

Autonome Exploration von Innenräumen mit der Multisensorik-Plattform *MARVIN**

Stefan Blum¹, Darius Burschka¹, Christof Eberst¹, Tobias Einsele¹,
Alexa Hauck¹, Peter Leven², Norbert O. Stöffler¹, Georg Färber¹

¹ Lehrstuhl für Prozeßrechner
Technische Universität München
roviwma@lpr.e-technik.tu-muenchen.de
² The Beckman Institute
University of Illinois at Urbana-Champaign
p-leven@uiuc.edu

Zusammenfassung Für Transport- und Überwachungsmissionen in unpräparierten Innenräumen, benötigt ein AMR eine geeignete Kombination aus hochentwickelter sensorischer Erfassung und internem Modell der Einsatzumgebung. In diesem Beitrag wird ein allgemeine und dynamisch erweiterbare Systemstruktur definiert und deren prototypische Implementierung in Hard- und Software auf der mobilen Multisensorik-Plattform *MARVIN* beschrieben.

1 Einleitung

Damit ein mobiler Serviceroboter in unpräparierten Innenräumen Transport- und Überwachungsmissionen ausführen kann, benötigt er eine geeignete Kombination aus hochentwickelter sensorischer Erfassung und internem Modell seiner Einsatzumgebung.

Ziel des beschriebenen Projekts ist, anhand von autonom durchgeführten Explorationsfahrten ein generalisiertes Umgebungsmodell aufzubauen, das alle für Sensordateninterpretation und Missionsplanung nötigen Abstraktionsebenen enthält. Hierfür wurde eine allgemeine und dynamisch erweiterbare Systemstruktur definiert (Abbildung 1).

Bei der eingebundenen Sensorik wurde der Schwerpunkt auf Standard-Videokameras gelegt, da sie bei moderaten Kosten den nötigen Detailreichtum auch für komplexe Perzeptionen zur Verfügung stellen können. Weitere auf *MARVIN* installierte Sensorikkomponenten sind ein abstandsgebender Laserscanner sowie Odometrie.

Die unterste Ebene der *logischen Sensoren* umfaßt all die Schritte der Sensordatenverarbeitung, die ohne explizites Modellwissen durchgeführt werden

* Die vorliegende Arbeit wurde im Rahmen des Sonderforschungsbereichs *Informationsverarbeitung in autonomen, mobilen Handhabungssystemen* (SFB 331), Teilprojekt Q5, und des Projekts *Exploration von Innenräumen mit optischen Sensoren auf mehreren, aufgabengerechten Abstraktionsebenen* (Förderungsnummer Fa109/14-1) von der Deutschen Forschungsgemeinschaft (DFG) gefördert.

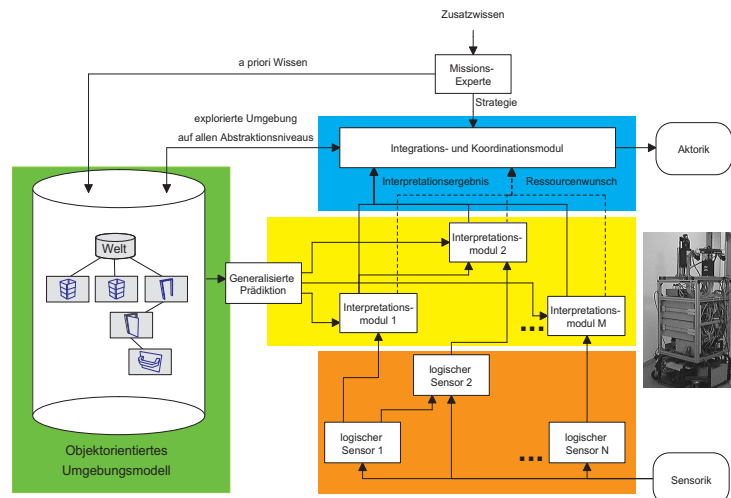


Abbildung 1. Systemstruktur

können. In der aktuellen Realisierungsstufe (siehe Kapitel 2) kommen zunächst logische Sensoren zum Einsatz, die aus dem von der Sensorik gelieferten Datenstrom einfache Merkmale wie zwei- und dreidimensionale Linien oder optische Flußvektoren extrahieren. Darüberhinaus wird auch aggregierte Information wie dreidimensionale Struktur, Eigenposition und Bewegungsparameter in dieser Schicht ermittelt.

Die Ebene der *Interpretationsmodule* interpretiert die ihr zur Verfügung stehende Information anhand eines dreidimensionalen, objektorientierten Umgebungsmodells. Dies geschieht auf Basis von generischer, im Modell geeignet abgelegter a priori Information über die in Innenräumen typischerweise vorhandenen Objekte und Raumstrukturen (siehe Kapitel 3). Modellzugriffe reichen von einfachen Merkmalsprädiktionen bis zum Abruf ganzer Objektklassen.

Als Interpretationsmodule realisiert sind zur Zeit mehrere Objekt- und Strukturerkennungsalgorithmen, die auf unterschiedlichen Abstraktionsniveaus einen spezifischen Vergleich von aus den Sensordaten extrahierter und anhand des Modells prädizierter Information leisten (siehe Kapitel 4).

Die Kommandierung der Sensorik und die Auswahl des Abstraktionsgrades der Interpretation erfolgt aktiv, in räumlich-zeitlicher Abhängigkeit von für die jeweilige Missionsphase relevantem Modellwissen (siehe Kapitel 5.2).

2 Logische Sensoren

2.1 Relative Eigenlokalisierung

Auf Basis eines abstandsgebenden Panorama-Laserscanners arbeitet ein logischer Sensor, der die relative Eigenlokalisierung der Experimentierplattform

MARVIN leistet. Relative Eigenlokalisierung bedeutet hier die fortlaufende Angabe eines Differenzvektors $(\Delta x, \Delta y, \Delta \varphi)$ zwischen der aktuellen und der bei der Initiierung des logischen Sensors eingenommenen Fahrzeugposition.

Abbildung2. Struktur des Eigenlokalisierungssystems

Aufgabe der Vorverarbeitung ist es, die einzelnen Meßpunkte innerhalb eines Rundum-Scans durch Strecken zu approximieren. Der hier favorisierte Ansatz nutzt einen aus der Bildverarbeitung stammenden Kantenextraktionsalgorithmus [9] mit anschließender Verfeinerung durch orthogonale Regression. Ein derart zu einer Sequenz von Strecken aufbereiteter Scan wird nun unter Verwendung des aus der Mustererkennung bekannten DP-Algorithmus [1] einem anderen, typischerweise unmittelbar zuvor aufgenommenen Scan qualitativ zugeordnet, indem Korrespondenzen zwischen den einzelnen Strecken etabliert werden. Die geometrische Auswertung all dieser Korrespondenzen führt auf ein lineares Gleichungssystem für die Vektorkomponenten Δx und Δy . Die Vektorkomponente $\Delta \varphi$ des Differenzvektors berechnet sich als Mittelwert aller bei der Etablierung der Korrespondenzen sich ergebenden Winkelunterschiede [5].

2.2 Bewegungsinformation aus dem optischen Fluß

Ein weiterer logischer Sensor, gestützt auf Standard-Videosensorik, ist das MPEG-Prozessor basierte System zur Detektion und Auswertung von Bewegung in Videobildfolgen.

Zentrales Element des MPEG Kompressionsverfahrens für Bildfolgen ist die Verfolgung konstanter Bildausschnitte über mehrere Bilder, um nur die Verschiebungsvektoren übertragen zu müssen. Diese extrem aufwendige Operation wird von speziellen Korrelationsprozessoren (*Motion Estimation Processor, MEP*) für ganze Bilder in Realzeit durchgeführt.

Eine Videokamera bildet das dreidimensionale Geschwindigkeitsfeld von Bewegungen in der Szene auf ein zweidimensionales Verschiebungsfeld (oft als *optischer Fluß* bezeichnet) ab. Um dieses Verschiebungsfeld zu rekonstruieren, entstand ein MEP-basiertes Bildverarbeitungsboard, das in der Lage ist, schritthaltend zum Ausgangsdatenstrom einer Videokamera bis zu 525 Verschiebungsvektoren pro Bild zu berechnen, wobei zusätzlich für jeden Vektor ein Konfidenzmaß bestimmt wird, welches eine Siebung des Vektorfeldes vor der Weiterverarbeitung ermöglicht [11].

Geht man davon aus, daß das Restfeld der hochkonfidenten Vektoren mit dem Verschiebungsfeld übereinstimmt, so sind aus diesem Feld die dreidimensionalen Bewegungen und die Tiefen der Szene bis auf einen Skalierungsfaktor rekonstruierbar. Dies ist auch dann gegeben, wenn Kamera und Objekte in der

Szene gleichzeitig bewegt werden, wobei für diesen Fall zuerst die zu den Objekten gehörenden Bildbereiche zu identifizieren sind, bevor dann für jeden dieser Bereiche die Bewegungsparameter separat bestimmt werden können [12].

2.3 Rekonstruktion von 3D-Kanten

Für die Tiefenberechnung aus einer binokularen Stereoanordnung werden die durch gerade Liniensegmente approximierten Helligkeitssprünge im Grauwertbild benutzt, die von einem schnellen Kantenfolger schritthaltend, ohne spezialisierte Hardware berechnet werden [8]. Die Extraktionszeit für eine Stereoanordnung ist von der Anzahl der gefundenen Kanten abhängig und beträgt in unseren Laborräumen max. 0.5s für die Stereo-Anordnung bei einer maximalen Segmentanzahl von ca. 350 gefundenen Kantensegmenten.

Der *logische Sensor* „Linienbasierte 3D-Rekonstruktion“ liefert als Ergebnis seiner Verarbeitung dreidimensionale Liniensegmente, die Objektbegrenzungen und Helligkeitssprünge auf Objektflächen repräsentieren. Die Rekonstruktion basiert auf der Interaktion der Stereo-Zuordnungsroutine mit der *Dynamischen Lokalen Merkmalskarte* (DLM) (Kap. 3.1), in der das resultierende Drahtmodell des lokalen Umgebungsbereiches abgelegt wird. Damit wird für die Rekonstruktion sowohl das a priori Wissen, wie auch Ergebnisse vorangehender Verarbeitungszyklen mitbenutzt. Alle bereits bekannten Merkmale in der DLM werden lediglich verifiziert, wodurch der aufwendige und fehleranfällige Zuordnungsprozeß nur auf die verbleibenden, noch unbekanntenen Kantensegmente beschränkt bleibt. Die Zuordnung von Bildmerkmalen wird auch durch die Hinzunahme der aus der Struktur- und Objekterkennung stammenden *hypothetischen Merkmale* (Kap. 4.3) erleichtert. Damit lassen sich sichere Zuordnungen bei schwierigen Beleuchtungsverhältnissen oder komplexen Raumstrukturen leichter herstellen. Der zusätzliche Hinweis auf eine Zuordnung zu einem hypothetischen Merkmal geht in die Kostenfunktion der Stereo-Zuordnungsroutine mit ein.

Die Interaktion mit der DLM erlaubt zusätzlich die durch Vibrationen der Kameras verursachten Fehler zu minimieren. Zuverlässige Kanten aus der DLM werden dazu benutzt, die externen Parameter der Kameras zu schätzen, um den Fehler in der Berechnung der Lage der dreidimensionalen Merkmale zu minimieren (Abb. 3).

2.4 Bildung von 3D-Kanten-Clustern

Eine deutliche Reduktion des Aufwandes der weiteren Verarbeitung¹ der 3D-Sensor-kanten in den Interpretationsmodulen (siehe Abschnitt 4), eine Verringerung der Sensor-kantenanzahl sowie die Möglichkeit zur dortigen parallelen Verarbeitung durch Mehrfachinstanziierung wird durch ein auf die DLM aufsetzendes Clustering-Modul erreicht. Hierbei werden je nach Konfiguration disjunkte bzw. sich nur leicht überschneidende Kantengruppen (Positionscluster) gebildet.

¹ Matching mit Modellkanten

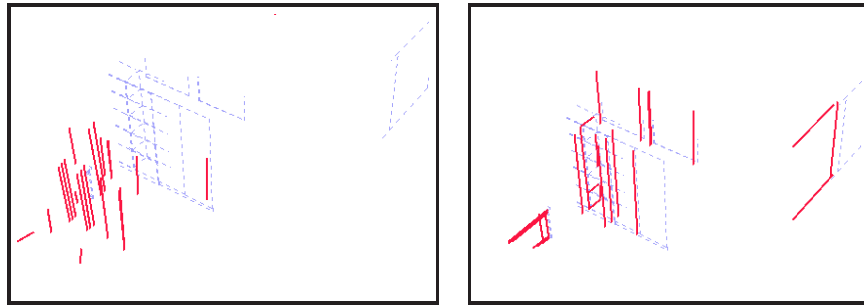


Abbildung 3. Rekonstruierte Daten bei einem Eingangsfehler von $\pm 1.5^\circ$ ohne und mit Orientierungskorrektur aus [2].

Als Abstandskriterium dient dabei der geometrische Abstand der Kanten im 3D-Sensordatenraum. Lange Kanten, die ungewollt einzelstehende Objekte zu Clustern verschmelzen würden², werden nach topologischen Regeln ggf. mehreren Clustern zugeordnet. Große Cluster werden in einem Nachbearbeitungsschritt nach einem Redundanz-Gridverfahren zerbrochen. Die entstandenen Positionskluster werden abschließend einzeln noch einem Richtungsclusteringverfahren unterzogen, wobei als Clustermaß die Kantenrichtung zugrunde gelegt ist.

Die Einfachheit des Verfahrens birgt die Fehlermöglichkeit, daß Objekte, deren Sensoranten z.B. aufgrund von Verdeckungen in verschiedene Cluster fallen. Abhilfe schafft hier ein globales Positionskluster, das nur dem Richtungsclusteringverfahren unterzogen und von einem mit niedriger Priorität gestarteten Interpretationsmodul weiterverarbeitet wird.

2.5 Optische Stoßstange

Aus dem physikalischen Sensor *binokulare Stereokamera* kann der *logische Sensor* „optische Stoßstange“ realisiert werden, der die Kenntnis des Bodenverlaufes voraussetzt, aus dem sich die Verschiebung der Merkmale (Disparität) in beiden Bildern ermitteln läßt. Dieser Sensor liefert eine Sektorkarte, in der die Entfernung zu möglichen Hindernissen in einem bestimmten Winkelbereich eingetragen wird. Wichtig ist hierbei, Bodenmuster und kleine Unebenheiten von tatsächlichen Hindernissen zu unterscheiden.

Aus den geometrischen Beziehungen kann die erwartete Disparität für Bodenmerkmale berechnet werden. Alle Merkmale, die eine abweichende Disparität aufweisen, werden als Hindernisse an die koordinierenden Instanzen weitergemeldet. Für diese Verarbeitung werden direkt die detektierten Pixelketten entlang der Helligkeitssprünge benutzt, um beliebig geformte Hindernisse detektieren zu können.

² z.B. Bodenkanten oder Fehlkorrespondenzen

3 Repräsentation der Umgebung

Repräsentationen der Umgebung treten in zwei Formen auf: Logische Sensoren legen die extrahierten Merkmale in lokalen Zwischenspeichern wie z.B. der *Dynamischen Lokalen Merkmalskarte* ab (siehe Kap. 3.1). Ziel der Exploration ist dagegen eine globale Beschreibung, welche durch Integration verschiedenartiger Information gewonnen wird. Für die Erkennung missionsrelevanter Objekte müssen zudem Referenzmodelle bereitgestellt und Sensorsichten präzisiert werden. Letztere Punkte werden durch das im Rahmen des SFB 331 entwickelte Umgebungsmodell **GEM** [10, 6, 14, 7] abgedeckt (siehe Kap. 3.2).

3.1 Räumlich-zeitliche Stabilisierung in der DLM

Einige, der in Kapitel 2 beschriebenen *logischen Sensoren*, die für die Rekonstruktion der dreidimensionalen Umgebungsbeschreibung benutzt werden, liefern zwar eine bestimmte Merkmalsart mit einer hohen Rate, aber häufig mit einer geringen Zuverlässigkeit. Als Beispiel kann hier die Rekonstruktion der 3D-Kanten (Kap. 2.3) dienen, die aufgrund von Fehlzuordnungen und Mehrdeutigkeiten auch falsche Merkmale erzeugt, die in dem vorgestellten System durch Verifikation von anderen Standorten ausgefiltert werden müssen (Abb. 4).

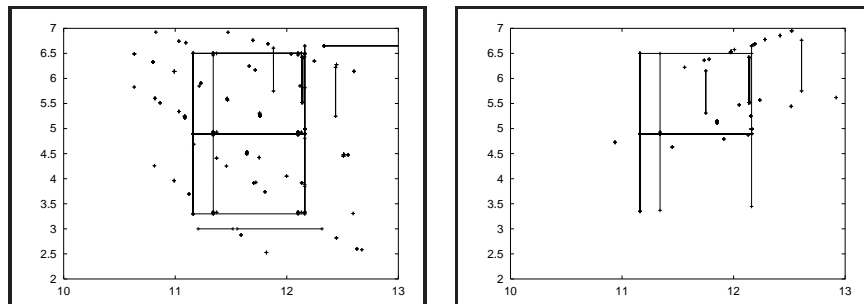


Abbildung 4. Normalprojektion (von oben) der gefundenen Merkmale einer Tischgruppe: (a) ohne Vorwissen, (b) mit Vorwissen aus vorangehenden Schritten.

Im Falle der erfolgreichen Verifikation eines Merkmals wird gleichzeitig seine Lage aufgrund der beiden Beobachtungen verfeinert, wodurch eine Genauigkeitszunahme erreicht wird [2]. Die DLM führt bei jedem Eintrag von neuer Information eine geometrische Fusion der neuen Daten mit dem bereits gespeicherten Wissen durch. Damit können kleine Fehler, die durch Rauschen und Detektionsfehler verursacht werden, kompensiert werden.

Eine zweite wichtige Aufgabe der DLM ist die Entkopplung des logischen Sensors von den Interpretationsmodulen. Während in dem Datenkreis zwischen DLM und dem logischen Sensor Daten mit einer hohen Rate, aber geringer

Zuverlässigkeit ausgetauscht werden, werden auf der Schnittstelle zum Interpretationsmodul zuverlässige Daten gefordert. Die teilweise hohe Komplexität der in den Interpretationsmodulen ablaufenden Algorithmen ist mit einer geringeren Taktrate auf dieser Schnittstelle verbunden. Die DLM entkoppelt die beiden Datenkreise und stellt das Wissen über eine bestimmte Merkmalsart u.U. mehreren Interpretationsmodulen zur Verfügung, ohne dass diese ihrerseits diese Information zwischenspeichern müssen. Damit vereinfacht sich die Konsistenzhaltung der Information im Gesamtsystem.

Um die Zugriffe der Interpretationsmodule zu optimieren und die Verarbeitung der logischen Sensoren robuster zu machen, wird die DLM mit dem Wissen aus dem globalen Modell GEM initialisiert. Damit stellt die DLM die Quelle für das aktuellste Wissen über die Umgebung dar. Man findet dort einerseits sichere Information, die aus dem globalen Modell stammt oder bereits mehrfach verifiziert werden konnte, und andererseits unsichere Information in Form von Hypothesen (Kap. 4.3) und noch nicht verifizierten Merkmalen, die z.B. Fehlkorrespondenzen darstellen können oder auf die der Aufmerksamkeitsbereich der folgenden Schritte konzentriert werden sollte [4]. Die Information wird direkt auf dem Abstraktionsniveau des daran angeschlossenen logischen Sensors gehalten, um eine einfache und schnelle Fusion der Daten zu ermöglichen.

3.2 Globales Modell GEM

Ein wesentlicher Aspekt bei der Konzeption von GEM war, ein einheitliches Modell für unterschiedlichste Sensoren zu verwenden. Dafür wird die Beschreibung der 3D-Strukturen der Umgebung von der Beschreibung der *Merkmale*, die für die verschiedenen Sensorklassen sichtbar sind, getrennt. Die 3D-Strukturen werden durch Polyeder approximiert und durch die so entstandenen Oberflächen beschrieben. Dabei wird unterschieden zwischen Modellen einzelner, missionsrelevanter *Objekte*, die auch mit bestimmten Weltpositionen verknüpft werden und auch aus starren, aber relativ zueinander beweglichen Komponenten zusammengesetzt sein können, und dem nicht weiter klassifizierten „Rest“ der Umgebung, der sogenannten *Kulisse*. Neben 3D-Modelllinien werden von GEM auch sogenannte *aggregierte Merkmale* unterstützt, die eine Gruppierung von Merkmalen erlauben, und mit zusätzlichen Merkmalsattributen versehen werden können.

Auf Modellinformation kann auf Klassen-, Objekt- oder Welt-Ebene zugegriffen werden. Der wichtigste Lesezugriff ist die Merkmalsprädiktion, also die Erzeugung synthetischer Sensorsichten auf Merkmalsebene. Darüberhinaus können auch Attribute und Zustand eines Objektes abgefragt werden, z.B. der aktuelle Öffnungswinkel einer Tür. Schreibzugriffe ermöglichen die Aktualisierung der Modellinformation; sie reichen von der Aktualisierung eines Attributs oder der geometrischen Objektbeschreibung bis zum Eintrag neu erkundeter Umgebungselemente.

4 Interpretationsmodule

Die bisher entstandenen Interpretationsmodule lassen sich in zwei Gruppen untergliedern: Die in den ersten drei Unterpunkten beschriebene hierarchische Struktur einer initialen Lokalisation (4.4) beruht auf einer langsamen, dafür aber genauen Objektidentifizierung (4.2), die alle sechs Freiheitsgrade im 3D-Raum berücksichtigt. Hierzu werden zur Beschleunigung Kantentripel aus der Objektbasis von GEM (??) gebildet (4.1). Davon zu unterscheiden ist eine schnelle Struktur- und Objekterkennung (??), die allerdings nur drei Freiheitsgrade berücksichtigt.

4.1 Bildung von Kantentripeln

4.2 Objekterkennung

Die Objekterkennung erfolgt zweistufig. Mittels Interpretationsbäume werden den Kantentripeln Sensorkanten zugeordnet. Die Blätter des Interpretationsbäume stellen sog. *Prähypothesen* dar. Durch die Zuordnung lässt sich die räumliche Lage vermuteter Objekte bestimmen. Eine Bild-Modell-Vergleich generiert ggf. Objekthypothesen, die mit einer Plausibilitätsmaß attributiert sind. Durch das Acquirieren neuer Sensorinformation überschreitet das Plausibilitätsmaß ggf. einen Schwellwert, wodurch ein Objekt als erkannt eingestuft wird.

4.3 Struktur- und Objekterkennung

Die stabilisierten 3D-Sensormerkmale werden als Strukturen und Objekte interpretiert. In einem ersten Schritt werden die stabilisierten Merkmale, für Laser- und Videosensor separat, gegeneinander getestet um primitive 3D-Strukturen, *Relationen* zu finden. Typische Relationen sind Ecken, T-Schnittpunkte und parallele Anordnungen der Merkmale

sowie Kantentripel die sich in Merkmalsclustern der DLM befinden.

Durch schrittweise Verknüpfung plausibler, widerspruchsfreier Relationen und der involvierten Sensormerkmale werden Strukturen steigender Komplexität identifiziert und zu Merkmals-Clustern zusammengefasst. Die Sensordaten der verschiedenen Sensoren werden auf Ebene von Strukturen und Merkmalen zu gemeinsamen Strukturen fusioniert und als Cluster gespeichert.

Aufbauend auf den gefundenen Merkmalsrelationen, erkannten komplexen Strukturen sowie den bestimmten Kantentripeln, werden Objekte mittels redundanter Schlüsselmerkmale indiziert. Schlüsselmerkmale

Die Komplexität üblicher Objekterkennungsverfahren verhält sich meist quadratisch zur Anzahl der Objektmerkmale n . Daher werden die Merkmale für eine Generation von Objekt-Prähypothesen auf im vorliegenden Fall auf Gruppen von jeweils drei charakteristischen Kanten (Kantentripel) eingeschränkt; für ein Objekt existieren dabei mehrere Tripel. Die Selektion der Kanten erfolgt mehrstufig: Zunächst werden verschiedene Projektionen der Objekte erzeugt, um die Modellkanten mit der Sichtbarkeit zu attributieren. Hierbei wird der übliche

Ansatz der Isokaderteilung verwendet []. Eine Relevanzbewertung der einzelnen Kanten erfolgt heuristisch aus der Sichtbarkeithäufigkeit, einem längenabhängigen Faktor sowie Informationsgehalt der Kante, in die eine Verwechslungsgefahr mit benachbarten Modellkanten eingeht. In einer zweiten Stufe werden mögliche Kantentripel bewertet, wobei die Relevanz der Einzelkanten, deren räumliche Lage³ und ein Maß der Eindeutigkeit eingeht. Die Auswahl der Kantentripel geschieht iterativ, wobei ausgewählt Kantentripel sich möglichst wenig überschneiden und alle zuvor generierten Projektionen abgedeckt werden. Die Zahl der entstandenen Tripel ist dabei minimal.

SCHLUESSEL + TRIPPLEI

Diese diversitäre Generierung der Objekthypothesen erlaubt es unterschiedliche Bedingungen mit resultierenden Unterschieden in der Merkmalsextraktion und Interpretation auszugleichen.

Zur Unterstützung der Merkmalsstabilisierung und zur Beschleunigung der Verifikation der Objekt- und Strukturhypothesen werden Merkmale der Hypothesen, die nicht detektiert wurden, als hypothetische 3D Merkmale in die DLM eingetragen. Die Merkmale bleiben in der DLM gespeichert bis sie verworfen werden.

4.4 Initiale Lokalisation

Zur Initialen Lokalisation ist Modellwissen⁴ notwendig, um die Freiheitsgrade $(x, y, \varphi)^T$ eines Weltkoordinatensystems zu fesseln. Erkannte Objekte generieren für die initiale Eigenlokalisierung eine oder mehrere Positionshypothesen, die jeweils eine affine Transformation zwischen Startkoordinatensystem des Roboters und Weltkoordinatensystem beinhalten. Eine Auflösung von Mehrdeutigkeiten erfolgt durch Clustering der Positionshypothesen (ähnlich [13]), wobei durch Mittelung Positionshypothesen in einem Cluster fusioniert und in ihrer Plausibilität erhöht werden. Bei Überschreitung eines Schwellwertes wird die ermittelte affine Transformation dem System gemeldet [3].

5 Integrations- und Koordinationsmodule

5.1 Integration

Die Ergebnisse der verschiedenen Verfahren zur Struktur- und Objekterkennung müssen zu einer konsistenten Interpretation des Raumes zusammengefaßt werden. Ergebnisse der verschiedenen Objekterkennungen werden Widersprüche hin getestet und gegebenenfalls fusioniert oder verworfen. Beide parallel operierenden Objekterkennungen hemmen sich gegenseitig automatisch in der Generierung strikt redundanter Hypothesen. Dies geschieht durch Testen der Relationen gegen bereits bestehende Hypothesen vor der Generierung neuer Objekthypothesen und dem Löschen von Sensormerkmale erkannter Objekte aus den Clustern. Letzteres verhindert deren erneute Indizierung.

Die Prädiktionen verschiedener Erkennungsstrategien werden zur Reduktion der Kommunikation vor ihrer Einspeisung in die DLM gegeneinander verglichen und gegebenenfalls fusioniert oder verworfen.

³ Die drei Kanten sollen dabei einen möglichst großen Raum aufspannen.

⁴ natürliche Landmarken, in GEM (siehe Kap. 3.2) gespeichert.

Zuverlässig erkannte Objekte und Strukturen können in das globale Modell GEM eingetragen werden.

5.2 Koordination

Die bisher beschriebenen Module stellen dem Gesamtsystem Dienste zu Verfügung, die für eine optimale Exploration der Umgebung koordiniert werden müssen. Der Einsatz von bestimmten Interpretationsmodulen erzwingt indirekt die Aktivierung der damit verbundenen logischen Sensoren, die eine bestimmte Betriebsart des zugrundeliegenden physikalischen Sensors erfordern. Diese Tatsache und die beschränkten Ressourcen eines mobilen Systems erzwingen einen situationsabhängigen Einsatz der Interpretationsmodule.

Die bereits implementierte, prototypische Variante der Koordinationsinstanz nutzt die Information der „Linienbasierten 3D-Rekonstruktion“ (Kap. 2.3), die in den beiden Interpretationsmodulen „3D-Clustering“ und „Freiraumbestimmung“ aufbereitet wird, um daraus die Folgeaktionen des Fahrzeugs zu bestimmen [?,?]. Diese Module sind in dem Modul „Navigator“ (Abb. ??) vereinigt. Es werden als mögliche Folgeziele: Cluster mit unsicherer Information, Cluster mit hypothetischen Merkmalen und Punkte am Rand des bekannten Freiraumes in betracht gezogen und abhängig von dem erwarteten Informationsgewinn und der zur Verfügung stehenden Zeit ausgewählt.

Da die 3D-Rekonstruktion nur im Stillstand des Fahrzeugs vorgenommen wird, wird während der Fahrt zwischen den einzelnen Positionen das „Selbstschutz“-Modul des Fahrzeugs aktiviert, das auf der in Kap. 2.5 beschriebenen „optischen Stoßstange“ aufsetzt.

In diesem Beispiel wird der physikalische Sensor „binokulare Stereokamera“ mit unterschiedlichen Blickrichtungen betrieben, die einen gleichzeitigen Einsatz der beiden logischen Sensoren „Linienbasierte 3D-Rekonstruktion“ und „optische Stoßstange“ ausschliessen. Eine Koordination des Einsatzes wird erforderlich [?,?,?,?].

6 Experimente

Die Exploration eines lokalen Umgebungsbereiches kann mit oder ohne a priori Wissen erfolgen. In dem durchgeführten Experiment wurde die Mission ohne Vorwissen gestartet. Als Ziel wurde dem System ein Punkt am rechten oberen Rand der Navigationskarte (Abb. 5b) angegeben.

In einer initialen Phase, in der in einem Rundumblick die DLM mit Information gefüllt wird, wurde die in Abb. 5a dargestellte Situation erfaßt. Daraus wurde mit Hilfe des Diffusionsalgorithmus ein kollisionsfreier Weg in Richtung auf das Ziel geplant (Abb. 5b).

Durch die Gewichtung der ermittelten Cluster (Kap. 5.2), die in Abb. 5b als Kreise dargestellt wurden, wird das nächste zu untersuchende Ziel bestimmt. Es werden bei den Missionen zwei Fälle unterschieden: freie Exploration, während der das System sich beliebig im lokalen Bereich bewegen kann, und Fahrauftrag, bei dem das angegebene Ziel direkt angesteuert wird. Wenn freie Exploration

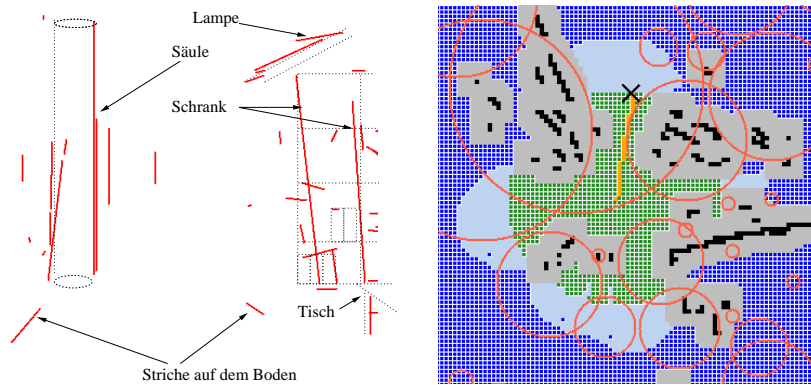


Abbildung 5. (a) Rekonstruktionsergebnis ohne Vorwissen, (b) geplanter Weg in der Navigationskarte

zugelassen wurde, dann versucht das System ein möglichst vollständiges Wissen über die Umgebung zu gewinnen. Eine typische Strecke während der Exploration eines lokalen Bereiches ist in Abb. 6 dargestellt.

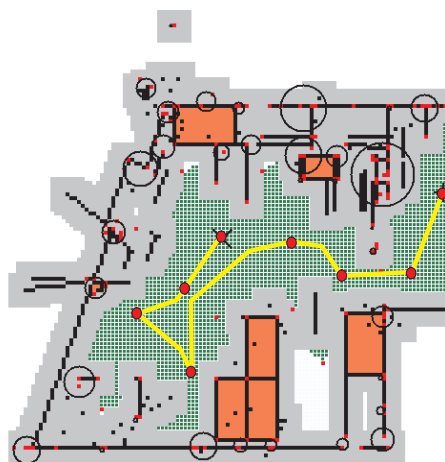


Abbildung 6. Beispiel für eine vollständige Exploration eines lokalen Bereiches

7 Zusammenfassung und Ausblick

- Schnittstellen & Konzept: EvII!
- Topologiegraph

- Integration und Modellrückfluß
Der Aufbau einer konsistenten Repräsentation der Umgebung wird sichergestellt durch eine online Bewertung der von den Interpretationsmodulen gelieferten Ergebnisse. Hierzu werden eventuell vorhandene Widersprüche zwischen den einzelnen Modulen aufgedeckt, durch eine geeignete Auswahl bzw. De-/Aktivieren von Interpretationsmodulen aufgelöst und die Ergebnisse fusioniert.
- mehr Module
- Koordination in echt

Literatur

1. R. E. Bellman and S. E. Dreyfus. *Applied Dynamic Programming*. Princeton University Press, 1962.
2. D. Burschka. *Videobasierte Umgebungsexploration am Beispiel eines binokularen Stereo-Kamerasystems*. PhD thesis, TU München, Fakultät für Elektro- und Informationstechnik, 1998.
3. D. Burschka and S. Blum. Identification of 3d Reference Structures for Video-Based Localization. In *Proc. 3rd Asian Conf. on Computer Vision (ACCV'98)*, volume 1, pages 128–135, 1998.
4. D. Burschka and G. Färber. Active Controlled Exploration of 3D Environmental Models Based on a Binocular Stereo System. *Proc. Int. Conf. on Advanced Robotics (ICAR'97)*, pages 971–977, July 1997.
5. T. Einsele. Real-Time Self-Localization in Unknown Indoor Environments using a Panorama Laser Range Finder. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'97)*, pages 697–703, Grenoble, France, Sept. 1997.
6. A. Hauck, S. Lanser, and C. Zierl. Hierarchical Recognition of Articulated Objects from Single Perspective Views. In *Proc. Computer Vision and Pattern Recognition (CVPR'97)*, pages 870–883. IEEE Computer Society Press, 1997.
7. A. Hauck and N. O. Stöffler. A Hierarchical World Model with Sensor- and Task-Specific Features. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'96)*, pages 1614–1621, 1996.
8. G. Magin and C. Robl. A Single Processor Real-Time Edge-Line Extraction System for Feature Tracking. In *IAPR Workshop on Machine Vision Applications (IAPR MVA'96)*, 1996.
9. T. Pavlidis and S. Horowitz. Segmentation of planar curves. *IEEE Transactions on Computers*, (C-23):860–870, 1974.
10. A. Ruß. *Sensornaher Umgebungsmodellierung mit echtzeitfähigen Zugriffsfunktionen*. PhD thesis, TU München, 1994.
11. N. O. Stöffler and G. Färber. An Image Processing Board with an MPEG Processor and Additional Confidence Calculation for Fast and Robust Optic Flow Generation in Real Environments. In *Proc. Int. Conf. on Advanced Robotics (ICAR'97)*, pages 845–850, Monterey, California, USA, July 1997.
12. N. O. Stöffler and Z. Schnepf. An MPEG-Processor-based Robot Vision System for Real-Time Detection of Moving Objects by a Moving Observer. In *Proc. 14th Int. Conf. on Pattern Recognition*, Brisbane, Australia, Aug. 1998. To appear.
13. G. Stockman. Object Recognition and Localization via Pose Clustering. *Computer Vision*, 40:361–387, 1987.

14. N. O. Stöfler, A. Hauck, and G. Färber. Ein geometrisch-symbolisches Umgebungsmodell zur Unterstützung verschiedener Perzeptionsaufgaben autonomer, mobiler Systeme. In G. Schmidt and F. Freyberger, editors, *Autonome Mobile Systeme*, Informatik aktuell, pages 108–117. Springer-Verlag, 1996.