

Automatischer Weißabgleich für eine omnidirektionale Kamera

T. Wilhelm, J. Kludas, H.-J. Böhme and H.-M. Gross
Technische Universität Ilmenau, Fachgebiet Neuroinformatik
PF 100565, 98684 Ilmenau

{torsten.wilhelm,hans.joachim.boehme,horst-michael.gross}@tu-ilmenau.de

Zusammenfassung

Eine weitgehende Farbkonstanz stellt eine wesentliche Voraussetzung für die visuelle Informationsverarbeitung in der moderner Servicerobotik dar. Als Beispiel seien hier die visuelle Nutzerdetektion in realen Umgebungen mit bewegten Kamerasystemen auf Basis von Hautfarbdetektion und die visuelle Selbstlokalisierung des Roboters genannt. Diese Verfahren sind sehr anfällig gegen Änderungen der Beleuchtungsbedingungen. Es wird ein System vorgestellt, das in der Lage ist, unter realen Bedingungen mit variierenden Beleuchtungsbedingungen eine weitgehende Farbkonstanz zu erreichen. Dazu wurde eine omnidirektionale Kamera mit einer Weißreferenz ausgestattet, mit der ein automatischer Weißabgleich der Kamera möglich wird.

1 Einleitung

Der Einsatz von Servicerobotern in realen Umgebungen gewinnt in den letzten Jahren immer mehr an Bedeutung. Die meisten dieser Roboter stellen allerdings eine bestimmte Dienstleistung bereit, ohne dabei aktiv mit einem Nutzer zu kommunizieren. Ein Grund hierfür mag die Schwierigkeit der visuellen Detektion von Personen unter realen Einsatzbedingungen sein. Trotzdem ist es für eine natürliche Mensch-Maschine-Kommunikation unerlässlich, Personen auch auf mobilen Systemen während der Ausführung einer Serviceleistung robust zu detektieren und zu verfolgen. Falls der Nutzer unaufmerksam ist, sich anderen Dingen zuwendet oder die Kommunikation ganz abbricht, sollte der Roboter geeignete Aktionen ausführen und nicht stupide weiter seine Serviceaufgabe ausführen.

Ein ganz anderes Problem stellt die visuelle Selbstlokalisierung des Roboters dar. Er muss zu jedem Zeitpunkt wissen, an welcher Stelle im Markt er sich befindet, wobei er sich an visuellen Eindrücken seiner Umgebung orientiert.

Die Realisierung dieser Fähigkeiten ist wesentlich für Serviceroboter, die in realen Einsatzumgebungen arbeiten sollen. Ein entscheidendes Problem sind dabei die stark veränderlichen Beleuchtungsbedingungen, mit denen bei mobilen Systemen zwangsläufig gerechnet werden muss.

Unser Serviceroboter PERSES wird für den Einsatz in einem Baumarkt entwickelt und soll seine Nutzer zu von ihnen gewünschten Produkten lotsen [7]. Die Beleuchtungsverhältnisse können hier von reinem Kunstlicht bis zu reiner natürlicher Beleuchtung variieren. Für die genannten Verfahren ist es essentiell, dass trotzdem eine weitgehende Farbkonstanz gewährleistet werden kann. In den folgenden Abschnitten wird unser automatischer Weißabgleich im Zusammenhang mit der hautfarbbasierten Nutzerdetektion vorgestellt.

2 Perses

Bei unserem Serviceroboter PERSES handelt es sich um einen B21 Roboter der Firma RWI. Ein Touchdisplay, Sprachsynthese und ein auf einer Pan-Tilt-Unit angebrachter Kopf mit einem ansteuerbaren Gesicht aus LED-Arrays ermöglichen eine intuitive Interaktion mit Nutzern. Für die Nutzerdetektion und die visuelle Selbstlokalisierung wurde er mit einer omnidirektionalen Kamera ausgestattet, siehe Abbildung 1.



Abbildung 1: Serviceroboter PERSES mit omnidirektionaler Kamera (ganz oben) für die Nutzerdetektion und die visuelle Selbstlokalisierung. Mit einem solchen Objektiv ist es dem Roboter möglich, zu jedem Zeitpunkt sein gesamtes Umfeld zu analysieren.

3 Hautfarbe

Hautfarbe ist ein häufig verwendeter Cue bei der Suche nach Personen in Bildern. Sie bietet den Vorteil, unabhängig von Eigenbewegungen des Roboters berechnet werden zu können und eignet sich dadurch besonders für den Einsatz auf mobilen Systemen. Allerdings funktioniert sie nur dann zufriedenstellend, wenn die Beleuchtungsverhältnisse bei der Aufnahme der Trainingsdaten und bei der Anwendung hinreichend ähnlich sind, was in der Regel nicht gewährleistet werden kann.

Eine mögliche Lösung für dieses Problem besteht in der kontinuierlichen Adaption des Farbmodells an die jeweiligen Beleuchtungsbedingungen [3] [8]. Allerdings ist diese Vorgehensweise nach unserer Erfahrung problematisch, da das jeweilige Ziel, in diesem Fall das Gesicht eines Nutzers, für die Adaption in jedem Zeitschritt absolut sicher detektiert werden muss. Bei nur sehr geringfügigen Positionsfehlern neigen solche adaptiven Modelle dazu, von der Beschreibung der eigentlichen Zielregion wegzudriften. Ein zu starkes Wegdriften des Farbmodells kann verhindert werden, indem eine Adaption nur innerhalb eines allgemeinen Hautfarbmodells zugelassen wird, welches unter verschiedenen Beleuchtungsbedingungen aufgenommen wurde [1] [4].

Es gibt aber auch Bemühungen, die Lokalisation der Hautfarbregionen im Bild durch die Verwendung zusätzlicher Features zu stabilisieren. So wird in [1] ein Gesichtsdetektor eingesetzt und das Farbmodell immer nur dann adaptiert, wenn an der entsprechenden Stelle auch ein Gesicht erkannt wurde. Solche Bestrebungen können das genannte Problem zwar verringern, nicht jedoch beseitigen, denn kein Gesichtsdetektor hat eine Falsch-Positiv-Rate von null. Auf der anderen Seite erhöhen sie die Berechnungskomplexität, was sich wiederum negativ auf ein kontinuierliches Tracken der Hautfarbregionen auswirken muss.

Ein anderer Weg, Beleuchtungseinflüsse zu beseitigen, besteht in der Vorverarbeitung des Bildes mit Farbkonstanz-Algorithmen. Die Aufgabe solcher Algorithmen besteht in der Erzeugung eines Bildes, wie es bei einer sogenannten kanonischen Beleuchtung entstanden wäre, aus einem Bild, das bei einer beliebigen unbekanntem Beleuchtung entstanden ist. In [2] wurden verschiedene Farbkonstanz-Algorithmen auf ihre Tauglichkeit im Zusammenhang mit einer farbbasierten Objekterkennung getestet. Dabei konnte zwar eine signifikante Steigerung der Erkennungsraten durch die Anwendung von Farbkonstanz-Algorithmen festgestellt werden, trotzdem war keine robuste Erkennung unter variablen Beleuchtungen möglich.

Im Gegensatz zu den bisher genannten Verfahren werden in dieser Arbeit die speziellen Abbildungseigenschaften eines omnidirektionalen Objektivs ausgenutzt. Dadurch kann ein automatischer Weißabgleich durchgeführt und weitgehende Farbkonstanz erreicht werden, so dass mit einem nicht-adaptiven Farbmodell und ohne aufwendige Vorverarbeitung gearbeitet werden kann.

3.1 Farbraum

Für unsere Untersuchungen verwenden wir den helligkeitsnormierten dichromatischen r-g-Farbraum [3], da er sehr gut geeignet ist, Hautfarbe in einem großen Bereich unterschiedlicher Beleuchtungsverhältnisse zu repräsentieren [8]. Im Prinzip ist aber jeder Farbraum geeignet, bei dem Helligkeits- und Farbinformationen dekorreliert vorliegen.

$$r = R/(R + G + B) \quad (1)$$

$$g = G/(R + G + B) \quad (2)$$

3.2 Farbmodell

Das verwendete Farbmodell besteht aus einer Lookup-Tabelle mit manuell als Hautfarbe klassifizierten Pixeln im r-g-Farbraum. Ein ähnlicher Ansatz, allerdings im YUV-Farbraum, wurde in [5] vorgestellt. Das verwendete Farbmodell ist in Abbildung 2 zu sehen. Auf eine Approximation der Hautfarbregion mit einer bivariaten Gaußverteilung wie in [3] wurde bewußt verzichtet, da diese die Genauigkeit der Abbildung reduziert. Allerdings muss bei dieser Vorgehensweise darauf geachtet werden, dass die Stichprobe hinreichend groß ist, so dass die entstehende Hautfarbregion keine „Löcher“ aufweist.

Die Varianzen dieser Verteilung hängen zum einen von den Variationen der Hautfarben der verschiedenen Probanden ab, aber auch von den Beleuchtungsschwankungen während der Aufnahme der Trainingsdaten. Nur wenn letztere möglichst gering gehalten werden, kann ein spezifisches Hautfarbmodell erzeugt werden. Verglichen mit der Hautfarbverteilung in [1] hat die Hautfarbverteilung in Abbildung 2 eine ähnliche Form (Skin locus) [6], aber eine wesentlich geringere Ausdehnung. Dies ist darauf zurückzuführen, dass die Daten für dieses Modell unter Verwendung des im nächsten Abschnitt beschriebenen automatischen Weißabgleichs aufgenommen wurden.

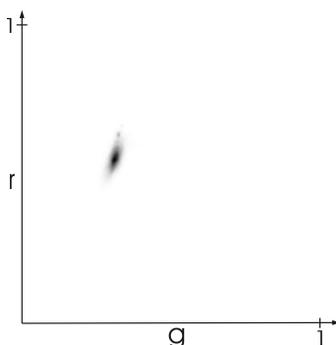


Abbildung 2: Lookup-Tabelle für Hautfarbe im dichromatischen r-g-Farbraum.

4 Automatischer Weißabgleich

4.1 Weißreferenz

Abbildung 3 verdeutlicht die drastischen Unterschiede der Kamerabilder bei verschiedenen Beleuchtungsbedingungen. Um diesem Problem zu begegnen, wurde ein automatischer Weißabgleich entwickelt, der auf den Bildern der omnidirektionalen Kamera arbeitet. Hierzu wurde die Kamera mit einem weißen Ring ausgestattet, der als Weißreferenz im Bild erscheint. Abbildung 4 zeigt die Kamera mit omnidirektionalem Objektiv und Weißreferenz und ein mit dieser Kamera aufgenommenes Bild, das die Referenz auf einem inneren Radius enthält. Die Oberfläche der Weißreferenz ist nicht horizontal und flach, sondern weist eine leichte konvexe Krümmung auf. Auf diese Weise wird auch von der Seite kommendes Licht berücksichtigt.

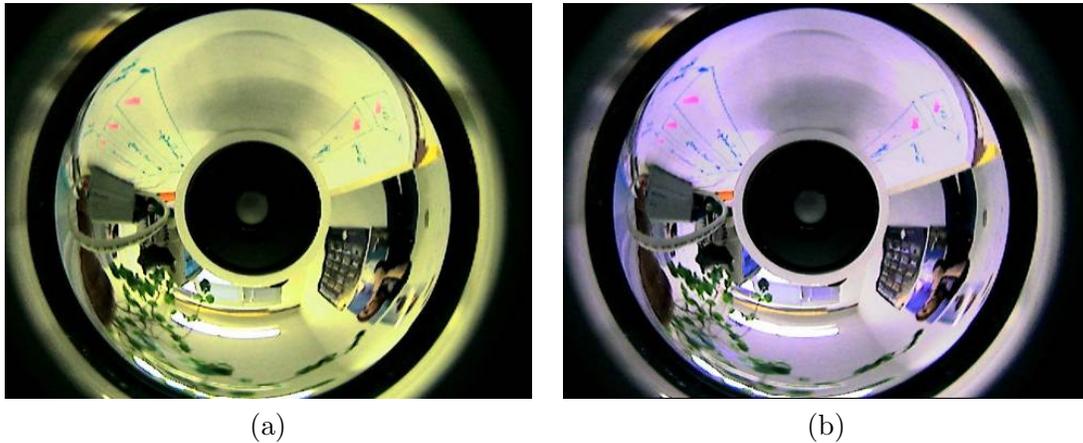


Abbildung 3: Zwei Bilder der omnidirektionalen Kamera, (a) mit weitgehend natürlicher Beleuchtung und (b) mit weitgehend künstlicher Beleuchtung.

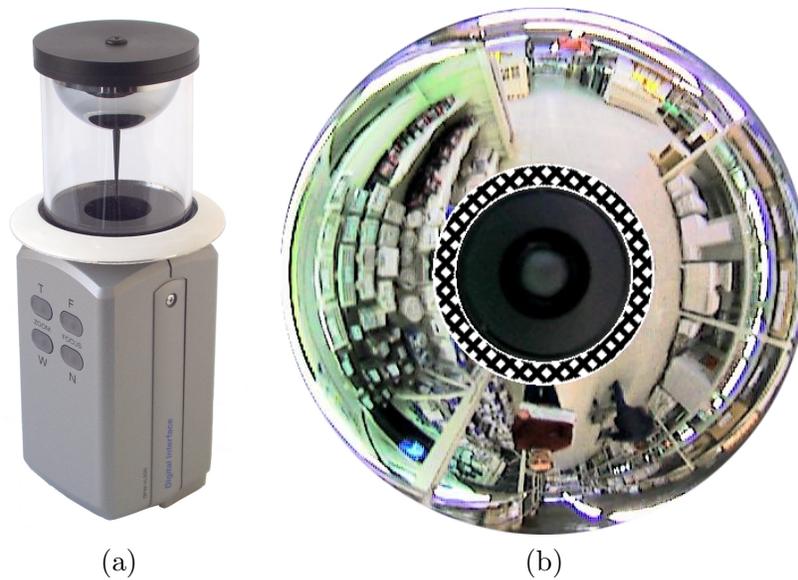


Abbildung 4: (a) Weißreferenz zwischen Kamera und Objektiv. (b) Mit dieser Kamera aufgenommenes Bild, wobei die Weißreferenz in der Nähe des Zentrums des Bildes erscheint und mit einer Schraffur gekennzeichnet wurde. Dieser Bildbereich entspricht dem Fußboden in der unmittelbaren Umgebung des Roboters und ist für eine Nutzerdetektion irrelevant.

4.2 Regelung

Es wird der Mittelwert für R , G und B über alle Pixel berechnet, die sich innerhalb der Weißreferenz befinden. Dieser Mittelwert wird dann in den YUV -Farbraum transformiert. Mit der Differenz zu den Sollwerten $U = 0$ und $V = 0$ werden mit Hilfe von zwei separaten PID-Reglern die Stellgrößen für die Verstärkungsfaktoren für U und V für die verwendete Digitalkamera (SONY DFW VL500) bestimmt. Außerdem wird der mittlere Wert von Y verwendet, um die Iris der Kamera so zu steuern, dass eine annähernd gleichbleibende Helligkeit im Bild erreicht wird. Abbildung 5 zeigt den Aufbau der Regelung, der in den nächsten Abschnitten im Detail beschrieben wird.

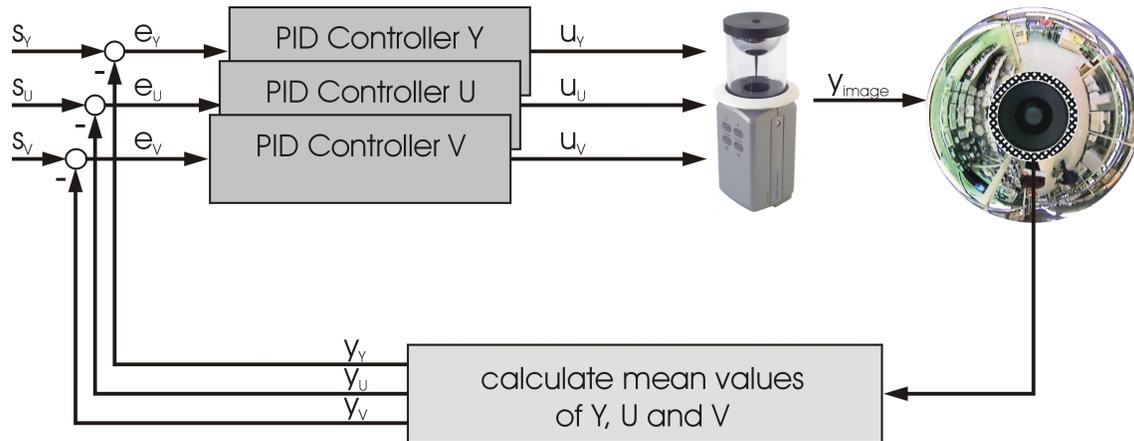


Abbildung 5: Struktur des digitalen Regelkreises.

4.3 Identifikation der Regelstrecke

Für die Bestimmung der Parameter der Regelstrecke wurde die Sprungantwort $h(t)$ für die Komponenten Y, U und V auf der Weißreferenz aufgezeichnet, siehe Abbildung 6. Dazu wurde das System durch An- und Ausschalten des Lichtes mit einem Sprung beaufschlagt.

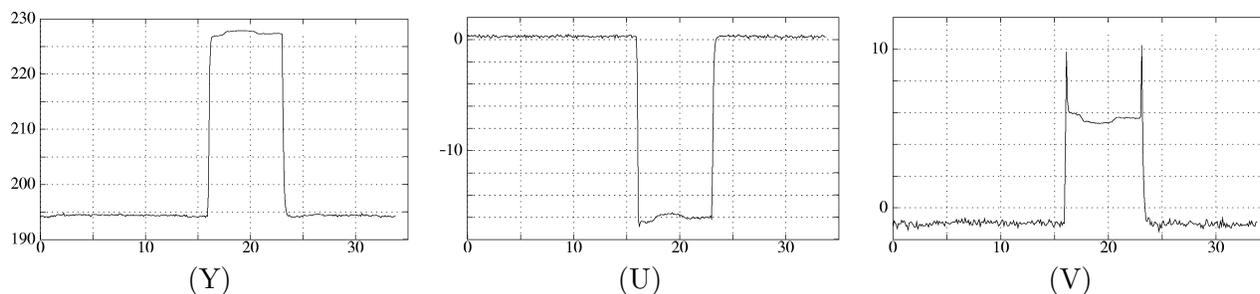


Abbildung 6: Sprungantworten für Y, U und V. Es wurde mit einer Abtastzeit von $T = 9Hz$ gearbeitet. Die Messwerte sind über die Zeit in Sekunden aufgetragen.

Diese Kurvenverläufe wurden unter der Annahme, dass es sich um PT2-Glieder handelt, mit Origin6 approximiert, siehe Abbildung 7. Hierzu wurden die Messwertkurven auf eine Höhe von 1 normiert, da die Verstärkung des Systems nur bestimmt werden kann, wenn die Höhe des Eingangssprunges bekannt ist. Das Überschwingen von V wurde nicht mit berücksichtigt. Mit dieser Approximation ergeben sich die Zeitkonstanten aus Tabelle 1.

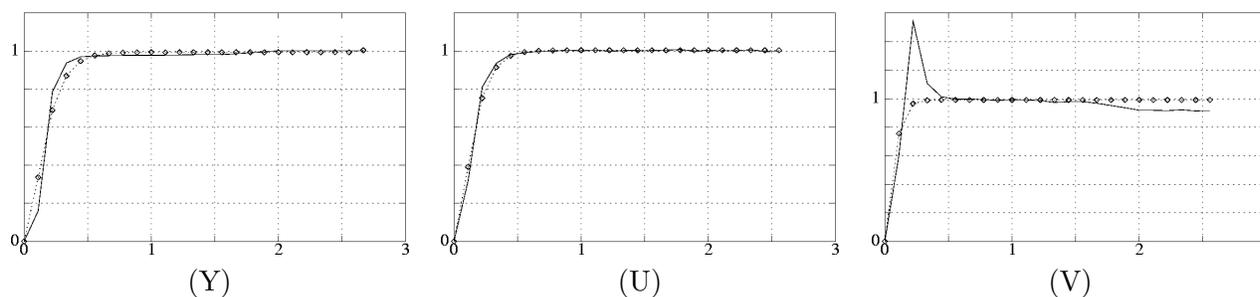


Abbildung 7: Fitting der Sprungantworten mit PT2-Gliedern. Die Messwerte sind über die Zeit in Sekunden aufgetragen.

4.4 Bestimmung der Reglerparameter

Der Regler wird nun so eingestellt, dass seine Nullstellen die Polstellen der Strecke kompensieren. Die Übertragungsfunktion der Strecke $G_S(s)$ und die des Reglers $G_R(s)$ ist:

$$G_S(s) = \frac{K}{(1 + T_1 s)(1 + T_2 s)} \quad (3)$$

$$G_R(s) = \frac{K}{1 + \frac{1}{T_N s} + T_V s} \quad (4)$$

Damit ergeben sich die Reglerparameter aus Tabelle 1.

4.5 Bestimmung der Reglerverstärkungen

Für die Bestimmung der Reglerverstärkungen wurde die Verstärkung eines einzelnen Reglers variiert, während die der anderen Regler fest auf einen kleinen Wert eingestellt wurden. Dabei wurde das System mit Sprüngen beaufschlagt und eine Verstärkung gesucht, bei der das System möglichst schnell, aber stabil ist. Im Gesamtsystem mit allen drei Reglern können die Verstärkungswerte so nicht verwendet werden, da sonst Instabilitäten auftreten. Deshalb wurden die Reglerverstärkungen sukzessive verkleinert, bis das Gesamtsystem stabil blieb. Abbildung 8 zeigt die Sprungantworten für verschiedene Verstärkungsfaktoren, die letztendlich gewählten Verstärkungsfaktoren sind in Tabelle 1 abgebildet.

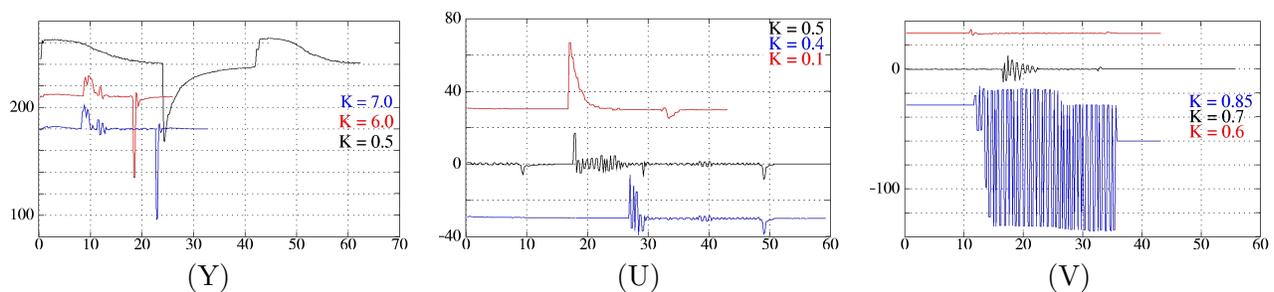


Abbildung 8: Sprungantworten für verschiedene Verstärkungsfaktoren.

Strecke	K	T_1	T_2	Regler	K	T_N	T_V
Y	0.99408	0.09218	0.09218	Y	6.0	0.184	0.046
U	1.00443	0.08243	0.08243	U	0.4	0.1648	0.0412
V	0.99134	0.04025	0.04025	V	0.6	0.0805	0.0201

Tabelle 1: Zeitkonstanten der Regelstrecken und des Reglers.

5 Ergebnisse

5.1 Hautfarberkennung

Abbildung 9 zeigt die Auswirkungen des automatischen Weißabgleiches auf die Hautfarbdetektion. Die Stellgrößen für den Weißabgleich sind direkt nach dem Einschalten des Systems mit Vorgabewerten belegt. Nach nur zehn Bildern (rund 1 Sekunde) ist die Regelabweichung nahezu null.



Abbildung 9: Zunehmende Detektionen des Hautfarbdetektors durch den automatischen Weißabgleich. Es handelt sich um das erste, das fünfte und das zehnte Bild einer Sequenz. Neben der eigentlichen Hautfarbregion entstehen auch mehr falsch-positive Detektionen auf holzfarbenen Objekten.

5.2 Systemverhalten

Für einen Test der Stabilität des Gesamtsystems wurde die Kamera jeweils mit einer gelben und einer blauen Folie zur Simulation von Farbstichen abgedeckt. Trotz der extremen Farbänderung blieb das System in beiden Fällen stabil, siehe Abbildung 10.

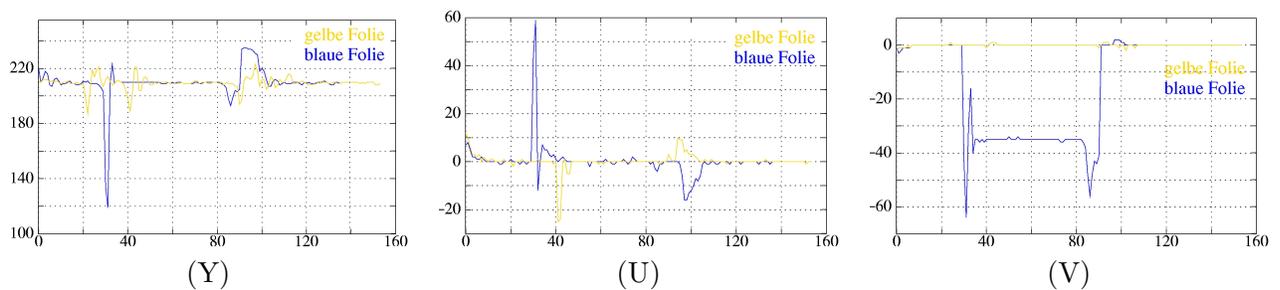


Abbildung 10: Systemverhalten bei Abdeckung der Kamera mit Farbfolien. Bei der blauen Farbfolie konnte die Regelabweichung für V nicht ausgeregelt werden, da die Stellgröße in die Sättigung lief.

6 Zusammenfassung und Ausblick

Es wurde ein System für einen automatischen Weißabgleich für eine Kamera mit omnidirektionalem Objektiv vorgestellt. Dieses Verfahren ermöglicht eine robuste Hautfarberkennung unabhängig von den aktuellen Beleuchtungsbedingungen. Ein Nachteil des Verfahrens besteht darin, dass die Beleuchtung am Roboter, an dem die Messung durchgeführt wird, nicht die selbe wie am Nutzer ist. Dieser Sachverhalt wirkt sich aber bei einer weitgehend homogenen Beleuchtung nur sehr geringfügig aus.

Weiterführende Untersuchungen könnten sich mit eventuellen Abhängigkeiten zwischen den Größen Y , U und V beschäftigen. Gegebenenfalls könnten mit einem Mehrgrößenregler noch bessere Ergebnisse erzielt werden.

Literatur

- [1] Fritsch, J., Lang, S., Kleinhagenbrock, M., Fink, G.A., and Sagerer, G. Improving adaptive skin color segmentation by incorporating results from face detection. In *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication (ROMAN), Berlin, Germany, 2002*.
- [2] Funt, B., Barnard K., and Martin, L. Is machine colour constancy good enough? *Lecture Notes in Computer Science*, 1406:445–459, 1998.
- [3] H.-J. Böhme, U.-D. Braumann, A. Brakensiek, A. Corradini, M. Krabbes, and Gross, H.-M. User localisation for visually-based human-machine interaction. In *Proc. 1998 IEEE Int. Conf. on Face and Gesture Recognition, Nara, Japan*, pages 486–491, 1998.
- [4] Jang, G.-J. and Kweon, I.-S. Robust object tracking using an adaptive color model. In *Proc. of the IEEE Int. Conference on Robotics and Automation*, pages 1677–1682, 2001.
- [5] S. Feyrer and A. Zell. Detection, tracking, and pursuit of humans with an autonomous mobile robot. In *International Conference on Intelligent Robots and Systems (IROS '99)*, pages 864–869, 1999.
- [6] Störring, M., Andersen, H.J., and Granum, E. Physics-based modelling of human skin colour under mixed illuminants. *Robotics and Autonomous Systems*, 34(3-4):131–142, 2001.
- [7] Wilhelm, T., Böhme, H.-J., and Gross, H.-M. Looking closer. *Proceedings of the European Conference on Mobile Robots*, 2003, to appear.
- [8] Yang, J. and Waibel, A. Skin-color modeling and adaptation. *Lecture Notes in Computer Science*, 1352:687–694, 1998.