: Systemaufbau

III Prozessor mit 550 MHz, der mit einem RGB-Framegrabber ausgestattet ist. Dieser PC wird ausschließlich als Bildverarbeitungsrechner verwendet, um auch komplexere Algorithmen in Echtzeit ausführen zu können und läuft unter dem Betriebssystem Windows NT. Zur Informationsverarbeitung und Steuerung des Roboters verwenden wir ein Toshiba Libretto 100 Notebook mit dem Betriebssystem Linux. Abbildung 3 zeigt eine farbige Darstellung des Gesamtsystems.

3.2 Farbsegmentierung mit dem Color Structure Code

Der CSC (Color Structure Code) ist ein bei uns entwickeltes Verfahren zur Farbsegmentierung, das sich bereits in mehreren Projekten sehr gut bewährt hat, insbesondere bei der Erkennung von Verkehrszeichen aus einem fahrenden Fahrzeug [4]. Der CSC zeichnet sich dadurch aus, dass er äußerst zuverlässige Ergebnisse in natürlichen Szenen liefert und dabei noch sehr effizient arbeitet. Dies wird erreicht durch eine neuartige Kombination von lokaler und globaler Information und einem speziell entwickelten Farbähnlichkeitskriterium im HSV-Farbraum. Damit werden typische Variationen in der Helligkeit und Sättigung in bunten Regionen besser toleriert. Das Ergebnis der Segmentierung mit dem CSC bleibt auch bei sich ändernden Beleuchtungsbedingungen weitgehend stabil. Details zum CSC finden sich z. B. in [5] oder [6].

3.3 Anwendungsorientierte Auswertung

Das Ziel der Auswertung ist die Erkennung aller relevanten Objekte. Nach der Segmentierung durch den CSC liegt eine Liste von Farbreionen vor, die zur Bedeutungszuordnung weiteren Analysen unterzogen wird. Direkt nach der Segmentierung stehen Informationen über die durchschnittliche Farbe, die Fläche und die Position zur Verfügung. Da es mehrere Objekte einer Farbe geben kann, müssen weitere Kriterien herangezogen werden, um ein Objekt sicher zu erkennen. Für die Erkennung des Balles ist beispielsweise die Kreisförmigkeit eine wichtige Eigenschaft. Im folgenden wird ein schnelles und robustes Verfahren zur Formanalyse vorgestellt, das auf einer Approximation der konvexen Hülle basiert.

Eine gute Approximation der konvexen Hülle ist ein Polygon P_{24} , welches das Objekt vollständig umschließt, mit 24 Seiten e_i , deren Summe minimal ist. Alle benachbarten Seiten sind dabei um jeweils 15 Grad gegeneinander verdreht. Die Seiten 0 und 12 verlaufen parallel zur X-Achse eines kartesischen Koordinatensystems (siehe Abbildung 2). Dieses Polygon ermöglicht eine einfache und schnelle Formanalyse [4]. Für einen idealen Kreis sind die Seitenlängen alle gleich. Die Güte eines Kreises ergibt sich damit zu

$$P_{circle} = \frac{1}{1 + \frac{\sum_{i=0}^{23} (\mu - e_i)^2}{100\mu^2}}, \text{ mit } \mu = \frac{1}{24} \sum_{i=0}^{23} e_i.$$

Für ein Rechteck, dessen Seiten parallel zum Koordinatensystem ausgerichtet sind, ergibt sich die Güte durch das Verhältnis der Seiten zueinander:

$$P_{rectangle} = \left(\frac{\sum_{i \in R} e_i}{\sum_{i=0}^{23} e_i} \right)^2$$

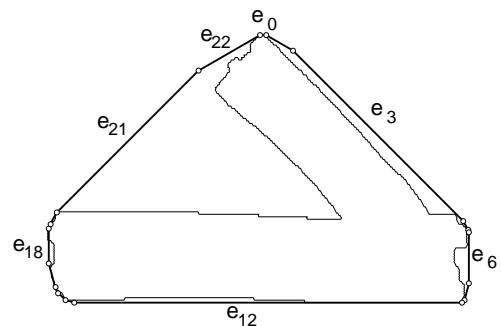


Abbildung 2: 24-seitiges Polygon

mit $R = \{0, 1, 5, 6, 7, 11, 12, 13, 17, 18, 19, 23\}$. Die Seiten in R sind dabei die Hauptseiten des Rechtecks P_{24} (0, 6, 12, 18) und alle Nachbarseiten (23, 0, 5, 7, 11, 13, 19). Dadurch ist die Erkennung robust gegenüber leichten Verdrehungen des Rechteckes. Andere Formmerkmale lassen sich auf ähnliche Art berechnen.

Für die weitere Verarbeitung wird für alle Regionen, in Abhängigkeit von der Farbe, die Güte der erwarteten Form des Objektes bestimmt. Für orange Objekte wird z. B. eine Kreis- oder Halbkreisform erwartet. Durch starke Schattenbildung, die insbesondere beim Ball, begünstigt durch die geometrische Form, auftritt, kann das Objekt bei der Segmentierung in mehrere Regionen zerfallen (siehe Abbildung 5). Benachbarte Regionen ähnlicher Farbe werden deshalb zusammengefaßt und anschließend einem Formtest unterzogen. Ergibt sich dabei eine besserer Formgüte als beim Test einzelner Regionen, werden die Bereiche zusammengesetzt und für die Weiterverarbeitung als ein Objekt behandelt.

In Abbildung 4 ist der prinzipielle Ablauf bei der Auswertung dargestellt. Nach dem Formtest wird ein zusätzlicher Plausibilitätstest durchgeführt. Je nach Objekt werden dabei Nachbarschaftsrelationen, Größe oder mögliche Positionen im Bild überprüft. Wird mehr als eine Region gefunden, die der Modellvorstellung eines Objektes entspricht (z. B. orange und kreisförmig), liefert die Methode mehrere Objekte mit dazugehörigen Wahrscheinlichkeiten zurück, die mit dem vorhandenen Wissen, beispielsweise über die Position im letzten Bild, weiterverarbeitet werden können. Ausgehend von einer Liste von Farbreionen wird also eine Liste der erkannten Objekte mit den Wahrscheinlichkeiten für die Zugehörigkeit zu einer bestimmten Objektklasse erzeugt. Abbildung 6 zeigt eine aufgenommene Szene aus der Sicht des Roboters, das Segmentierungsergebnis durch den CSC und die in der Szene erkannten Objekte.

4 Ergebnisse und Zusammenfassung

Das von uns entwickelte System besteht aus einer 1 CCD Farbkamera mit RGB-Ausgang und einem kleinen, kompakten PC mit einem Pentium III Prozessor mit 550 MHz. Als Basis-Software setzen wir den bereits in vielen Projekten bewährten CSC (Color-Structure-Code) ein. Der CSC segmentiert Farbbilder sehr effizient und liefert insbesondere in natürlichen Szenen robuste Ergebnisse durch die Nutzung eines Farbähnlichkeitskriteriums im HSV-Farbraum. Die vom CSC gelieferten Farbregionen bilden die Basis für die weiteren Analysen. So wird z.B. gezielt nach orangen, kreisförmigen Regionen gesucht. Im Gegensatz zur Farbklassifikation kann hier die Farbe orange großzügiger definiert werden. Zerfallene, farblich ähnliche Regionen können auf dieser Ebene wieder zusammengefaßt werden. Ähnlich wird bei der Suche nach den rechteckigen Toren mit den Farben blau und gelb vorgegangen.

Das System hat sich im Einsatz als sehr zuverlässig erwiesen. Auch unter stark wechselnden Beleuchtungsverhältnissen erfolgt durch den Einsatz der Farbsegmentierung eine sichere Ballerkennung. Es ist schnell adaptierbar, da zur Anpassung an wechselnde Bedingungen nur ein korrekter Weißabgleich erforderlich ist.

Zur Zeit werden auf dem Pentium III Prozessor mit 550 MHz bei einer Auflösung von 360 mal 254 Pixel etwa 8 Farbbilder pro Sekunde verarbeitet. Durch Reduzierung der Auflösung in x- und y- Richtung um den Faktor 2 werden sogar 19 Bilder pro Sekunde erreicht. Allerdings reicht bei weit entfernten Objekten diese Auflösung nicht aus, so dass bei Bedarf auf die höhere Auflösung umgeschaltet werden muss. Eine einfache Aufrüstung des Prozessors ließe jedoch auch bei voller Auflösung bereits Verarbeitungsfrequenzen von 12 Bildern pro Sekunde erreichen. Das System erfüllt also die beiden Anforderungen, Effizienz und Robustheit, einer Farbbildverarbeitung für fußballspielende Roboter, was sich leicht in kleinen Trainingsspielen demonstrieren lässt.

Literatur

- [1] M. Asada und H. Kitano (Eds.): *RoboCup-98: Robot Soccer World Cup II*. Lecture Notes in Artificial Intelligence, Springer-Verlag, Berlin, 1999.
- [2] M. Jamzad: CS-Sharif ROCS99 team in middle-sized robots league, *RoboCup-99 Team Descriptions*, Linköping University Electronic Press, S. 118-126, 1999.
- [3] C. A. Poynton: *A Technical Introduction to Digital Video*, John Wiley & Sons, Inc., New York, 1996.
- [4] L.Priese, J. Klieber, R. Lakmann, V. Rehrmann, R. Schain: New Results on Traffic Sign Recognition, *Proceedings of the Intelligent Vehicles Symposium*, S. 249-154, Paris, Okt. 1994.
- [5] V. Rehrmann: *Stabile, echtzeitfähige Farbbildauswertung*, Verlag Fölbach, Koblenz, 1994.
- [6] V. Rehrmann, L. Priese: Fast and Robust Segmentation of Natural Color Scenes, *3rd Asian Conference on Computer Vision*, Hongkong, Januar 1998.
- [7] W.-M. Shen et al.: DREAMTEAM 99: Team Description Paper, *RoboCup-99 Team Descriptions*, Linköping University Electronic Press, S. 146-149, 1999.
- [8] A. Wright et al.: *Cognachrome Vision System User's Guide*, Newton Research Labs, Renton, 1996.

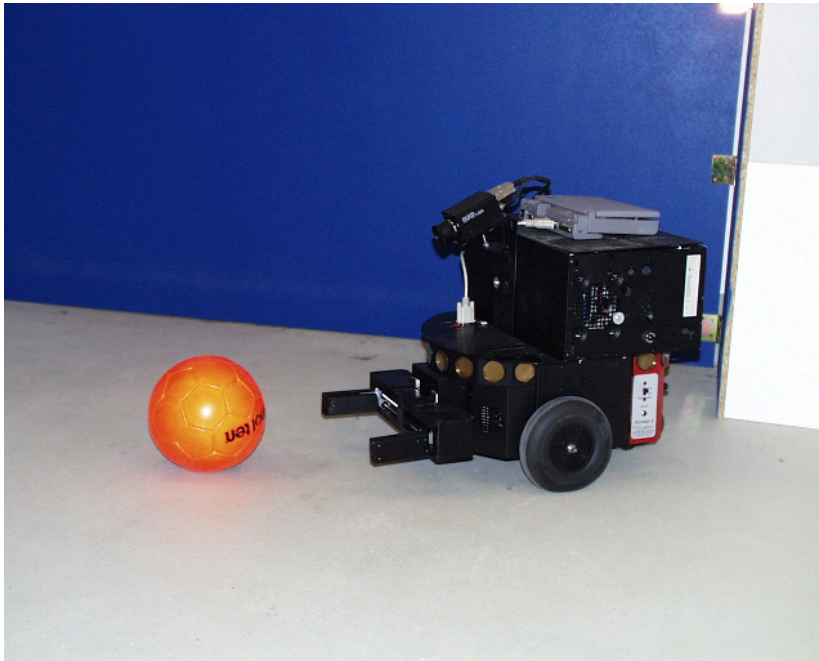


Abbildung 3: Unser autonomer Roboter mit aufgebautem PC, Kamera und Notebook

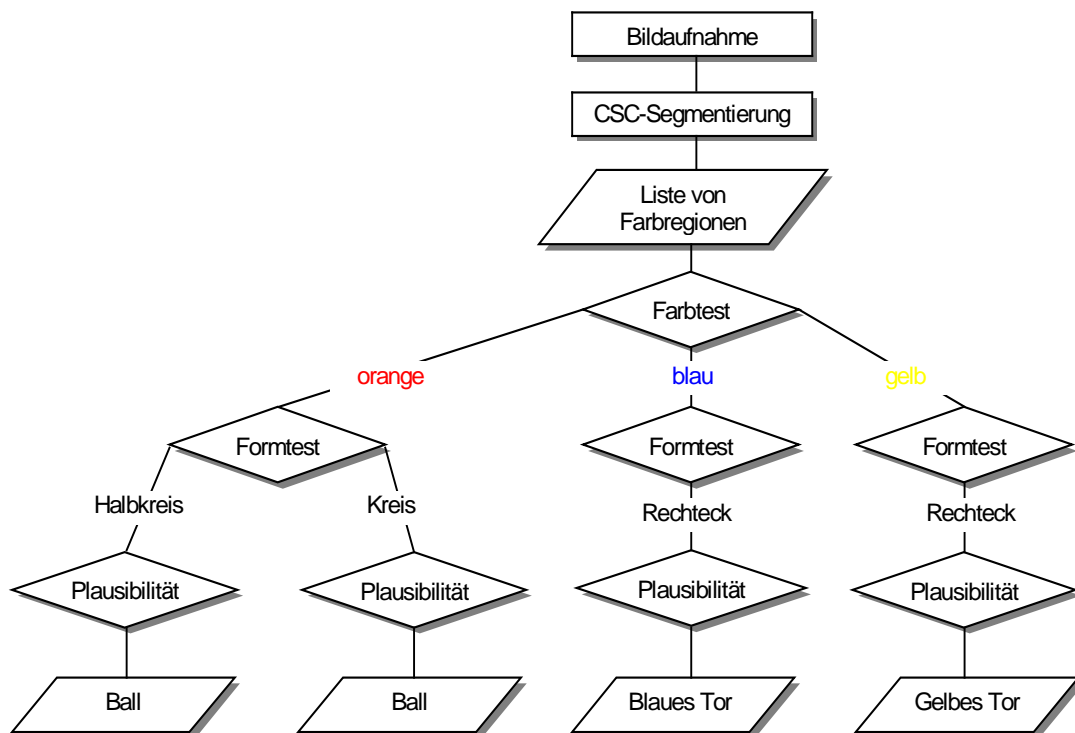


Abbildung 4: Prinzipieller Ablauf der Farbbildverarbeitung: Nach der Segmentierung mit dem CSC liegt eine Liste von Farbregionen vor, die in Abhängigkeit von der Farbe weiter analysiert werden.



(a)

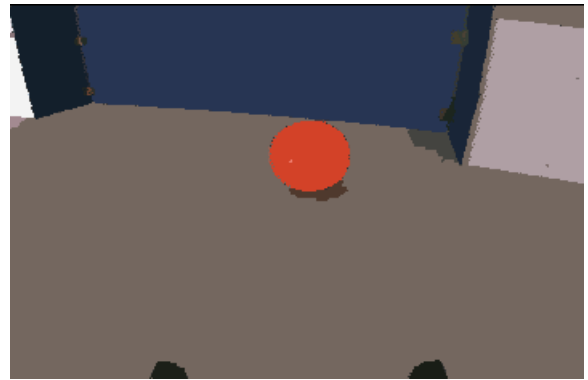


(b)

Abbildung 5: Durch die starke Schattenbildung auf dem Ball (a), zerfällt das Objekt nach der Segmentierung in zwei Regionen (b). Für die weitere Auswertung können benachbarte Regionen ähnlicher Farbe zusammengefasst werden.



(a)



(b)



(c)

Abbildung 6: Farbsegmentierung mit dem CSC: (a) Originalbild aus der Sicht der Roboterkamera; (b) Ergebnis der Segmentierung mit dem CSC; (c) die nach der Auswertung erkannten und lokalisierten Objekte „Ball“ und „Blaues Tor“