

Computergrafik 2: Organisatorisches

Prof. Dr. Michael Rohs, Dipl.-Inform. Sven Kratz

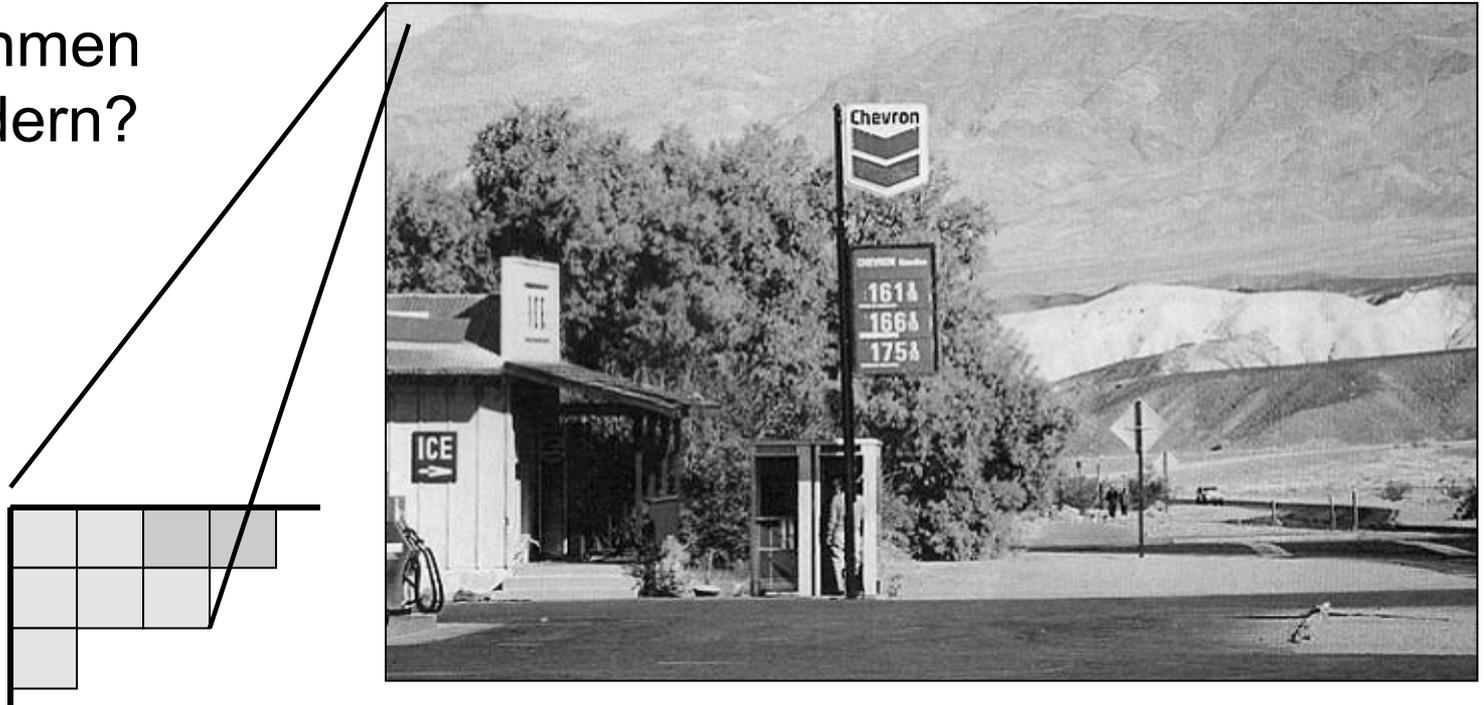
michael.rohs@ifi.lmu.de

MHCI Lab, LMU München

Folien teilweise von Andreas Butz, sowie von Klaus D. Tönnies
(Grundlagen der Bildverarbeitung. Pearson Studium, 2005.)

Digitale Bilder

- Zwei-dimensionale Anordnung von Bildelementen (Pixel)
- Datenstruktur: 2-dimensionales Feld (Array, Matrix)
- Dateneinheit: 1 Pixel
- Algorithmen auf Bildern?



Quelle: Tönnies: Grundlagen der Bildverarbeitung.

Bildverarbeitung und Bildanalyse

- Bildverarbeitung
 - Verbesserung der Bildqualität für die menschliche Wahrnehmung
 - Trennung von Informationen und Artefakten (Störungen)
- Bildanalyse
 - Identifikation von Bedeutung tragenden Merkmalen
 - Berechnung von Merkmalswerten
 - Zuordnung von Bedeutung
- Algorithmische Verarbeitung von Informationen in Bildern
 - Restauration, Verbesserung, Segmentierung, Verstehen

Informationen in Bildern

- Videokompression
 - $720 \times 576 \times 3 = 1,2\text{MB}$ pro Bild
 - $24 \text{ Bilder/s} \times 1,2\text{MB/Bild} \times 60\text{s/min} = 1,8\text{GB/min}$
 - Kapazität DVD ca. 8.5 GB: knapp 5min Video unkomprimiert
 - ca. 120min Video komprimiert \rightarrow 96% der Daten weggeworfen
- Was sind die „wesentlichen“ 4%?
 - Definition von „wesentlich“
 - „Unwesentliches“ herausnehmen
 - Unabhängig vom Inhalt
- Wie macht man das?
 - Redundanz ausnutzen, invariante Merkmale bestimmen
 - Methoden: Algorithmen, Mathematik

Übung: Anzahl möglicher Bilder

- Gegeben: $n \times n$ Grauwertbild (256 Grauwertstufen)
- Wie viele mögliche Bilder?
- Ab welchem n ist die Anzahl möglicher Bilder größer als die Anzahl Atome im Universum?

Themen der Vorlesung

- Grundlagen der Bildverarbeitung
 - Entfernen von Bildstörungen
 - Helligkeits- und Kontrastanpassung
 - Filterung im Ortsraum und Frequenzraum
- Grundlagen der Bildanalyse
 - Kantendetektion
 - Segmentierung
 - Merkmalsbasierte Bilderkennung
- Praktische Anwendung
 - Python, NumPy & matplotlib

Aufnahme digitaler Bilder

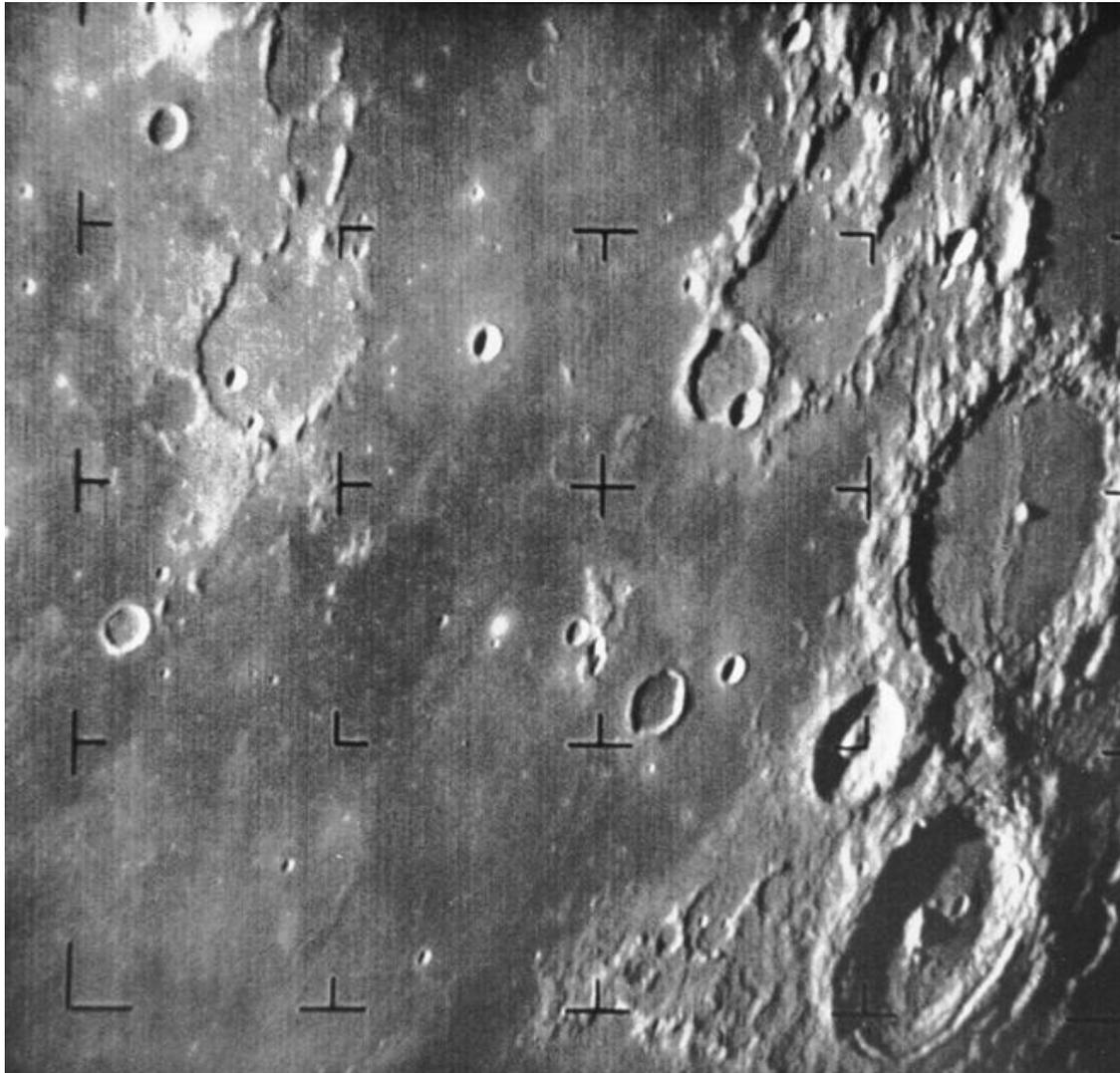
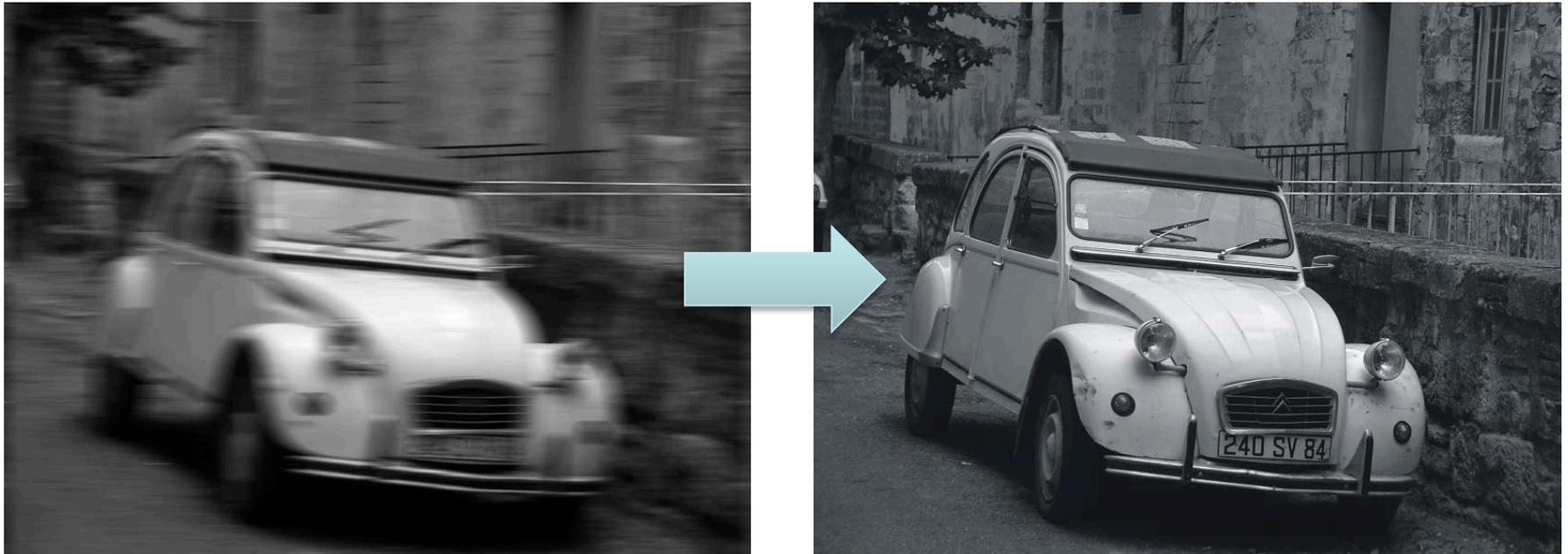


FIGURE 1.4 The first picture of the moon by a U.S. spacecraft. *Ranger 7* took this image on July 31, 1964 at 9 :09 A.M. EDT, about 17 minutes before impacting the lunar surface. (Courtesy of NASA.)

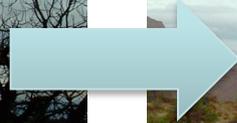
© R. C. Gonzalez & R. E. Woods, Digital Image Processing

Bildrestauration



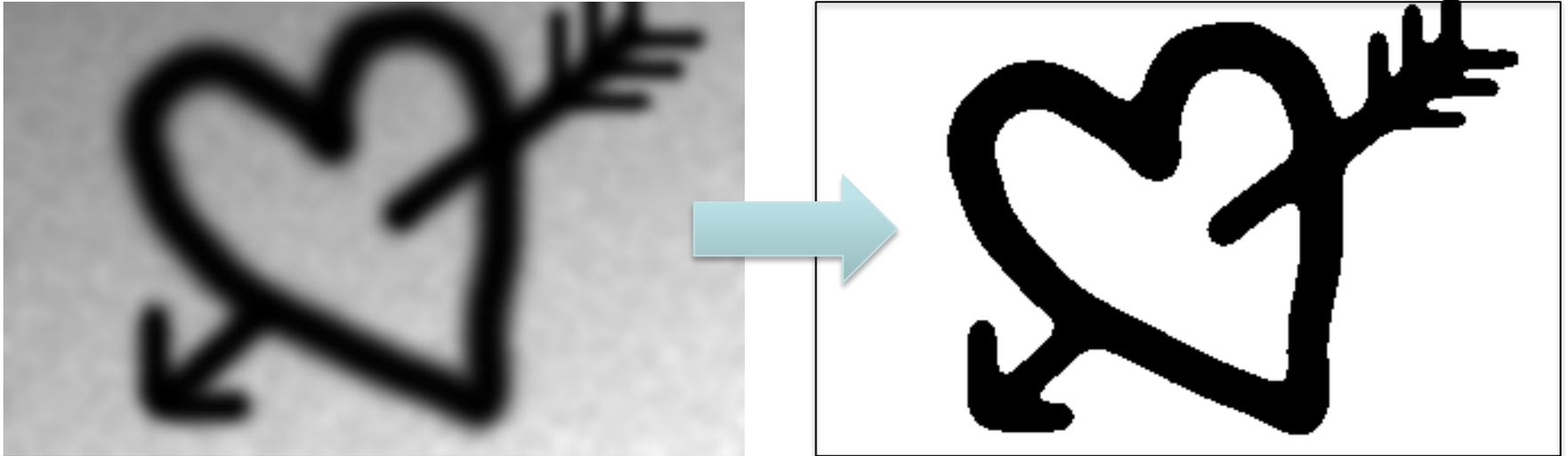
Bewegungsunschärfe eliminieren

Bildverbesserung



Helligkeit und Kontrast anpassen

Bildsegmentierung



- Hypothese: Gruppen von Pixeln sind Informationsträger
- Zusammenfassen von Pixeln nach bestimmten Kriterien
- Modellwissen?
- Algorithmen?

Bildverstehen

- Hypothese: Die visuellen Informationen reichen aus, um die korrekte Bedeutung zu finden
 - Stimmt das?
- Fragestellungen
 - Was sind kennzeichnende Merkmale und wie können sie berechnet werden?
 - Wie können Merkmale aus dem Bild mit dem erwarteten Inhalt verglichen werden?

Bedeutung

- Anwendungsabhängig: ein Bild kann viele Bedeutungen haben
 - Sind Menschen abgebildet?
 - Wo sind Fahrzeuge zu sehen?
 - Ist Mr. X auf dem Bild?
 - Wie spät ist es?
 - Ist ein Straßenschild im Bild?
 - Wo wurde die Aufnahme gemacht?
- Erster Schritt einer Bildanalyse
 - Spezifikation der gesuchten Bedeutung



Die semantische Lücke

Bild

- Helligkeit
- Farbe
- Musterung



Interpretation

- Name
- Existenz/Nichtexistenz von Objekten
- Eigenschaften von Objekten



Auflösung, Kontrast, Größe,
Perspektive, etc. unterschiedlich



„Mount Fuji“

Bedeutung
gleich

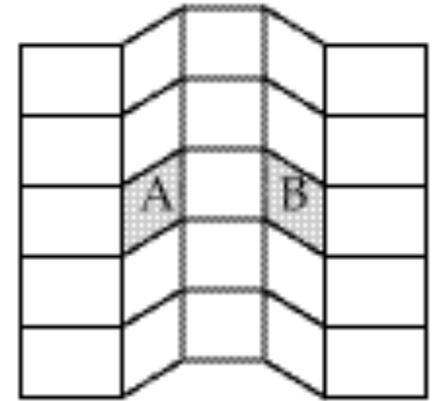
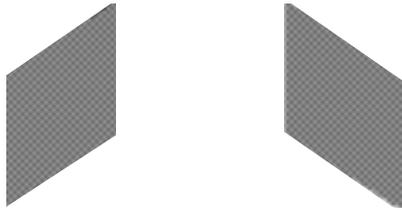
Bedeutungstragende Merkmale?

- Charakteristische Farben
 - Problem: Variation durch Beleuchtung und Spiegelungen
- Charakteristische Musterung (Textur)
 - Problem: Variation mit Perspektive (Distanz, Skalierung, Rotation, etc.)



Quelle: http://www.soval.de/gallery-pixeleye/v/Autos/roter_oldtimer.jpg.html

Bedeutungsträger Pixelhelligkeit?



© D.D. Hoffman, Visual Intelligence

Modell

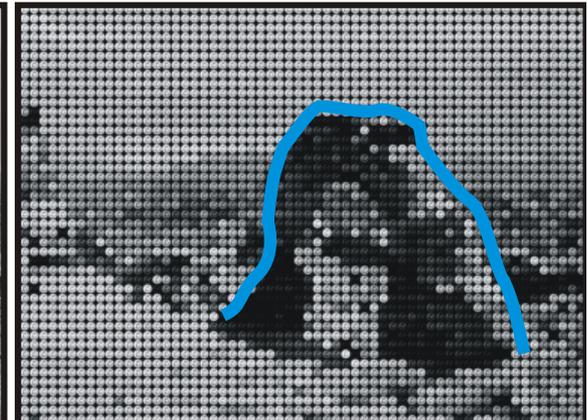
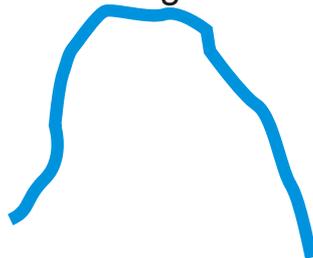
- Modell beschreibt a-priori Wissen über Bildinhalt
 - Erwartung möglicher Bildinhalte
 - abhängig von Anwendung bzw. Fragestellung
 - kann unvollständig oder falsch sein
- Aufgabe: Finde wahrscheinlichste Erklärung der Daten unter gegebenem Modell
- Hilft beim Sprung über die semantische Lücke

Bottom-Up vs. Top-Down-Strategie



von den Daten zum Modell

Modell des Bogens



vom Modell zu den Daten

Einbringen von Kontextwissen

- Top down (Modell-getrieben)
 - Abhängig vom Wissen im Langzeitgedächtnis und Aktivierung im Kurzzeitgedächtnis
- Bottom up (Daten-getrieben)
 - Abhängig von visuellen Attributen des Objekts



Visuelle Verarbeitung beim Menschen

- Visuelle „Rohdaten“ im ikonischen Speicher
 - Rohdaten analog zu physikalischen Stimuli
 - Sehr kurze Speicherzeit: 200 [90, 1000] ms
 - Zeitliche Auflösung: 100 [50, 200] ms (10 Hz [5,20 Hz])
- Symbolische Repräsentation im Kurzzeitgedächtnis (KZG)
 - Kognitive Operationen nur im KZG möglich
- Wissen im Langzeitgedächtnis (LZG)
- Unterscheidung von Figur und Hintergrund
 - Top-down Verarbeitung (Konzept-getrieben)
 - Abhängig vom Wissen im LZG und Aktivierung im KZG
 - Bottom-up Verarbeitung (Daten-getrieben)
 - Abhängig von visuellen Attributen des Objektes



Einbringen von Kontextwissen



Einbringen von Kontextwissen



Einbringen von Kontextwissen

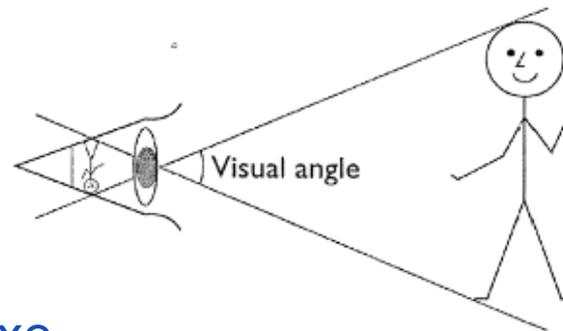
The quick brown
fox jumps over the
the lazy dog.

Wahrnehmung von Größe und Tiefe

- Konstanz wahrgenommener Größe trotz Änderung der Winkelgröße (visual angle)
 - Wissen über die Welt

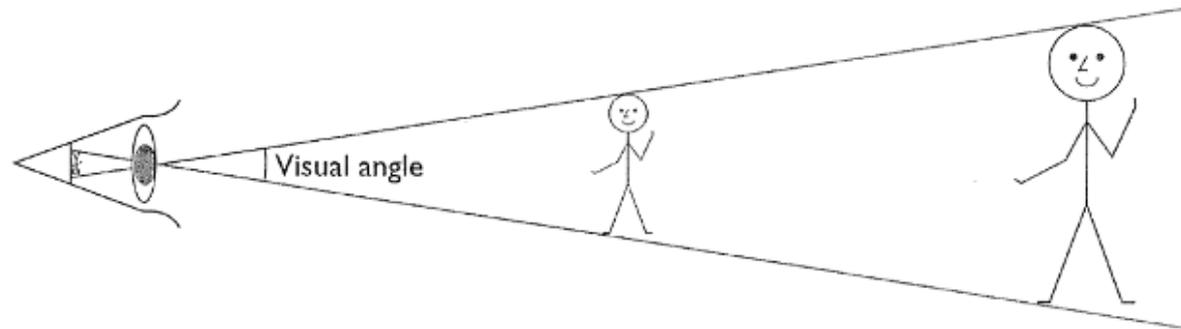
- Hinweise für Tiefe

- Verdeckung
- Textur
- Bewegungsparallaxe



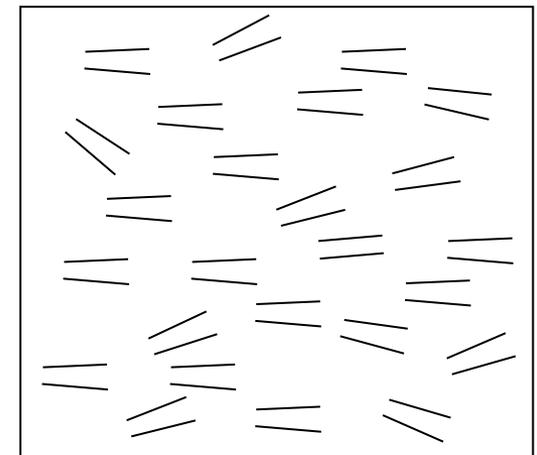
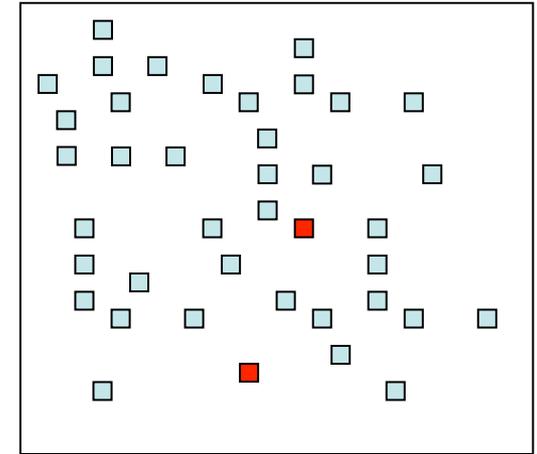
Objects of the same size at different distances have different visual angles

Objects of different sizes and different distances may have the same visual angle

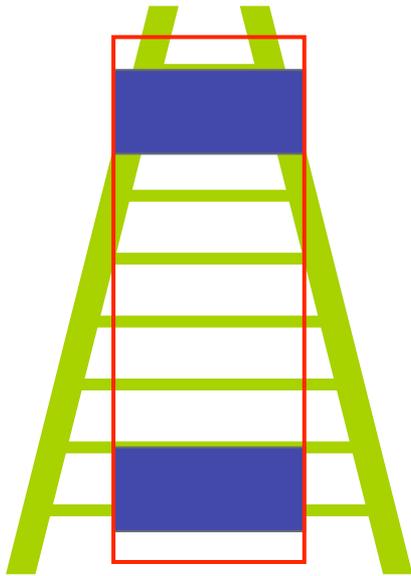


Visuelle Suche beim Mensch

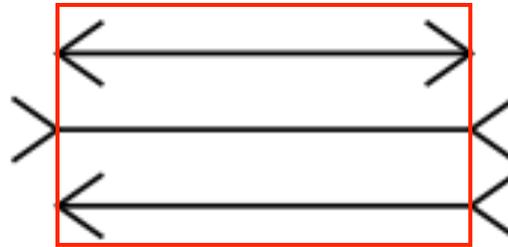
- Suchmenge = Zielreize + Ablenker
 - Ziele unterscheiden sich in einem Attribut von den Ablenkern
 - Beispiel: Suche nach einer Menüoption
- Automatische visuelle Suche (pop out)
 - „Unmittelbare“ Erkennung des Zielreizes
 - Suchzeit unabhängig von Suchmengengröße
 - Attribute: Farbe, Blinken, Bewegung
- Kontrollierte visuelle Suche
 - Serielle Suche durch die Suchmenge
 - Suchzeit proportional zur Suchmengengröße
 - Attribute: Parallelität



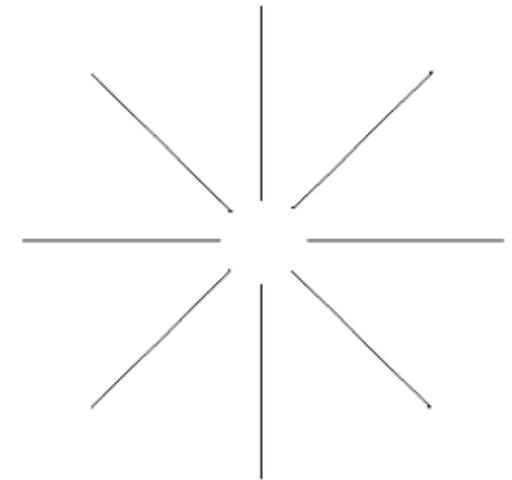
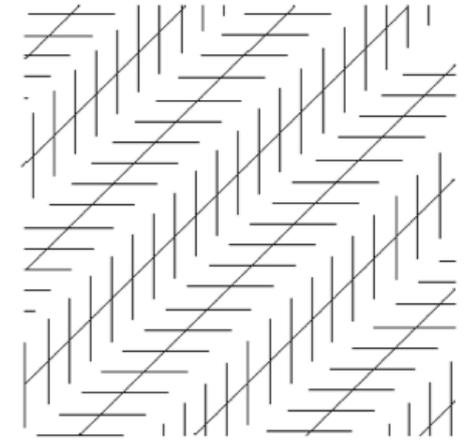
Wahrnehmung von Objekten



Ponzo-Illusion



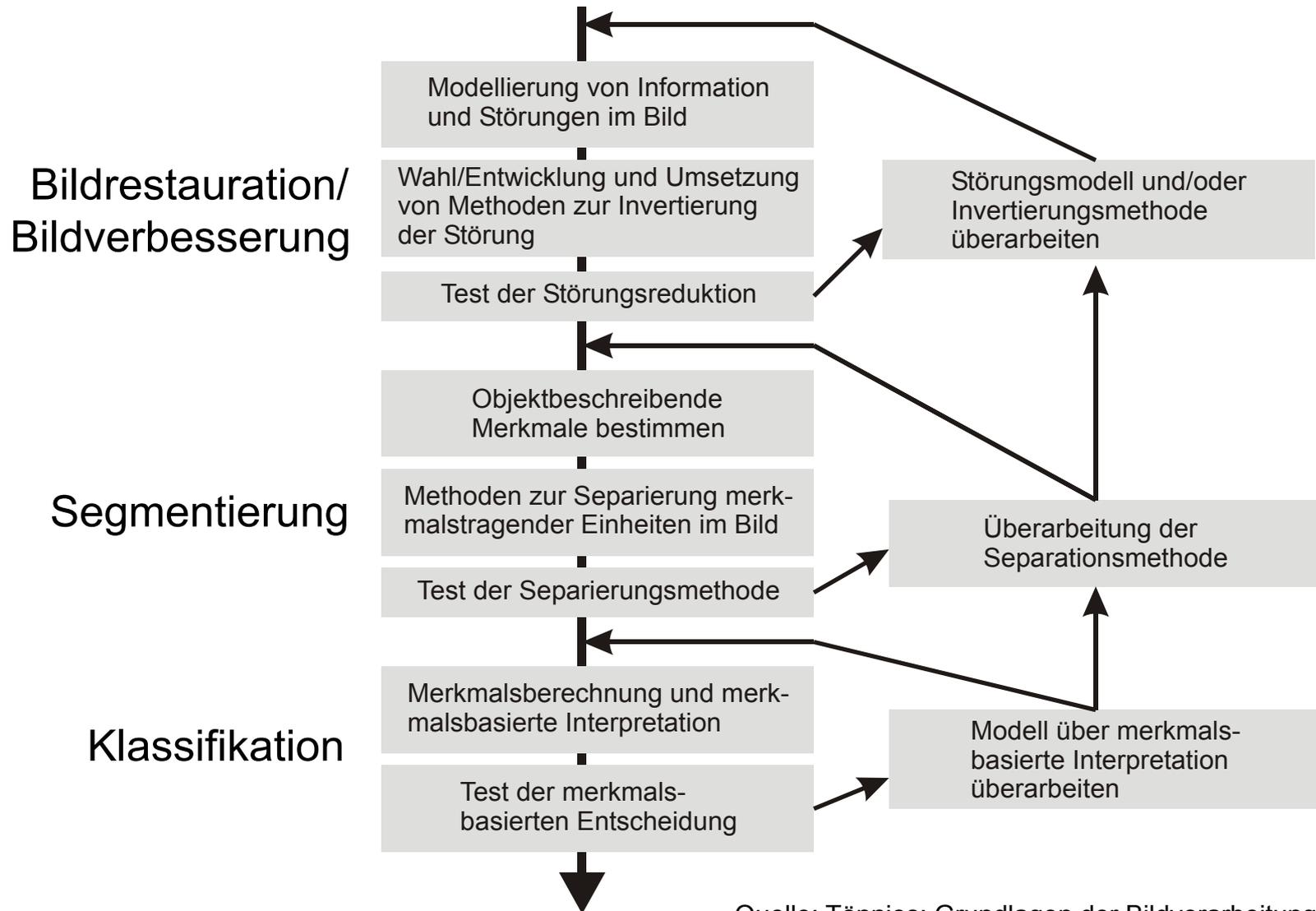
Müller-Lyer-Illusion



Leistungen des menschlichen visuellen Systems

- Größen- und Tiefenwahrnehmung
 - konstante Größe bekannter Objekte
- Farbwahrnehmung
 - konstante Farbe trotz variierender Lichtverhältnisse
- Szenenerkennung
 - schnelle Erkennung räumlicher Konfigurationen von Objekten
- „Bildstabilisierung“
 - Kompensation für die Bewegungen von Augen und Kopf

Prozesse der Bildverarbeitung



Quelle: Tönnies: Grundlagen der Bildverarbeitung.

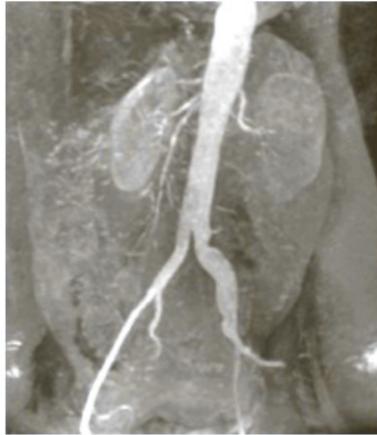
BEISPIELE & ANWENDUNGEN

Röntgenbilder

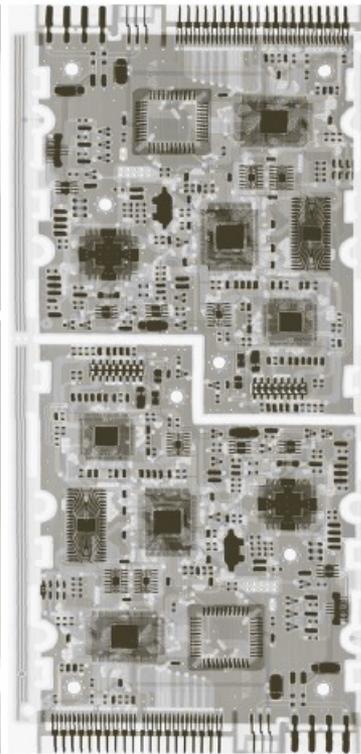
Röntgenbild



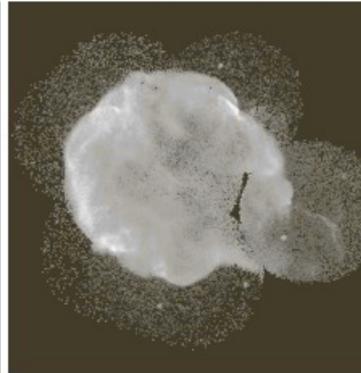
Angiografie



Computertomographie



Produkt-
prüfung



Astronomie

a	d
b	e
c	e

FIGURE 1.7 Examples of X-ray imaging. (a) Chest X-ray. (b) Aortic angiogram. (c) Head CT. (d) Circuit boards. (e) Cygnus Loop. (Images courtesy of (a) and (c) Dr. David R. Pickens, Dept. of Radiology & Radiological Sciences, Vanderbilt University Medical Center; (b) Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School; (d) Mr. Joseph E. Pascente, Lixi, Inc.; and (e) NASA.)

© R. C. Gonzalez & R. E. Woods, Digital Image Processing

Mikroskopie

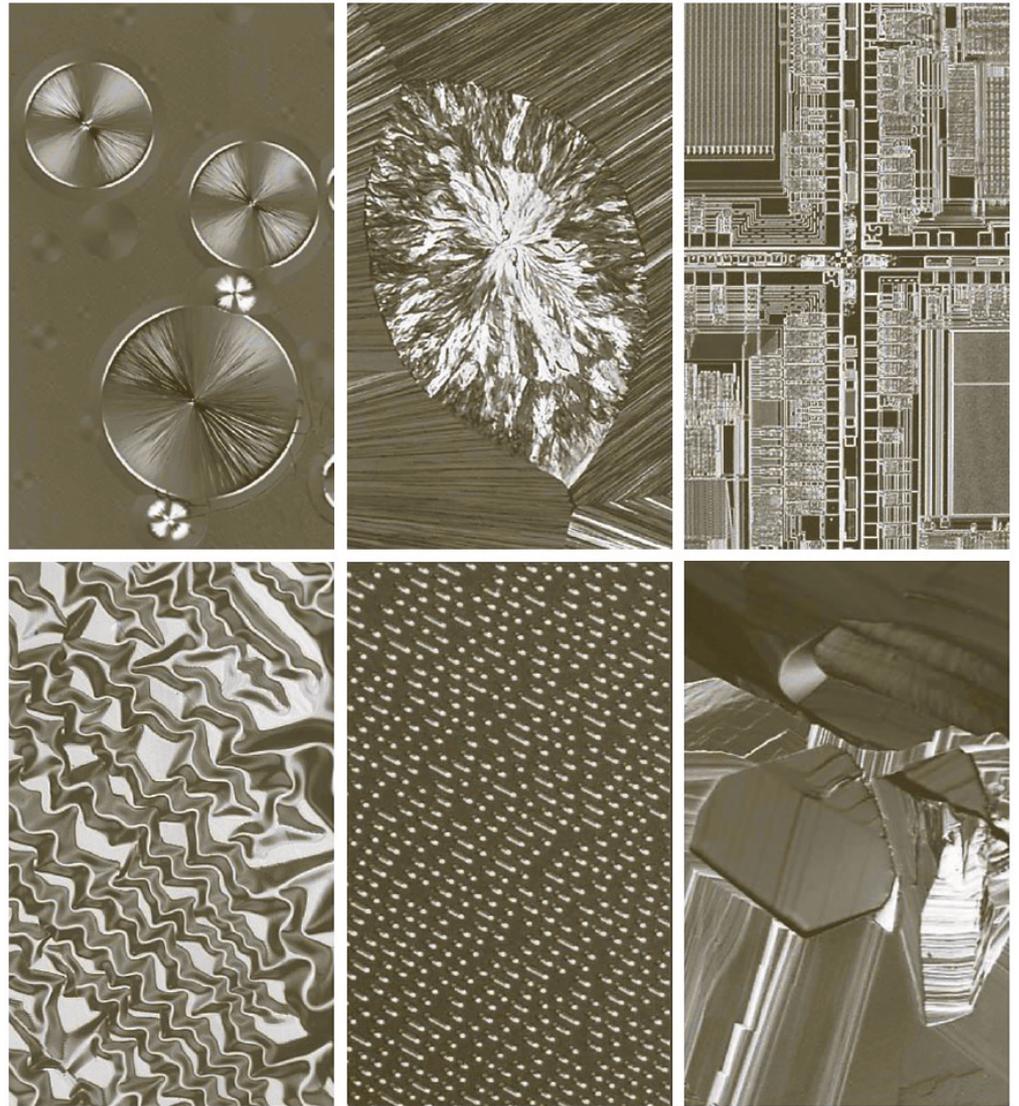


FIGURE 1.9 Examples of light microscopy images. (a) Taxol (anticancer agent), magnified 250 \times . (b) Cholesterol—40 \times . (c) Microprocessor—60 \times . (d) Nickel oxide thin film—600 \times . (e) Surface of audio CD—1750 \times . (f) Organic superconductor—450 \times . (Images courtesy of Dr. Michael W. Davidson, Florida State University.)

© R. C. Gonzalez & R. E. Woods, Digital Image Processing

Satellitenbilder zur Erdbeobachtung

- Verschiedene Wellenlängenbereiche sind für unterschiedliche Zwecke geeignet

1. 450-520nm (blau)
2. 520-600nm (grün)
3. 630-690nm (rot)
4. 760-900nm (nah-IR)
5. 1550-1750nm (IR)
6. 1040-1250 (IR)
7. 2080-2350nm(IR)

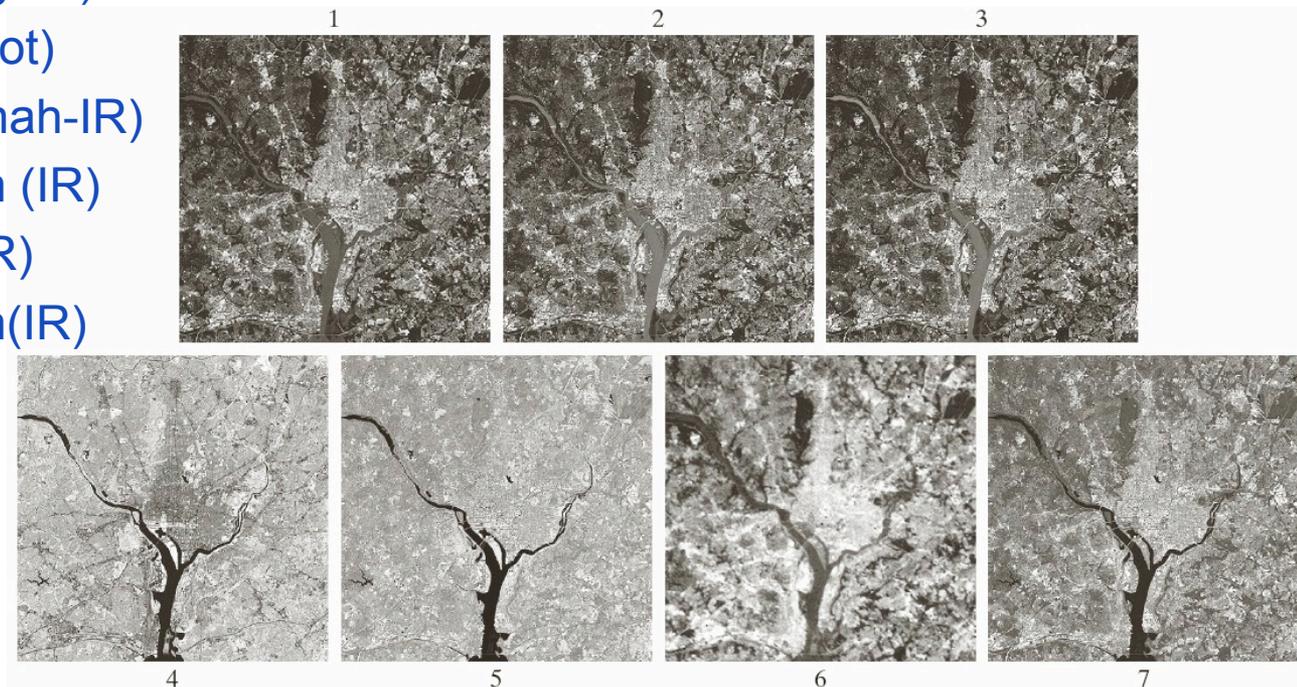


FIGURE 1.10 LANDSAT satellite images of the Washington, D.C. area. The numbers refer to the thematic bands in Table 1.1. (Images courtesy of NASA.)

© R. C. Gonzalez & R. E. Woods, Digital Image Processing

Satellitenbilder zur Wetterbeobachtung

- Wirbelsturm
(Katrina, 2005)

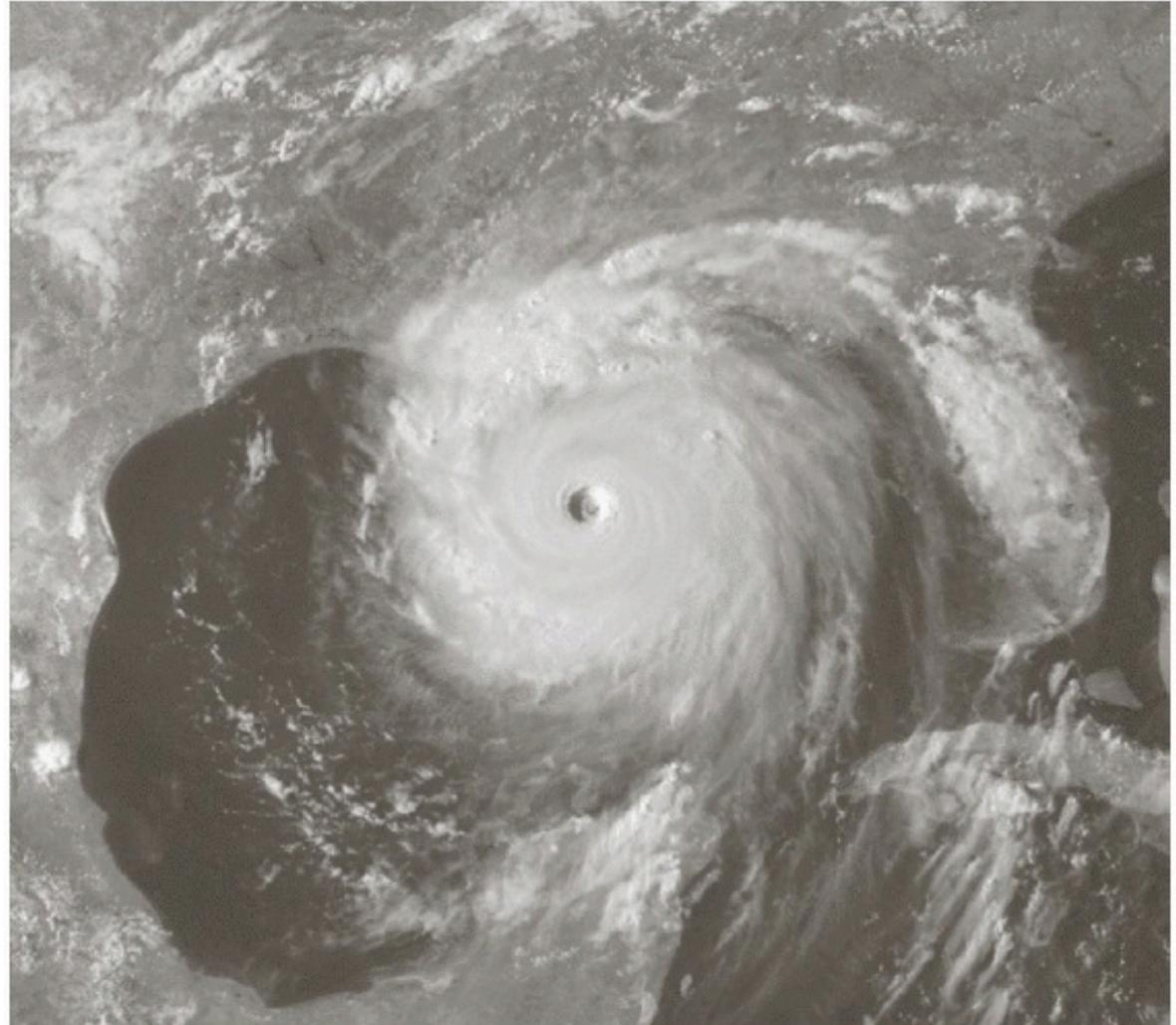


FIGURE 1.11
Satellite image
of Hurricane
Katrina taken on
August 29, 2005.
(Courtesy of
NOAA.)

© R. C. Gonzalez & R. E. Woods, Digital Image Processing

Population, Energieverbrauch, etc.



FIGURE 1.12

Infrared satellite images of the Americas. The small gray map is provided for reference.

(Courtesy of NOAA.)

Population, Energieverbrauch, etc.

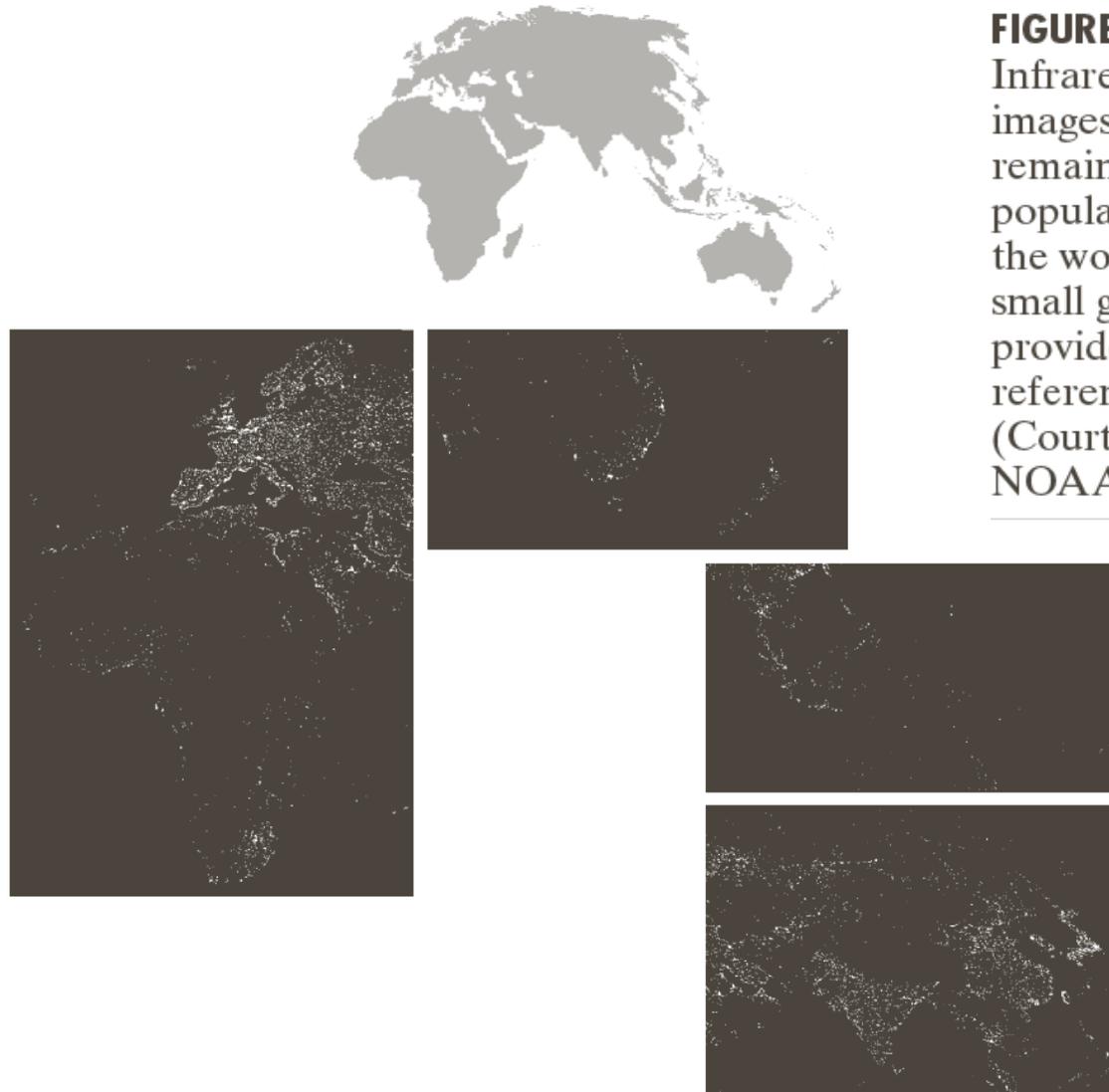
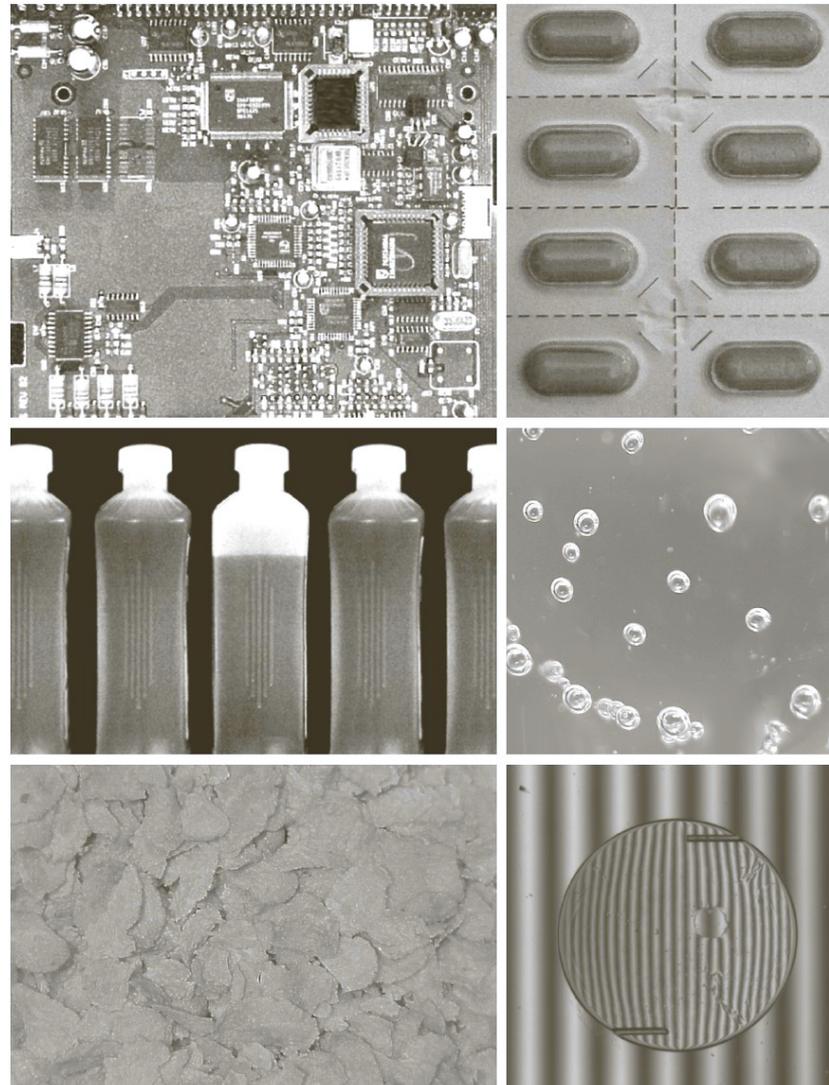


FIGURE 1.13

Infrared satellite images of the remaining populated part of the world. The small gray map is provided for reference.

(Courtesy of NOAA.)

Qualitätskontrolle von Produkten



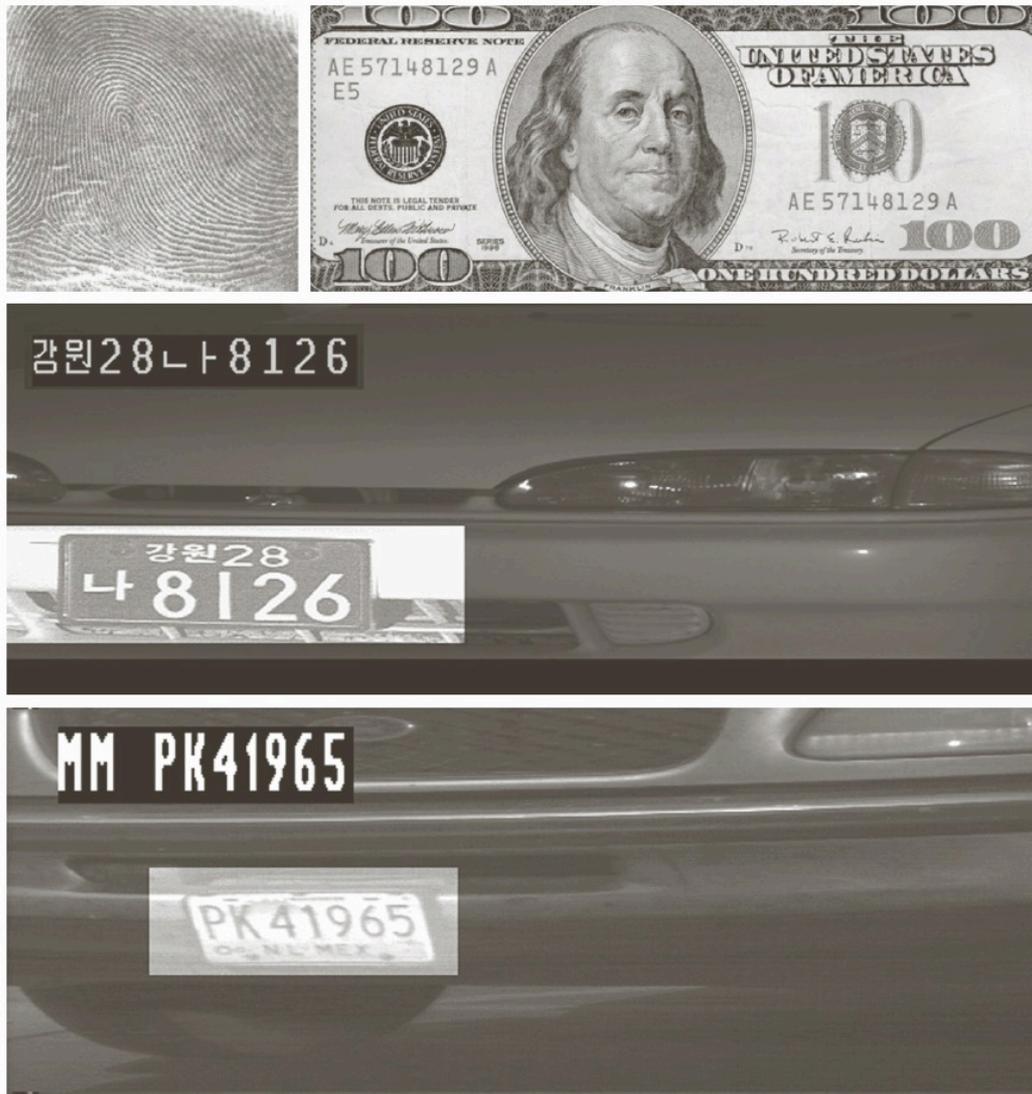
a	b
c	d
e	f

FIGURE 1.14

Some examples of manufactured goods often checked using digital image processing.

- (a) A circuit board controller.
 - (b) Packaged pills.
 - (c) Bottles.
 - (d) Air bubbles in a clear-plastic product.
 - (e) Cereal.
 - (f) Image of intraocular implant.
- (Fig. (f) courtesy of Mr. Pete Sites, Perceptics Corporation.)

Erkennung von Symbolen und Mustern



a b
c
d

FIGURE 1.15 Some additional examples of imaging in the visual spectrum. (a) Thumb print. (b) Paper currency. (c) and (d) Automated license plate reading. (Figure (a) courtesy of the National Institute of Standards and Technology. Figures (c) and (d) courtesy of Dr. Juan Herrera, Perceptics Corporation.)

Astronomie

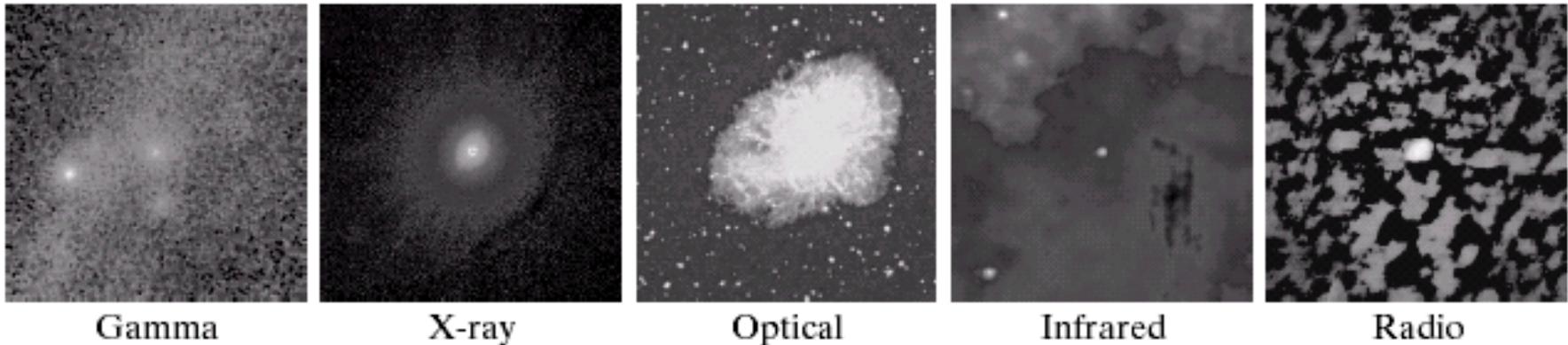


FIGURE 1.18 Images of the Crab Pulsar (in the center of images) covering the electromagnetic spectrum. (Courtesy of NASA.)

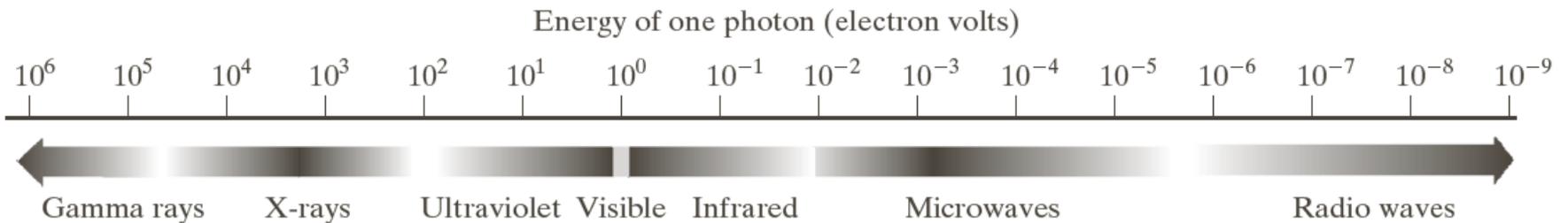
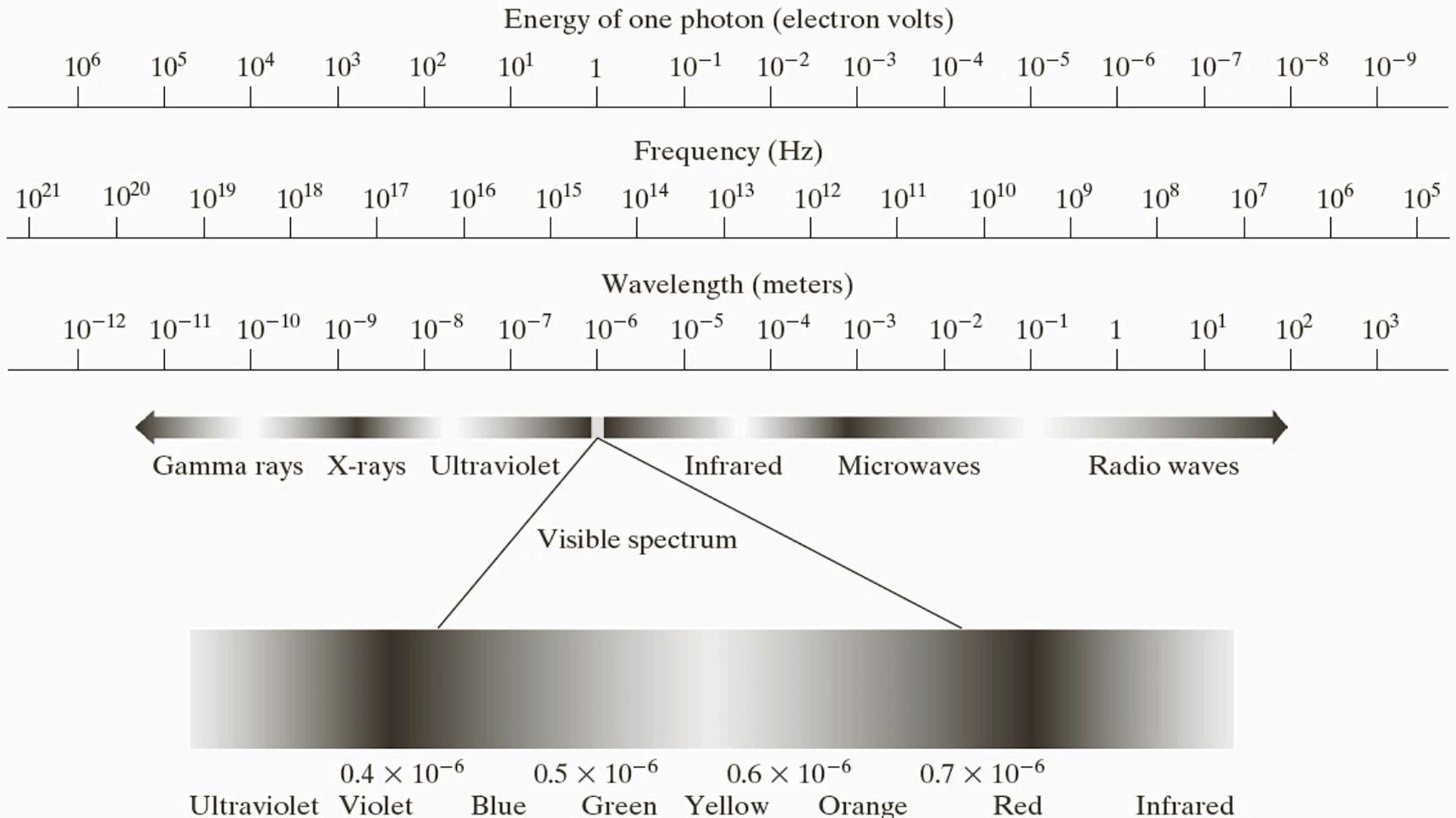


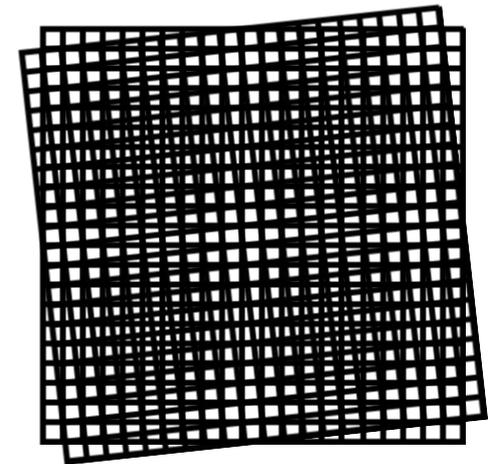
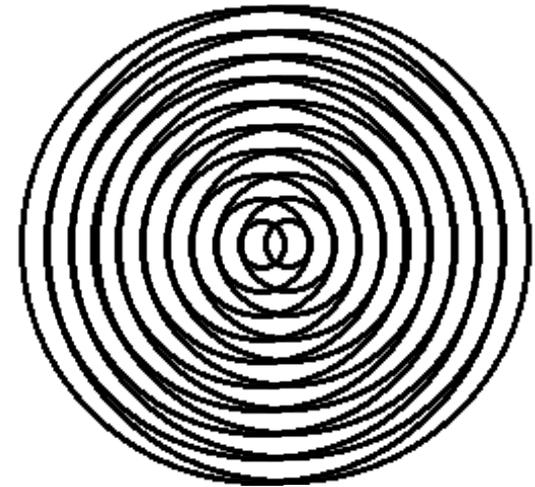
FIGURE 1.5 The electromagnetic spectrum arranged according to energy per photon.

Elektromagnetisches Spektrum

- Sichtbares Licht nur kleiner teil des EM Spektrums



Moiré Muster vermindern

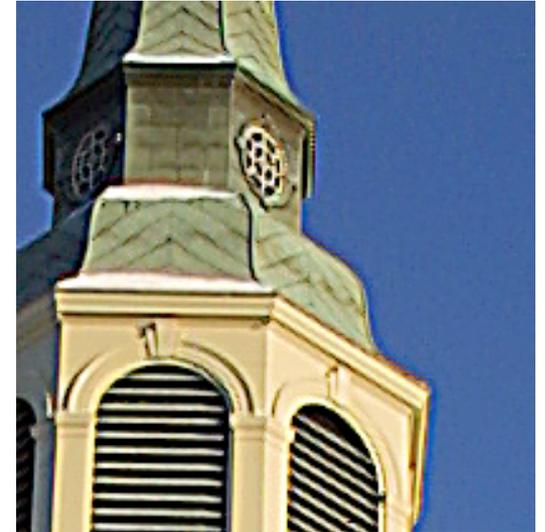
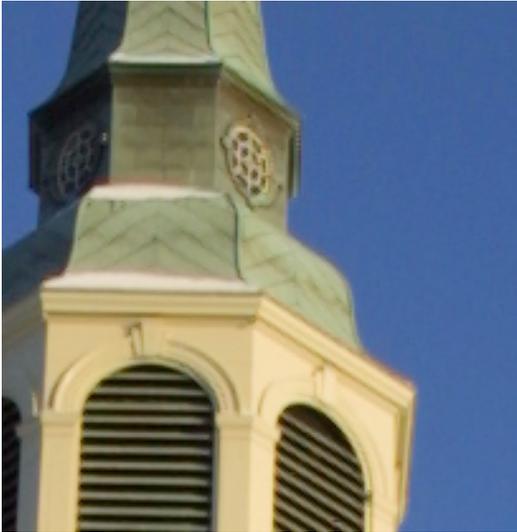


Quelle: <http://www.daube.ch/docu/glossary/moiree.html>

Rauschen reduzieren

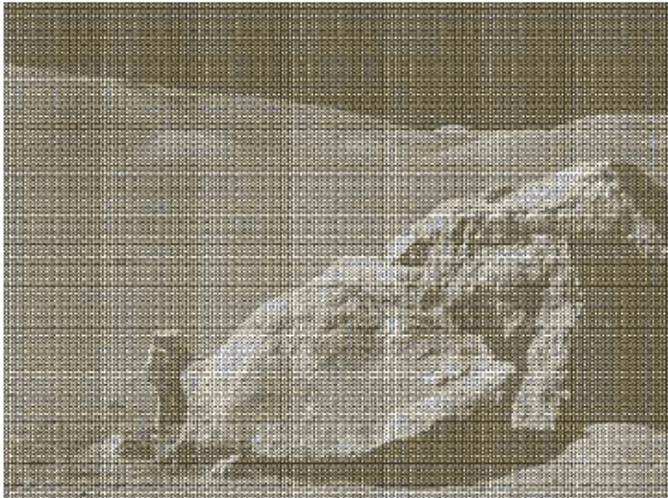


Schärfe erhöhen



- Grundidee: Lokalen Kontrast erhöhen
- Verschiedene Methoden
- Automatisch in Digitalkameras

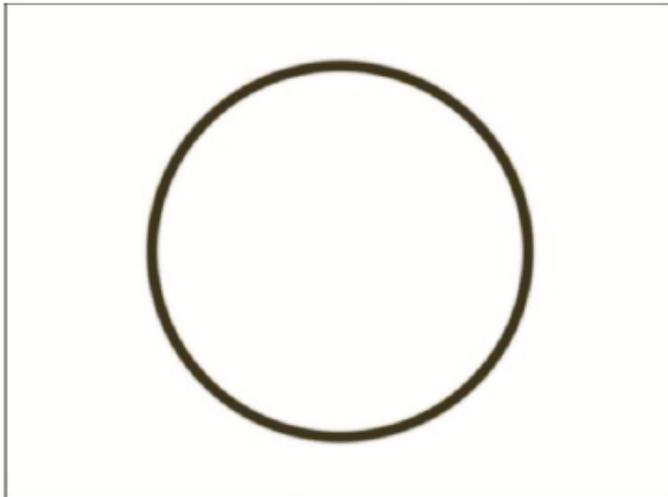
Artefakte entfernen



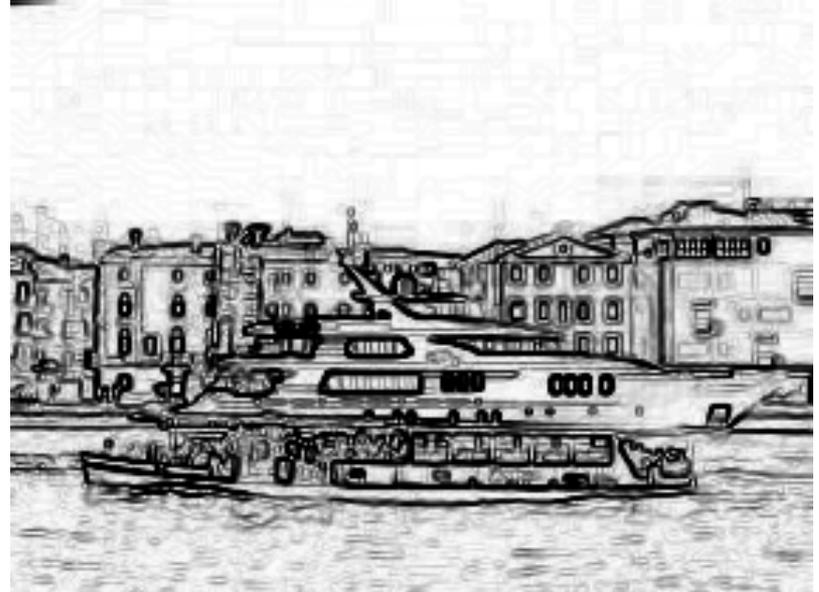
a b
c d

FIGURE 2.40

(a) Image corrupted by sinusoidal interference. (b) Magnitude of the Fourier transform showing the bursts of energy responsible for the interference. (c) Mask used to eliminate the energy bursts. (d) Result of computing the inverse of the modified Fourier transform. (Original image courtesy of NASA.)



Kantenerkennung



- Grundidee: Finden starker Intensitäts- oder Farbänderungen
- Basis für Objekterkennung

Morphologische Operatoren

Test 1 2 3

Test 1 2 3

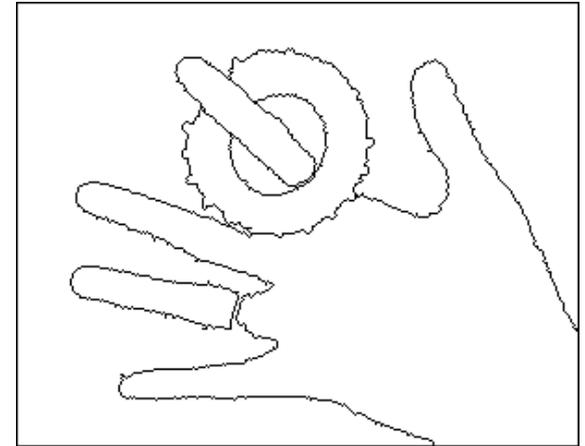
Test 1 2 3

Test 1 2 3

Test 1 2 3

- Boolesche Operationen zwischen Pixeln im Bild und Pixel eines „strukturierenden Elements“
- „Erosion“ und „Dilatation“ von Bereichen
- Schließen von Löchern
- Beseitigung von Störungen

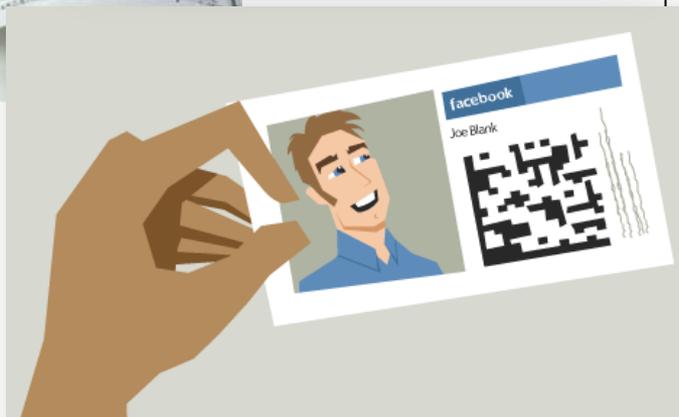
Segmentierung



- Gruppierung von Pixeln zu größeren Einheiten
- keine Annahmen über Bedeutung (bottom up)
- vollständig und überdeckungsfrei
- Grundlage für die Erkennung von Formen

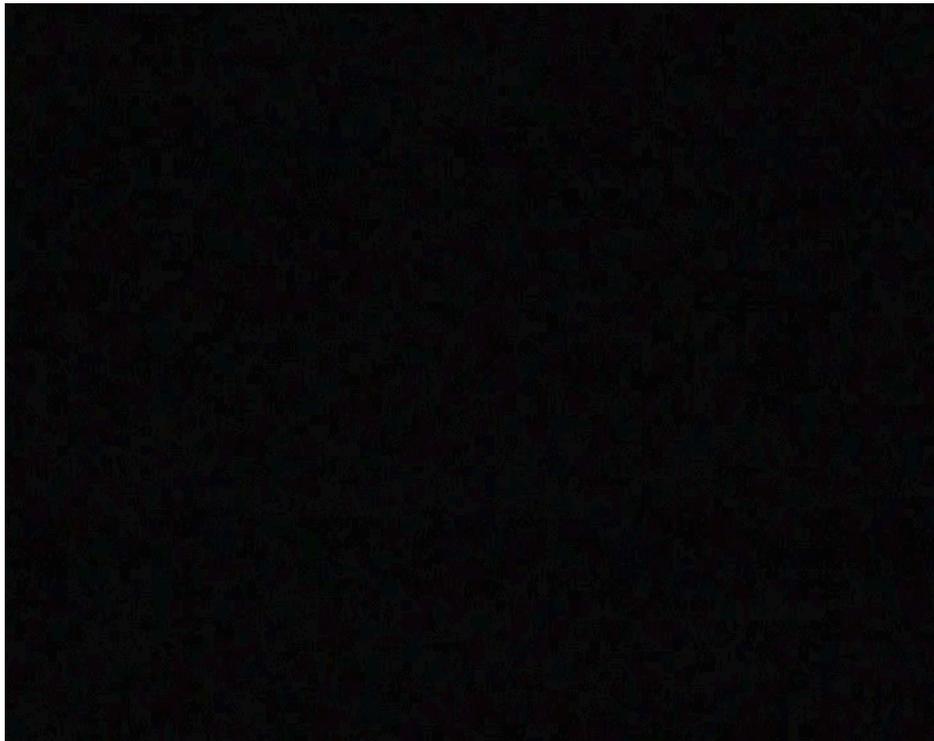
Mobile Tagging: QR Codes

Newspapers, business cards, coupon flyers, etc.



Visual Code Widgets

- Printable user interface elements
 - Embedded in user's environment
 - Camera phone as “see-through tool“



Handheld Augmented Reality Games on Cereal Boxes: “Penalty Kick”

Augmented Reality Games on Product Packages

Michael Rohs, Jean-Daniel Merkli
Institute for Pervasive Computing
© 2005 ETH Zurich, Switzerland



Server-based Image Recognition



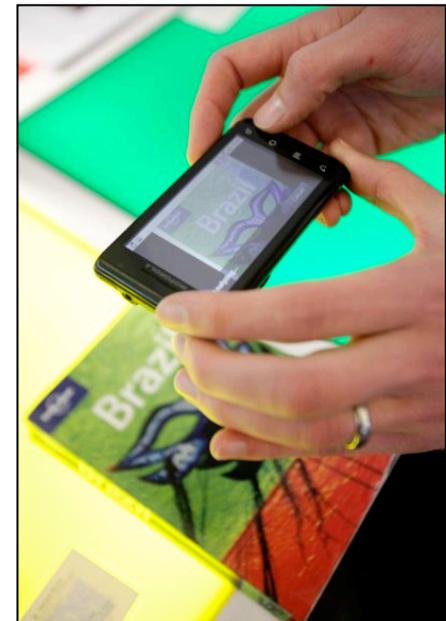
Source: Rahul Swaminathan, T-Labs

Google Goggles

- Visual search queries for the Web

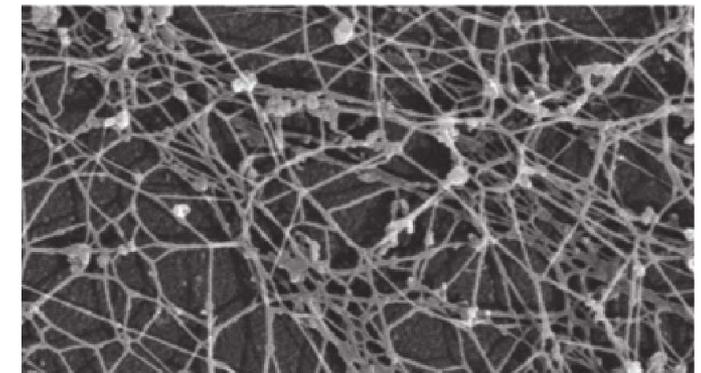
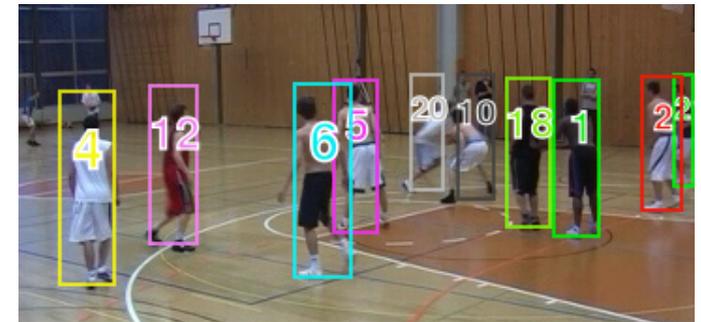
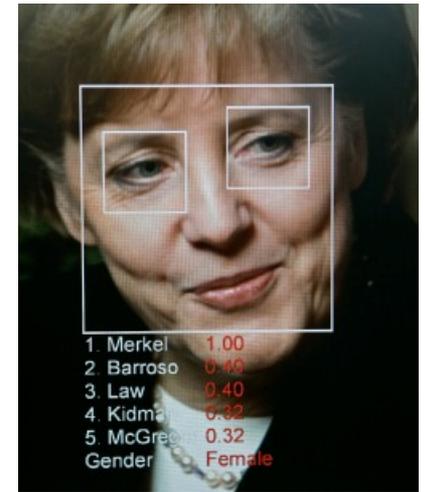


- Recognizes wide range of artifacts
 - Landmarks, barcodes, logos, book covers, etc.
- www.google.com/mobile/goggles



und vieles mehr...

- Gesichtserkennung
- Tracking von Personen, Überwachung
- Schrift- und Zeichenerkennung
- Erkennung Objekt/Szene
- Erkennung von Tumorzellen



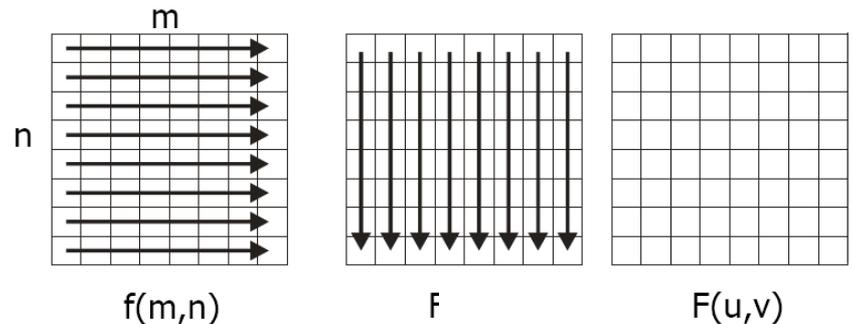
VORLESUNG & ÜBUNG

Vorlesungen

Datum	Thema
24.4.	Einführung, Organisatorisches (Übungen, Klausur)
1.5./8.5.	keine Vorlesungen (wegen 1. Mai und CHI-Konferenz)
15.5.	Abtastung von Bildern, Punktbasierte Verfahren der Bildverbesserung
22.5.	Licht, Farbe, Farbmanagement
30.5.	Kontrastverbesserung, Konvolution, Filterung im Ortsraum (Verschiebung wegen Pfingstdienstag)
5.6.	Fouriertransformation: Grundlagen
12.6.	Filterung im Frequenzraum
19.6.	Kanten, Linien, Ecken
26.6.	Segmentierung
3.7.	Segmentierung, Morphologische Operationen
10.7.	Klassifikation
17.7.	Image Matching
23.-27.7.	Klausur

Separierbarkeit

$$\begin{aligned}
 F(u, v) &= \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \exp\left[-i \frac{2\pi}{N} (um + vn)\right] = \\
 &= \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \exp\left[-i \frac{2\pi}{N} um\right] \exp\left[-i \frac{2\pi}{N} vn\right] = \\
 &= \frac{1}{N} \sum_{m=0}^{N-1} \left(\frac{1}{N} \sum_{n=0}^{N-1} f(m, n) \exp\left[-i \frac{2\pi}{N} vn\right] \right) \exp\left[-i \frac{2\pi}{N} um\right] = \\
 &= \frac{1}{N} \sum_{m=0}^{N-1} \exp\left[-i \frac{2\pi}{N} um\right] \left(\frac{1}{N} \sum_{n=0}^{N-1} f(m, n) \exp\left[-i \frac{2\pi}{N} vn\right] \right) = \\
 &= \frac{1}{N} \sum_{m=0}^{N-1} \exp\left[-i \frac{2\pi}{N} um\right] F_u(m)
 \end{aligned}$$



- Vorgehensweise: $F_u(m)$ für alle Spalten m berechnen und dann bei den Zeilen verwenden.

Übungen

- Praktische Umsetzung der Konzepte aus der Vorlesung
 - Praktische Erfahrung mit grundlegenden Algorithmen der Bildverarbeitung
- Ort und Zeit
 - Mittwoch 14-16 Uhr, Donnerstag 14-16 Uhr
 - Theresienstraße 39, Raum B 133
- Tools
 - Python: docs.python.org/tutorial/index.html
 - NumPy: www.scipy.org/Tentative_NumPy_Tutorial
 - matplotlib: matplotlib.sourceforge.net/index.html

Numpy Beispiel: FT anzeigen

```
import numpy as np
import matplotlib.pyplot as pp

a = pp.imread('/.../ente.png')

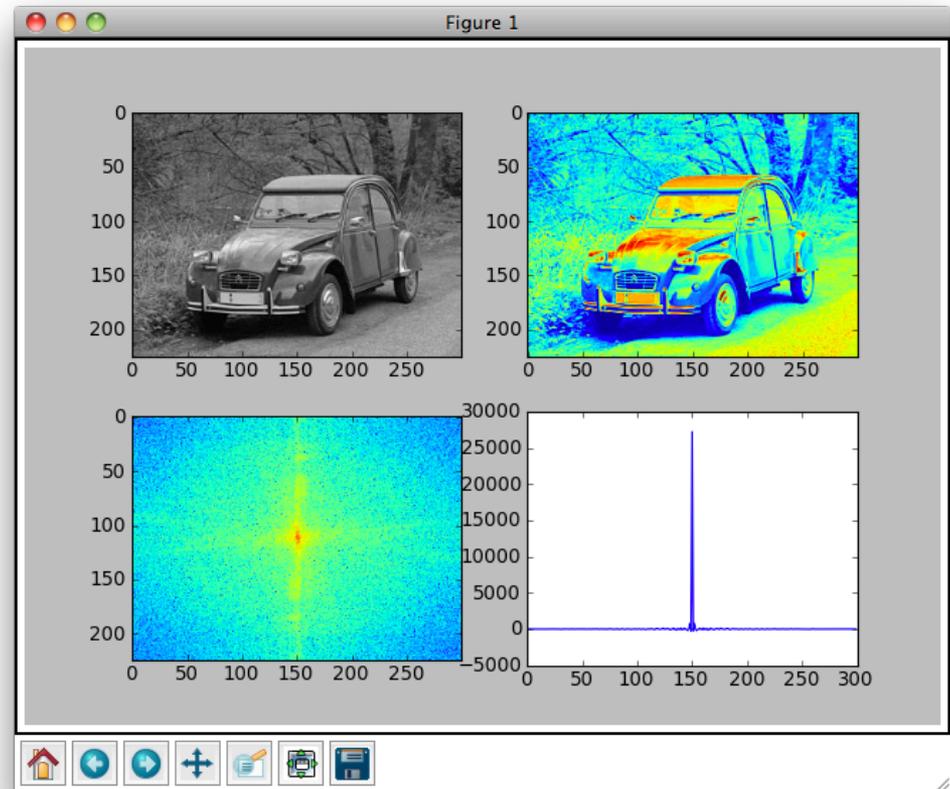
pp.figure(1)
pp.subplot(221)
pp.imshow(a, cmap=pp.get_cmap('gray'))
pp.subplot(222)
pp.imshow(a)

b = np.fft.fft2(a)
b = np.fft.fftshift(b)

pp.subplot(223)
pp.imshow(np.log(np.abs(b)))

s = np.shape(b)
x = np.arange(0,s[1])
y = b[s[0]/2, x]
pp.subplot(224)
pp.plot(x, y)

pp.show()
```



Credit Points

- Umfang: 2 SWS Vorlesung, 2 SWS Übung
 - ECTS-Credits: 4 (Vorlesung) + 2 (Übung)
 - Modul: WP10: Computergrafik II (für Master Medieninformatik)
- wöchentliche Übungen
 - Teilnahme freiwillig
 - Bearbeitung zu 2-3 Studierenden
 - empfohlen als Vorbereitung auf die Klausur
- Klausur
 - Stoff ist Inhalt der Vorlesung und der Übung

Klausur

- am Ende des Semesters bzw. Woche danach
- Stoff ist Inhalt der Vorlesung **und** der Übung
- geprüft wird, ob das Handwerkszeug sitzt
 - Verständnis, nicht auswendig lernen!
 - Daumenregel: Wer die Übungsblätter geschafft hat, braucht die Klausur nicht zu fürchten
- Wiederholungsklausur wie vorgeschrieben
 - vermutlich kurz vor Beginn des folgenden Semesters

Service

- Vorlesungsfolien vor der Vorlesung im Web
 - www.medien.ifi.lmu.de/lehre/ss12/cg2
- Übungsblätter und Folien zur Übung im Web
- UniWorX: Computergrafik 2
 - <https://uniworx.ifi.lmu.de/?action=uniworxCourseWelcome&id=59>
- Austausch/Diskussion
 - <http://www.die-informatiker.net/forum/Medieninformatik>
- Sprechstunde
 - Nach der Vorlesung oder
 - Dienstag 13-14 Uhr

Nicht Bestandteil der Vorlesung...

- 2D & 3D Vektorgrafik
 - Prof. Butz: Computergrafik 1
- Bildkomprimierung
 - Prof. Hussmann: „Medientechnik“ + „Digitale Medien“
- 3D Computer Vision
 - eigenes Forschungsfeld

Literatur

- Klaus D. Tönnies: Grundlagen der Bildverarbeitung.
Pearson Studium, 2005.
ISBN 3-8273-7155-4
- R.C. Gonzalez, R.E. Woods:
Digital Image Processing.
Prentice Hall, 3rd ed., 2007.
ISBN 0-13-168728-X

