Rainer Mautz,
Sebastian Tilch

Innenraumpositionierung mit optischen Methoden

ALLGEMEINE VERMESSUNGS-NACHRICHTEN

Dieser Beitrag wurde nach Begutachtung als
PEER REVIEWED PAPER
angenommen.
Tag der Einreichung: 21. Mai 2010

In diesem Beitrag werden die aktuellen optischen Verfahren vorgestellt, die der Positionierung und Navigation in Innenräumen dienen. Aufgrund von verbesserter CCD/CMOS-Sensor-Technologie haben sich kamerabasierte Positions- und Navigationssysteme in den letzten Jahren in vielen Anwendungsbereichen etabliert. Die Anwendungen erstrecken sich über ein breites Spektrum, das sich von der Industriemesstechnik mit Genauigkeitsanforderungen im Submillimeterbereich über die Ingenieur-, Roboter- bis hin zur Fußgängernavigation erstreckt. Der Schwerpunkt des Beitrages bildet ein optisches Innenraum-Positionierungssystem, dass an der ETH Zürich entwickelt wird. Dessen Prinzip beruht darauf, dass projizierte Laserpunkte von einer Kamera erkannt werden und sich dadurch die Position der Kamera bestimmen lässt. Erste Messergebnisse mit einem Prototyp lassen den Schluss zu, dass sich das System für industrielle Anwendungen eignen könnte.

1. Problemstellung und Zielsetzung

Das Thema Indoor-Positionierung ist aktueller denn je – werden doch im Zuge der Automatisierung der rotoberunterstützen Fertigungsindustrie Positions- und Orientierungsdaten von Werkzeugen und Montageobjeken benötigt. In der Lagerhaltung und Logistik wünscht man sich eine lückenlose Produktverfolgung und der Bedarf Personen zu lokalisieren besteht in Krankenhäusern, beim Einsatz der Feuerwehr oder zur Assistenz von Sehbehinderten. Insbesondere stützen sich moderne Anwendungen auf den Ortsbezug, wie z. B. Augmented Reality (Erweiterte Realität), Computer Vision (Rechnergestütztes Sehvermögen), Physical Gesture Games (körpergebundene Spiele), Wireless Sensor Networks (Sensornetze), Location Based Systems (Standortbezogene Dienste). Diese nicht abschließende Liste mit Anwendungen impliziert eine große Nachfrage für die Innenraumpositionierung, der zurzeit nicht ausreichend nachgekommen werden kann. Bei der Erklärung dafür, warum die heutige Technik den Bedarf nicht abzudecken vermag wird richtigerweise mit dem Fehlen von satellitengestützten Verfahren innerhalb von Gebäuden argumentiert. In der Tat liegt die Ursache vor allem am komplexen Anforderungskatalog der Anwender sowie an der Vielfallt der Umgebungsbedingungen, bei denen oftmals ungünstige Sicht- und Signalausbreitungsverhältnisse bestehen.

Ein Lösungsansatz, der alle Genauigkeitsklassen umfasst, ist die Verwendung der optischen Sensorik. Dieser Beitrag beschränkt sich auf die optoelektronischen bzw. videometrischen Verfahren, die Einzug in alle Anwendungsgebiete erhalten haben. Der Erfolg begründet sich in der Effizienzsteigerung und Miniaturisierung der Aktoren (z. B. Laser), vor allem aber der Detektoren (z. B. CCD-Sensoren). Simultan haben sich auch das Leistungsvermögen bei der Übertragung großer Datenmengen, deren rechentechnische Verarbeitung und die Algorithmik in der Bildverarbeitung gesteigert.

2. Optische Innenraum-Positionierungssysteme

Entscheidend für die Einteilung der Systeme ist, ob ein bewegter Sensor (Kamera) positioniert werden soll (auch als Problem der Egomotion bekannt) oder ob ein statischer Sensor dazu dient, Objekte im Bildraum zu positionieren. Zusätzlich spielt die verwendete Sensorik, bzw. die resultierende(n) Messgrösse(n) eine Rolle bei der Einteilung. Es muss auch zwischen Sensoren, die im Infrarotbereich, im sichtbaren Licht oder anderen Wellenlängen empfindlich sind unterschieden werden.

Das grundsätzlich zu lösende Problem eines Kamera-basierten Systems besteht darin, sich mit Hilfe der primären Messinformation – der 2D Positionsdaten auf einem CCD-Sensor – in einer 3D Welt zu positionieren und zu orientieren. Für eine Transformation vom Bildraum in den Objektraum fehlt der Maßstab, bzw. die Entfernungsinformation zu den Objekten. Dieser lässt sich mit Zweikamerasystemen bei einer bekannten Basis mit Hilfe der Zweibildauswertung aus den stereoskopischen

Bildern ableiten. Grob kann die Distanz auch mit Hilfe der Stellung eines Autofokus abgeschätzt werden. Bei der Verwendung von SIFT (Scale-Invariant Feature Transform, Lowe 2004) lässt sich die relative Entfernung für jedes lokale Bildmerkmal ableiten. Alternativ kann die absolute Distanz mit zusätzlichen Sensoren, wie beispielsweise mit Laser-Scannern oder Range Imaging Kameras gemessen werden. Letztere geben bei einer Bildfrequenz von bis zu 100 Hz einen Distanzwert für jeden Pixel eines 320×240 Bildes zurück.

Im Folgenden führen wir die aktuellen Arbeiten im Bereich der optischen Navigation auf, wobei die Einteilung anhand der verwendeten Referenzinformation erfolgt. Eine ausführliche Zusammenstellung älterer Arbeiten bietet DeSouza und Kak (2002).

a) Referenz durch vektorielle Geoinformationsdaten

Diese Klasse der Positionierungsmethoden basiert auf die Erkennung von Objekten durch den Sensor und den Vergleich mit Datenbanken, in denen Positionsinformationen abgelegt sind. Der entscheidende Vorteil dieser Methoden liegt darin, dass keine lokale physikalisch installierte Referenzinfrastruktur benötigt wird, d. h. an die Stelle von Passpunkten tritt eine virtuelle Passpunkt-Koordinatendatei. Die Kosten für eine Systemskalierung sind dementsprechend gering.

Kohoutek et al. (2010) nutzen das semantisch-geometrische Innenraummodell CityGML 3D mit höchstem Detailierungsgrad, um die Position einer Range-Imaging-Kamera zu bestimmen. In einem zweistufigen Verfahren wird zunächst der Raum eindeutig zugeordnet, in dem sich die Kamera befindet. Dabei werden aus der Punktwolke feste Raumbestandteile wie Fenster oder Türen erkannt und deren geometrische Merkmale mit einer Datenbank verglichen. In einem zweiten Schritt wird eine Feinpositionierung mit Hilfe eines Ausgleichungsalgorithmus durchgeführt.

Hile und Borriello (2008) vergleichen einen Gebäudeplan mit dem aktuellen Bild einer Handykamera. Die aus dem Bild extrahierten Features werden nach Korrespondenzen im Plan durchsucht. Nach einem erfolgreichen Matching kann die Pose des Handys auf Dezimetergenauigkeit berechnet und ortsbezogene Information auf dem Display eingeblendet werden.

Kitanov et al. (2007) vergleichen ein vektorielles 3D Modell mit den Kanten, die aus den Kamerabildern eines Roboters abgeleitet werden. Durch wiederholte Kameraorientierung aufeinanderfolgender Bilder wird die Position und Richtung der Kamera während einer Fahrt übertragen. Es ergibt sich eine Drift, die mit Hilfe eines Odometers möglichst gering gehalten wird.

Das Computer Vision Module von Schlaile et al. (2009) basiert ebenso auf dem Prinzip des Feature Trackings aus einer Bildsequenz, dient jedoch lediglich der Positionskorrektur innerhalb eines integrierten Navigationsmodules zur Steuerung eines Indoor Quadrocopters.

Sjö et al. (2009) navigieren einen Roboter mit Hilfe einer geringauflösenden Kamera mit Zoomobjektiv anhand eines 2D Gebäudeplans, der im Vorfeld aus Laserscanner Punktwolken abgeleitet wurde. Zur Abschätzung der Distanz zu Objekten berechnen sie aus dem aktuellen Kamerabild RFC-Histogramme (Receptive Field Coocurrence Histogramms) und vergleichen sie mit Histogrammen, die in einer vorangegangenen Lernphase aufgenommen wurden. Daher kann dieses Verfahren ebenso in die Kategorie b) Referenz durch Bildfolgen eingeordnet werden.

b) Referenz nur durch Bildfolgen

Bei diesem Ansatz wird das aktuelle Bild des Sensors mit vorangegangenen Aufnahmen verglichen. Die größte Herausforderung ist dabei die rechentechnische Bearbeitung in navigationskonformer Echtzeit, insbesondere deshalb, weil das Korrespondenzproblem durch das Fehlen von passiven wie auch aktiven optischen Zielmarken verschärft wird.

Muffert et al. (2010) berechnen die Trajektorie eines sphärischen Videokamerasystems aus dessen Bildsequenz mit Hilfe der relativen Bildorientierung aus jeweils zwei aufeinanderfolgenden Bilden. Ohne zusätzliche Passpunktinformation driftet dabei die Orientierung ähnlich wie bei der Koppelnavigation entlang der Trajektorie ab.

Basierend auf dem Prinzip der optischen Odometrie, dass bei modernen Computermäusen Verwendung findet, entwickelte Maye et al. (2006) ein System, bei dem ein optischer Maussensor umgerüstet wird, um Einkaufswagen zu lokalisieren. Der Sensor befindet sich 5 cm über dem Boden und driftet bei einer Geschwindigkeit von bis zu 2 m/s etwa 1 %. Zur Stabilisierung der Orientierung wird ein elektronischer Kompass eingesetzt. Zusätzlich wird eine Positionskorrektur beim Passieren von installierten Marken vorgenommen.

Zur Navigation eines humanoiden Roboters durch Bürogebäude führen Ido et al. (2009) ein sogenanntes Template Matching durch. Dabei wird in einer Lernphase eine Bildfolge mit vier im Roboter verbauten Kameras aufgenommen, komprimiert und gespeichert. In der autonomen Navigationsphase wird dann die Korrelation zwischen den gespeicherten und den aktuellen Bildern berechnet. Liegt der Korrelationskoeffizient über einer vorgegebenen Schranke, kann die aktuelle Position und der horizontale Versatz zur Sollposition auf wenige Zentimeter genau angegeben werden. Durch die Ausnutzung der Bildreihenfolge und Verwendung eines robusten und schnellen Matching Algorithmus wird eine Anwendung in Echtzeit erst möglich.

c) Verwendung von kodierten, fest installierten Referenzmarken

Visuelle Positionierungsverfahren, die sich bei der Referenzierung auf die Erkennung von natürlichen Merkmalen in Bildern verlassen, erreichen oftmals bei variierenden Lichtverhältnissen nicht die gewünschte Robustheit. Diese kann durch die Verwendung von kodierten, fest installierten Referenzmarken erreicht werden. Ferner lassen sich vereinfachte Algorithmen für die automatische Punkterkennung und eindeutige Identifizierung einsetzen, was den Rechenaufwand wesentlich reduziert. Zusätzlich lässt sich über bekannte Referenzpunkte der Systemmaßstab bestimmen, der sich sonst aus reinen Messungen von natürlichen Merkmalen in Bildern nicht ermitteln ließe.

Für die Lokalisierung von Gabelstaplern in Lagerhallen hat Sky-Trax Inc. (2010) ein System entwickelt, bei dem kodierte Referenzmarken an Hallendecken befestigt werden. Auf dem Dach eines Gabelstaplers oder eines mobilen Fahrzeuges befindet sich ein optischer Sensor,

der ein digitales Bild der Referenzmarken aufnimmt. Die Bilder werden mit einer Funkverbindung zu einer Zentralstation gesendet und dort ausgewertet. Mit dem System lässt sich die Orientierung und die Position eines Fahrzeugs auf 2 – 30 cm genau bestimmen.

Mulloni et al. (2009) stellt ein kostengünstiges Navigationssystem für Besucher einer Konferenz oder Ausstellung vor, das mit einfachen Handykameras betrieben werden kann. Wird eine Kamera auf ein quadratisches Barcodefeld gerichtet, kann die relative Position der Kamera auf einige Zentimeter genau ermittelt werden. Die Positionierung erfolgt in Echtzeit, allerdings nur punktuell dort, wo Codes angebracht wurden.

Für hochgenaue industrielle Anwendungen im Bereich der Crashfahrzeug-Vermessung bietet die Firma AICON 3D Systems (2010) das System ProCam an, bei der eine Mobile Kamera auf ein Messpunkt-Festfeld ausgerichtet wird. Ein in der Kamera integriertes Infrarot-Blitzlicht sorgt für eine aktive Beleuchtung der Festpunkte zu Unterstützung der Punkterkennung. Die zu messenden Punkte werden taktil mit einem handgeführten, an der Kamera fest montierten Messtaster, einer sogenannten Probe, aufgenommen.

Auch das System StarGazer von Hagisonic (2008) zur Positionierung von Robotern basiert auf der Verwendung von retroreflektierenden Referenzpunktfeldern, die an der Decke angebracht werden. Eine auf Infrarot empfindliche Kamera beobachtet verschiedene Anordnungen von Punktmustern, die durch einen Infrarotsender aktiv beleuchtet werden. Aufgrund einer unsymmetrischen Anordnung der Punktmuster kann neben der Position auch die Orientierung des Rovers bestimmt und verschiedene Räume eindeutig zugeordnet werden.

Frank (2008) beschreibt das System stereoScan-3D der Firma Breuckmann zur Aufnahme von Oberflächen, bei dem zwei hochauflösende Kameras mit fester Basis ein industrielles Messobjekt erfassen. In einem Kalibrierungsschritt werden Passmarken auf das Objekt geklebt und photogrammetrisch aufgenommen. Die eigentliche Datenerfassung besteht aus der Projektion verschiedener Streifen und Gitterstrukturen auf das Messobjekt und deren Aufnahme mit den beiden Kameras für eine Oberflächenbestimmung.

d) Indirekte, projizierte Referenzinformation

Das Projizieren von Passpunkten oder Referenzmustern – auch unter dem Begriff der aktiven Triangulation bekannt – hat gegenüber einer physikalischen Aufhängung den Vorteil, dass die ausgesendete Signalstärke, bzw. Lichtmenge bekannt und regelbar ist, wodurch die Empfindlichkeit gegenüber Fremdlichteinflüssen gesenkt werden kann. Die Signifikanz des Signals nimmt gegenüber dem Hintergrund zu und damit auch die algorithmische Robustheit bei der Punkterkennung. Einige Anwendungen erlauben keine Anbringung von Signalisierungen, teils auch in wirtschaftlicher Hinsicht. Bei der Lichtstrahlprojektion kann der Sehstrahl einer Kamera durch den Projektionsstrahl im Sinne einer inversen Kamera ersetzt werden (Wollnack 2009). Das im Kapitel 3 detailliert beschriebene Industriemessverfahren CLIPS (Mautz 2009) verwendet das Verfahren der Passpunkt Projektion mit dem inversen Kamera-Prinzip.

Mit einem Prototypen bestehend aus einem Projektor und einer einfachen Webkamera erreichen Köhler et al. (2007) an der ETH Zürich mit ihrem TrackSense-System eine Positionsgenauigkeit von 4 cm. Mit Hilfe des Projektors wird ein Gitternetz auf eine Wand projiziert, dessen Ko-

Tabelle 1: Optische Verfahren zur Innenraumpositionierung

Name	Koordinaten-referenz	Genauigkeits-bereich	Ab-deckung	CCD Größe [Pixel]	Mess-frequenz	Objekt/Kamera Positionierung	Kamera Kosten	Markt-reife
Tappero	keine	dm – m	30 m^2	356 × 292	3 Hz	Obj., OV6620	20 $ US	Vorschlag
Maye	Bilder/Barken	1 % vom Weg	skalierbar	16 × 16	2300 Hz	Kam., ADNS-2051	1.35 €	Entwicklung
Sjö	Scan/Bilder	sub m	skalierbar	320 × 240	30 Hz	Kam., VC-C4R	700 €	Entwicklung
Kohoutek	GIS	dm	skalierbar	176 × 144	54 Hz	Kam., SR 4000	9000 $	Vorschlag
Ido	Bilder	dm	skalierbar	320 × 240 × 4	30 Hz	Kam., IEEE1394		Entwicklung
Kitanov	Kantenmodell	dm	skalierbar	752 × 585	10 Hz	Kam., EVI-D31	£ 245	Entwicklung
Sky-Trax	Kodierte Marken	2 – 30 cm	skalierbar			Kam.		Produkt
Hile	Gebäudeplan	30 cm	skalierbar	640 × 480	0.1 Hz	Obj., Handy	100 $	Entwicklung
TrackSense	Projektion	4 cm	25 m^2	640 × 480	15 Hz	Obj.,CamPro4000	200 $	Entwicklung
Mulloni	Kodierte Marken	cm – dm	skalierbar	176 × 144	15 Hz	Kam., Handy	gering	Produkt
StarGazer	Kodierte Marken	cm – dm	skalierbar		20 Hz	Kam.	980 $	Produkt
Schlaile	Kantenmodell	1 dm/min	skalierbar	752 × 582	50 Hz	Kam., VC-PCC48P	175 €	Entwicklung
Muffert	Bilder	0.15 gon/min	raumweit	1616 × 1232 × 6	15 Hz	Kam., Ladybug3	>10000 $	Entwicklung
CLIPS	Projektion	0.5 mm	36 m^2	1032 × 778	30 Hz	Kam.,GuppyF80	1000 €	Entwicklung
AICON ProCam	Kodierte Marken	0.1 mm	Fahrzeug	1628 × 1236	7 Hz	Kam., ProCam	hoch	Produkt
naviSCAN3D	Projektion	50 μm	1.5 – 10 m	2448 × 2048 × 2	1 Hz	Obj., steroSCAN	hoch	Produkt

ordinaten sich bei einer bekannten relativen Orientierung zwischen Projektor und Kamera bestimmen lassen.

e) Verfahren ohne Referenzpunkte

Tappero (2009) schlägt ein extrem kostengünstiges System mit einer fest montierten Kamera vor, um bewegte Objekte auf Dezimetergenauigkeit zu verfolgen. Die bei der Verwendung von low-cost Komponenten mit geringen Speicherkapazitäten notwendige algorithmische Effizienz wird dadurch erreicht, dass lediglich aufeinanderfolgende Differenzbilder ausgewertet werden, in denen das bewegte Objekt mit einfachen Algorithmen lokalisiert werden kann.

In Tabelle 1 werden die wichtigsten Kenngrößen einiger optischer Positionsverfahren zusammengefasst.

3. Prinzipien und Aufbau des Systems CLIPS

CLIPS steht für Camera and Laser Indoor Positioning System und wird als Forschungsprojekt an der ETH Zürich entwickelt. Mit diesem Projekt soll ein System geschaffen werden, das

- ohne anzubringende Zielmarken auskommt
- keine komplizierte Mechanik beinhaltet
- kostengünstig ist (Bauteile < 1500 €) und
- das Potential für hohe Genauigkeit im Sub-Millimeter Bereich bietet.

In Abb. 1 ist das Prinzip von CLIPS veranschaulicht. Die Hauptbestandteile des Systems bestehen aus einer Digitalkamera (Guppy GF 080 C, 1032×778 Pixel, 30 Hz Bildfrequenz) und einer mit 16 Laserdioden bestückten Hartplastikkugel, die fortan als Laserigel bezeichnet wird. Durch den Laserigel werden sechzehn Laserstrahlen unter der Annahme eines gemeinsamen Zentrums radial in den Raum projiziert. Da die Ausrichtung der Laserstrahlen im Raum mittels einer vorangegangenen Kalibrierung durch Tilch (2009) auf $0.05°$ bekannt ist, kann mit Hilfe des Laserigels das Prinzip einer inversen Kamera realisiert werden. Mit der Digitalkamera wird das projizierte Referenzpunktfeld bestehend aus 16 Laserpunkten aufgenommen, wobei die Laserpunkte nicht zwangsweise auf eine Ebene projiziert werden müssen, sondern beliebig im Raum zu liegen kommen können.

Bevor eine Berechnung der relativen Orientierung möglich wird, müssen die Laserpunkte im Bild als solche erkannt und eindeutig den korrespondierenden Laserpointern zugeordnet sein. Um das zu erreichen, werden die Laserpunkte mit Hilfe von Schwellwert- und Segmentierungsoperatoren als Regionen extrahiert und die Regionszentren bestimmt. Da dieser Ansatz abhängig von günstigen Beleuchtungsbedingungen ist, soll in Zukunft die Punkterkennung durch ein Feature Matching erfolgen. Für die Identifizierung der einzelnen Laserpunkte im Bild wurden in Tilch (2010) drei verschiedene Methoden vorgeschlagen.

Bei der ersten Variante erfolgt die Kodierung zeitlich. Dabei wird die eindeutige Zuordnung der Bildpunkte zu den einzelnen Laserpointern mit Hilfe der Projektion unterschiedlicher, sich abwechselnder Muster von Laserpunkten ermöglicht. Allerdings ist dieser Ansatz für eine bewegte Kamera problematisch, da für die Identifizierung mehrere Bilder ausgewertet werden müssen. Für die Positionsbestimmung der Kamera in naher Echtzeit muss die Identifikation der Laserpunkte in nur einem Bild realisiert werden. Zwei alternative Konzepte wurden zu diesem Zweck erwogen. Zum einen können verschiedenfarbige Laserpointer eingesetzt werden, um die resultierenden Laserpunkte mit Hilfe einer Analyse der Farbkanäle (RGB) zu identifizieren. Allerdings sind Laser für das gesamte Farbspektrum kaum verfügbar, technisch aufwendig und entsprechend mit hohen Kosten verbunden. Die dritte Möglichkeit einer Punktidentifikation besteht darin, die Laserpointer in mehreren, eindeutigen Gruppen so anzuordnen, dass verschiedene Punktmuster auf eine Oberfläche projiziert werden. In Abhängigkeit der räumlichen Lage der Oberfläche so wie der Laserstrahlen kommt es dabei zu einer geometischen Verzerrung, die eine Identifizierung der Laserpunkte erschwert.

Nach einer erfolgreichen Erkennung und Identifizierung der einzelnen Laserpunkte im Bild kann mit Hilfe der Methode der relativen Orientierung die Pose, d. h. die Position und die Orientierung der Kamera bezüglich des Laserigels bestimmt werden.

Um das Prinzip der inversen Kamera zu veranschaulichen, kann eine virtuelle Bildebene gemäß Abb. 2 oberhalb des Laserigels konstruiert werden, die senkrecht zur Hauptachse des Laserigels orientiert ist. Die kalibrierten Raumwinkel der Laser werden mittels stereographischer Projektion in das Bildkoordinatensystem der virtuellen Bildebene transformiert. Das Projektionszentrum der Laserstrahlen entspricht dabei dem Projektionszentrum O, der inversen Kamera. Abb. 3 veranschaulicht die Lage der beiden Projektionszentren und Bildebenen zueinander. Die relative Orientierung der Kamera bezüglich des Laserigels lässt sich mit Hilfe der Koplanaritätsbedingung

$$\underline{r}'^T(\underline{b} \times \underline{r}'') = 0 \qquad (1)$$

bestimmen, die sich aus der epipolaren Geometrie ergibt. (1) sagt aus, dass die Projektionsstrahlen r', r, und der

Tabelle 2: Verfahren zur Punktidentifikation (Tilch, 2010)

Ansatz	Prinzip	Vorteile	Nachteile
1. Zeitliche Kodierung		+ bereits realisiert + einfache Realisierung	– zeitkritisch – mehrere Bilder notwendig
2. Farbliche Kodierung		+ nur ein Bild notwendig + Echtzeitfähig	– Kosten und mangelnde Verfügbarkeit farbiger Laser
3. Punktgruppen			+ nur ein Bild notwendig + Echtzeitfähig

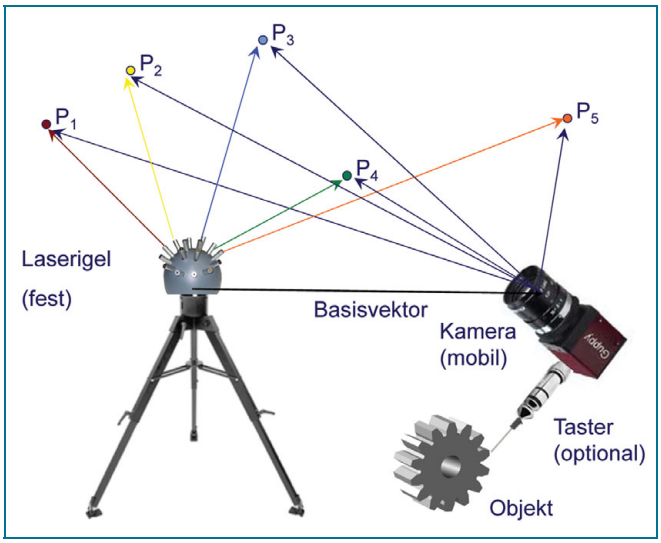

Abb. 1: Funktionsprinzip von CLIPS

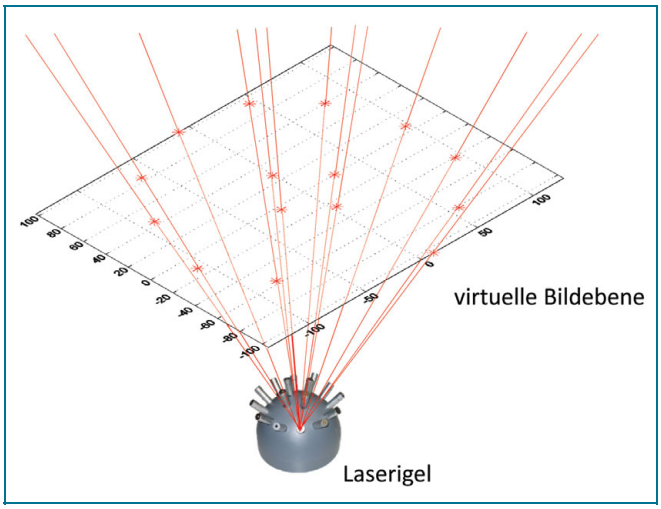

Abb. 2: Virtuelle Koordinatenebene des Laserigels

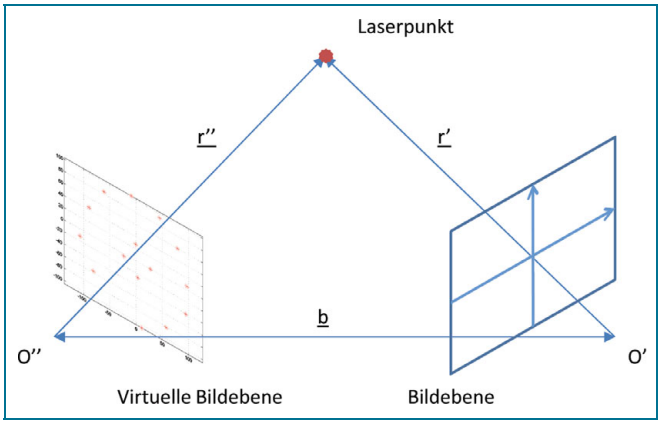

Abb. 3: Epipolare Geometrie bezüglich Laserigel und Kamera

Basisvektor b in einer Ebene liegen. Durch jede Koplanaritätsgleichung ergibt sich pro Laserpunkt je eine Bedingung. Für die Bestimmung der relativen Orientierung der Kamera sind die 3 Komponenten des Basisvektors sowie die 3 Richtungskomponenten einer Rotationsmatrix zu bestimmen. Da der Maßstab des Systems mit dieser Methode nicht bestimmt werden kann, lässt sich eine der 6 Unbekannten festhalten, so dass die Durchführung der relativen Bildorientierung mit 5 Laserpunkten möglich wird. Die Berechnung der relativen Orientierung erfolgt nach Stewénius (2005) mit Hilfe eines 5-Punkte-Algorithmus. Es sei dabei angemerkt, dass dieser Algorithmus bis zu 10 Lösungen liefert, da ein Polynom vom Grad 10 gelöst werden muss. Um die richtige Lösung ohne Kenntnis von Näherungswerten zu identifizieren, wird ein MSAC (M-Estimator SAmple Consensus) Algorithmus eingesetzt. Es müssen hierbei aus verschiedenen Minimalkonfigurationen (von jeweils 5 Punkten) unter Verwendung aller Daten bewertet werden, wodurch letztlich mehr als 5 Punkte benötigt werden. In einem abschließenden Schritt wird die Lösung der Orientierungsparameter mit Hilfe eines iterativen, linearen Newtonverfahrens ausgeglichen.

Der Maßstabsfaktor des Systems muss zusätzlich bestimmt werden. Für dessen Einführung stehen grundsätzlich verschiedene Methoden zur Verfügung, wie beispielsweise das Anmessen eines Probekörpers oder die Verwendung von mehreren Laserigelzentren mit bekannter Basis. Im Rahmen einer ersten Realisierung des Gesamtsystems in Tilch (2010) erfolgte die Bestimmung des Maßstabes durch tachymetrische Messung der Basis zwischen Laserigel und Kamera.

4. Erste Messergebnisse mit dem System CLIPS

Da die räumliche Lage des Projektionszentrums im Gehäuse der CLIPS-Kamera weder vom Hersteller angegeben noch ohne Aufwand bestimmbar war, wurde ein erster Systemtest nicht über die Kameraposition, sondern über die 3D-Koordinaten der Laserpunktzentren durchgeführt. Letztere ergeben sich nach erfolgter Kameraorientierung durch Berechnung von räumlichen Vorwärtsschnitten. Zwecks Genauigkeitsabschätzung wurden die Zentren der Laserpunkte ein weiteres Mal unabhängig mit einer Totalstation bestimmt. Nach diesem Schritt lagen die 3×16 Laserpunktkoordinaten in zwei verschiedenen Koordinatensystemen vor, die mit einer überbestimmten 3D-Ähnlichkeitstransformation überführt wurden. Es ergab sich ein Maßstabsfaktor von 0.9981 , was etwa 2 mm/m entspricht. Die starke Abweichung von 1 erklären wir uns daraus, dass bei der tachymetrischen Maßstabsbestimmung die Lage des Projektionszentums der Kamera nur grob geschätzt werden konnte. Die 1-σ Standardabweichung der Gewichtseinheit, die sich aus der Ähnlichkeitstransformation ergab, betrug 0.6 mm. Dieses Resultat kann lediglich als Genauigkeitsabschätzung bewertet werden, zumal die Vergleichsmessung nicht mit einem Theodolit-Messsystem, sondern provisorisch mit einem Tachymeter reflektorlos erfolgte und daher nicht den Genauigkeitsanspruch einer Absolutmessung erheben kann.

Die Bestimmung der Präzision, in unserem Fall die Wiederholgenauigkeit bei der Berechnung eines Kamerastandpunktes ergab bei 20 Messungen pro Standpunkt

eine mittlere 2-σ Standardabweichung von $\sigma_x = 0.16$ mm, $\sigma_y = 0.16$ mm und $\sigma_z = 0.35$ mm. Die größere Streuung in der Höhe führen wir auf die Geometrie der Laserpunkte zurück, die sich ausschließlich an der Raumdecke befanden.

5. Schlussfolgerungen und Ausblick

Die ermutigenden Testergebnisse im Submillimeter-Genauigkeitsbereich sollen nicht darüber hinwegtäuschen, dass diese Versuche unter Laborbedingungen durchgeführt wurden und weit von einer praktikablen Systemlösung entfernt sind. So gelingt eine zuverlässige Punkterkennung derzeit nur bei ausgeschalteter Raumbeleuchtung und die Punktidentifikation nur dann, wenn sich alle Punkte im Sichtfeld der Kamera befinden. Auch muss noch eine geeignete Lösung für die Maßstabsbestimmung gefunden werden.
Dennoch weisen die Tests daraufhin, dass CLIPS das Potential für einen Einsatz in der Ingenieur- oder auch in der Industrievermessung besitzt. Für Anwendungen zur Navigation ließe sich eine Low-Cost Variante entwerfen. Zunächst einmal gilt es, das System zur Reife zu bringen, sodass eine Anwendung realisiert werden kann.

Literatur

[1] AICON 3D Systems (2010), http://www.aicon.de, zuletzt besucht März 2010.

[2] Breuckmann stereoScan-3D (2010), http://www.breuckmann.com/, zuletzt besucht März 2010.

[3] Desouza, G. und Kak, A. (2002): Vision for Mobile Robot Navigation: A Survey, IEEE Transaction on Pattern Analysis and Machine Intelligence, Band 24, Nr. 2, Seiten 237-267.

[4] Frank, M. (2008): Eine Idee nimmt Gestalt an 3D-Daten von physischen Modellen und Bauteilen erzeugen, QZ. Qualität und Zuverlässigkeit, Band 53, Nr. 12, Seiten 53-55.

[5] Hagisonic (2008): User's Guide Localization system StarGazerTM for Intelligent Robots, http://www.hagisonic.com/, zuletzt besucht März 2010.

[6] Hile, H. und Borriello, G. (2008): Positioning and Orientierung in Indoor Environments Using Camera Phones, IEEE Computer Graphics and Applications, Juli/August 2008, Seiten 32-39.

[7] Ido, J., Shimizu, Y., Matsumoto, Y. and Ogasawara, T. (2009): Indoor Navigation for a Humanoid Robot Using a View Sequence, The International Journal of Robotics Research, Band 28, Nr. 2, Seiten 315-325.

[8] Kitanov, A., Bišević, S., Petrović, I. (2007): Mobile robot self-localization in complex indoor environments using monocular vision and 3D model. Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Zürich, Switzerland.

[9] Köhler, M., Patel, S., Summet, J., Stuntebeck E. und G. Abowed (2007): TrackSense: Infrastructure Free Precise Indoor Positioning Using Projected Patterns, in Pervasive Computing, LNCS, Band 4480, Seiten 334-350.

[10] Kohoutek, T., Donaubauer, A. and Mautz, R. (2010): „CityGML als Grundlage für die Indoor Positionierung mittels Range Imaging", Tagungsband des 15. Münchner Fortbildungsseminars Geoinformationssysteme.

[11] Lowe, D. (2004): Distinctive Image Features from Scale-Invariant Keypoints In: International Journal of Computer Vision. Band 60, Nr. 2, Seiten 91-110.

[12] Maye, O., Schaeffner, J. und Maaser, M. (2006): An Optical Indoor Positioning System for the Mass Market, Tagungsband des WPNC'06, Seiten 111-115.

[13] Mautz, R. (2009): Camera and Laser Indoor Positioning System (CLIPS) (unveröffentlicht), Förderungsgesuch an den Schweizer Nationalfond.

[14] Muffert, M., Siegemund, J., Förstner, W. (2010): The estimation of spatial positions by using an omnidirectional camera system. Tagungsband der 2nd International Conference on Machine Control & Guidance, Seiten 95-104.

[15] Mulloni, A., Wgner, D., Schmalstieg, D., Barakonyi, I. (2009): Indoor Positioning and Navigation with Camera Phones, Pervasive Computing, IEEE, Band 8, Seiten 22-31.

[16] Schlaile, C., Meister, O., Frietsch, N., Kessler, C., Wendel, J., Trommer, G.F. (2009): Using natural features for vision based navigation of an indoor-VTOL MAV, Aerospace Science and Technology Band 13, Nr. 7, Seiten 349-357.

[17] Sjö, K., López, D., Paul, C., Jensfelt, P. and Kragic D. (2009): Object Search and Localization for an Indoor Mobile Robot, Journal of Computing and Information Technology, Band 17, Nr. 1.

[18] Sky-trax inc. (2010): http://www.sky-trax.com/, zuletzt besucht März 2010.

[19] Stewénius, H. (2005): Groebner Basis Methods for Minimal Problems in Computer Vision, Dissertationsschrift, Centre for Mathematical Sciences, Lund Institute of Technology.

[20] Tappero, F. (2009): Low-cost optical-based indoor trakking device for detection and mitigation of NLOS effects, in Procedia Chemistry, Band 1, Nr. 1, Seiten 497-500.

[21] Tilch, S. (2010): Entwicklung eines optischen Innenraum-Positionierungssystems, Masterarbeit, Institut für Geodäsie und Photogrammetrie, ETH Zürich, 2010.

[22] Tilch, S. (2009): Kalibrierung der Laser für CLIPS (unveröffentlicht), Masterprojektarbeit, Institut für Geodäsie und Photogrammetrie, ETH Zürich, 2009.

[23] Wollnack, J. (2009): 3D/6D-Visionsysteme und Lasermessverfahren in der Robotik und Fertigungstechnik der Luftfahrtindustrie, AVN, Nr. 5, 2009, Seiten 175-187.

Anschrift der Autoren:
Dr.-Ing. Rainer Mautz
MSc ETH Sebastian Tilch
Swiss Federal Institute of Technology, Zurich
Institute of Geodesy and Photogrammetry
Wolfgang-Pauli-Str. 15
CH-8093 Zuerich
E-Mail: mautz@geod.baug.ethz.ch