

Untersuchung alternativer Farbmerkmale für die visuelle Monte-Carlo-Selbstlokalisierung für mobile Roboter

Alexander König, Simon Daniel Moser, Horst-Michael Gross
Technische Universität Ilmenau, Fachgebiet Neuroinformatik
PF 100565, 98684 Ilmenau
{alexander.koenig,horst-michael.gross}@tu-ilmenau.de

Zusammenfassung

In Fortführung zu [1] wird gezeigt, wie die visuell-basierte Monte-Carlo-Selbstlokalisierung unter Verwendung von Farbhistogrammen noch robustere und genauere Lokalisierungsergebnisse liefern kann. Als Ausgangspunkt dient die Monte-Carlo-Lokalisierung (MCL) von Fox u.a. [2], ein samplebasiertes Verfahren zur Zustandsschätzung der Position und Orientierung eines mobilen Systems. Dieses in den letzten Jahren populär gewordene Verfahren, welches typischerweise abstandsmessende Sensoren verwendet, führt in unserer realen Umgebung (Baumarkt) zu sensorischen Mehrdeutigkeiten im Gangsystem und wurde daher so modifiziert, dass es mit Bildern einer omnidirektionalen Kamera arbeitet. Diese liefert durch die farbige Panoramaansicht der Umwelt ein deutlich signifikanteres Maß zur Charakterisierung der Umgebung. Das bisher entwickelte visuell-basierte MCL-Verfahren [1] verwendet dazu sehr einfache Merkmale (gemittelte RGB-Werte pro Bildsegment), die aus dem Kamerabild extrahiert werden. Damit wurden bislang bereits gute Lokalisierungsergebnisse erzielt. Probleme ergaben sich allerdings bei Beleuchtungsänderungen (Helligkeit und Farbtemperatur), was zu ungenaueren Positionsschätzungen bis hin zu Fehllokalisationen führte. Aus diesem Grund wurde nach szenenbeschreibenden Merkmalen gesucht, die robuster gegenüber veränderlichen Beleuchtungsbedingungen sind. Dazu wurden Merkmale auf Basis von Histogrammen in unterschiedlichen Farbräumen und Pixelklassifikationsmethoden experimentell untersucht und anhand der gewonnenen Ergebnisse ihre Tauglichkeit für die visuelle MCL bewertet. Die mit den histogrammbasierten Merkmalen erzielten Ergebnisse stellen eine deutliche Verbesserung im Vergleich zum bisherigen Ansatz dar. Die Verfahren und Ergebnisse werden im Beitrag beschrieben.

1 Einführung und Motivation

In unserem PERSES-Projekt [3] entwickeln wir einen Serviceroboter (PERSES, Abb. 1a), welcher als interaktiver Einkaufs-Assistent [4] innerhalb einer Markt-Umgebung dienen soll. Typische Dienstleistungen, die angeboten werden sollen, beinhalten beispielsweise das Führen des Kunden zu einem gesuchten Artikel oder das Begleiten als mobiles Informationsterminal. Die Akzeptanz eines solchen Systems hängt dabei entscheidend von einer möglichst intuitiven Interaktion mit den Kunden ab. Zur Lösung dieser Aufgabe werden insbesondere visuell-basierte Methoden eingesetzt, da diese einerseits in der Lage sind, viele Informationen über die Umgebung zu vermitteln, die mit anderen Sensoren nicht wahrgenommen werden können und andererseits zu einer möglichst natürlich wirkenden Kommunikation mit menschlichen Nutzern führen sollen. Zu diesem Zweck ist der Roboter unter anderem mit einem omnidirektionalen Kamerasystem ausgestattet, welches aus einer digitalen Farbkamera und einem sphärischen Spiegel besteht (Abb. 1b). Neben der Interaktion müssen eine Reihe von Navigationsaufgaben gelöst werden. Der Roboter muss eine Repräsentation seiner Einsatzumgebung (Umgebungsmodell) besitzen und in der Lage sein, bestimmte Punkte darin selbstständig anzufahren. Dazu muss er zu jedem Zeitpunkt Kenntnis von seiner eigenen Position innerhalb der Umwelt haben (Selbstlokalisierung).

Unser Lokalisationsverfahren basiert auf der in den letzten Jahren sehr populär gewordenen MonteCarlo-Lokalisierung nach Fox u.a. [2]. Hierbei werden in der Regel zur sensorischen Erfassung der Umwelt entfernungsmessende Systeme, wie z.B. Laser- und Ultraschallscanner, benutzt. Warum diese Vorgehensweise in

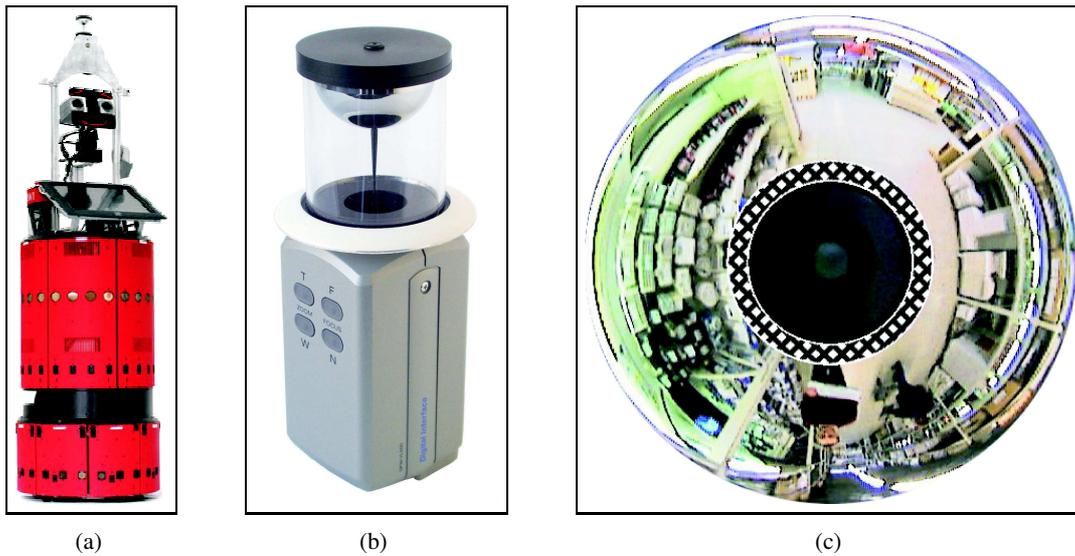


Abbildung 1: (a) Interaktiver Shopping Assistent PERSES (B21 Roboter); (b) omnidirektionale Kamera mit weißem Referenzring für eine referenzbasierte Blendenregelung und Farbumstimmung; (c) omnidirektionale Ansicht der Umwelt mit weißem Referenzring (schraffierter Bereich)

unserem Fall problematisch ist, kann man in Abb. 2 erkennen. Die für den interaktiven Shopping-Assistenten PERSES vorgesehenen Einsatzgebiete (Marktumgebungen) weisen oft sehr symmetrisch angeordnete Regalreihen mit gleichen Abständen auf. Für einen Laserscanner bedeutet das, dass er an sehr vielen Positionen im Markt die gleichen sensorischen Eindrücke aus der Umwelt erhält. Solche Mehrdeutigkeiten erschweren die Lokalisation des Roboters erheblich. Aus diesem Grund verwenden wir die schon erwähnte omnidirektionale Kamera, welche eine farbige Panoramaansicht der Umwelt liefert (Abb. 1c). Da sich die Regalreihen auf Grund ihrer Struktur und Befüllung vor allen farblich unterscheiden lassen, kann die Verteilung von Farbe in der Umgebung des Roboters zur Beschreibung der aktuellen Situation herangezogen werden. Im nächsten Abschnitt wird zunächst kurz die Monte-Carlo-Lokalisation vorgestellt. Hier erfolgt eine kurze Beschreibung der klassischen und der visuellen MCL, sowie ein Beschreibung des Gesamtsystem und der Funktionsweise. Danach wird ein Überblick über die verwendeten Verfahren zur Merkmalsextraktion und die experimentellen Untersuchungen geben. Als Merkmale kamen unterschiedliche Farbhistogramme in unterschiedlichen Farbräumen zum Einsatz. In den Experimenten sollte ihre Eignung für die Monte-Carlo-Lokalisation untersucht werden. Am Ende stehen die Zusammenfassungen und der Ausblick.

2 Monte-Carlo-Lokalisation

Allgemein: Die Monte-Carlo-Lokalisation schätzt die Wahrscheinlichkeitsverteilung für die tatsächliche Position des Roboters mittels eines **Partikel-Filters**. Dieser Partikel-Filter besteht aus einem Satz von N gewichteten Partikeln. Jedes dieser Partikel (oder auch Sample) s^i gibt durch den ihm zugeordneten **Importance Factor** w^i für eine bestimmte Position $\mathbf{x}^i = (x^i, y^i, \varphi^i)$ (auch **Zustand**) im Raum die Wahrscheinlichkeit an, dass sich der Roboter dort befindet. Bei initial unbekannter Position entspricht die zu approximierende Wahrscheinlichkeitsverteilung einer Gleichverteilung über alle möglichen Zustände des Roboters, daher werden alle Samples zufällig im Zustandsraum verteilt und erhalten den gleichen Importance Factor $\frac{1}{N}$. Während der Bewegung des Roboters werden die Positionen der Samples und ihrer zugehörigen Importance Factors schrittweise aus der vorherigen Verteilung neu berechnet. Dazu werden zunächst Samples mit niedriger Importance entfernt und in der Umgebung von Samples mit hoher Importance eingefügt. Die Sample-Verteilung adaptiert sich also derart, dass an den Positionen, die mit hoher Wahrscheinlichkeit als tatsächliche Position eingeschätzt werden, die Dichte der Samples erhöht wird. Als nächstes wird die gemessene Eigenbewegung

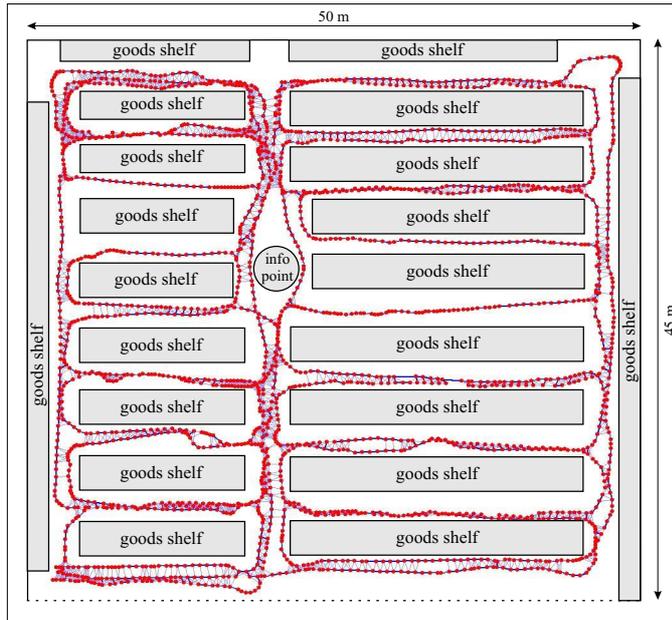


Abbildung 2: Grundriss des Marktes mit graphenbasiertem Umweltmodell; Knoten zwischen den Regalreihen stellen die gespeicherten Referenzansichten dar, mit deren Hilfe der zu erwartende Umwelteindruck für (fast) jede beliebige Position im Markt geschätzt werden kann. Das Umweltmodell wird während einer einmaligen Teach-Fahrt mittels Joystick unter Verwendung eines visuellen Odometrie-korrekturverfahrens [5] gelernt.

des Roboters auf die Samples übertragen.

Die Berechnung des Importance Factors für ein Sample erfolgt, indem der aktuelle Sensoreindruck o_t (z.B. Laser- oder Ultraschallscan) aus der Umgebung mit dem für das Sample s^i zu erwartenden Sensoreindruck o_t^i verglichen wird. Je besser die Erwartung mit der Realität übereinstimmt, desto größer wird das Gewicht w_t^i und damit die Wahrscheinlichkeit das sich der Roboter dort befindet. Für diesen Vergleich wird ein **Umgebungsmodell** benötigt, welches die dem Roboter bekannte Umwelt beschreibt und darüber zu jeder möglichen Position x^i die erwarteten Beobachtungen o_t^i (Sensoreindrücke) liefern kann.

Diese Berechnungen werden zyklisch wiederholt und nach und nach konzentrieren sich die Samples immer stärker an den wahrscheinlichsten Positionen. Durch Berücksichtigung von Eigenbewegung und Umweltbeobachtung wird, ausgehend von einer völlig unbekanntem Position, innerhalb weniger Schritte die tatsächliche Position des Roboters ermittelt. Die Verwendung des Partikel-Filters verringert den Zeitbedarf gegenüber anderen Verfahren und ermöglicht gleichzeitig durch die hohe Auflösung an interessanten Punkten eine sehr hohe Genauigkeit.

Visuell: Unser Ansatz besteht darin, visuelle Merkmale der Umwelt für die Lokalisation zu verwenden. Durch die Verwendung von Farbmerkmalen, welche aus dem Bild der Omnikamera extrahiert werden, wird eine robustere Lokalisation erreicht, da es gegenüber Entfernungsprofilen weniger Mehrdeutigkeiten gibt. In Abbildung 3 wird ein Überblick über das entwickelte Gesamtsystem der visuellen MCL gegeben.

Der Roboter (Abb. 3 links) liefert zwei für die MCL wichtige Informationen, den aktuellen sensorischen Umwelteindruck o_t (Kamerabild) und die Bewegungsdaten u_t (Odometrie). Um die szenenbeschreibenden Merkmale aus dem Bild zu extrahieren, wird das kreisförmige omnidirektionale Kamerabild polartransformiert (Abb. 3 oben). Das Ergebnisbild wird anschließend für die Merkmalsextraktion in n vertikale Segmente unterteilt (typischer Weise 10). In jedem dieser Segmente werden m Merkmale berechnet. Damit führt man eine Reduktion des Eingaberaums von $h \cdot w$ Bildpunkten auf $n \cdot m$ Merkmale durch, wobei grobe Nachbarschaftsinformationen im Bild durch die Segmentierung erhalten bleiben.

Um den Vergleich zwischen der tatsächlichen Beobachtung und der für ein Sample erwarteten Beobachtungen durchführen zu können, ist ein Umgebungsmodell (Abb. 2 und 3 rechts) nötig. Dazu wird in einer „Teachphase“ ein Modell der Umwelt aufgebaut. Dies geschieht, indem der Roboter die Umgebung exploriert und an „wichtigen“ Positionen eine Referenzansicht der Umwelt aufnimmt. Diese Teach-Fahrt wird dem Roboter mittels Joystick per Hand vorgegeben. Die Referenzansicht o_r (nur der extrahierte Merkmalsvektor) wird zusammen mit den aktuellen Positionsdaten x_t als Knoten in einer graphenbasierten Karte gespeichert.

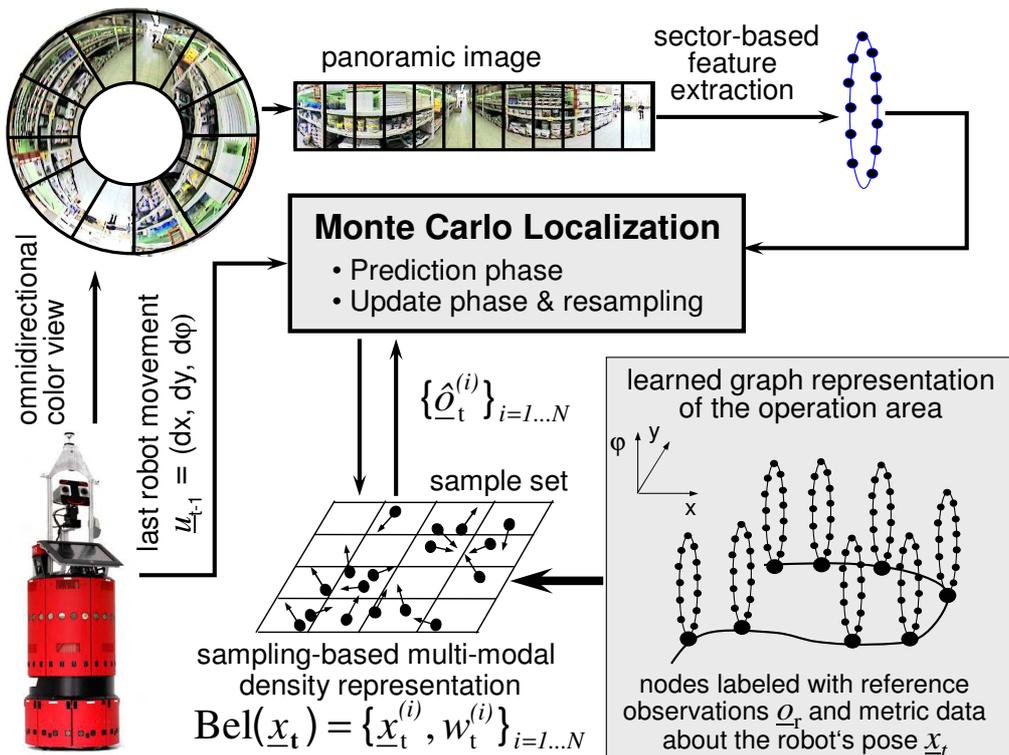


Abbildung 3: Grundidee unserer visuellen MCL; Grundlage bildet ein graphenbasiertes Umweltmodell. Der Roboter liefert visuelle und Odometrieinformationen, woraus im MCL-Verfahren die aktuelle Position und Orientierung (Pose) geschätzt wird. (Bild aus [6])

Die exakte Referenzposition wird während der Teach-Fahrt aus den Odometriedaten in Verbindung mit einem visuellen Odometriekorrekturnverfahren gewonnen, das die Fliesenstruktur des Marktes zur Korrektur der Orientierung des Roboters verwendet. Details hierzu sind [5] zu entnehmen. In Abb. 2 ist eine solche Karte zu sehen, die im Einsatzfeld aufgenommen wurde. Die Größe der Umgebung beträgt $45 \cdot 50$ Meter ($\approx 2200\text{m}^2$). In der Abbildung sind die Knoten (≈ 3000), an denen die Referenzansicht gespeichert sind, und die Regalreihen des Marktes in der Karte zu erkennen. Mit Hilfe der Informationen q_r , u_t und dem Umgebungsmodell schätzt das MCL-Verfahren die aktuelle Position \underline{x}_t des Roboters in der Umwelt.

Die Bestimmung des Gewichtes w^i für ein Sample s^i hängt vom verwendeten Fehler- bzw. Ähnlichkeitsmaß ab. Bisher [1] wurden die mittleren RGB-Werte pro Segment als Merkmale benutzt (im Weiteren **RGB-Mean** genannt), woraus sich ein Merkmalsvektor mit 30 Komponenten ergibt, wenn das Kamerabild typischerweise in 10 Segmente unterteilt wird. Der Importance Factor wurde hierbei aus der Euklidischen Distanz zwischen dem aktuellen Merkmalsvektor und dem für ein Sample erwarteten Vektor bestimmt. Bei dieser sehr einfachen Merkmalsgewinnung entstehen die schon angesprochenen Probleme bei unterschiedlichen Beleuchtungsbedingungen, da diese Merkmale sehr helligkeitsvariant sind, wie man auch in den Experimenten des nächsten Abschnitts sehen kann. Aus diesem Grund wird eine helligkeitsinvariantere Beschreibung der aktuellen Beobachtung o_t des Roboters gesucht.

3 Untersuchung alternativer Farbmerkmale

Ähnlichkeitsuntersuchung: Für die Lokalisation ist es wichtig, dass sich verschiedene Positionen durch ihre visuellen Merkmale voneinander unterscheiden lassen. Dies soll auch bei schwankenden Beleuchtungsbedingungen der Fall sein. Aus diesem Grund wurde zur Bewertung alternativer Verfahren zur Merkmalsextraktion für die visuelle MCL wie folgt vorgegangen. Zunächst wurden drei Referenzansichten aus dem



Abbildung 4: Drei Aufnahmen in einem gangartigen Bereich ähnlich dem späteren Einsatzfeld zur Untersuchung der Merkmale; Oben und Mitte identische Position aber unterschiedliche Beleuchtung; Oben und Unten identische Beleuchtung aber unterschiedliche Positionen (ca. 1,5 Meter von einander entfernt)

typischen Einsatzfeld (Gangstruktur) aufgenommen. Das erste Bild wurde als Modellbild (*PIH*, Abb. 4 oben) bestimmt und weist eine „normale“ Beleuchtung (mittlere Helligkeit $\approx 70\%$) an Position 1 auf. Das zweite Bild (*PID*, Abb. 4 Mitte) wurde an der selben Position 1 aufgenommen, jedoch mit einer deutlich geringeren Beleuchtung (mittlere Helligkeit $\approx 40\%$). Das letzte Bild (*P2H*, Abb. 4 unten) wurde an einer etwa 1,5 Meter entfernten Position 2 aufgenommen (gleiche Beleuchtung wie das Bild *PIH*). Anhand dieser Bilder wurden die Verfahren auf ihre Robustheit gegenüber Beleuchtungsschwankungen und auf ihre Fähigkeit, benachbarte Positionen möglichst effektiv voneinander unterscheiden zu können, untersucht. Geeignete Merkmale sollten sich dadurch auszeichnen, dass die Ähnlichkeit sim_1 (Robustheit gegenüber Beleuchtungsänderungen) zwischen Bild *PIH* und *P1D* möglichst groß und die Ähnlichkeit sim_2 (Unterscheidbarkeit unterschiedlicher Positionen anhand ihrer Merkmale) zwischen Bild *PIH* und *P2H* deutlich geringer als sim_1 ist. Der Vollständigkeit halber sei noch angemerkt, dass ein großer Teil der auftretenden Beleuchtungsänderungen durch ein zusätzliches referenzbasiertes Korrekturverfahren zur Steuerung der Luminanz und der UV-Signale der verwendeten digitalen Kamera (Sony DFW-VL 500) kompensiert [7] wird. Der in Abb. 1b zu erkennende weiße Ring unterhalb des Spiegels wird dazu verwendet, die Parameter für die Blende und für den Weißabgleich der Kamera zu bestimmen. Damit können deutliche Schwankungen der Beleuchtung und der Farbtemperatur bereits zum frühestmöglichem Zeitpunkt, d.h. noch während der Bildgewinnung, kompensiert werden. Die oben beschriebenen Bilder wurden zwar mit dieser Regelung aufgenommen, aber die Beleuchtungsänderung war größernoch stärker als durch diese Regelung maximal kompensiert werden konnte (hier speziell die Helligkeit).

Die zu untersuchenden Farbmerkmale wurden jeweils durch Histogramme repräsentiert. Als Ähnlichkeitsmaß zum Vergleich zweier Histogramme diente die **Histogramm Intersection (HI)**. Die HI hat einen sehr geringen Berechnungsaufwand, was für die Echtzeitfähigkeit des Verfahrens vorteilhaft ist. Es wurden noch andere Histogrammvergleichsmethoden untersucht (*sum of square differences*, χ -square und *weighted euclidian distance*), aber HI zeigte die besten Resultate. Für die weiteren Untersuchungen wurde deshalb die HI verwendet. Es wurden Merkmale auf Basis folgender Farbräume untersucht, wobei bei jedem dieser

Farbräume die Achsen in N Bins unterteilt wurden. Diese Quantisierungsstufen sind in Anlehnung an die Literatur [8, 9, 10] wie folgt gewählt wurden:

| Farbraum | Stufen | Bins pro Segment | Bins pro Bild (10 Segmente) | Ähnlichkeit $sim_1(P1H, P1D)$ | Ähnlichkeit $sim_2(P1H, P2H)$ |
|----------|---------|------------------|-----------------------------|-------------------------------|-------------------------------|
| RGB | 8x8x8 | 512 | 5120 | 0,120 | 0,504 |
| OCS | 16x16x8 | 2048 | 20480 | 0,129 | 0,510 |
| HSV | 18x4x4 | 288 | 2880 | 0,133 | 0,377 |
| rg | 8x8 | 64 | 640 | 0,906 | 0,794 |
| RGB-Mean | - | 3 | 30 | 0,605 | 0,938 |

Im Vergleich zu RGB-Mean [1] erkennt man, dass sich die Anzahl der Daten zum Teil stark erhöht hat. Obwohl hier schon absehbar ist, dass nicht mit allen Verfahren zur Merkmalsextraktion eine echtzeitfähige Lokalisation möglich ist, wurden diese trotzdem untersucht. Anhand der Ähnlichkeiten erkennt man hier, warum es bei RGB-Mean zu Problemen mit der Lokalisation bei schwankenden Beleuchtungsbedingungen kam. Die Ähnlichkeit sim_1 zwischen Bildern an ein und derselben Position bei unterschiedlicher Beleuchtung ist deutlich geringer als die Ähnlichkeit sim_2 zwischen zwei Bildern an benachbarten Positionen bei gleicher Beleuchtung. Damit kann der Roboter seine exakte Position bei Beleuchtungsänderungen nicht korrekt wiedererkennen. Noch schlechter sehen die Ergebnisse für die RGB-, HSV- und OSC-histogrammbasierten Merkmale aus. Als einziger Ansatz erfüllt der helligkeitsnormierte rg-Farbraum die Bedingung, bei unterschiedlichen Positionen eine geringere und bei unterschiedlichen Beleuchtungen eine größere Ähnlichkeit zu erreichen. Damit kommt letztlich nur der rg-Farbraum für das MCL-Verfahren in Frage.

Alternativ dazu wurde eine weitere Möglichkeit zur Merkmalsextraktion untersucht, welche das Bild in Anlehnung an [11] in einen chromatischen und achromatischen Teil aufteilt. Dabei werden die Pixel des Bildes bzw. Segmentes anhand des HSV-Farbraums entweder in ein Farbwinkel- (16 Bins) oder in ein Grauwert-Histogramm (4 Bins) eingeordnet. Es entsteht also ein Histogramm mit 20 Komponenten, anders als das oben beschriebene HSV-Histogramm mit insgesamt 288 Bins. Hierbei soll der helligkeitsinvariante Farbwinkel **Hue** des HSV-Farbraums stärker genutzt werden, um die Robustheit gegenüber Intensitätsschwankungen zu verbessern. Des Weiteren wurde diese Methode zur Merkmalsgewinnung in einigen Details verändert. Das Verhältnis chromatischer zu achromatischen Pixeln wurde für einen gewichteten Histogrammvergleich benutzt. Die Klassifizierung der Pixel in chromatische und achromatische wurde nicht anhand des Farbwinkels sondern mit Hilfe eines Schwellwertgebirges nach [11] durchgeführt. Dies geschah, weil sich der Farbwinkel für kleine Sättigungen und Intensitäten im HSV instabil verhält. Die erste Methode (herkömmliche Pixelklassifizierung anhand des Farbwinkel Hue) zur Merkmalsgewinnung wird im Weiteren als **TPC** (traditional pixel classification) und die veränderte Methode (Pixelklassifizierung anhand des Schwellwertgebirges und gewichteter Histogrammvergleich) als **MPC** (modified pixel classification) bezeichnet. Es zeigten sich folgende Ergebnisse:

| Verfahren | Bins pro Segment | Bins pro Bild (10 Segmente) | Ähnlichkeit $sim_1(P1H, P1D)$ | Ähnlichkeit $sim_2(P1H, P2H)$ |
|-----------|------------------|-----------------------------|-------------------------------|-------------------------------|
| TPC | 16+4 | 200 | 0,494 | 0,484 |
| MPC | 16+4 | 200 | 0,586 | 0,497 |

Die beiden Merkmale TPC und MPC haben eine deutlich geringere Anzahl an zu berechnenden Daten, womit eine Echtzeitfähigkeit gegeben wäre. Auch die an die Merkmale gestellte Bedingung, dass die Ähnlichkeit sim_1 (gleicher Ort, unterschiedliche Beleuchtung) größer sein soll als die Ähnlichkeit sim_2 (benachbarter Ort, identische Beleuchtung), erfüllen sie recht zufriedenstellend, wobei MPC mit den Veränderungen leicht bessere Ergebnisse erzielen kann. Damit stehen drei geeignete Kandidaten als zur Merkmalsextraktion für die Monte-Carlo-Lokalisation zur Verfügung.

AMP-Untersuchung: Die aussichtsreichsten Verfahren (rg, TPC, MPC) wurden einer weiteren Untersuchung mit Hilfe der *Average-Match-Percentile* (AMP nach [8]) unterzogen. Mit 61 Bildern aus dem Baumarkt wurde untersucht, ob sich die korrekten Ähnlichkeiten der Bilder untereinander ergeben. Durch die in realen Umgebungen auftretenden Beleuchtungsschwankungen und dynamischen Verdeckungen, z.B. durch umherlaufende Personen, wurden die Bilder unterschiedlich stark beeinflusst. Von diesen 61 Bildern wurden 5 an (fast) ein und derselben Position mit entsprechenden Beleuchtungsänderungen ($\approx 10\%$) aufgenommen. Abwechselnd wurde eines dieser 5 Bilder als Modellbild definiert und die Reihenfolge der restlichen 4 Bilder (an gleicher Position) und des Modellbildes selbst im Ähnlichkeitsranking aller 61 Bilder bewertet. Es wurden damit 25 Durchläufe (5 Modellbilder mal 5 Suchbilder) ausgewertet und der AMP-Wert (100 = perfektes Matching; 50 = zufällig) bestimmt:

| Merkmal | Correct Match Placement | | | | | | AMP |
|------------------|-------------------------|-----|-----|-----|-----|-------|--------|
| | 1st | 2nd | 3rd | 4th | 5th | > 5th | |
| TPC, 16+4 Bins | 5 | 2 | 2 | 2 | 1 | 13 | 92,866 |
| rg, 8x8 Bins | 5 | 1 | 4 | 0 | 1 | 14 | 86,066 |
| MPC, 16 + 4 Bins | 5 | 0 | 0 | 2 | 3 | 15 | 84,800 |
| RGB-Mean | 5 | 3 | 0 | 1 | 0 | 15 | 77,600 |

Alle Verfahren erreichen für den ersten Platz jeweils fünf korrekte Platzierungen, da hier das Modellbild sich selber immer am ähnlichsten ist. Keines der Verfahren erzielte jedoch unter den schwankenden Beleuchtungsbedingungen korrekte Platzierungen. Auf den Plätzen 2-5 sollten sich die restlichen 20 Bilder der identischen Position befinden. RGB-Mean erzielte das schlechteste Ergebnis. Das bedeutet, dass die neuen Verfahren zur Merkmalsgewinnung eine bessere Resistenz gegenüber dynamischen Veränderungen speziell in der Beleuchtung besitzen. Interessant ist allerdings, dass der TPC-Ansatz, der unter den starken Intensitätsschwankungen im vorherigen Versuch noch die schlechtesten Ergebnisse unter den neuen Verfahren geliefert hatte, hier mit Abstand die beste Performanz zeigte.

MCL-Untersuchung: Als Abschluss der Untersuchungen wurden die Verfahren zur Merkmalsextraktion in realen Lokalisationsexperimenten vergleichend untersucht. Dafür wurde für jedes Verfahren das in Abb. 2 dargestellte Umgebungsmodell ($\approx 2200 \text{ m}^2$) aufgebaut. Die Experimente wurden nicht im Baumarkt selbst, sondern mit real aufgezeichneten Bild- und Bewegungssequenzen (vier Sequenzen mit Länge zw. 20 und 400 Metern) aus dem Baumarkt durchgeführt. Die Lokalisation konnte so mit unterschiedlichen Ausgangssituationen (z.B. variable initiale Partikelverteilung oder Startposition), aber mit immer den gleichen Bild- und Bewegungsdaten berechnet werden. Dies ermöglicht eine Vergleichbarkeit der Lokalisationsergebnisse für alle Verfahren, da die Experimente alle auf den gleichen Eingabedaten durchgeführt wurden. Die Merkmalsextraktion und der MCL-Algorithmus (10000 Samples) wurden jeweils für jedes Verfahren neu auf den Bild- und Bewegungssequenzen berechnet, um so eine Aussage über die benötigte Berechnungszeit zu erhalten. Die Ergebnisse der Lokalisation wurden gemittelt und ergaben folgendes:

| | RGB-Mean | TPC | rg | MPC |
|--------------------------|----------|----------|----------|----------|
| Lokalisationsfehler | 61,85 cm | 52,80 cm | 50,25 cm | 57,95 cm |
| Lokalisationssicherheit | 99,70% | 100,00% | 100,00% | 99,81% |
| Rechenzeit pro Schätzung | 238 ms | 378 ms | 3469 ms | 410 ms |

Wie erwartet zeigt das alte Verfahren RGB-Mean die schlechtesten Ergebnisse. Dies gilt für die Lokalisationsgenauigkeit wie auch für die Lokalisationssicherheit, bei der gemessen wurde, wie viele Schätzungen über einem maximal tolerierbaren Fehler ($> 2,5$ Meter) lagen. Leider enttäuschte auch das MPC-Verfahren trotz des guten Abschneidens in den ersten Untersuchungen. Auch hier kam es zu Lokalisationsverlusten. Die besten Ergebnisse erzielten mit etwa 50 cm Lokalisationsfehler die TPC- und rg-Methoden, bei denen keine Lokalisationsunsicherheiten auftraten. Betrachtet man jedoch den Berechnungsaufwand (Messung auf einer 1,5 GHz CPU) von 3469 ms für eine Schätzung, so erkennt man, dass dieses Verfahren derzeit nicht in Echtzeit eingesetzt werden kann.

4 Zusammenfassung und Ausblick

In Weiterführung zu [1] wurde hier gezeigt, wie mit verbesserten Verfahren zur Merkmalsextraktion die Genauigkeit und Robustheit der visuellen Monte-Carlo-Lokalisation für mobile Roboter weiter gesteigert werden konnte, insbesondere in Bezug auf schwankende Beleuchtungsbedingungen. Zum einen wird dies durch die angesprochene Kameraregelung erreicht und zum anderen durch die stärker helligkeitsinvarianten Merkmale. Die im bisherigen Verfahren auftretenden Lokalisationsverluste wurden durch die Verwendung des TPC-Verfahrens (traditional pixel classification), welches den HSV-Farbraum zur Einteilung der Pixel des Bildes in einen chromatischen und achromatischen Teil zum Aufbau eines Farbwinkel- und eines Grauwert-histogramms verwendet, beseitigt. Die hier erreichten Lokalisationsgenauigkeiten sind für den Einsatz der Roboter als Shopping-Assistent mit Lotsenfunktion zu Artikelgruppen völlig ausreichend. Weitaus wichtiger ist hierbei die Robustheit und Zuverlässigkeit des Verfahrens, was mit den erfolgreichen Untersuchungen zur Lokalisationssicherheit verdeutlicht werden kann. Des Weiteren erfüllt das TPC-Verfahren auch die Echtzeitanforderungen, im Gegensatz zum rg-Verfahren, welches die zehnfache Rechenzeit benötigt und daher derzeit nicht eingesetzt werden kann. In Zukunft sollen dynamische Veränderungen der Umwelt, insbesondere der Regalbefüllung, welche einer ständigen Veränderung unterliegt, untersucht werden. Dazu soll geklärt werden, wie allgemeine Veränderungen der Umweltstruktur am besten im Umgebungsmodell repräsentiert werden können, um so auch über einen langen Zeitraum eine genaue Lokalisation zu gewährleisten.

Literatur

- [1] SCHRÖTER, Ch. ; KÖNIG, A. ; GROSS, H.-M.: Farbbild-basiertes Monte-Carlo-Lokalisationsverfahren für mobile Roboter. In: *8. Workshop Farbbildverarbeitung*, ZBS Ilmenau, ISSN 1432-3346, 2002, S. 111–118
- [2] FOX, D. ; BURGARD, W. ; DELLERT, F. ; THRUN, S.: Monte Carlo Localization: Efficient Position Estimation for Mobile Robots. In: *Proceedings of the AAAI, National Conference on Artificial Intelligence, AAAI*. Orlando, Florida, 1999
- [3] GROSS, H.-M. ; BOEHME, H.-J.: PERSES - A Vision-based Interactive Mobile Shopping Assistant. In: *Proceeding IEEE International Conference on Systems, Man and Cybernetics*, 2000, S. 80–85
- [4] BÖHME, H.-J. ; WILHELM, T. ; SCHRÖTER, Ch. ; GROSS, H.-M.: PERSES - ein interaktiver Einkaufsassistent. In: *Conference Robotics 2002*, VDI-Berichte 1679, VDI-Verlag, 2002, S. 641–646
- [5] SCHRÖTER, C. ; BÖHME, H.-J. ; GROSS, H.-M.: Extraction of orientation from floor structure for odometry correction in mobile robotics. In: *DAGM'03, 25th Pattern Recognition Symposium, Magdeburg*, 2003
- [6] GROSS, H.-M. ; KOENIG, A. ; SCHROETER, C. ; BOEHME, H.-J.: Omnivision-based Probabilistic Self-localization for a Mobile Shopping Assistant Continued. In: *Proceedings of IROS'2003, IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas*, 2003
- [7] WILHELM, T. ; BÖHME, H.-J. ; GROSS, H.-M.: User Detection with an Omnidirectional Camera, 2003. – submitted to: *9. Workshop Farbbildverarbeitung*
- [8] SWAIN, M.J. ; BALLARD, D.H.: Color Indexing. In: *International Journal of Computer Vision* 7:1 (1991), S. 11–32
- [9] FURHT, B. ; SAKSOBHAVIVAT, P. *A fast content-based multimedia retrieval technique using compressed data*. Proceedings of SPIE Symposium on Multimedia Storage and Archiving Systems. 1998
- [10] TICO, M. ; HAVERINEN, T. ; KUOSMANEN, P.: A Method of Color Histogramm Creation for Image Retrieval. In: *Proceedings of NORSIG 2000, IEEE Nordic Signal Processing Symposium*. Kolmården, Sweden, June 13-15 2000
- [11] BALTHASAR, D. ; REHRMANN, V.: Robustes histogrammbasiertes Farbmatching. In: *Proceedings of 5. Workshop Farbbildverarbeitung '99*, Ilmenau, Germany, 1999