

# Automatisierte Objektaufnahme für Bilddatenbanken

Wolfram Hans und Dietrich Paulus

Universität Koblenz-Landau  
Aktives Sehen  
{hans,paulus}@uni-koblenz.de,  
<http://www.uni-koblenz.de/agas>

**Zusammenfassung** Die Gewinnung von Trainingsdaten für Objekterkennungsverfahren bedeutet die Erstellung von umfangreichem Bildmaterial der zu lernenden Objekte. Soll ein Objekt mit mehreren Ansichten unter verschiedenen Beleuchtungen mit unterschiedlichen Kameras aufgenommen werden, ergeben sich so viele Kombinationsmöglichkeiten, dass diese manuell nur mit hohem zeitlichen Aufwand durchlaufen werden können. Durch Automatisierung lässt sich nicht nur der Aufwand einschränken, sondern bei der Speicherung von Objektbildern mit deren Zusatzinformationen auch die Fehleranfälligkeit durch manuelle Dateneingabe verringern. Zu diesem Zweck wird die Schnittstelle zu einer Datenbank genutzt, die nach speziellen Anforderungen der Bildverarbeitung entworfen wurde. In ihr können neben den eigentlichen Bildinformationen Objektmasken, Lichtquellen und weitere spezifische Daten gespeichert werden.

## 1 Einleitung

Viele Methoden der Objekterkennung benötigen umfangreiche Trainingsdaten für die Ermittlung von Objekteigenschaften. Neben den zum Lernen benötigten Bildinhalten sind auch Informationen über den Bildentstehungsprozess [2] nützlich. Dies sind vor allem Kameraparameter, Kalibrierdaten und Angaben zu den Beleuchtungsverhältnissen, die bei der Objektaufnahme vorliegen. Um von einem Objekt mehrere Ansichten zu gewinnen, wird es aus verschiedenen Blickwinkeln aufgenommen. Eine gängige Methode, die z. B. auch in [1] verwendet wird, ist der Einsatz eines Drehtellers. Damit wird der manuelle Aufwand, Objektansichten in großer Zahl zu gewinnen, deutlich reduziert. Ein ähnlicher Ansatz wird auch hier verfolgt.

Im Gegensatz zu anderen existierenden Bilddatenbanken wie COIL-100<sup>1</sup> etc. werden hier automatisch zusätzliche Informationen direkt beider Bildgewinnung erzeugt und verwendet. Neben Objektwinkel sind dies vor allem Angaben zur Lichtquelle und Objektmasken. Die Nutzung eines Farbprüfmusters ermöglicht später auch eine Farbnormierung der Objekte.

<sup>1</sup> <http://tev.fbk.eu/DATABASES/objects.html>

Neben der Automatisierung, die den manuellen Aufwand deutlich reduziert, ist vor allem die Kombination mit der Bilddatenbank, die die zusätzlichen Informationen aufnimmt, von Bedeutung.

## 2 Bilddatenbank für Farbobjekterkennung

Basis des Systems ist eine Datenbank, die speziell auf die Bedürfnisse der farb-basierten Objekterkennung ausgerichtet ist. Eine wesentliche Anforderung dabei ist, möglichst viele Informationen über den Bildentstehungsprozess zu speichern.

Über sogenannte Katalogtabellen werden beispielsweise Angaben zur verwendeten Kamera oder zur vorherrschenden Beleuchtung gemacht, die später bei der Weiterverarbeitung genutzt werden können. Die Bilder selbst werden Projekten und diese einzelnen Szenen zugeordnet. Zusätzliche Daten werden in sogenannten EAV-Tabellen (*Entity Attribute Value*) abgelegt. Diese Daten sind nahezu beliebig und können beispielsweise aus Binärmasken, Histogrammen oder ähnlichem bestehen. Desweiteren besteht die Möglichkeit, die Bildinformation in unterschiedlichen Auflösungsstufen abzulegen. Genutzt wird diese Funktionalität von der Anwendung selbst, die beim Import von Bildern eine Vorschau (Thumbnail) anlegt.

Im Folgenden wird zwischen einem Bild-Datensatz (dem *Bild*) als zentralem Element und den reinen Bildinformation, d.h. den Binärdaten, unterschieden (*Bilddaten*). Ein *Bild* in der Datenbank enthält also sämtliche verfügbaren Informationen über Aufnahmebedingungen, Objekte und Binärdaten. Letztere beschränken sich nicht nur auf die Bildinhalte sondern können auch Masken sein, die dem *Bild* zugeordnet sind. Dies ist besonders dann sinnvoll, wenn einzelne, alleinstehende Objekte aufgenommen werden. Die Objektposition selbst lässt sich über Rahmen (bounding boxes) ebenfalls speichern und kann so nachfolgenden Prozessen die Suche erleichtern oder als Referenz für Objekterkennungsalgorithmen dienen. So lässt sich beispielsweise eine Farbnormalisierung vereinfachen, wenn die Lage des Objektes "ColorChecker<sup>®</sup>" bekannt ist. Eine Besonderheit der bounding boxes ist, dass sie über relative Koordinaten gespeichert werden. Dies ermöglicht die genaue Lokalisierung der umschlossenen Objekte in jeder Auflösungsstufe.

## 3 Systemaufbau

Das System zur Automatisierten Objektaufnahme besteht aus mehreren Komponenten, die miteinander verbunden sind und zentral gesteuert werden. Der Drehteller als übliches Hilfsmittel zur Erzeugung mehrerer Objektansichten wird in einer Lichtbox mit drei definierten Leuchtquellen (D65, Leuchtstoffröhre und Glühlicht) positioniert. Einen Überblick über den Aufbau gibt Abbildung 1(a). Die Anwendung bietet darüber hinaus die Möglichkeit mehrere Kameras nacheinander zur Aufnahme zu verwenden. Mit der Verwendung eines Greifarms können letztlich auch verschiedene Objekte nacheinander auf dem Drehteller platziert werden.

### 3.1 Drehteller

Als gängige Methode, Objekte aus verschiedenen Blickwinkeln aufzunehmen, wird auch hier ein Drehteller verwendet. In der Anwendung kann dabei der Drehwinkel entweder über die Anzahl der aufzunehmenden Bilder oder aber direkt angegeben werden. Die Genauigkeit der Ansteuerung übertrifft im vorliegenden Fall mit 27 Schritten pro Grad die Anforderung einer präzisen Objektrotation. Die Angaben über den Drehwinkel bei der Aufnahme werden über die Anwendung dann als Information mit dem Bild in der Datenbank gespeichert.

### 3.2 Lichtbox

Die Basis der verwendeten Lichtbox bildet eine konventionelle *JUST Pantone*<sup>®</sup> *Color Viewing box* mit drei Lichtquellen. Dabei werden die Schalter der einzelnen Lichtquellen durch die Ausgänge einer vierfach-Relaiskarte ersetzt. Der verbleibende vierte Ausgang steht für eine zusätzliche Lichtquelle zur Verfügung. Die Kanäle der Karte selbst werden über einen I/O-Warrior (s. Abbildung 1(c)) angesteuert, der eine USB-Schnittstelle zum Steuerrechner bietet. Über die Anwendung können die Lichtquellen einzeln angesteuert werden. Dabei ist die Kombination mehrerer Leuchtmittel zu einem definierten Mischlicht möglich. Hierfür ist eine Zuordnung der in der Datenbank vorhandenen Lichtquellen zu den Schaltmöglichkeiten der Relaiskarte nötig, d.h. die jeweilige Licht-ID in der Datenbank wird fest mit einem Zustand der Relais verknüpft.

### 3.3 Kameras

Die Anwendung sieht vor, mehrere Kameras einzusetzen. Diese werden nacheinander aktiviert und für die Bildaufnahme genutzt. Dabei werden sowohl Firewire- als auch USB-Kameras unterstützt. Die Zuordnung der Kameras zu den in der Datenbank vorhandenen Einträgen wird über die System-ID der Kamera realisiert. So können evtl. vorhandene Kalibrierinformationen oder Sensitivitätskurven der Kamera genutzt werden, wenn sie in der Datenbank vorliegen. Diese Nutzung mehrerer Kameras ermöglicht eine weitere Variation der Aufnahmesituation und damit die automatische Gewinnung umfangreicher Bilddaten. Je nach geforderter Genauigkeit ist jedoch zu berücksichtigen, wie die Kameras angeordnet sind. Eine horizontale Anordnung bedeutet bei geringem Abstand schon eine merkliche Änderung des Betrachtungswinkels, der damit fehlerhaft in der Datenbank abgelegt wird. Für einfache Genauigkeit sollte eine vertikale Anordnung genutzt werden. Eine präzise Positionierung erreicht man aber letztlich nur durch den Einsatz eines weiteren rechnergesteuerten Systems, wie z. B. einem Linearschlitten, der die Kameras jeweils in die korrekte Position bewegt.

### 3.4 Greifarm

Eine weitere Automatisierung ist durch den Einsatz eines Greifarms erzielt. Durch seine Verwendung werden die zu erfassenden Objekte nacheinander auto-

matisch auf dem Drehteller platziert. Dies ermöglicht die unbeaufsichtigte Aufnahme mehrerer Objekte. Die zu erfassenden Objekte werden dabei einer Position zugeordnet und in einer Liste gespeichert, die dann vom Greifarm abgearbeitet wird, nachdem jeweils alle Varianten von Drehwinkel, Beleuchtung und Kamera abgearbeitet sind. Gesteuert wird er über eine achtfach-Relaiskarte (Abbildung 1(d)), die eine USB-Schnittstelle zum Rechner bereitstellt. Jeweils zwei Kanäle werden dabei für einen Freiheitsgrad des Greifarms verwendet. So z. B. öffnen und schließen des Greifers, drehen im und gegen den Uhrzeigersinn usw. Mit acht Kanälen sind damit nur vier Achsen des Roboters steuerbar. Bei dem verwendeten fünf-achsigen Greifarm wird auf die Rotation des Greifers verzichtet, da die Objekte immer senkrecht (oder waagrecht) erwartet werden. Eine mögliche Anordnung der Objekte zeigt Abbildung 1(b).

## 4 Graphische Benutzerschnittstelle

Die Benutzeroberfläche der Anwendung gliedert sich in drei Teile. Der erste ermöglicht eine Teilsicht auf die Einträge in der Datenbank und stellt u.a. die Katalogtabellen für Kamera und Lichtquelle zur Verfügung. Der zweite Teil bietet Zugriff auf die Basisfunktionalitäten, während der letzte der Konfiguration und mitunter Steuerung der Peripherie dient. Eine Übersicht der graphischen Benutzerschnittstelle zeigt Abbildung 3

### 4.1 Anwendungsbeschreibung

Über den Konfigurationsbereich ist z. B. das Bild der Kamera zu erreichen. Die Wahl des Bildausschnittes und optische Kontrolle der aufzunehmenden *Bilddaten* ist über den ersten Reiter *Live Image* zu erreichen.

Den zweiten Reiter (*Image mask settings*) des Konfigurationsteils aktiviert man zur Erzeugung von Binärmasken, was in Abschnitt 5 genauer beschrieben wird. Dieser Bereich wird für die manuelle Schwellwertbestimmung einmal vor dem geplanten Durchlauf aufgerufen.

Im Bereich *Grabbing device* werden die Kameras angezeigt, die am System angeschlossen sind. Die für den Bildaufnahmeprozess vorgesehenen Kameras können hier ausgewählt werden. Eine Zuordnung zu den in der Datenbank vorhandenen Kameras erfolgt über die Identifikationsnummer, mit der die Kamera sich am Computer anmeldet.

Die Lichtquellen werden im Bereich der Katalogtabellen der Datenbank angezeigt und können dort ausgewählt werden. Eine manuelle Steuerung der Relais kann jedoch über *Lightsources* erfolgen. Dies ist z. B. bei der Einrichtung der Lichtbox oder Platzierung des Drehtellers hilfreich.

Dieser selbst, oder genauer sein Verhalten, wird über den allgemeinen Teil der Benutzeroberfläche konfiguriert. Jedoch kann auch hier über *Turntable* eine manuelle Justage zur Einrichtung des Aufbaus erfolgen.

Hinter dem Reiter *Settings* verbirgt sich die Einstellung, welche Datenbank verwendet werden soll oder ob die Daten in eine Datei gespeichert werden.

Im letzten Teil des Konfigurationsteils befindet sich unter *Robot arm settings* die Konfiguration des Greifarms. Dort wird die Position des Drehtellers eingegeben, sowie die Koordinaten der Objekte, die auf ihm plaziert werden sollen. Dies erfolgt in einer Liste, die aus den ausgewählten Objekten der Datenbank erzeugt wird.

## 4.2 Vorgehensweise

Im Folgenden wird ein Ablauf beschrieben, der für die Gewinnung einer Bildserie mit dem vorgestellten System beispielhaft ist.

Zunächst wählt der Anwender ein Projekt aus, für das der geplante Durchlauf vorgesehen ist, sowie eine Szene, der die Konfiguration zugeordnet wird. Aus weiteren Tabellen kann das verwendete Kameramodell ausgewählt werden, wenn es nicht durch das System automatisch erkannt wird.

Bei der Auswahl der Lichtquelle(n) spielt die Reihenfolge eine Rolle. Denn soll eine Binärmaske erzeugt werden, wird dafür die erste ausgewählte Lichtquelle genutzt.

Die Objekte selbst, die in dem Durchlauf aufgenommen werden sollen, sind ebenfalls auszuwählen. Sollte dabei ein Eintrag fehlen, da das Objekt, das aufgenommen werden soll, in der Datenbank noch nicht vorhanden ist, kann dieses ergänzt werden.

Bei einer Mehrfachauswahl wird anschließend bei der Greifarmsteuerung eine Liste erzeugt, in die die Positionen der Objekte eingetragen werden müssen. Ein Draufsicht auf eine beispielhafte Positionierung der Objekte zeigt Abbildung 1(b). Die Positionsangaben erfolgen metrisch mit dem Koordinatenursprung im Zentrum der Roboterplattform. Sie umfassen die zwei Achsen in der Ebene der Roboterplattform, sowie die Greifhöhe darüber. Die Position des Drehtellers oder genauer, der Ort, an dem das Objekt abgesetzt werden soll, wird auf die gleiche Weise einmalig in einem separaten Feld eingetragen. Exemplarisch zeigt Abbildung 2 die Ablage eines Objekte auf dem Drehteller.

Ein weiterer Schritt ist die Konfiguration des Drehtellers. Neben der individuellen Steuerung über den Bereich, der über den Reiter *Turntable* zu erreichen ist, bietet der statische Teil der Anwendung eine einfache Eingabe des gewünschten Verhaltens. So kann der Drehwinkel durch Angabe der Ansichten vorgegeben werden. D.h. für den Fall, dass man z. B. ein Objekt aus zwölf Ansichten erhalten möchte, gibt man diese Zahl in das dafür vorgesehene Feld ein. Alternativ kann man den Drehwinkel, in diesem Fall  $30^\circ$ , angeben. Der Drehbereich selbst ist außerdem einschränkbar. Wenn man beispielsweise nur Bilder der Vorderseite, also einen Rotation um nur  $180^\circ$  machen möchte, ist dies auch einfach möglich.

Der schwierigste Teil ist die Erstellung der Maske. Hierzu wird die Option *make binary mask* ausgewählt und der Reiter *Image mask settings* aktiviert. Durch drücken der Taste *Background* wird ein Bild des Hintergrundes erzeugt. Dabei wird wie oben bereits erwähnt die erste selektierte Lichtquelle verwendet. Nach der Aufnahme steht das Bild zur Kontrolle zur Verfügung. Ein Klick auf *Foreground* nimmt entsprechend das Vordergrundbild auf. Über einen Schieberegler lässt sich dann ein Schwellwert einstellen, der für die aktuelle Beleuchtung eine

möglichst gute Binärmaske erzeugt, die unterhalb der beiden Ansichten angezeigt wird. Dies ist beispielhaft auch in Abbildung 3 zu sehen. Eine weiterführende Darstellung der Vorgehensweise bietet Abschnitt 5.

### 4.3 Ablauf

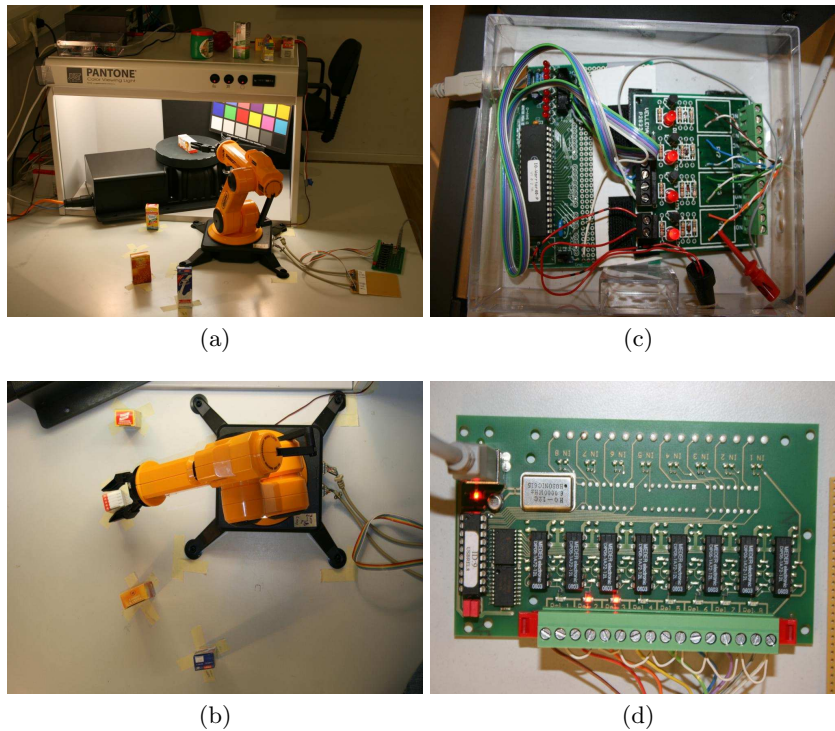
Die Sequenz selbst wird durch drücken des Knopfes *Turn* angestoßen und arbeitet danach die Permutationen von Beleuchtung, Kamera, Objekt und Rotationswinkel ab. Durch die vorangegangene Information wird dem *Bild* bereits ein Großteil der Daten zugeordnet. Den Rest, z. B. die Angabe des Drehwinkels, erfolgt durch die Anwendung und wird automatisch in der Datenbank abgelegt. Wie eine weitere Information, die Position des Objektes in den *Bilddaten*, wird im folgenden beschrieben.

## 5 Informationsgewinnung

Die *Bilddaten* sollen nicht einfach in einer Datenbank gespeichert werden, sondern zusätzliche Informationen enthalten, um für nachgelagerte Anwendungen in der Farbbildverarbeitung Nutzen zu bieten. Durch die einmalige Selektion der verschiedenen Elemente aus den Katalogtabellen in der Vorbereitung, werden diese Informationen automatisch für jedes *Bild* mit abgelegt, welches im Rahmen des Durchlaufs erzeugt wird.

Die Information über die Ansicht wird durch die Verwendung des Drehtellers mit Kenntnis seines Drehwinkels vollautomatisch ermittelt.

Eine zusätzliche Information des Objektes, nämlich seine Position in den *Bilddaten* wird durch die Erzeugung eines Differenzbildes gewonnen. Dazu wird wie in Abschnitt 4.2 beschrieben das Hintergrundbild vom Vordergrundbild abgezogen. Da durch Sensorrauschen und andere Störgrößen im Differenzbild noch fehlerhafte Maskenwerte existieren, werden verschiedene öffnende und schließende Operatoren angewendet. Außerdem wird mit einem Regler unter optischer Kontrolle das Ergebnis der Maske optimiert. Dieser Schwellwert wird manuell einmalig angepasst, um so genauere Objektmasken zu erzeugen. Der initial ermittelte Schwellwert wird dann auch für die weiteren Differenzbilder – die letztlich binarisiert werden – auch unter anderen Drehwinkeln verwendet. Die Binärmaske selbst wird ebenfalls in der Datenbank abgelegt und dem ausgewählten Objekt zugeordnet. Dies vereinfacht z. B. die Erzeugung von Farbhistogrammen der Objekte, da nur diese selbst im Bereich der Maske berücksichtigt werden. Mit der Binärmaske schließlich lässt sich, wie oben erwähnt, auch die Position des Objektes im Bild ermitteln, welche selbst in Form einer bounding box in der Datenbank abgelegt wird. Für die Objektbilder unter zusätzlichen, d.h. anderen Beleuchtungen werden die selben Masken aus der ersten Serie verwendet, da durch die Präzision des Drehtellers die Objektlagen identisch sind.



**Abbildung 1.** a) Systemaufbau, b) Draufsicht, c) IOWarrior mit Schaltkarte, d) 8fach-Relaiskarte.

## 6 Ausblick und Zusammenfassung

In dem Beitrag wird eine Anwendung beschrieben, die bei der Gewinnung umfangreicher Bilddaten für Farbobjekterkennungsverfahren hilft. Dabei werden mehrere Objektansichten sowie Kameras und Lichtquellen kombiniert. Eine Besonderheit ist, dass die Objekte mit ihrer Position, die automatisch ermittelt wird, und der zugehörigen Binärmaske in einer Bilddatenbank abgelegt werden. Die Verwendung unterschiedlicher Lichtquellen sowie der Einsatz verschiedener Kameras macht den Aufbau besonders für Anwendungen in der Farbbildverarbeitung interessant, da hier ein direkter Vergleich von verschiedenen Einflüssen unter kontrollierten Bedingungen möglich ist. Die Automatisierung, auch bei der Bestückung des Drehtellers, erlaubt die einfache Gewinnung umfangreicher Bilddatensätze für verschiedene Zwecke. Eine geplante Ergänzung ist die Lokalisation des ColorChecker<sup>®</sup> im Bild. Dadurch wird die Anwendung im Bereich der Farbbildverarbeitung noch interessanter, da damit die Bilder der Datenbank sofort z. B. für Farbnormalisierungsverfahren geeignet sind. Damit geht der vorgestellte Ansatz weit über die übliche Objektaufnahme mit Einsatz von Drehteller und Lichtkasten hinaus.

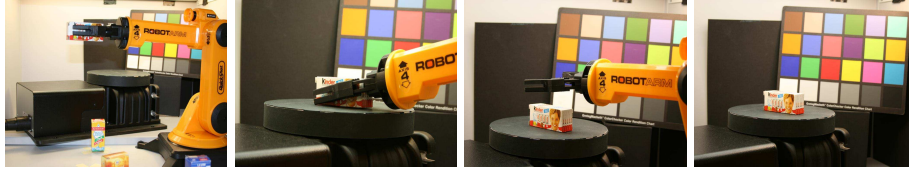


Abbildung 2. Bestückung des Drehtellers durch den Greifarm.

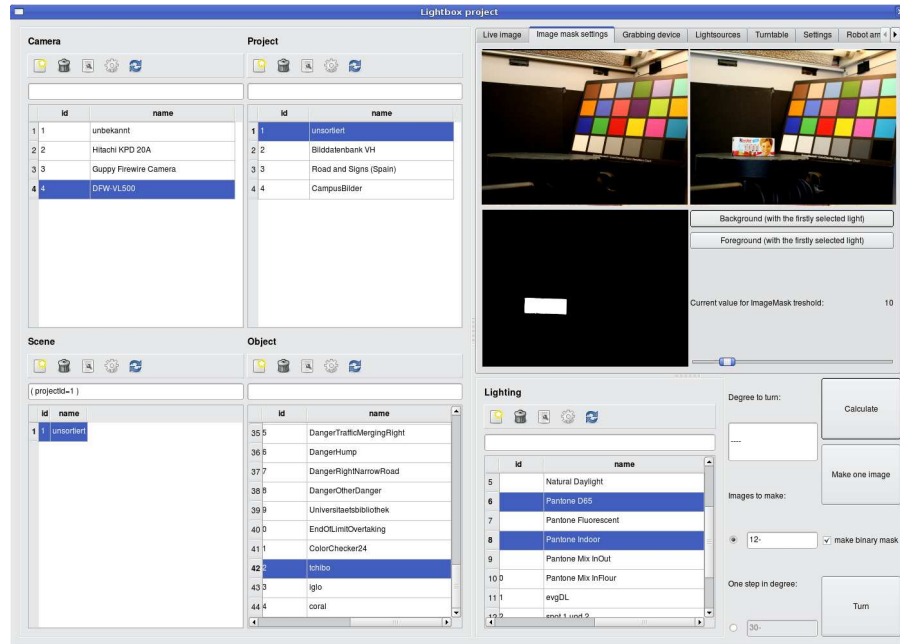


Abbildung 3. Graphische Benutzeroberfläche am Beispiel der Maskenerzeugung.

## Literatur

1. Grzegorzek, M.: Appearance-Based Statistical Object Recognition Including Color and Context Modeling. PhD thesis, Technische Fakultät der Universität Erlangen-Nürnberg (2007)
2. Hans, W., Grosch, T., Feldmann, T., Paulus, D., Müller, S.: Modell der bildentstehung mit hdr-kameras. In: 12. Workshop Farbbildverarbeitung. (10 2006) 97–108