

Dickmanns Dirk (1997). Rahmensystem für die visuelle Wahrnehmung veränderlicher Szenen durch Computer. Diss. UniBw München, Fak. Informatik

Zusammenfassung

Als Ziel der Arbeit sollte durch Abstraktion von speziellen Anwendungen ein flexibles, in vielfältigen Anwendungsgebieten einsetzbares Systemkonzept für visuelle dynamische Lagebestimmung und die Erfassung anderer Veränderungen von mit bildgebenden Sensoren erfaßbaren Objekten entworfen und realisiert werden.

Dazu wurden bestehende Systeme, insbesondere die an der UniBwM entstandenen, analysiert und eine Architektur entwickelt, in der die anwendungsspezifischen Modellbasen von den allgemein anwendbaren Verfahren und Methoden streng getrennt sind.

Das wesentliche Konzept zur Trennung von anwendungsunabhängigem Rahmensystem und anwendungsspezifischen Modellbasen war die Einführung von Schemata für generische Objektmodelle und für hierarchische Szenenmodelle als Organisationsstruktur (Szenenbaum) für Instanzen der Objektmodelle. In den Szenenbäumen können mehrere Instanzen von Objekt- und Kameramodellen in frei wählbaren Bezugssystemen relativ zueinander angeordnet werden.

Die Methoden zur dynamischen Lagebestimmung wurden so verallgemeinert, daß sie von den speziellen Inhalten der Modellbasen unabhängig sind und auf den in den Schemata dargestellten Modellen operieren können. Dazu waren folgende Erweiterungen und Neuentwicklungen gegenüber den zuvor existierenden Systemen erforderlich:

1. *Kronos*: An die Eigenschaften von Bildkanten anpaßbare, Subpixel-genaue Bildmeßverfahren mit hoher Winkelauflösung für Videobildfolgen (auf Parallelrechnern implementiert);
2. aspektabhängige Auswahl von vermeßbaren Merkmalen und Ermitteln geeigneter Parameter für die Meßverfahren;
3. Kombination von 2D-Bild- und 3D-Objekt Modellen für
 - schwellwertfreie zyklische Merkmalzuordnung und
 - Subpixel Auswertung mit Schätzung von systematischen Abweichungen zwischen vorhergesagten und tatsächlichen Meßwerten (*offset*-Schätzung);
4. Als Teil der Schätzverfahren wird die Jacobi Matrix (Matrix der partiellen Ableitungen der Meßgrößen nach den zu schätzenden Größen) für beliebige Szenenbäume durch mehrfache Anwendung des Meßmodells numerisch ermittelt; damit ist die Berechnung einer Jacobi Spalte zu jedem beliebigen Parameter des Meßmodells möglich. Durch Aufnahme von Form- und Kameraparametern in die Parameterliste des Meßmodells können neben den Zustandsgrößen auch diese Größen geschätzt werden (autonome Kalibrierung von Kameras und Kameraanordnungen).

Als fachliche Grundlagen wurden dabei folgende Techniken aus den Gebieten CAD, Computergrafik, Robotik und Systemdynamik eingesetzt:

1. homogene Koordinatentransformationen und Szenenhierarchien;
2. parametrisierte Formmodelle;
3. 4D-Ansatz zur Bildfolgenauswertung:
 - Bildmessungen: Anwendung von Masken entlang Suchpfaden;
 - dynamische Modelle und Meßmodelle zur Meßwertvorhersage;
 - rekursive Schätzverfahren: Kalman-Filter.

In den Schemata können eine große Vielfalt von Objekten und Szenen modelliert und die Bewegungen der Objekte dann von dem Rahmensystem erfaßt werden. Die Modellbasen können leicht erstellt und ausgetauscht werden. Damit ist, wie in 3.3 gefordert, die Methodik des 4D-Ansatzes von anwendungsspezifischem Wissen über Objekte und Szenen getrennt und die Aufgabe der visuellen Zustandsschätzung bewegter Objekte in komplexen Szenen in sehr allgemeiner Form gelöst worden. Die Unabhängigkeit der Methoden und Verfahren von dem jeweiligen

Anwendungsgebiet wurde durch Modellierung und visuelle dynamische Zustandsschätzung folgender sehr unterschiedlicher Objekte und Szenen gezeigt:

- verschiedene Objekte für Labor-Experimente,
- der Freiflieger des Roboter-Technologie Experiments ROTEX der DLR im Rahmen der *Spacelab-D2*-Mission an Bord der Raumfähre Columbia, der anhand der ermittelten Lagegrößen als Premiere im All automatisch eingefangen werden konnte; sowohl die Roboterhand und damit die Kamera als auch das zu fangende Objekt bewegten sich dabei mit sechs Freiheitsgraden im Raum;
- Flugplätze mit Landebahnen, Hubschrauber-Landeflächen und Rollbahnen zur autonomen Landung mit Flächenflugzeugen und Drehflüglern (in Echtzeit-Simulationen mit echter Hardware für den "künstlichen Gesichtssinn" im Kreis); dabei wurden für Navigationsaufgaben auch Straßenkreuzungen und andere Merkmale der Umgebung erfaßt;
- menschliche Gesichter zur Bestimmung von Kopfbewegungen und Blickrichtung.

Neben der Lage- und Bewegungsschätzung von mehreren Objekten in frei wählbaren Bezugssystemen ist mit der numerischen Berechnung der Jacobi-Matrix die Schätzung von Form- und anderen Objektparametern möglich; alle Größen und Parameter, die sich auf die Meßgrößen auswirken, können geschätzt werden, wobei die systemdynamische Beobachtbarkeit vorausgesetzt wird.

Die *offset*-Schätzung für Meßwerte erlaubt die Nutzung von hochgenauen Meßwerten und damit genauere Schätzwerte auch bei weniger genauen 3D-Formmodellen; dabei werden genaue Modelle gewissermaßen als Kombination der 3D-Formmodelle und der *offsets* im Bild *online* ermittelt und so ein erster Schritt zu einfachen Lernvorgängen getan. Ohne *offset*-Schätzung könnte der Vorteil der Subpixelvermessung nur mit sehr genauen Kamera- und 3D-Formmodellen genutzt werden; eine derart genaue Modellierung ist aber in der Praxis nur sehr selten möglich.

In dem Schema für Szenenmodelle können auch gegliederte Objekte als Teilszenen dargestellt werden (zum Beispiel ein Roboter-Arm); elastische Objekte könnten mit den Mitteln der generischen Formmodelle und Transitionsmatrizen modelliert werden. Besonders interessant sind die

Konsequenzen, die sich aus der Behandlung von Kameras als Bestandteil der Szenenbäume ergeben:

1. Kameras können sich an beliebiger Stelle des Szenenbaums befinden, also auch in mobilen autonomen Systemen und/oder auf steuerbaren Kameraplattformen.
2. Die Lage und Bewegung der Kameras kann ebenso wie die anderer Objekte fortlaufend geschätzt werden. Damit ist die Eigenzustandsbestimmung von autonomen Fahrzeugen möglich.
3. Es kann eine wählbare Anzahl von Kameras für multi-okulare Stereoauswertung mit beliebigen Stereobasen (*skew stereo*) und/oder unterschiedlichen Brennweiten für multifokales (foveales und peripheres) Sehen eingesetzt werden.
4. Auch Kameraparameter können geschätzt werden. Das wird, wie bei den Formparametern von Objekten, durch die Integration dieser Parameter in das Meßmodell und damit in die Berechnung der Jacobi Matrix ermöglicht. Zusammen mit der Lageschätzung von Kameras ist damit eine Kalibrierung von einzelnen Kameras und von Anordnungen mehrerer Kameras *online* möglich.

Insbesondere die Eigenzustandsschätzung, die Stereobildfolgenauswertung und die Kamerakalibrierung wurden bisher fast stets als eigenständige Aufgaben betrachtet; eine Ausnahme ist [Wünsche 87]. Mit dem Rahmensystem *Realis* können sämtliche Schätzaufgaben *simultan* ausgeführt werden.

Weitere besonders zu erwähnende Eigenschaften des Rahmensystems sind:

1. Für die dynamische Lagebestimmung werden eindimensionale Meßwerte verwendet, nämlich die Abstände der Schnittpunkte von Bildkanten und Suchpfaden von den Mittelpunkten der Suchpfade. Diese Meßwerte können mit geringem Rechenaufwand von *Kronos* ermittelt werden. Das damit verbundene Apertur Problem (die Verschiebung einer Kante kann nur in Richtung des Suchpfads, nicht aber längs der Kante selbst erfaßt werden) wird durch das Verfahren zur numerischen Berechnung der Jacobi Matrix und das Zusammenspiel der Ergebnisse mehrerer Messungen gelöst.

2. Das Rahmensystem ermöglicht die Bewertung und Überprüfung von Hypothesen, indem es Gütemaße für die Übereinstimmung von Modell und Meßdaten bereitstellt. Damit wird eine Anbindung von Systemen für initiale (statische) Lagebestimmung ermöglicht.

Die Implementierung des Systems erfolgte ursprünglich in **occam** und in **C++**. Die ständig steigende Größe des Systems veranlaßte den Autor zur Suche nach einer für komplexe Systeme besser geeigneten Sprache. Mehrere Sprachen wurden hinsichtlich verschiedener Kriterien verglichen; schließlich wurde Ada 95 der Vorzug gegenüber Eiffel 3 und Modula 3 gegeben. Mit der Reimplementierung in Ada wurde hohe Zuverlässigkeit erreicht, ohne im Vergleich zu der C++-Version und einer zwischenzeitlich entstandenen ANSI-C-Version nennenswerte Einbußen im Laufzeitverhalten (<20%) hinnehmen zu müssen.

Das Rahmensystem konnte sehr klein gehalten werden (ca. 350 KB Ada 95 für *Realis* und ca. 250 KB **occam** für *Kronos*), was wiederum eine weitgehende Optimierung bei vertretbarem Arbeitsaufwand ermöglichte. Das System ist insgesamt so effizient, daß es auf modernen PCs oder Workstations ohne zusätzliche Spezialhardware für die oben angegebenen Anwendungen Zykluszeiten von 40 ms erreicht und damit für diese Anwendungen echtzeitfähig ist.

Eine wesentliche Erweiterung des Rahmensystems wäre die weitergehende Integration des dritten grundlegenden Konzepts zur Modellbildung, also der radiometrischen Eigenschaften elektromagnetischer Strahlung. Damit sind detailliertere Objektmerkmal-Modelle, genauere Meßwertvorhersagen und der Einsatz von Farbkameras möglich. Mit der Hinzunahme der Parameter derartiger Beleuchtungs- und Ausbreitungsmodelle in das Meßmodell könnten auch für diese Parameter Jacobi-Spalten berechnet und eine Schätzung dieser Größen versucht werden, zum Beispiel die Lageschätzung von Lichtquellen oder die Schätzung von Reflexionsparametern. Die *offset* Schätzung könnte dann auch für gemessene Intensitätswerte eingesetzt werden.